

研究报告

环境胁迫和非胁迫条件下类植物乳杆菌转录组分析

闫洪波 张香美* 许冬倩

(河北经贸大学生物科学与工程学院 河北 石家庄 050061)

摘要:【背景】类植物乳杆菌 L-ZS9 是一株在食品发酵与保鲜方面具有潜在应用价值的产细菌素益生菌。【目的】深入了解环境胁迫影响类植物乳杆菌 L-ZS9 细菌素合成的重要相关调节基因的信息。【方法】通过高通量测序技术对类植物乳杆菌转录组进行测序, 对所有转录本进行 COG (Clusters of Orthologous Groups)、GO (Gene Ontology)和 KEGG (Kyoto Encyclopedia of Genes and Genomes)分类和 Pathway 注释。【结果】转录组测序显示 2 个样品基因覆盖度都在 90% 以上, 差异表达基因 927 个, 其中 744 个上调, 183 个下调。KEGG 分析结果表明, 649 个差异表达基因中 68 个集中在“ABC 转运途径”, 占 10.48%, 其中 3 个基因表达上调超过 16 倍, 1 个基因表达下调超过 1/16, 暗示类植物乳杆菌 L-ZS9 细菌素合成与“ABC 转运途径”密切相关。另外多个孤儿基因表达变化也超过 16 倍, 有的甚至表达上调达万倍。【结论】进一步拓展了类植物乳杆菌的基因信息, 为类植物乳杆菌细菌素代谢途径和逆境反应研究提供了坚实的基础。

关键词: 乳杆菌, 转录组, 基因表达, 环境胁迫, 细菌素

Transcriptome analysis of *Lactobacillus paraplantarum* L-ZS9 under environmental stress and control conditions

YAN Hong-Bo ZHANG Xiang-Mei* XU Dong-Qian

(Bioscience and Bioengineering School, Hebei University of Economics and Business, Shijiazhuang, Hebei 050061, China)

Abstract: [Background] *Lactobacillus paraplantarum* L-ZS9, which could produce bacteriocin, would be used in food preservative and fermentation. [Objective] To understand expression characters of regulated genes of bacteriocin synthesis under environment stress conditions for *Lactobacillus paraplantarum* L-ZS9. [Methods] We used high-throughput sequencing technology to sequence its transcriptome. Based on the Clusters of Orthologous Groups (COG), Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, we carried out functional annotation and classification, pathway annotation. [Results] The results suggested that distribution of genes' coverage is beyond 90% of two samples, the total genes with expression difference were 927, of which 744 were up regulation and 183 down regulation. KEGG pathway analysis showed that 68 out

Foundation item: National Natural Science Foundation of China (31471707)

*Corresponding author: E-mail: xzmwshw@sohu.com

Received: March 15, 2017; **Accepted:** July 05, 2017; **Published online** (www.cnki.net): August 11, 2017
基金项目: 国家自然科学基金(31471707)

*通信作者: E-mail: xzmwshw@sohu.com

收稿日期: 2017-03-15; 接受日期: 2017-07-05; 网络首发日期(www.cnki.net): 2017-08-11

of the 649 genes with different expression levels were classed into ABC transport pathway, accounting for 10.48% and three genes were up regulated beyond 16 times and one was down regulated beyond 1/16. In addition, expression of some orphan genes changed with 16 times up or down-regulation, some even could increase more than 10 000 times. [Conclusion] This study further improved the gene information and laid the foundation of bacteriocin metabolic pathway and stress respond research for *Lactobacillus paraplantarum*.

Keywords: *Lactobacillus paraplantarum*, Transcriptome, Gene expression, Environmental stress, Bacteriocin

类植物乳杆菌是一类重要的益生菌,在食品发酵和食品保鲜等方面有着广泛的应用^[1]。类植物乳杆菌 L-ZS9 (*Lactobacillus paraplantarum* L-ZS9)是自发酵肉制品中分离的一株细菌素产生菌^[2],该菌能有效抑制肠杆菌科微生物的生长,具有开发成优良益生菌制剂或食品益生菌添加剂的潜力,其细菌素的合成受双组分群体感应系统 *plnB/plnCD* 调控^[3]。张旭的研究结果显示:类植物乳杆菌 L-ZS9 用于鲟鱼肉香肠及鱼肉猪肉混合香肠的发酵能够更为迅速地启动发酵,较好地抑制肠杆菌的生长,同时又能够提高发酵香肠的感官品质^[3]。Wang 等将类植物乳杆菌 L-ZS9 用于鲟鱼肉香肠的发酵,大大提高了其安全性和功能性^[4]。环境胁迫影响乳杆菌细菌素的合成,已成为限制产细菌素乳杆菌应用的一个重要因素^[5-6]。目前类植物乳杆菌 L-ZS9 已经完成了基因组测序^[7],测序结果已提交 GenBank (<https://www.ncbi.nlm.nih.gov/Traces/wgs/?val=JPEB01#contigs> 以及 https://www.ncbi.nlm.nih.gov/nucore/NZ_CP013130.1),但是环境胁迫影响其细菌素合成的分子机制尚需深入研究。

