

# 乳酸菌基因组学研究现状

颜翼 伦永志

乳酸菌(*Lactobacillus*)广泛分布在人和动物的消化道、阴道以及乳制品和发酵的动植物制品中<sup>[1]</sup>,可参与宿主的营养、吸收、代谢、免疫及抗肿瘤、抗感染过程,被认为是微生物生态学研究的重心<sup>[2]</sup>。

根据伯杰细菌鉴定手册,乳酸菌目前主要分为23个属,主要有乳杆菌属(*Lactobacillus*)、乳球菌属(*Lactococcus*)、明串珠菌属/白联球菌属(*Leuconoszo*)、片球菌属(*Pediococcus*)、链球菌属(*Streptococcus*)、肠球菌属(*Enterococcus*)、双歧杆菌属(*Bifidobacterium*)和肉食菌属(*Caroobacterium*)等<sup>[3]</sup>。从2001年完成乳酸乳球菌 IL1430 即第一株乳酸菌的全基因组测序<sup>[4]</sup>开始,目前已完整公布全基因组序列的乳酸菌共有14种,包括嗜热链球菌、约氏乳杆菌、乳酸乳球菌、肠系膜明串珠菌等。乳酸菌基因组全长约为1.8~3.2 Mbp, (G+C)%普遍偏低,编码基因的百分比在80%左右,基因数量和可预测的蛋白质数量相差较大(见表1)。多数菌属都存在假基因,从肠系膜明串珠菌的19个到嗜热链球菌的206个,数量相差较大。尽管乳酸菌各菌属RNA结构基因数量不一,但均含有rRNA操纵子和tRNA编码基因,某些乳酸菌还含有质粒与噬菌体编码基因。

## 一、基本功能基因

1. rRNA操纵子和tRNA编码基因(表2):各菌属的rRNA操纵子数量变化较大,其中乳杆菌属的rRNA操纵子数量较多。tRNA编码基因数量为43~98个,同样乳杆菌属偏多。通过分析rRNA操纵子的长度,16S rRNA操纵子为1464~1580 nt,23S rRNA操纵子为2623~3051 nt,5S rRNA操纵子则在119 nt上下波动,只有肠系膜明串珠菌属相对较低。rRNA操纵子排列具有一定顺序性,通常为16S rRNA、23S rRNA、5S rRNA,而且在其周围通常会出现tRNA<sup>Ala</sup>、tRNA<sup>Ile</sup>、tRNA<sup>Lys</sup>等编码基因。在亲缘关系相近的菌属中,16S rRNA中有相同的突变位点<sup>[7]</sup>,因此16S rRNA的同源性比较常用于推断物种间的亲缘关系。

2. 糖转化酶相关基因:乳酸菌糖转化酶相关基因归属于11个基因簇,其编码蛋白和酶对特异性细菌噬菌体的敏感性相关<sup>[8]</sup>。研究发现,乳球菌噬菌体的识别受体为细胞壁糖蛋白,而糖转化酶恰好与之相关,但目前关于其功能报道尚少见<sup>[9]</sup>。乳酸乳球菌 MG 1363 有28个糖转化酶相关基因,其中有13个可划归为4个基因簇,3个含有3个糖转化酶基因,另一个含有4个糖转化酶基因<sup>[10]</sup>。

3. 糖代谢相关基因:大部分乳酸菌中都含有糖酵解相关基因,以满足乳酸菌

获得能量的需求。此类基因数量较多,如1,4- $\alpha$ -葡萄糖分支酶、保护蛋白、纯化蛋白酶、糖转化酶、6-磷酸- $\beta$ -葡萄糖酶和 $\beta$ -葡萄糖酶等<sup>[8]</sup>。这些基因下游常存在部分糖酵解物质基因,编码产生的酶可以将三磷酸甘油醛分解成丙酮酸酯,再将丙酮酸酯转化为2,3-二丁醇,还可以将乙酸酯转变成乙酰辅酶A,参与戊糖磷酸酯的反应途径。虽然这些酶不足以单独完成乳酸菌的新陈代谢,但是存在或者缺少这些关键酶均会影响菌体发酵过程<sup>[3]</sup>。

