

DOI:10.13350/j.cjpb.220404

• 论著 •

鼠伤寒沙门菌 STM14_3199 过表达菌株的构建及转录组测序分析*

王琦, 李冰清**, 岳盈盈**

(山东第一医科大学(山东省医学科学院)基础医学院(基础医学研究所)病原生物学系, 山东济南 250000)

【摘要】 **目的** 通过转录组测序技术比较鼠伤寒沙门菌未知功能蛋白 STM14_3199 过表达菌株(WT_3199)和鼠伤寒沙门菌 ATCC14028 株(WT)各基因转录的差异,为进一步探究 STM14_3199 的功能提供方向。**方法** PCR 扩增 STM14_3199 基因,插入表达载体 pBAD24,电转化至 WT 中,获得 STM14_3199 过表达菌株 WT_3199,将其与 WT 在相同的营养条件下培养至对数生长期,制样后进行转录组测序分析。**结果** 经 PCR、SDS-PAGE 电泳及测序验证证明过表达菌株构建成功,对转录组数据进行质控显示样本组数据合格,统计结果中的差异基因,发现共有 1 906 个表达量出现差异的基因,其中上调 887 个,下调 1 019 个。将差异基因按照功能分类,并注释到 GO 及 KEGG 数据库中富集分析,发现这些差异基因涉及许多通路,影响体内多种功能的发挥。**结论** 成功构建了 STM14_3199 过表达菌株 WT_3199。WT_3199 中表达量上调的基因主要集中在碳源利用、精氨酸的降解及合成等途径,部分下调基因在柠檬酸盐的利用以及精氨酸的分解代谢过程中发挥功能。GO 富集分析表明差异基因主要与代谢过程、胞内组分以及一些酶的催化活性等有关,KEGG 富集分析表明差异基因主要富集在与碳代谢、氧化磷酸化和三羧酸循环等相关通路中。这些分析均提示我们 STM14_3199 可能在这些通路中发挥作用,为进一步研究其功能奠定了基础。

【关键词】 鼠伤寒沙门菌;STM14_3199 蛋白;过表达;转录组学

【中图分类号】 R378.23

【文献标识码】 A

【文章编号】 1673-5234(2022)04-0393-07

[*Journal of Pathogen Biology*. 2022 Apr;17(4):393-399,405.]

Construction of STM14_3199 overexpression strain in *Salmonella typhimurium* and transcriptome analysis

WANG Qi, LI Bing-qing, YUE Ying-ying (Department of Pathogen Biology, School of Basic Medicine, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan 250000, China)**

【Abstract】 **Objective** Comparing the transcription differences of *Salmonella typhimurium* ATCC14028 strain and the STM14_3199 overexpression strain by transcriptome sequencing technology. In order to provide directions for further research of STM14_3199 protein. **Methods** STM14_3199 gene was amplified by PCR, inserted into expression vector pBAD24, and electro transformed into WT strain to obtain STM14_3199 overexpression strain WT_3199, then culture them to logarithmic growth stage under the same nutritional conditions. After sample preparation, transcriptome sequencing was performed. **Results** After verification by PCR, SDS-PAGE electrophoresis and sequencing, the overexpression strain was successfully constructed, and quality control of the transcriptome data showed that the sample set data were qualified. The differential genes in the results were counted, and a total of 1 906 genes with differential expression were found, of which 887 were up-regulated and 1 019 were down-regulated. The differential genes were classified according to function and annotated into GO and KEGG databases for enrichment analysis, and it was found that these differential genes were involved in many pathways, affecting multiple functions *in vivo*. **Conclusion** We successfully constructed the STM14_3199 overexpression strain WT_3199. The up-regulated genes in WT_3199 were mainly in carbon utilization, degradation and synthesis of arginine, while some of the down-regulated genes functioned in citrate utilization and catabolism of arginine. The GO enrichment analysis showed that the differential genes were mainly related to metabolic processes, intracellular components and catalytic activities of some enzymes. KEGG enrichment analysis showed that the differential genes were mainly enriched in the pathways related to carbon metabolism, oxidative phosphorylation and tricarboxylic acid cycle. These analyses suggest that STM14_3199 may play a role in these pathways, laying the foundation for further studies of their function.