转录组是指某一生理条件下细胞内所有转录产物的集合,包括 mRNA 和非编码 RNA 的所有 RNA 总和^[8]。转录组分析能够对转录产物分类、基因组功能元件、转录结构、转录后修饰以及胁迫条件的差异表达进行研究^[9]。高通量测序技术又称“下一代”测序技术(“Next-generation” sequencing technology),根据参考基因组对测序结果进行组装,具有灵敏度高、无需设计探针、单碱基差异分辨率高以及重复性好等优点,目前,该技术已大量应用在动物、植物和微生物的转录组研究中^[10-12]。

利用高通量测序技术对类植物乳杆菌 L-ZS9 进行转录组测序,获得大量的基因信息,并对测序组装的转录本表达量变化、GO (Gene Ontology)功能分类、KEGG (Kyoto Encyclopedia of Genes and Genomes)代谢通路等进行研究,旨在从转录组层面探究环境胁迫条件影响 L-ZS9 细菌素合成的重要相关调节基因。

1 材料与amp;方法

1.1 材料

1.1.1 菌株

类植物乳杆菌 L-ZS9 由中国农业大学食品科学与营养工程学院李平兰教授赠送。

1.1.2 主要试剂和仪器

实验所用普通化学试剂和生化药品,生工生物工程(上海)股份有限公司;RNA 提取试剂 Trizol, Invitrogen 公司;Ribo-Zero rRNA Removal Kit, Illuminia 公司;MiniElute PCR Purification 试剂盒、QIAquick 试剂盒和酶, QIAGEN 公司。台式高速冷冻离心机, Sigma 公司;2100 Bioanalyzer, Agilent 公司。

1.2 方法

1.2.1 材料处理

将 *L. paraplantarum* L-ZS9 以 1% 的比例接种到 MRS^[13] 液体培养基中(终浓度为 10^7 CFU/mL), 37 °C 静置培养 20 h, 然后 4 °C、10 000 r/min 离心 10 min, 收集菌体沉淀备用(试样 1)。在 45 °C、6% 乙醇(MRS 液体培养基中终浓度)、0.10 g/L 的 NaNO₂ (MRS 液体培养基中终浓度)条件下分别培养 *L. paraplantarum* L-ZS9, 分别收集各个条件下培

养 20 h 菌体, 按菌体量 1:1:1 混合, 提取 RNA 备用 (试样 2)。

1.2.2 RNA 的提取和文库构建

总 RNA 提取采用 Trizol 试剂方法, 经过检测符合转录组 RNA 检测标准(RNA 总量 ≥ 6 g, OD_{260}/OD_{280} 在 1.8–2.1 范围, rRNA 比率 23S:18S $\geq 1.5:1$), 通过甲醛变性胶检测总 RNA 的完整性。

样品总 RNA 用 Ribo-Zero rRNA Removal Kit 去除 rRNA, 然后将 mRNA 打断成短片段, 以 mRNA 为模板, 用六碱基随机引物合成第一条 cDNA 链, 加入缓冲液、dATP、dGTP、dCTP、dUTP、RNase H 和 DNA polymerase I 合成第二条 cDNA 链。在经过 QIAquick 试剂盒纯化后做末端修复。加 PolyA 并连接测序接头, 然后用 UNG 酶消化第二条 cDNA 链, 连接产物经 MiniElute PCR Purification 试剂盒纯化后进行 PCR 扩增, 建好测序文库, 质检合格后进行测序。

1.2.3 测序数据分析

测序结果的原始数据(Raw reads)中有些是低质量的, 使用 Q20 (碱基准确率达到 99.9%)的质控标准。使用短 Reads 比对软件 SOAPaligner/soap2^[14] 将 Clean reads 分别比对到参考基因组和参考基因序列, 并统计出比对结果。

