

## 植物乳杆菌编码复制起始蛋白天然质粒的系统进化分析

孙大庆\* 李洪飞 杨健

(黑龙江八一农垦大学国家杂粮工程技术研究中心 黑龙江 大庆 163319)

**摘要:**【目的】探究复制起始蛋白(Replication initiation protein, Rep)是否可以作为天然质粒系统进化关系研究的分子标记。【方法】以植物乳杆菌天然质粒编码的 Rep 为研究对象,通过构建 Rep 系统进化树详细分析和讨论这些质粒的系统进化关系。【结果】植物乳杆菌 45 个编码 Rep 天然质粒可以划分为 5 个进化关系紧密的家族和 1 个独立进化质粒 pG6302, 家族 1–4 质粒可以进一步划分为 10 个进化关系更近的亚家族类群,因此这些质粒可能起源于 6 个祖先质粒。【结论】Rep 氨基酸序列显示了适度的保守性和变异性,是植物乳杆菌编码 Rep 质粒进化研究理想的分子标记,为植物乳杆菌天然质粒系统进化研究提供了一种简单、有效的分析方法和标准,并为植物乳杆菌或其他乳酸菌天然质粒系统进化研究提供了分子水平的佐证和依据。

**关键词:** 植物乳杆菌, 质粒, 系统进化树, 复制起始蛋白

## Phylogenetic analysis of natural plasmids encoding the replication initiation protein in *Lactobacillus plantarum*

SUN Da-Qing\* LI Hong-Fei YANG Jian

(National Coarse Cereals Engineering Research Center, Heilongjiang Bayi Agricultural University, Daqing, Heilongjiang 163319, China)

**Abstract:** [Objective] In order to explore whether replication initiation protein (Rep) can be used as a molecular biomarker for the phylogenetic relationship study of natural plasmids. [Methods] The Reps of natural plasmids of *Lactobacillus plantarum* as a target, the phylogenetic relationships of these plasmids were analyzed and discussed in detail by constructing Rep phylogenetic tree. [Results] A total of 45 plasmids encoding Rep in *Lactobacillus plantarum* could be clustered into 5 closely related families and 1 single plasmid pG6302, of which family 1 to 4 plasmids could be further subdivided into 10 subfamily groups with closer evolutionary relationships, suggesting that these plasmids may originate from the 6 ancestral plasmids. [Conclusion] Since the amino acid sequences of Reps show suitable conservation and variability, it may be an ideal molecular biomarker for phylogenetic relationship study of plasmids encoding Rep in *Lactobacillus plantarum*. The results

**Foundation item:** Youth Science Foundation of Heilongjiang Province (No. QC2014C020)

\*Corresponding author: E-mail: sundaqing1979@163.com

**Received:** December 22, 2016; **Accepted:** February 23, 2017; **Published online** (www.cnki.net): February 27, 2017  
基金项目: 黑龙江省青年科学基金项目(No. QC2014C020)

\*通讯作者: E-mail: sundaqing1979@163.com

收稿日期: 2016-12-22; 接受日期: 2017-02-23; 优先数字出版日期(www.cnki.net): 2017-02-27

could provide a simple and effective method and standard, as well as the evidence and basis at the molecular level for phylogenetic evolution study of natural plasmids encoding Rep in *Lactobacillus plantarum* or other lactic acid bacteria.

**Keywords:** *Lactobacillus plantarum*, Plasmid, Phylogenetic tree, Replication initiation protein

乳杆菌属是乳酸菌中菌种多样性十分丰富的群体,目前已鉴定菌种数量达 221 个(<http://www.bacterio.net/lactobacillus.html>),它们广泛分布于动物、植物、人类及食品等碳水化合物丰富的环境中。然而由于基因丢失和环境选择压力,乳杆菌中许多菌种属于高度特异进化的物种,它们只在有限的生境中存活,例如德氏乳杆菌已高度适应乳环境;嗜酸乳杆菌、约氏乳杆菌、罗伊氏乳杆菌和鼠李糖乳杆菌则是典型的胃肠道“居民”。与此相反,植物乳杆菌在自然界分布广泛,对多种不同的生境具有高度适应性<sup>[1]</sup>,这可能主要依赖于植物乳杆菌具有乳杆菌属中较大的基因组;另一方面,我们经过统计和整理分析发现,植物乳杆菌是乳杆菌属中天然质粒最丰富和多样的菌种,是植物乳杆菌染色体外天然基因库,其中不乏多种重要的功能基因,例如耐药基因、细菌素和胞外多糖合成基因<sup>[2]</sup>。因此,丰富的天然质粒在植物乳杆菌生长、发育、遗传和进化过程中可能同样扮演着十分重要的角色。

以往乳杆菌天然质粒的研究主要集中于功能基因元件的分离和鉴定<sup>[3-4]</sup>及克隆和表达载体构建等<sup>[5-7]</sup>,而关于质粒系统进化的研究很少,已有的研究仅局限于个别质粒的进化研究<sup>[8]</sup>。近年,随着 DNA 测序技术的飞速发展,乳杆菌质粒基因组序列数据快速增加。目前,NCBI (National Center for Biotechnology Information)的 RefSeq 数据库中收录乳杆菌属质粒完整基因组序列 253 个,其中植物乳杆菌质粒 84 个。据统计,植物乳杆菌不仅是乳杆菌属中天然质粒最多的菌种,也是乳酸菌中天然质粒最丰富的菌种之一,因此植物乳杆菌质粒是乳杆菌属或乳酸菌天然质粒遗传进化研究十分理想的研究群体。本研究通过构建编码 Rep 质粒进化树,详细分析和讨论了植物乳杆菌天然质粒系统进化关系,以及可能经历的进化过程和起源,旨在为植