表1 常见乳酸菌的基因组特征

菌属	基因长度(nt)	(G + C) %	编码基因(%)	基因数量	编码蛋白数量	RNA 基因数量	假基因数量
肠系膜明串珠菌属	2,038,396	37	88	2073	1970	85	19
乳脂乳球菌属	2,438,589	35	81	2610	2384	82	144
戊糖片球菌属	1,832,387	37	87	1847	1755	72	20
浑球红细菌属	3,188,609	69	88	3092	3022	48	22
酒类酒球菌属	1,780,517	37	82	1864	1691	51	122
嗜酸乳杆菌属	1,993,560	34	88	1936	1862	74	0
干酪乳杆菌属	2,895,264	64	82	2909	2751	76	82
短乳杆菌属	2,291,220	46	84	2314	2158	82	49
乳酸乳球菌属乳脂亚种	2,529,478	35	82	2597	2434	81	82
嗜热乳酸链球菌属	1,856,368	39	76	2003	1710	87	206
加氏乳杆菌属	1,894,360	35	88	1898	1755	98	48
德氏乳杆菌属	1,856,951	49	77	2040	1721	127	192
米酒乳杆菌属	1,884,661	41	86	1963	1879	84	0

表2 常见乳酸菌的RNA基因结构特征

菌属	rRNA 操纵子 基因数量	tRNA 编码 基因数量	16 s rRNA 基因长度(nt)	23 s rRNA 基因长度(nt)	5 s rRNA 基因长度(nt)
肠系膜明串珠菌属	4	71	1549	2914	89
乳脂乳球菌属	6	62	1548	2901	83
戊糖片球菌属	5	55	1580	2923	117
浑球红细菌属	1	43	1464	2883	115
酒类酒球菌属	2	43	1567	2888	36
嗜酸乳杆菌属	4	60	1572	3051	117
干酪乳杆菌属	9	98	1574	2920	120
短乳杆菌属	5	65	1563	2918	119
乳酸乳球菌属乳脂亚种	6	61	1549	2902	116
嗜热乳酸链球菌属	6	67	1550	2902	124
加氏乳杆菌属	6	78	1579	2632	119
德氏乳杆菌属	9	98	1561	2560	103
米酒乳杆菌属	7	61	1571	2919	117

4. 磷酸酶基因系统:根据氨基酸序列特征,原核生物磷酸酶可分为4个蛋白质超家族<sup>[11]</sup>。第一个是磷蛋白磷酸酶(PPP)家族,主要存在于真核生物细胞中,

在细菌中很少存在。第二个是  $\text{Mn}^{2+}$ 、 $\text{Mg}^{2+}$  依赖性磷酸酶 (PPM) 家族, 主要分布于跨膜区域, 在蛋白相互作用和细胞信号传递中起作用。第三个家族主要包括 DNA 磷酸酶和仅在酵母、乳杆菌等中存在的组氨酸磷酸酶 (PHP), 后者在酵母进行氨基酸生物合成时可除去组氨酸中的磷酸盐<sup>[12-14]</sup>。该家族特征性模序结构中均含有组氨酸和天门冬氨酸残基, 这可能与金属依赖性水解磷酸结合酶活性有关, 酪氨酸激酶编码基因与 PHP 编码基因相邻, 有时可能替代 PHP 活性<sup>[12]</sup>, 最后一个是酪氨酸磷酸酶 (PTP) 家族, PTP 能催化蛋白质分子中特定位点的磷酸化酪氨酸残基脱磷酸, 在控制细胞磷酸化酪氨酸水平上, 蛋白酪氨酸磷酸酶起着高度特异性的积极作用<sup>[15]</sup>。

## 二、特殊功能基因

1. 质粒编码基因: 嗜热乳酸链球菌、肠系膜明串珠菌和某些乳杆菌具有质粒编码基因, 它们编码基因的数量不等, 但一般无假基因和 RNA 编码基因。乳酸乳球菌 *ssp* 可编码 5 种质粒, 编码基因多达 129 个, 而嗜热链球菌可编码两个质粒, 编码基因仅为 6 个。德氏乳杆菌的质粒是一个长度为 6220 bp 的环状 DNA 分子, (G + C)% 为 44.6%, 含有 6 个开放式阅读框架, 编码 3 种膜镶嵌蛋白、1 个转移酶和 1 个螺旋复制蛋白<sup>[16]</sup>。较为典型的是乳酸乳球菌乳脂亚种 MG1363, 其编码质粒的基因长度为 59 kb, 有 2 个质粒复制蛋白基因残基 HsdRMS 模版系统基因和 *opp-pepo* 操纵子, 该序列还包括 3 个重复序列, 其中 1 个与溶解基因本身的重连接位点高度相似<sup>[10]</sup>, 这种重复序列在其它细菌的质粒编码基因中也有发现, 包括嗜热链球菌基因组中的“水平基因转移多发位点”。水平基因转移是指在单个细胞内的细胞器之间进行的遗传物质交流, 而水平基因转移多发位点就是在细菌基因中存在的容易发生水平基因转移的基因位点<sup>[17]</sup>。