【Key words】 *Salmonella typhimurium*; STM14_3199 protein; overexpression; transcriptome

* **【基金项目】** 国家自然科学基金项目(No. 31800054, 32170034)。

** **【通讯作者】** 李冰清, E-mail: bingqingsdu@163.com; 岳盈盈, E-mail: yyy0513@163.com

【作者简介】 王琦(1997-), 女, 山东人, 硕士在读。主要研究方向: 微生物致病蛋白的结构和功能。E-mail: 15963743209@163.com

沙门菌是一类以人和动物为宿主的肠道致病菌,在自然界中分布极其广泛,可在世界范围内引起严重的公共卫生问题^[1]。根据沙门菌感染后致病的情况不同,可以将其大体分为伤寒沙门菌和非伤寒沙门菌。其中,非伤寒沙门菌主要引起食源性疾病,而伤寒沙门菌则会导致伤寒,包括伤寒和副伤寒^[2]。鼠伤寒沙门菌(*Salmonella typhimurium*, STM)是沙门菌的典型致病血清型,是一种重要的人畜共患病病原菌,同时也是临床上常见的致病菌^[3]。鼠伤寒沙门菌主要经口传播感染人体,人群普遍易感,易造成暴发性食物中毒和医院内感染,可导致严重并发症并且危及生命。因此,针对鼠伤寒沙门菌的研究无论在医学领域还是公共卫生事业领域都具有十分重要的意义^[4]。

一般情况下,细菌被巨噬细胞吞噬后会与溶酶体融合进而被杀伤^[5],而鼠伤寒沙门菌则可通过形成沙门菌囊泡 SCV(*Salmonella containing vacuole*)^[6-7]的方式免于被攻击,从而生存下来^[8-9]。研究发现,鼠伤寒沙门菌不仅可以在巨噬细胞中生长繁殖,还可以巨噬细胞为宿主向肠外组织播散^[10],进而累及全身,这也是鼠伤寒沙门菌出现系统性感染和严重并发症的重要原因^[11]。

然而,在被巨噬细胞吞噬后,鼠伤寒沙门菌应对巨噬细胞内的种种不利因素的机制仍不清楚。为了进一步了解鼠伤寒沙门菌在体内的生存机制,作者前期对被巨噬细胞吞噬前后的鼠伤寒沙门菌进行了转录组测序分析,发现被吞噬后细菌中许多基因的转录水平发生了不同程度的上调或下调^[12]。在这些基因编码的蛋白中,除了已经命名的蛋白外,仍然存在许多功能未知的蛋白。通过对这些未知功能蛋白的筛选,最终确定了几个在沙门菌入侵后广泛表达的蛋白进行后续实验,希望从中找到与鼠伤寒沙门菌的感染及致病有较大相关性的蛋白,为感染的临床治疗提供新的思路与方法。其中,STM14_3199 基因及其编码的蛋白倍受关注。

STM14_3199 蛋白的相对分子质量仅为 7 ku,是一个只有 62 个氨基酸、无同源结构的微小蛋白(microprotein)。在沙门菌入侵细胞后,STM14_3199 基因拷贝数由入侵前的 191.093 上升至 1 810.59,提高近 10 倍。有研究表明,编码 microprotein 的序列广泛存在于动植物基因组中,但目前对于这类蛋白知之甚少,相关研究还处于探索阶段。在鼠伤寒沙门菌 ATCC 14028 菌株中,STM14_3199 有一个在氨基酸序列水平上同源性高达 75% 的蛋白 STM14_1447,然而有关 STM14_1447 蛋白的功能也不十分清楚。有趣的是,与 STM14_3199 相邻的几个基因所编码的功能已知蛋白如 STM14_3194、STM14_3195、STM14_