物乳杆菌或其他乳酸菌天然质粒的系统进化研究提供一种新的研究思路和方法。

## 1 材料与方法

### 1.1 材料

NCBI RefSeq 数据库收录的 84 个植物乳杆菌质粒完整基因组(数据截止到 2016 年 8 月 30 日),金黄色葡萄球菌质粒 pC194 基因组(NC\_002013.1)和 pAD1 基因组<sup>[9]</sup>,无乳链球菌质粒 pMV158 基因组(NC\_010096.1)和 pIP501 基因组<sup>[10]</sup>,嗜盐四联球菌质粒 pUCL287 基因组(X75607.1)。

### 1.2 方法

**1.2.1 质粒 Rep 查找、统计和分析:**在 NCBI 的 RefSeq 数据库中检索植物乳杆菌质粒,在每个质粒的基因组注释信息中人工查找 Rep,下载质粒编码 Rep 氨基酸序列,对每个 Rep 进行 BLAST 比对分析,鉴定其保守结构域,确定其所属蛋白家族。

**1.2.2 编码 Rep 质粒系统进化分析:**利用 DNAMAN 5.5 软件进行多序列同源性分析。利用 MEGA 7.0<sup>[11]</sup>软件构建 Rep 系统进化树,系统进化树均采用 Neighbor-Joining 模型,并采用 Bootstrap 1 000 计算置信度。

## 2 结果与分析

### 2.1 植物乳杆菌质粒基因组基本特征

经序列比对和统计分析,NCBI RefSeq 数据库收录 84 个植物乳杆菌质粒基因组序列。其中 17 个质粒基因组中存在大片段(>10 kb)重复区域,推测很可能存在基因组拼接错误,因此舍弃,不做后续分析。其余 67 个质粒基因组基本特征及编码 Rep 情况见表 1。由表 1 可知,植物乳杆菌 67 个质粒基因组大小为 1.80–74.17 kb,G+C%含量为 33.37%–42.56%,其中 45 个质粒编码 Rep,质粒 pMR15.2 和 pZL3 编码 2 个 Rep,22 个质粒没有编码已知的 Rep。

表 1 植物乳杆菌质粒基因组基本特征  
Table 1 General characteristics of plasmid genome in *Lactobacillus plantarum*

序号 No.	质粒名称 Plasmid name	大小 Size (kb)	G+C 含量 G+C content (%)	Rep 大小 Rep size (aa)	序号 No.	质粒名称 Plasmid name	大小 Size (kb)	G+C 含量 G+C content (%)	Rep 大小 Rep size (aa)
1	pLPI8	1.80	37.49	237	35	pMD5057	10.87	36.22	311
2	pWCFS101	1.91	39.49	319	36	pCMPG5300.02	10.94	38.59	309
3	pCAUH201	1.94	37.81	317	37	pCAUH203	11.32	36.97	N
4	pLP2000	2.06	38.23	317	38	LZ206p3	11.65	34.29	N
5	pLR1	2.06	37.75	318	39	pZL3	11.89	37.66	304; 311
6	p200	2.06	38.22	317	40	ZJ95p2	12.48	36.37	N
7	pLD1	2.11	37.78	318	41	pZL4	12.51	39.38	311
8	pLP2111	2.11	38.32	317	42	Lp16I	13.34	37.91	307
9	pCMPG5300.01	2.12	38.03	317	43	pLP-ZJ101	15.16	40.17	311
10	pC30i1	2.14	36.96	317	44	LBPp7	15.17	39.68	N
11	pLP2140	2.14	38.27	N	45	LBPp5	16.10	42.14	155
12	pLTK2	2.29	38.56	317	46	Lp16C	27.28	39.86	366
13	pWCFS102	2.36	34.33	219	47	pCMPG5300.03	27.83	40.55	373
14	pA1	2.82	34.96	196	48	LZ206p2	33.00	39.93	367
15	pPB1	2.89	37.74	208	49	pLTK13	34.52	39.04	102
16	pXY3	2.96	38.24	219	50	LZ227p3	35.46	39.61	102
17	pR18	3.21	35.81	317	51	pWCFS103	36.06	40.83	102
18	pZL2	3.26	38.24	318	52	LBPp4	37.04	41.24	N
19	pM4	3.32	38.73	318	53	Lp16D	37.09	40.79	N
20	pG6301	3.51	37.26	318	54	LZ227p5	38.21	39.80	N
21	pLFE1	4.03	34.43	212	55	pLP-ZJ102	39.11	38.69	N
22	pMRI5.2	5.21	35.79	215; 314	56	LBPp3	39.46	42.29	N
23	pLJ42	5.52	42.41	308	57	Lp16E	40.14	41.36	367
24	LZ206p1	6.11	37.47	N	58	pLP-ZJ103	41.50	39.50	N
25	pCAUH202	6.40	33.68	311	59	LBPp1	45.41	39.67	327
26	Lp16L	6.46	34.67	N	60	ZJ95p1	48.55	39.05	N
27	p256	7.22	36.73	N	61	LBPp2	49.01	42.01	N
28	Lp16A	7.24	33.37	311	62	Lp16F	50.19	39.78	N
29	Lp16B	8.63	35.93	311	63	Lp16G	51.85	42.56	512
30	LBPp6	8.68	35.98	311	64	LZ227p2	52.14	41.83	N
31	pG6302	9.11	36.39	200	65	pST-III	53.56	38.69	N
32	p100	9.25	37.37	N	66	Lp16H	74.07	41.46	510
33	pLP9000	9.25	37.37	N	67	LZ227p1	74.17	41.33	512
34	pG6303	10.04	36.83	N					

注：N：没有编码已知 Rep.