转录本预测。利用 Trinity 和 Velvet 软件对过滤后的 Clean reads 进行从头组装^[15]。利用 Bowtie (<http://bowtie.bio.sour.cegenome.net/index.shtml>) 将 100 bp 能重叠的 Reads 拼接^[16], 拼接形成的片段即为 Contig 组装片段。利用 Paired-end reads 确定同一转录本的不同 Contig 及其距离, 将这些 Contig 连接, 获得两端不能延长的序列。挑选出长度大于 150 bp 且平均覆盖度大于 2 的转录区域, 再从中找出位于

基因间区域(一个基因 3'端下游 200 bp 到下一个基因 5'端上游 200 bp 之间的区域)的潜在转录本作为候选的新转录本。

功能注释和 ORF 预测。将转录本与 NCBI 中非冗余蛋白数据库(Nr) (<http://www.nlm.nih.gov/>)、去冗余蛋白数据库(SwissProt)、蛋白质直系同源簇^[17-18]、KEGG^[19]比对。利用 BLASTn 将转录本与 NCBI 非冗余核酸数据库(Nt)比对(E -value cutoffs $\leq 10^{-5}$), 得到相似度最高的蛋白信息对其进行注释。利用 InterProScan 软件对蛋白功能进行注释, BLAST2GO 将注释的蛋白映射到相应的 GO (基因本体)中^[20]。利用 WEGO 软件将 GO 功能进行分类^[21]。基因表达量的计算使用 RPKM 法^[22], 参照 Audic 等^[23]基于测序的差异基因检测方法, 利用基因与目标基因丰度的比值来判读基因表达升降。

2 结果与分析

2.1 测序特征评价

2.1.1 转录组与参考基因组比对

使用短 Reads 比对软件将两个样品的序列比对到类植物乳杆菌 L-ZS9 基因组中。其中能够完全比对上基因组的 Reads 被称为 Perfect match reads, 与基因组比对不超过 2 个碱基错配的 Reads 被称为 Mismatch reads, 这两者统称为 Mapped reads; 能够比对到基因组多个位置的被称为 Multi-position matched reads, 这种 Reads 可信度较低, 一般不用来作基因表达差异分析; 仅能够比对到基因组唯一位置的被称为 Unique matched reads, 可信度高。因此统计 Unique matched reads 用于后续分析, 以保证分析结果的可信度(表 1)。从表 1 可以看出 Multi-position match 比例非常低, 说明测序结果质量比较好。

表 1 样品和参考基因比对的统计结果

Table 1 Reads number based on the RNA-Seq data in two libraries

Sample	Total reads	Mapped reads	Multi-position match	Unique match
Sample 1	8 962 774	7 761 215 (86.59%)	958 (0.01%)	7 760 257 (86.58%)
Sample 2	8 985 876	7 792 327 (86.72%)	1 853 (0.02%)	7 790 474 (86.70%)

2.1.2 测序随机性评价

由于不同参考基因有不同长度,把 Reads 在基因上的位置标准化到相对位置,然后统计基因的 Reads 不同位置比对上的数。如果打断随机性好,在基因组各部位应分布得比较均匀。从图 1 可以看出,类植物乳杆菌 L-ZS9 两个样品的转录组文库打断的随机性较好,测序获得序列在基因组各部位分布较均匀,说明测序样品的随机性符合预期。

2.2 基因覆盖度分析

基因覆盖度是指每个基因被 Reads 覆盖的百分比,其值等于基因中 Unique mapping reads 覆盖的碱基数跟基因编码区所有碱基数的比值。类植物乳杆菌 L-ZS9 产细菌素的试样 1 有 7 760 257 个 Unique match reads 匹配到其基因组,共覆盖了 2 683 个基因,覆盖率为 90.2% (图 2); 不产细菌素的试样 2,有 7 790 474 个 Unique match reads 匹配到其基因组,共覆盖了 2 695 个基因,覆盖率为

94.3% (图 2)。这表明 RNA-Seq 测序技术具有较高的灵敏度,可以覆盖绝大多数类植物乳杆菌 L-ZS9 基因组,为研究细菌素合成调节机制和环境胁迫影响提供了可能。

2.3 新转录本预测

通过 Reads 的分布以及参考基因的注释集合,在类植物乳杆菌 L-ZS9 两个样品的转录组中分别检测出了 6 090 和 5 953 个新转录本(表 2 只列出 5 个)。