3196、STM14_3203、STM14_3207 以及 STM14_3208 等均为噬菌体蛋白,而与 STM14_1447 相邻的基因编码的蛋白中也有许多噬菌体蛋白。此外,STM14_3199 似与 IsrI 蛋白有着某种未知的联系。IsrI 是一种存在于鼠伤寒沙门菌 ATCC14028、SL1344 和 LT2 株 Gifsy-1 前噬菌体上的能够编码小 RNA 的蛋白质,预测 STM14_3199 是 IsrI 的开放阅读框,对 IsrI 编码小 RNA 的能力有着至关重要的影响^[13]。这些线索都为探索 STM14_3199 蛋白的功能提供了可靠的方向。为深入了解 STM14_3199 蛋白具体作用通路以及其影响这些通路的分子机制,本研究对过表达 STM14_3199 菌株 WT_3199 与标准株 WT 进行转录组学测序分析,以进一步探索其功能。

转录组是指在一定的生长阶段或条件中一个细胞或组织所有转录的集合,包括非编码 RNA、rRNA、tRNA 以及 mRNA^[14]。原核转录组学是在 RNA 水平上研究不同条件下基因表达情况的变化,从而分析不同条件对原核生物的影响^[15]。近年来基因测序技术迅猛发展,RNA-seq 凭借高通量、应用广泛、灵敏度高、分辨率高、成本较低等优点逐渐成为了转录组学研究的主要方法^[16-17]。通常先在样品中提取总 RNA,根据 RNA 种类分离纯化去除 rRNA 后得到 mRNA,再将其反转录成双链 cDNA,对 cDNA 进行末端修复,加尾接头后选择合适的 cDNA 片段进行 PCR 扩增,最终获得 cDNA 文库^[18],建库完成并检测合格后即可进行高通量测序^[19],测序的方法主要有 Illumina IG, HiSeq 2000, ABI SO1ID 等,本实验采用的是边合成边测序的 Illumina 法。

本研究通过同源重组技术,构建鼠伤寒沙门菌 STM14_3199 基因过表达菌株 WT_3199^[20],并通过转录组测序技术比较野生菌株 WT 与 WT_3199 在完全相同的营养条件下基因组转录水平的差异和富集分析,为探究 STM14_3199 基因的功能提供思路及研究方向。

材料与方法

1 材料

1.1 菌株与质粒 鼠伤寒沙门菌 ATCC 14028 购自美国 ATCC 菌种库;大肠埃希菌 DH5 α 及表达载体 pBAD24 均为本实验室保存。

1.2 主要仪器与试剂 ECM830 电穿孔仪购于美国 BTX 公司;GS-900 校准型光密度仪及 Powerpac basic 基础电泳仪购于美国 Bio-Rad 公司;紫外凝胶成像分析系统 Essential V6 购于美国 Uvitec 公司;DYY-6C 电泳仪购于北京六一生物科技有限公司;质粒小提试剂盒(TIANprep Mini Plasmid Kit),通用型 DNA 纯

化回收试剂盒(Universal DNA Purification Kit)及 TRNzol Universal 总 RNA 提取试剂(TRNzol Universal Reagent)购自天根生化科技(北京)有限公司; DL2000 DNA Marker 购自宝日医生物技术(北京)有限公司;限制性内切酶 EcoRI 和 HindIII, T4 DNA 连接酶, Veriti 96 孔热循环仪及 NanoDrop™ One^C 微量紫外-可见光分光光度计均购自赛默飞世尔科技(中国)有限公司; 180ku prestained protein marker 购自南京诺维赞生物科技有限公司; KOD-Plus-Neo 高保真 PCR 酶购自日本 TOYOBO 公司。

2 方法

2.1 WT₃₁₉₉ 菌株的构建

2.1.1 引物设计与合成 根据 STM14₃₁₉₉ 的基因序列设计一对特异性引物(表 1), 其中上游引物引入保护碱基和 EcoRI 酶切位点, 下游引物引入保护碱基和 HindIII 酶切位点, 由铂尚生物技术(上海)有限公司合成。

表 1 PCR 扩增引物
Table 1 Primers used in PCR amplification

引物名称 Segment	引物序列(5'-3') Sequence of primer	扩增用途 Amplified fragment
STM14 ₃₁₉₉ -1aa-EcoRI-5'	ATAGAATTCATGAAGAAGCCTTCTACTG	上游同源臂
STM14 ₃₁₉₉ -186aa-HindIII-3T	ATAAAGCTTTTAAATTGGAGGTCTAAACGGTCC	下游同源臂