Note：N: No known Rep has been encoded.

## 2.2 植物乳杆菌编码 Rep 质粒系统进化关系

45 个植物乳杆菌质粒和 5 个参照质粒的 Rep 系统进化树见图 1。根据进化树的拓扑结构和分支的置信度可知，植物乳杆菌 44 个质粒与 5 个参照质粒明显聚类，质粒 pG6302 单独成支。质粒 pMRI5.2 的 2 个 Rep 分别与家族 1 和家族 2 质粒聚

类，根据参照质粒可知，家族 1 和家族 2 质粒分别属于滚环复制(Rolling circle replication, RCR)质粒 pC194 和 pMV158 家族，因此质粒 pMRI5.2 很可能由这两个家族来源的 2 个质粒重组进化形成<sup>[12]</sup>。质粒 pZL3 的 2 个 Rep 均与家族 3 质粒聚类，根据参照质粒，家族 3 质粒属于 Theta 复制质粒 pUCL287

家族,因此质粒 pZL3 可能是 pUCL287 家族来源的 2 个质粒重组进化的结果。质粒 pG6302 的 Rep 含有不完整的 pfam01051 结构域,该结构域普遍存在于家族 3 质粒 Rep,但质粒 pG6302 与家族 3 质粒没有聚类,并在进化树中最近的分支处显示了较低的置信度(44%),因此质粒 pG6302 与 5 个家族质粒没有明显的进化关系。综上所述,经 Rep 系统进化

树分析,植物乳杆菌编码 Rep 的 45 个质粒可以划分为 5 个进化关系紧密的家族和 1 个独立质粒 pG6302,每个家族质粒可能具有共同的进化祖先和起源,质粒 pG6302 可能具有独立的进化过程和起源。