2.4 差异基因表达分析

参照 Audic 等^[23]基于测序的差异基因检测方法,利用基因与目标基因丰度的比值来判读基因表达升降。在错误校验率(False discovery rate)<0.001 条件下,统计表达量差异在 2 倍以上的基因,基因表达量的计算使用 RPKM 法。在各种条件下,产细菌素样品和不产细菌素样品相比有 927 个(34.5%)基因表达发生变化,其中 744 个(27.7%)上调,183 个(6.8%)下调(图 3)。

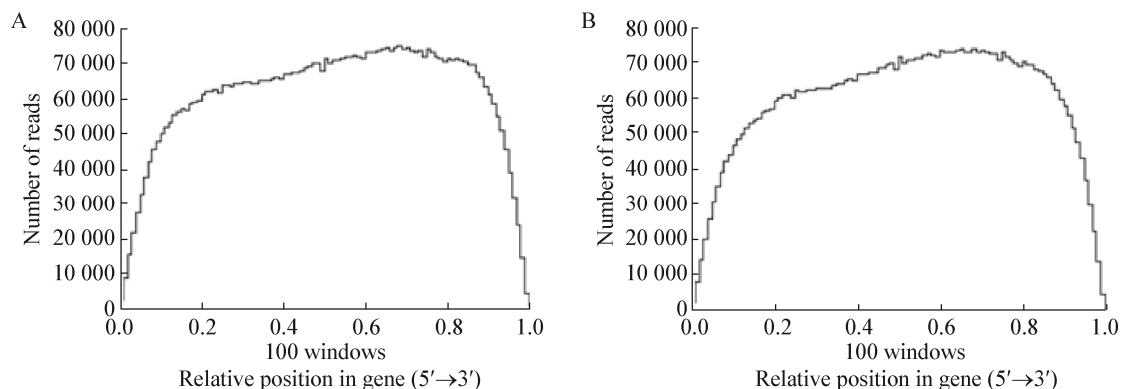


图 1 Reads 在参考基因组上的分布

Figure 1 Distribution statistics of reads mapped to genome

Note: A: Sample 1; B: Sample 2.

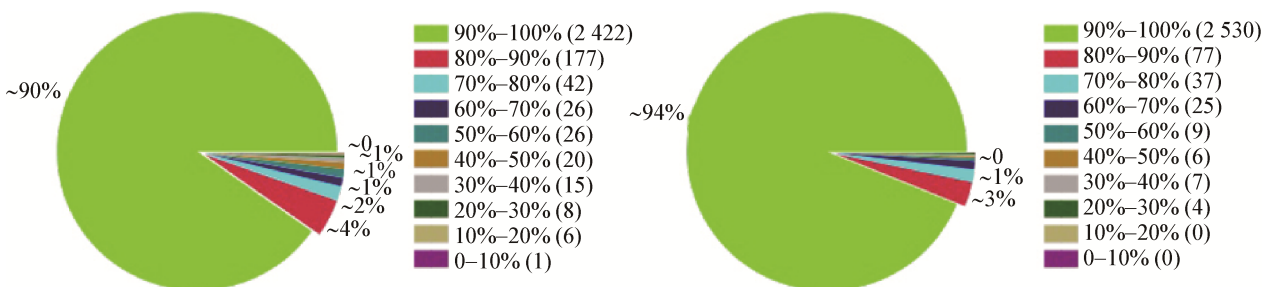


图 2 基因覆盖度结果

Figure 2 The result of distribution of genes' coverage

Note: A: Distribution of genes' coverage (sample 1); B: Distribution of genes' coverage (sample 2).

表 2 试样 1 中检测到的新转录本

Table 2 Novel transcripts detected in sample 1

新转录本编号 Novel TU ID	染色体 Chromosome	起始 Start	终止 End	长度 Length (bp)
TU1	KN173683.1	6	1 057	1 052
TU2	KN173683.1	7	994	988
TU3	KN173684.1	2	1 060	1 059
TU5	KN173685.1	1	1 113	1 113
TU7	KN173686.1	4	1 079	1 076

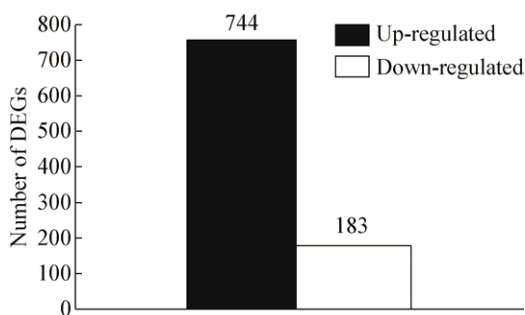


图 3 差异表达基因统计图

Figure 3 Statistical diagram of differentially expressed genes

注: 样品 1 与样品 2 相比.