2.1.2 目的基因 STM14₃₁₉₉ 片段的扩增 以鼠伤寒沙门菌 ATCC 14028 菌株核酸作为模板 PCR 扩增目的基因。反应条件: 95 °C 预变性 5 min; 95 °C 变性 30 s, 56 °C 退火 30 s, 72 °C 延伸 30 s, 共 30 个循环; 72 °C 再延伸 10 min。PCR 产物进行 1% 琼脂糖凝胶电泳鉴定, 使用通用型 DNA 纯化回收试剂盒切胶回收。

2.1.3 STM14₃₁₉₉-pBAD24 重组质粒的构建、转化及鉴定 将 pBAD24 质粒和 PCR 产物分别使用 EcoRI 和 HindIII 进行双酶切。酶切产物进行 1% 琼脂糖凝胶电泳并切胶回收, 回收后的 PCR 产物和 pBAD24 质粒用 T4 连接酶 16 °C 过夜连接, 连接产物转入大肠埃希菌 DH5 α 态细胞中, 转化菌均匀涂布在氨苄抗性的 LB 固体培养基平板上于 37 °C 恒温箱过夜培养。从平板上挑取单菌落至 1 ml 氨苄抗性 LB 培养基中, 37 °C 培养至 A₆₀₀ 为 0.5 时加入 L-阿拉伯糖, 37 °C 诱导 2 h 后取样进行 SDS-PAGE 电泳, 分析蛋白表达情况。阳性菌种送铂尚生物技术(上海)有限公司测序。

2.1.4 STM14₃₁₉₉-pBAD24 重组质粒的提取 将测序正确的菌种转接于带有氨苄抗性的 LB 液体培养基中 37 °C 摇菌培养。用质粒小提试剂盒提取 STM14₃₁₉₉-pBAD24 质粒, -20 °C 保存。

2.1.5 电转化 将鼠伤寒沙门菌 ATCC 14028 菌株

摇菌培养过夜后转接于高盐 LB 培养基(含 15g/L NaCl)中, 37 °C 振荡培养至 A₆₀₀ 为 0.3, 用冰冷无菌去离子水反复清洗 3~5 次, 获得感受态细胞。向感受态细胞中分别加入 pBAD24 及 STM14₃₁₉₉-pBAD24 质粒, 标记为 ATCC14028-pBAD24 和 ATCC14028p3199-pBAD24, 混匀后冰浴 2 min, 转至电转杯中进行电击(电压 2 300 V, 时间 130 μ s, 间隔 200 ms, 共 80 次), 恢复培养 1 h 后均匀涂布至带有氨苄抗性的固体 LB 培养基平板上, 37 °C 培养过夜。第 2 d 挑取单菌落进行阳性重组子鉴定, 阳性菌株置 -80 °C 保存。

2.2 转录组学制样 将鉴定阳性的 ATCC14028-pBAD24 重组菌株(WT)及 ATCC14028p3199-pBAD24 重组菌株(WT₃₁₉₉)摇菌培养过夜, 第 2 d 按 1:100 的比例将菌液转接于带有氨苄抗性的 LB 培养基中, 37 °C、200 r/min 培养 1 h 后, 向培养基中加入终浓度为 0.05% 的 L-阿拉伯糖继续培养至 A₆₀₀ 为 0.5。分别取 1 ml 菌液 12 000 r/min 离心 5 min, 弃上清, 加入 1 ml TRNzol Universal 充分裂解后液氮速冻, -80 °C 冻存。每种菌株设 3 个生物学重复(标记为 WT₁, WT₂, WT₃, WT_{3199_1}, WT_{3199_2}, WT_{3199_3})。

2.3 总 RNA 的提取 将上述样品解冻, 使用 TRNzol Universal 总 RNA 提取试剂盒提取总 RNA, 用 50 μ L RNase-Free ddH₂O 充分吹打溶解, 检测所得总 RNA 溶液的浓度和纯度, 再用 Agilent 2100 bio-analyze 对其完整性及总量进行质控。