为了进一步揭示 5 个家族内部各质粒之间的进化关系,基于 Rep 氨基酸序列同源性,构建了 5 个家族质粒 Rep 系统进化树,结果见图 2。

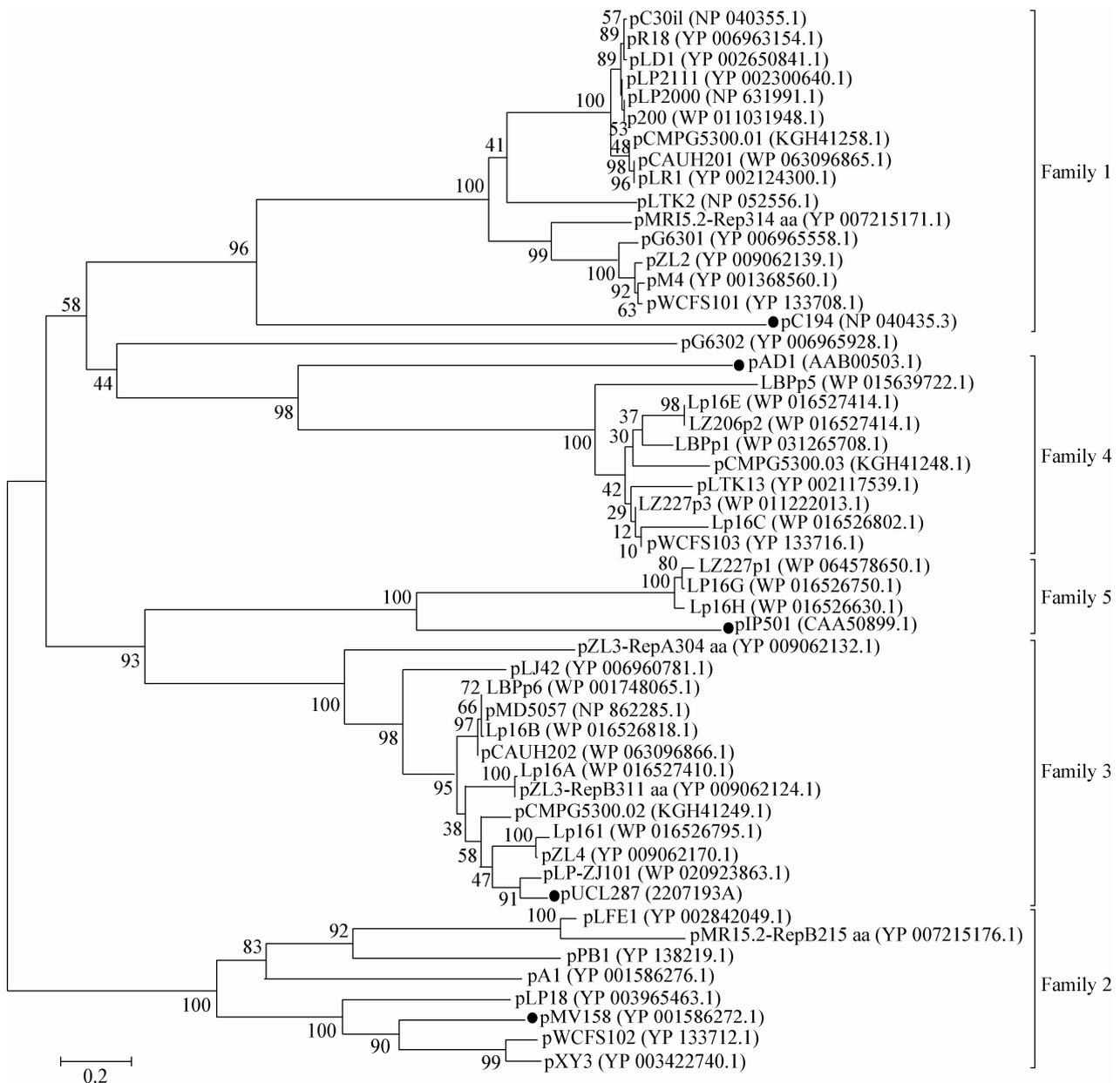
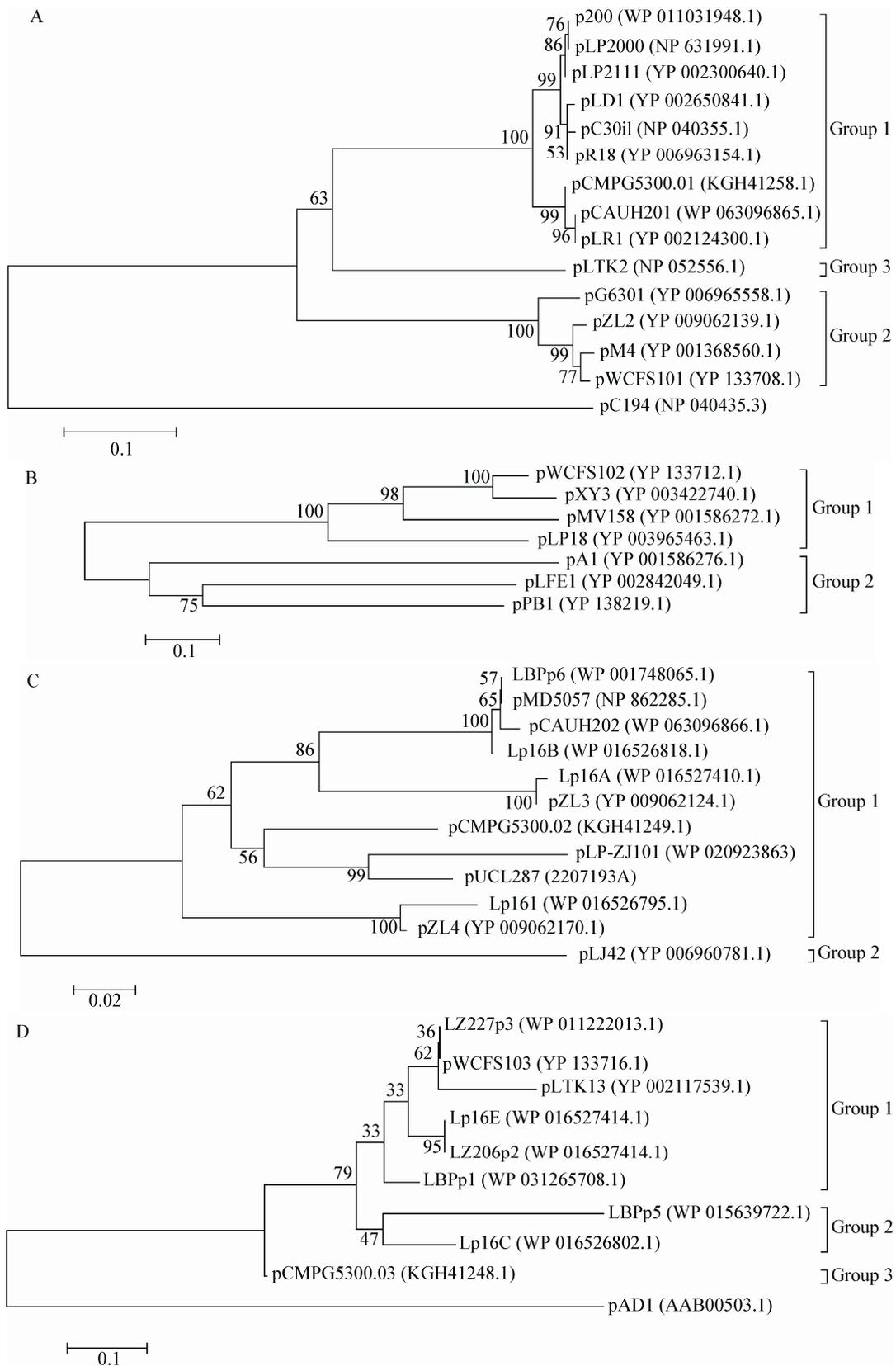


图 1 植物乳杆菌质粒 Rep 系统进化树

Figure 1 Phylogenetic tree of Rep amino acid sequences in *Lactobacillus plantarum* plasmids