Note: Sample 1 vs Sample 2.

2.5 GO 功能注释

利用 Gene Ontology 数据库 (<http://www.geneontology.org/>)分析所有差异表达基因项。GO 功能分析可分为三个部分: 分子功能(Molecular function)、生物过程(Biological process)和细胞组分(Cellular component)。为了获得基因完整的功能信息, 对差异表达基因进行 GO 功能分析, 显示在鉴定的 2 695 个类植物乳杆菌基因中, 有 961 个 GO 条目注释到 39 个功能类别(图 4)。如图 4 所示, 只发生基因表达上调的基因中有属于生物过程类的细胞黏连、发育过程、免疫应答过程、正负调节过程、多器官发生过程、生殖过程基因; 属于细胞组分的细胞外基质和细胞器组分基因; 属于分子功能类的 D-丙氨酰载体和电子传递载体基因; 而属于细胞组分的有关病毒粒基因只发生下调。其他类型基因都有上调发生和下调发生。

在生物体内, 不同基因相互协调行使其生物学功能, 通过 Pathway 显著性富集能确定差异表达基因参与的最主要生化代谢途径和信号转导途径。本实验中类植物乳杆菌的 2 个样品通过映射到参考标准途径得到差异基因显著性富集的 Pathway (表 3), 649 个差异表达基因匹配到多个 KEGG pathways, 其中最具代表性的富集性通路有“ABC 转运途径”(ko02010)。虽然“复杂环境下的微生物代谢”(ko01120)和双组分系统(ko02020) P 值大于 0.05, 但是这两个代谢途径却有大量基因表达发生变化, 其中“复杂环境下的微生物代谢”差异表达基因比例为 14.84%, 排在第一位; 而“双组分系统”差异表达基因比例为 6.41%, 列在第四位, 这些注释信息为研究类植物乳杆菌细菌素合成调节途径提供了有价值的资源。

2.6 类植物乳杆菌调节细菌素合成重要相关基因分析

类植物乳杆菌是研究群体感应和细菌素合成的一种重要微生物, 实验已经表明类植物乳杆菌(L-ZS9)试样 1 可以正常产生细菌素, 而试样 2 受环境条件影响不产生细菌素, 因此可以通过研究这两个试样基因表达的差异来分析细菌素合成的重要相关基因。对类植物乳杆菌转录组数据进行差异表达基因 Pathway 富集性分析, 结果表明(表 4), 有 68 条转录本归类于“ABC 转运系统”, 其中, HR47_00285、HR47_08345、HR47_08340 的表达量上调超过 16 倍, HR47_00775 基因表达量降低到原来表达量的 1/16 以下。为进一步分析差异表达基因具体变化情况, 统计了上调超过 16 倍的基因(表 4)。HR47_00190、HR47_13845、HR47_08350、HR47_10850、HR47_09160 等几个膜蛋白基因表达量上调超过了 16 倍, 同时发现多个编码未知功能蛋白的孤儿基因表达量也大量上调, 如 HR47_00435、HR47_00255、HR47_01180、HR47_10460、HR47_10475, 其中 HR47_00255 表达量上调达上万倍, 但是 HR47_00010 表达量降低超过了 1/16。