2. 噬菌体编码基因: 某些乳酸菌中还含有噬菌体编码基因, 但数量一般都很少, 仅为 1~2 个。乳酸乳球菌乳脂亚种 MG 1363 中的噬菌体编码基因均携带有 tRNA 编码基因, 此外还存在一个噬菌体校正序列, 其中包括一个长度为 712 个氨基酸的插入序列。该序列对菌体是否有影响还是个未知数, 而在其它细菌中却始终未被发现<sup>[8]</sup>。

## 三、未知功能基因

乳酸菌的很多基因是通过基因水平转移获得的, 但这些基因对细菌本身的影响是无法事先预测的, 这样在乳酸菌中就会出现很多的未知功能新基因, 尽管有时表现出某些酶的活性, 但却不易识别其完整的生理功能。例如乳杆菌的一个未知功能基因的编码蛋白可表现出烯醇化酶活性, 有研究表明它是通过基因水平转移而从其它亲缘关系较远的菌属中获得的, 并不是通过基因突变而进化来的<sup>[18]</sup>。此外, 乳酸菌的诸多未知功能基因都是单拷贝基因, 并且其编码蛋白多与酶相关。

## 四、展望

开展乳酸菌基因组学研究可拓宽对乳酸菌的认识, 从更微观的角度认识乳酸菌及其功能基因。应用基因组学方法可将乳酸菌基因结构与其生理功能建立联

系,从而促进其代谢组学、分泌蛋白和进化方向研究。对乳酸菌未知功能基因的研究可进一步拓宽乳酸菌的应用范围,对开发新型基因工程药物和功能性食品将会有个极大的推进。乳酸菌对机体产生的各种生态效应和生理作用的主要物质基础就是细菌蛋白,特别是分泌性蛋白。同其它细菌一样,乳酸菌在胁迫环境下包括高温或低温、极酸或极碱、高盐、氧化及各类营养限制等,通过诱导表达方式改变代谢类型与水平来适应环境。利用蛋白质化学方法可有效地分析获得乳酸菌的诱导蛋白,再结合基因组学与生物信息学方法开展蛋白功能研究,为进一步揭示乳酸菌的行为机理和生理作用提供新的认识。

### 参 考 文 献

- 1 康白. 微生物学. 大连:大连出版社,1988.
- 2 Cotter PD, Hill C, Ross RP. Bacteriocins: developing innate immunity for food. *Nat Rev Microbiol*,2005,3:777-788.
- 3 凌代文. 乳酸细菌的分类鉴定及实验方法. 北京:中国轻工业出版社,1999.
- 4 AL Exander, Bolotin PW. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *Lactis* IL1403. *Genome Res*,2001,11:731-753.
- 5 Botina SG, Sukhodolets VV. Speciation in bacteria: comparison of the 16S rRNA gene for closely related *Enterococcus* species. *Genetika*,2006,42:325-330.
- 6 Wegmann U, O'Connell-Motherway M, Zomer A, et al. Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J Bacteriol*,2007,189:3256-3270.
- 7 Dupont K, Janzen T, Vogensen FK, et al. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. *Appl Environ Microbiol*,2004,70:5825-5832.
- 8 Dabour N, Lapointe G. Identification and molecular characterization of the chromosomal exopolysaccharide biosynthesis gene cluster from *Lactococcus lactis* subsp. *cremoris* SMQ-461. *Appl Environ Microbiol*,2005,71:7414-7425.
- 9 Shi L. Manganese-dependent protein O-phosphatases in prokaryotes and their biological functions. *Front Biosci*,2004,9:1382-1397.
- 10 Grangeasse C, Cozzzone AJ, Deutscher J, et al. Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem Sci*,2007,32:86-94.
- 11 Mijakovic I, Petranovic D, Bottini N, et al. Protein-tyrosine phosphorylation in *Bacillus subtilis*. *J Mol Microbiol Biotechnol*,2005,9:189-197.
- 12 Mijakovic I, Poncet S, Boel G, et al. Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J*,2003,22:4709-4718.
- 13 LaPointe G, Atlan D, Gilbert C, et al. Characterization and site-directed mutagenesis of Wzb, an O-phosphatase from *Lactobacillus rhamnosus*. *BMC Biochem*,2008,9:1-12.
- 14 Lee JH, Halgerson JS, Kim JH, et al. Comparative sequence analysis of plasmids from *Lactobacillus delbrueckii* and construction of a shuttle cloning vector. *Appl Environ Microbiol*,2007,73:4417-4424.
- 15 Bolotin A, Quinquis B, Renault P, et al. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol*,2004,22:1554-1558.
- 16 Koonin EV. Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet*,2005,39:309-338.

(收稿日期:2008-09-22)

(本文编辑:孙荣华)

颜翼, 伦永志. 乳酸菌基因组学研究现状[J/CD]. 中华实验和临床感染病杂志:电子版,2010,4(2):193-196.