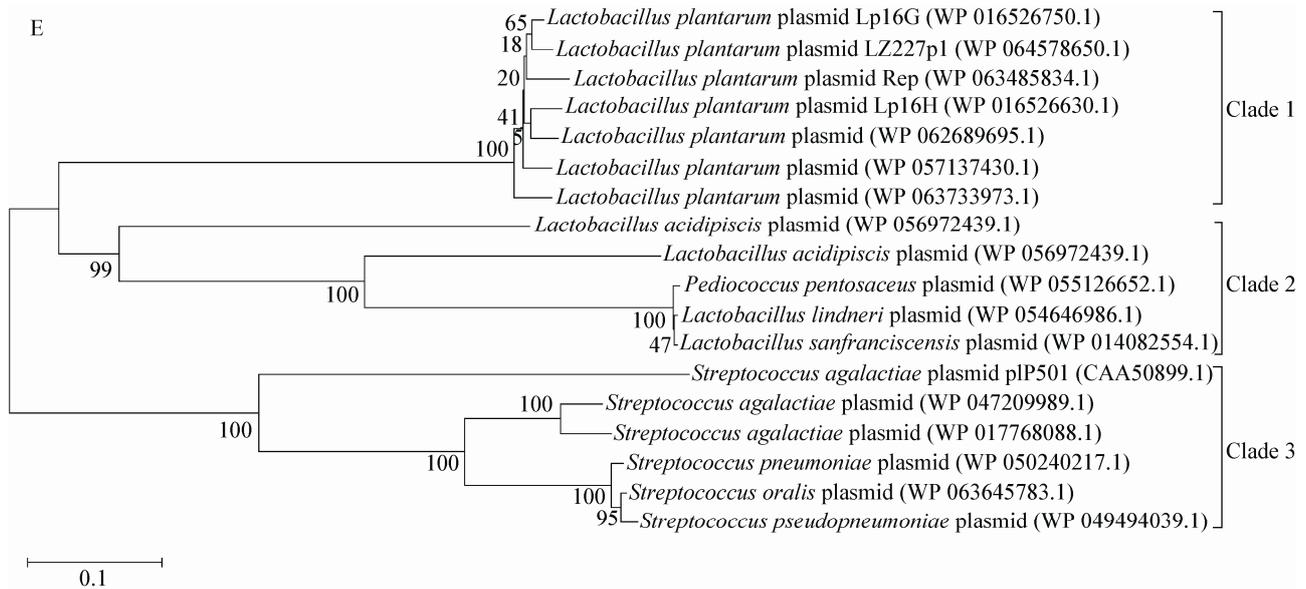


图 2 植物乳杆菌质粒家族 Rep 系统进化树

Figure 2 Phylogenetic tree of the Reps of plasmid family in *Lactobacillus plantarum*

注：A、B、C、D、E：植物乳杆菌质粒家族 1、2、3、4、5 Rep 系统进化树。

Note: A, B, C, D and E represent the Rep phylogenetic tree of *Lactobacillus plantarum* plasmid family 1–5, respectively.

### 2.3 家族 1 质粒的系统进化关系

经序列比对分析，家族 1 质粒 Rep 与质粒 pC194<sup>[13]</sup> Rep (232 aa) 序列大小差异显著，具有 20.94%–24.30% 的一致性，表明它们可能经历了漫长的进化时间，具有较远的进化关系。这一推论可以从它们宿主较远的亲缘关系(葡萄球菌和乳杆菌跨越分类学“目”的水平)得到进一步佐证。然而，这个家族的 Rep 功能没有改变，因为家族 1 质粒和质粒 pC194 的 Rep 均含有 3 个保守的 Motifs (表 2)，这些 Motifs 已被证明是 pC194 家族质粒复制起始的关键<sup>[14-15]</sup>。因此推测家族 1 质粒属于 RCR 质粒 pC194 家族，它们具有共同祖先，但分歧进化时间已经比较久远。

由图 1 和图 2A 可知，植物乳杆菌编码 Rep 质粒家族 1 包括 14 个质粒成员，质粒大小为 1.91–3.51 kb；G+C% 含量为 35.81%–39.49%，Rep 大小为 317–319 aa，属于 Rep\_1 蛋白家族(pfam 01446)。家族 1 质粒可以进一步划分为 3 个进化关系更近类群。类群 1 质粒 Rep 序列一致性 97.38%，类群 2 质粒 Rep 序列一致性 96.87%。质粒 pLTK2 Rep 单独成支，这

表明质粒 pLTK2 Rep 经历了一个独特的进化事件或过程，因此发生了一定程度的分歧进化。质粒 pC194 与家族 1 所有质粒进化距离较远，单独成支，这进一步表明它与家族 1 质粒分歧进化时间已经比较久远。

### 2.4 家族 2 质粒的系统进化关系

经序列比对分析，家族 2 质粒 Rep 与质粒 pMV158<sup>[16]</sup> Rep (210 aa) 具有 20.55%–62.56% 的一致性。质粒 pMV158 家族 Rep 含有 5 个保守的 Motifs，

表 2 家族 1 质粒 Rep 保守基序  
Table 2 Conservative motifs of plasmid Rep in the family 1

项目 Item	基序 I Motif I	基序 II Motif II	基序 III Motif III
Group 1	FLTLTVKN	QLHLVLL/VF	TAKYEVKSAD
Group 2	FLTLTVK/EN	HHMHVLLF	TAKYQVKSAD
Group 3	FLTLTEEN	QLHLVLLF	TAKYQVKSAD
pC194	FLTLTPN	PHFHVLI	MAKYSGKSD
Conserved site	FLTL**N	*H*HVL**	*AKY**K**D

注：\*：非保守性位点。

Note: \*: No conservative sites.