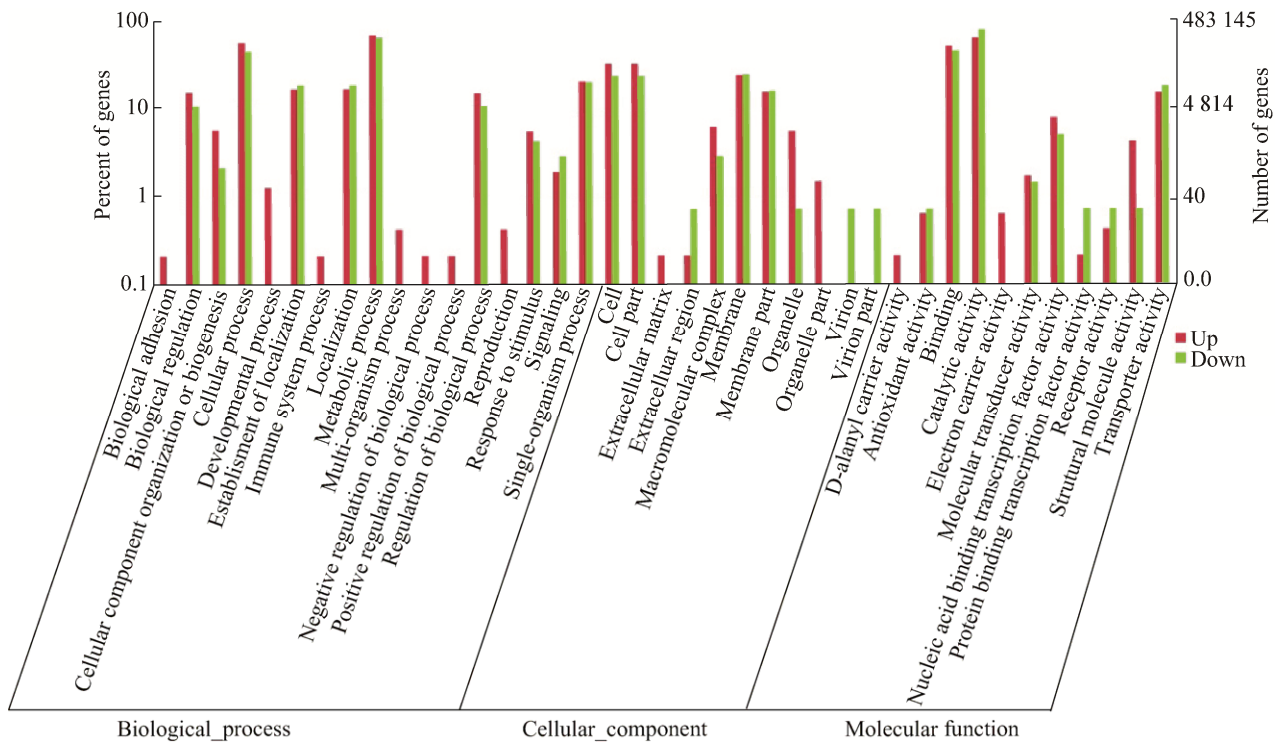


图 4 GO 功能分析结果

Figure 4 GO analyses of DEGs

表 3 显著性富集的代谢途径(P<0.05)

Table 3 List of pathway enrichment analysis (P<0.05)

KEGG pathway	DEGs with pathway annotation	P-value	Q-value	Pathway ID
ABC transporters	68 (10.48%)	0.000 126 273 0	0.015 026 49	ko02010
Valine, leucine and isoleucine biosynthesis	9 (1.39%)	0.027 049 430 0	0.702 572 60	ko00290
Valine, leucine and isoleucine degradation	6 (0.92%)	0.038 663 750 0	0.702 572 60	ko00280
<i>Staphylococcus aureus</i> infection	5 (0.77%)	0.041 774 650 0	0.702 572 60	ko05150
Beta-Alanine metabolism	4 (0.62%)	0.042 300 770 0	0.702 572 60	ko00410

表 4 表达量变化超过 16 倍的统计结果

Table 4 Statistics result of different gene expression level beyond 16 times

基因编号 Gene ID	变化倍数对数 Expression difference using log ₂ ratio	基因编号 Gene ID	变化倍数对数 Expression difference using log ₂ ratio
HR47_00190	4.418 134 339 544 83	HR47_08345	5.220 532 859 843 22
HR47_06130	4.781 836 091 164 48	HR47_06430	4.434 185 431 615 51
HR47_00285	4.129 330 850 087 09	HR47_13845	5.084 699 924 424 92
HR47_00435	4.849 222 930 894 35	HR47_00255	14.267 100 008 587 40
HR47_00795	5.517 601 439 803 14	HR47_01180	12.942 417 802 063 30
HR47_07915	13.424 880 389 355 90	HR47_08320	4.501 299 627 474 04
HR47_08350	4.353 224 515 097 90	HR47_09160	13.442 491 405 197 90
HR47_08340	5.367 147 313 223 75	HR47_09265	4.820 602 252 954 58
HR47_10425	14.520 974 415 135 70	HR47_10460	14.397 891 407 131 90
HR47_10475	5.565 429 964 893 76	HR47_10500	4.565 429 964 893 76
HR47_10710	4.175 723 755 571 54	HR47_10850	5.926 732 298 115 27
HR47_00010	-4.658 571 709 304 35	HR47_00775	-4.139 244 673 238 02