Motif III 和 Motif IV 对质粒前导链复制起始具有关键作用<sup>[17]</sup>。在家族 2 质粒中,除了 pA1 Rep 含有 3 个保守 Motifs 其余 5 个质粒 Rep 均含有 pMV158 家族 5 个保守 Motifs (表 3)。因此基于 Rep 大小、序列同源性,尤其 Motifs 序列一致性可以看出,虽然家族 2 质粒和 pMV15 的宿主亲缘关系跨越了分类学“科”的水平,但 Rep 的功能区域仍然是保守的,因此推测家族 2 质粒属于 RCR 质粒 pMV158 家族,它们可能起源于共同祖先。

由图 1 和图 2B 可知,植物乳杆菌编码 Rep 质粒家族 2 包括 6 个质粒成员,质粒大小为 1.80–4.03 kb, G+C% 含量为 34.33%–38.24%。Rep 大小为 196–237 aa,属于 Rep\_2 蛋白家族(pfam 01719)。与其他家族比较,家族 2 质粒之间遗传距离较远,分歧进化程度最高,这可能与同源质粒 pMV158 家族的广泛宿主性质相关。家族 2 质粒可以划分为 2 个进化关系更近的类群,这一结果从 Rep 5 个保守 Motifs 序列的差异性上也可以得出(表 3)。与参照质粒 pMV158 比较,类群 2 质粒可能经历了更复杂的分歧进化事件,而类群 1 质粒与 pMV158 显示具有更近的系统进化关系。

### 2.5 家族 3 质粒的系统进化关系

经比对分析,家族 3 质粒 Rep 与质粒 pUCL287 Rep (311 aa)序列具有 68.17%–91.32%的一致性,Rep 在大小和序列同源性上显示了高度保

守性,这表明家族 3 质粒很可能属于 Theta 复制 pUCL287 家族<sup>[18-19]</sup>,并且它们很可能具有共同祖先。

由图 1 和图 2C 可知,植物乳杆菌编码 Rep 质粒家族 3 包括 11 个质粒成员,质粒大小为 5.52–15.16 kb, G+C% 含量为 33.37%–42.41%。Rep 大小为 307–311 aa,属于 Rep\_3 家族(pfam 01051)。家族 3 质粒可以划分为 2 个进化关系更近的类群。与其他家族比较,家族 3 质粒(尤其是类群 1 质粒)之间的分歧进化程度最低,表明这些质粒可能具有很近的进化关系。参照质粒 pUCL287 首先与 pLP-ZJ101 聚类(置信度 99%),之后与类群 1 其他质粒聚类,这表明虽然质粒 pUCL287 的宿主(嗜盐四联球)和植物乳杆菌亲缘关系较远(跨越分类学“科”水平),但质粒 pUCL287 与家族 3 质粒显示了亲密的进化关系,因此推测这两个亲缘关系较远的宿主菌可能在近期发生过质粒水平转移事件,从而使起源质粒进一步分歧进化。这从另一方面表明家族 3 质粒可能与 pUCL287 家族质粒相似,具有 Theta 复制质粒中较宽泛的宿主范围。进化树中质粒 pLJ42 与家族 3 其他质粒遗传距离较远,单独成支,这表明质粒 pLJ42 与类群 3 其他质粒进化关系较远,可能经历了一个特殊且独立的进化过程。

### 2.6 家族 4 质粒的系统进化关系

经序列比对分析,家族 4 质粒 Rep 与质粒 pAD1 Rep (336 aa)显示了很低的一致性(4.44%–29.07%),

表 3 家族 2 质粒 Rep 保守性基序  
Table 3 Conservative motifs of plasmid Rep in the family 2

项目 Item	基序 I Motif I	基序 II Motif II	基序 III Motif III	基序 IV Motif IV	基序 V Motif V
pLFE1	WWIVVYPESL	AISPLHDKD	KKPHYHIVFN	I S GLTGAVRYLTH-MDN	YDN
pPB1	WSWIVYPESA	IESPLHDKD	KKHHWHII-IS	VASLQGAVQYLWHR-NN	YDK
pA1	WTFIVYPESA	IESPLHDKD	KKSHYHILLTF	-	-
pLPI8	F S F I I YPESI	AISPLHD	KKPHYHVL--Y	VDSVENTFKYLTHESKD	YAK
pWCFS102	F T F LLYPESI	AISPLHDKD	KKAHYHVI--Y	VQSMENMYLYLTHESKD	YSK
pXY3	F T F LLYPESI	AISPLHDKD	KKAHYHVI--Y	VQSMDN I YLYLTHESKN	YSK
pMV158	F T F LLYPESI	AISPLHDKD	KKAHYHVL--Y	VLNVENMYLYLTHESKD	YDK
Conserved site	*****YPES*	*SPLHD*D	KK*H*H*****	*****Y**H****	Y**

注: \*: 非保守性位点; -: 没有保守性基序。

Note: \*: No conservative sites; -: No conservative motif.

但家族 4 质粒较大的 Rep (>200 aa)与 pAD1 Rep N 端显示高度同源性;家族 4 质粒 4 个较小的 Rep (<200 aa)与 pAD1 Rep 几乎没有同源性,但家族 4 所有质粒 Rep C 端显示高度同源性;Rep N 和 C 末端之间序列显示了很低的同源性。家族 4 质粒 Rep 3 个区域的保守性特征与 RepA\_N 家族质粒 Rep 近年的研究结果一致,这与每个区域负责的复制起始功能密切相关<sup>[20-21]</sup>。这些分析结果表明家族 4 质粒可能属于 Theta 复制 RepA\_N 家族。

由图 1 和图 2D 可知,植物乳杆菌编码 Rep 质粒家族 4 包括 9 个质粒成员,质粒大小为 16.10–45.41 kb,G+C%含量为 39.04%–42.14%。Rep 大小为 102–373 aa,大于 200 aa 的 Rep 属于 RepA\_N 蛋白家族(pfam 06970)。家族 4 质粒可以进一步划分为 3 个类群。虽然家族 4 质粒 Rep 序列大小差异显著,导致序列整体同源性较低和进化树中各分支置信度较低,但由于小 Rep 与大 Rep C 端序列具有高度同源性,因此进化树中质粒之间的遗传距离仍然较近,含有不同大小 Rep 质粒明显聚类(见类群 1 和 2),这表明家族 4 质粒分歧进化程度较低,进化关系较近,很可能拥有共同祖先。进化树中参照质粒 pAD1 远离家族 4 所有质粒,单独成支,这表明家族 4 质粒与 pAD1 进化关系较远,它们的祖先可能很早以前就因为宿主的改变而发生了分歧进化。这些特征与 Weaver 等<sup>[20]</sup>对相关质粒的研究结果一致,很可能与 RepA\_N 家族质粒的广泛分布和狭窄宿主特征相关。

### 2.7 家族 5 质粒的系统进化关系

序列比对分析显示,家族 5 质粒 3 个 Rep 具有 97.98%一致性,与无乳链球菌质粒 pIP501 编码的 RepR 具有 31.47%–32.49%的一致性。此外,经 BLAST 比对分析发现,Lp16G Rep 与众多植物乳杆菌(90%以上)、其他乳杆菌和片球菌(40%–45%)、链球菌(30%–40%)质粒 Rep 具有不同程度的同源性。所有这些蛋白均含有保守的 PriCT-1 和 HTH 结构域。PriCT-1 是引发酶 C 末端结构域,可以引发复制体组装;HTH 则是最常见的 DNA 结合结构域。

因此含有这两个保守结构域的质粒在植物乳杆菌、片球菌和链球菌中广泛存在,它们可能同属于 pIP501 相似的 Theta 复制质粒家族<sup>[22]</sup>,并拥有共同的祖先。

由图 1 和图 2E 可知,植物乳杆菌编码 Rep 质粒家族 5 包括 3 个质粒成员,质粒大小为 51.85–74.17 kb,G+C%含量为 41.33%–42.56%,Rep 大小为 510–512 aa,含有保守结构域 PriCT-1 (pfam 08708)和 HTH (cl21459)。家族 5 及相关质粒明显聚类为 3 个进化支,进化支 1 包括家族 5 质粒与其他植物乳杆菌质粒,进化支 2 包括其他乳杆菌和片球菌质粒,进化支 3 包括所有链球菌质粒,所有质粒的进化关系与宿主的亲缘关系基本一致,这些结果表明,家族 5 相关质粒具有严格的宿主范围(可能为“物种”特异),主要与宿主菌共同进化,很少进行异源宿主菌之间的水平转移,因此该家族祖先质粒很早以前便随着宿主菌的进化开始了分歧演化。

### 3 结论与讨论

质粒 Rep 系统进化树(图 1)分析表明,植物乳杆菌 45 个编码 Rep 质粒可以与 5 个参照质粒明显聚类为 5 个进化关系紧密的家族和 1 个单独成支质粒 pG6302,因此推测这些质粒可能起源于 6 个祖先质粒,每个家族质粒与聚类的参照质粒可能具有共同祖先。每个家族质粒和参照质粒或同源质粒构建的 Rep 系统进化树(图 2)可以进一步揭示每个家族质粒之间(家族 1–4)或家族质粒与同源 Rep 质粒之间(家族 5)的系统进化关系。根据进化树(图 1 和 2)中质粒之间的遗传距离,RCR 质粒(家族 1–2)之间分歧进化程度较高,Theta 复制质粒(家族 3–5)之间分歧进化程度较低,这一结果与两种复制类型质粒的宿主范围特征一致<sup>[15,23]</sup>。以上分析结果表明,植物乳杆菌质粒 Rep 氨基酸序列具有适度的保守性和变异性,可以简单且有效地反映出编码 Rep 质粒之间的系统进化关系和演化过程,因此 Rep 是植物乳杆菌编码 Rep 质粒理想的遗传和进化分子标记,可为植物乳杆菌质粒的系统进化研究提供一种简

单、有效的分析方法和标准。我们正利用该方法进行更广泛宿主来源质粒的系统进化关系的分析和研究,理论上本方法可能适用于所有编码 Rep 质粒的系统进化研究,但对于非编码 Rep 质粒,本方法有明显的局限性,因此,如何改进和完善这个不足需要今后进一步的研究。