3 讨论与结论

目前虽然对类植物乳杆菌已经有多篇报道^[6,24], 但是从转录组水平对类植物乳杆菌研究的报道甚少。采用 Illumina HiSeqTM 2000 高通量测序技术对类植物乳杆菌转录组进行测序, 类植物乳杆菌 L-ZS9 两个样品的转录组文库打断随机性较好, 测序获得的基因各部位分布较均匀; Multi-position match 比例非常低, 仅分别在 0.01% 或 0.02% 的水平; Unique match reads 匹配到其基因组, 基因覆盖率都在 90% 以上。说明测序结果质量较好, 利用这些数据研究类植物乳杆菌环境胁迫影响细菌素合成相关调节基因的表达特征是可靠的。

虽然类植物乳杆菌 L-ZS9 完成了基因组测序, 但是各种胁迫条件影响细菌素合成的途径和分子机制还不清楚。我们通过对产细菌素样品和胁迫条件下不产细菌素样品的转录组比对, 发现有 927 个 (34.5%) 基因表达发生变化, 其中 744 个 (27.7%) 上调, 183 个 (6.8%) 下调, 说明类植物乳杆菌 LZS-9 在胁迫条件下利用多个基因表达变化来应答这一过程。其中基因表达上调主要发生在细胞黏连、免疫系统过程、生殖过程、细胞外基质、D-丙氨酰载体和电子传递载体等系统中。

进一步分析显示 649 个差异表达基因虽然可以匹配到多个 KEGG 代谢途径, 最具代表性的富集性通路是“ABC 转运途径”(ko02010), 其中 HR47_00285、HR47_08345、HR47_08340 三个基因的表达量上调超过 16 倍, HR47_00775 基因表达量降低到原来表达量 1/16 以下, 这与在乳酸菌细菌素合成过程中“ABC 转运途径”参与细菌素的加工与分泌^[25-26]相一致。虽然 KEGG pathways 方法分析“双组分途径基因”表达变化的 *P* 值在 0.05 以上, 但是“双组分途径”按差异表达基因比率计算排在第四位, 而且多个基因如 HR47_01535、HR47_07340、HR47_03375、HR47_00485 表达量上升也超过了 8 倍。从以上两个方面看出“ABC 转运途径”和“双组分途径”在类植物乳杆菌 L-ZS9 细菌素合成中发

挥着重要作用, 这为下一步研究这几个基因如何影响细菌素合成奠定了基础。另外多个孤儿基因如 HR47_00435、HR47_00255、HR47_01180、HR47_10460、HR47_10475 表达量也发生大幅度的变化, 有的变化幅度甚至达上万倍, 提示类植物乳杆菌 L-ZS9 细菌素合成调节或许还存在其他调节机制。这些表达变化差异很大的孤儿基因的功能尚需要进一步鉴定。