## 参 考 文 献

- [1] Siezen RJ, van Hylckama Vlieg JET. Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer[J]. *Microbial Cell Factories*, 2011, 10(S1): S3
- [2] Sun DQ, Li HF, Song DW, et al. Progress in research on natural plasmids in *Lactobacillus*[J]. *Food Science*, 2015, 36(11): 251-255 (in Chinese)  
孙大庆, 李洪飞, 宋大巍, 等. 乳杆菌属天然质粒研究进展[J]. *食品科学*, 2015, 36(11): 251-255
- [3] Fan J, Xi XD, Huang Y, et al. Isolation of a minireplicon of the plasmid pG6303 of *Lactobacillus plantarum* G63 and characterization of the plasmid-encoded Rep replication protein[J]. *Journal of Genetics*, 2015, 94(2): 177-186
- [4] Aleksandrak-Piekarczyk T, Koryszewska-Bagińska A, Grynberg M, et al. Genomic and functional characterization of the unusual pLOCK 0919 plasmid harboring the *spaCBA* pili cluster in *Lactobacillus casei* LOCK 0919[J]. *Genome Biology and Evolution*, 2016, 8(1): 202-217
- [5] Sun ZL, Kong J, Kong WT. Characterization of a cryptic plasmid pD403 from *Lactobacillus plantarum* and construction of shuttle vectors based on its replicon[J]. *Molecular Biotechnology*, 2010, 45(1): 24-33
- [6] Panya M, Lulitanond V, Tangphatsomruang S, et al. Sequencing and analysis of three plasmids from *Lactobacillus casei* TISTR1341 and development of plasmid-derived *Escherichia coli-L. casei* shuttle vectors[J]. *Applied Microbiology and Biotechnology*, 2012, 93(1): 261-272
- [7] Rattanachaikunsopon P, Phumkhachorn P. Construction of a food-grade cloning vector for *Lactobacillus plantarum* and its utilization in a food model[J]. *Journal of General Applied Microbiology*, 2012, 58(4): 317-324
- [8] Asteri IA, Papadimitriou K, Boutou E, et al. Comparative and evolutionary analysis of plasmid pREN isolated from *Lactobacillus rennini*, a novel member of the theta-replicating pUCL287 family[J]. *FEMS Microbiology Letters*, 2011, 318(1): 18-26
- [9] Francia MV, Haas W, Wirth R, et al. Completion of the nucleotide sequence of the *Enterococcus faecalis* conjugative virulence plasmid pAD1 and identification of a second transfer origin[J]. *Plasmid*, 2001, 46(2): 117-127
- [10] Thompson JK, Collins MA. Completed sequence of plasmid pIP501 and origin of spontaneous deletion derivatives[J]. *Plasmid*, 2003, 50(1): 28-35
- [11] Kumar S, Stecher G, Tamura K. MEGA 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874
- [12] Cho GS, Huch M, Mathara JM, et al. Characterization of pMRI 5.2, a rolling-circle-type plasmid from *Lactobacillus plantarum* BFE 5092 which harbours two different replication initiation genes[J]. *Plasmid*, 2013, 69(2): 160-171
- [13] Horinouchi S, Weisblum B. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance[J]. *Journal of Bacteriology*, 1982, 150(2): 815-825
- [14] Khan SA. Plasmid rolling-circle replication: highlights of two decades of research[J]. *Plasmid*, 2005, 53(2): 126-136
- [15] Ruiz-Masó JA, Machón C, Bordanaba-Ruiseco L, et al. Plasmid rolling-circle replication[J]. *Microbiology Spectrum*, 2015, 3(1): PLAS-0035-2014
- [16] van der Lelie D, Bron S, Venema G, et al. Similarity of minus origins of replication and flanking open reading frames of plasmids pUB110, pTB913 and pMV158[J]. *Nucleic Acids Research*, 1989, 17(18): 7283-7294
- [17] Moscoso M, del Solar G, Espinosa M. Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA[J]. *Journal of Biological Chemistry*, 1995, 270(8): 3772-3779
- [18] Benachour A, Frère J, Novel G. pUCL287 plasmid from *Tetragenococcus halophila* (*Pediococcus halophilus*) ATCC 33315 represents a new theta-type replicon family of lactic acid bacteria[J]. *FEMS Microbiology Letters*, 1995, 128(2): 167-175
- [19] Benachour A, Frère J, Flahaut S, et al. Molecular analysis of the replication region of the theta-replicating plasmid pUCL287 from *Tetragenococcus* (*Pediococcus*) *halophilus* ATCC33315[J]. *Molecular and General Genetics MGG*, 1997, 255(5): 504-513
- [20] Weaver KE, Kwong SM, Firth N, et al. The RepA\_N replicons of Gram-positive bacteria: a family of broadly distributed but narrow host range plasmids[J]. *Plasmid*, 2009, 61(2): 94-109
- [21] Schumacher MA, Tonthat NK, Kwong SM, et al. Mechanism of staphylococcal multiresistance plasmid replication origin assembly by the RepA protein[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(25): 9121-9126
- [22] Grohmann E, Goessweiner-Mohr N, Brantl S. DNA-binding proteins regulating pIP501 transfer and replication[J]. *Frontiers in Molecular Biosciences*, 2016, 3: 42
- [23] Lilly J, Camps M. Mechanisms of theta plasmid replication[J]. *Microbiology Spectrum*, 2015, 3(1): PLAS-0029-2014