REFERENCES

- [1] Liu CJ, Quan CS, Fan SD, et al. Properties of an antimicrobial substance from *Lactobacillus paraplantarum* and application of the fermented solution on pork preservation[J]. Food Research and Development, 2009, 30(8): 152-155 (in Chinese)
刘长建, 权春善, 范圣第, 等. 类植物乳杆菌抑菌物质的性质及其在猪肉保鲜中的作用[J]. 食品研究与开发, 2009, 30(8): 152-155
- [2] Zhang X, Zhao B, Zhang XM, et al. Screening of bacteriocin producing lactic acid bacteria and the analysis of bacteriocin related genes[J]. Journal of China Agricultural University, 2013, 18(4): 168-177 (in Chinese)
张旭, 赵斌, 张香美, 等. 产细菌素乳酸菌的筛选及细菌素相关基因的分析[J]. 中国农业大学学报, 2013, 18(4): 168-177
- [3] Zhang X. Action mode of quorum sensing regulated IIB bacteriocin and producing bacteria application on sturgeon sausage[D]. Beijing: Master's Thesis of China Agricultural University, 2014 (in Chinese)
张旭. 群体感应介导的 IIB 类细菌素的抑菌机理及产生菌在鲟鱼肠中的应用[D]. 北京: 中国农业大学硕士学位论文, 2014
- [4] Wang Y, Sun Y, Zhang X, et al. Bacteriocin-producing probiotics enhance the safety and functionality of sturgeon sausage[J]. Food Control, 2015, 50: 729-735
- [5] Zhou K, Zeng YT, Han XF, et al. Modelling growth and bacteriocin production by *Lactobacillus plantarum* BC-25 in response to temperature and pH in batch fermentation[J]. Applied Biochemistry and Biotechnology, 2015, 176(6): 1627-1637
- [6] Zhang XM, Li PL. Effects of PlnA on bacteriocin production in *Lactobacillus paraplantarum*: influence of environmental factors[J]. Microbiology China, 2013, 40(9): 1624-1630 (in Chinese)
张香美, 李平兰. 环境因素对 PlnA 诱导类植物乳杆菌产生细菌素效果的影响[J]. 微生物学通报, 2013, 40(9): 1624-1630
- [7] Liu L, Li PL. Complete genome sequence of *Lactobacillus paraplantarum* L-ZS9, a probiotic starter producing class II bacteriocins[J]. Journal of Biotechnology, 2016, 222: 15-16
- [8] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics[J]. Nature Reviews Genetics, 2009, 10(1): 57-63

- [9] Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities[J]. *Nature Reviews Genetics*, 2011, 12(2): 87-98
- [10] Wang F, Li Y, Wu XS, et al. Transcriptome analysis of coding and long non-coding RNAs highlights the regulatory network of cascade initiation of permanent molars in miniature pigs[J]. *BMC Genomics*, 2017, 18(1): 148
- [11] Hao M, Li A, Shi TW, et al. The abundance of homoeologue transcripts is disrupted by hybridization and is partially restored by genome doubling in synthetic hexaploid wheat[J]. *BMC Genomics*, 2017, 18(1): 149
- [12] Nagalakshmi U, Wang Z, Waern K, et al. The transcriptional landscape of the yeast genome defined by RNA sequencing[J]. *Science*, 2008, 320(5881): 1344-1349
- [13] de Man JC, Rogosa M, Sharpe ME. A medium for the cultivation of *Lactobacilli*[J]. *Journal of Applied Bacteriology*, 1960, 23(1): 130-135
- [14] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. *Bioinformatics*, 2009, 25(14): 1754-1760
- [15] Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. *Nature Biotechnology*, 2011, 29(7): 644-652
- [16] Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory efficient alignment of short DNA sequences to the human genome[J]. *Genome Biology*, 2009, 10(3): R25
- [17] Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families[J]. *Science*, 1997, 278(5338): 631-637
- [18] Tatusov RL, Fedorova ND, Jackson JD, et al. The COG database: an updated version includes eukaryotes[J]. *BMC Bioinformatics*, 2003, 4: 41
- [19] Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: new developments in KEGG[J]. *Nucleic Acids Research* 2006, 34(S1): D354-D357
- [20] Zdobnov EM, Apweiler R. InterProScan-an integration platform for the signature-recognition methods in InterPro[J]. *Bioinformatics*, 2001, 17(9): 847-848
- [21] Ye J, Fang L, Zheng HK, et al. WEGO: a web tool for plotting GO annotations[J]. *Nucleic Acids Research*, 2006, 34(S2): W293-W297
- [22] Mortazavi A, Williams BA, McCue K, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq[J]. *Nature Methods*, 2008, 5(7): 621-628
- [23] Audic S, Claverie JM. The significance of digital gene expression profiles[J]. *Genome Research*, 1997, 7(10): 986-995
- [24] Sun ZH, Harris HM, McCann A, et al. Expanding the biotechnology potential of *Lactobacilli* through comparative genomics of 213 strains and associated genera[J]. *Nature Communications*, 2015, 6: 8322
- [25] Nes IF, Diep DB, Håvarstein LS, et al. Biosynthesis of bacteriocins in lactic acid bacteria[J]. *Antonie van Leeuwenhoek*, 1996, 70(2/4): 113-128
- [26] Zhang XM, Li PL. Quorum sensing in class II bacteriocin-producing lactic acid bacteria and its application-a review[J]. *Acta Microbiologica Sinica*, 2011, 51(9): 1152-1157 (in Chinese)
- 张香美, 李平兰. 产II类细菌素乳酸菌群体感应及其应用[J]. *微生物学报*, 2011, 51(9): 1152-1157