2.4 转录组文库构建 由于原核生物 mRNA 无 polyA 尾巴, 所以需去除总 RNA 中的 rRNA 以保证后续试验所得数据的质量, 然后将纯化得到的 mRNA 随机打断成短片段, 逆转录合成 cDNA 并进一步纯化, 对纯化的 cDNA 进行末端修复, 加尾接头, 最后选择合适的片段进行 PCR 扩增, 获得 cDNA 文库, 质检合格后对样品进行转录组测序分析, 由北京诺禾致源生物信息科技有限公司完成。

结果

1 目的基因扩增

以 ATCC 14028 菌株核酸为模板, 用预先设计的特异性引物 PCR 扩增目的基因, 扩增产物进行 1% 琼脂糖凝胶电泳分析, 结果如图 1, 扩增片段约为 186 bp, 与 STM14₃₁₉₉ 基因大小一致。

2 基因过表达菌株 WT₃₁₉₉ 的鉴定

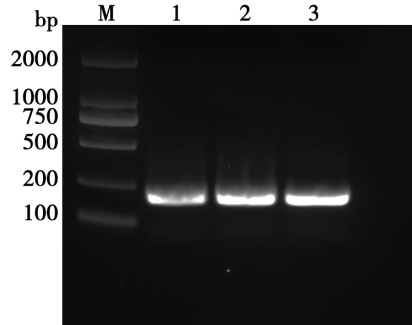
将构建的 WT₃₁₉₉ 重组菌株在普通 LB 中培养, 加入终浓度为 0.05% L-阿拉伯糖诱导 2 h 后离心收菌, 进行 SDS-PAGE 电泳检测, 结果如图 2, 在相应大

小位置有过表达蛋白带,结合测序结果判断 WT_3199 重组菌株构建成功。

2 转录组测序数据质量控制及参考基因组比对分析

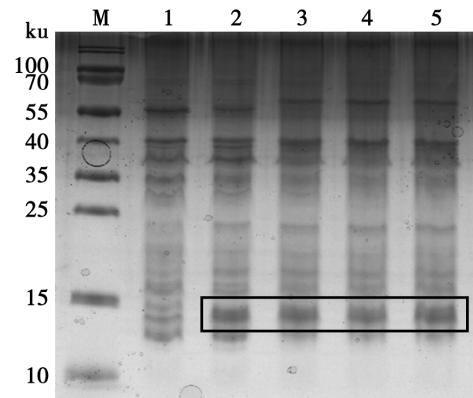
2.1 测序数据质量分析

对原始数据进行过滤、错误率检查、GC 含量检查后汇总,获得过滤后的 reads 数,用于后续分析^[21]。分析结果如表 2,碱基质量大于 20



M DNA 标志物 1~3 目的基因 PCR 产物
图 1 目的基因 PCR 产物 1%琼脂糖凝胶电泳分析
M DNA Marker 1-3 Target fragments
Fig. 1 Identification of target gene

(Q20)的均占 98%以上,碱基质量大于 30(Q30)的均占 95%以上,过滤后 reads 数中 G 和 C 的数量占总碱基数的 53%以上,表明从每个样品中提取的 RNA 序列质量良好。



M 蛋白分子质量标准 1 阴性对照 2~5 WT_3199 重组菌表达产物
图 2 WT_3199 重组菌表达产物的 SDS-PAGE 电泳分析

M Protein Marker 1 Negative control 2-5 WT_3199 recombinants
Fig. 2 The result of recombinants

表 2 样品测序数据质量
Table 2 Quality of sample sequencing data

样品名称 Sample name	原始数据中 reads 数 Raw reads	过滤后 reads 数 Clean reads	过滤后 碱基数 Clean bases	测序错误率 Error rate(%)	Phred 值>20 的碱基占比 Q20(%)	Phred 值>30 的碱基占比 Q30(%)	GC 含量 GC content(%)
WT_1	15798302	15676870	2.35G	0.02	98.40	95.15	53.24
WT_2	15187542	15087686	2.26G	0.02	98.52	95.40	53.23
WT_3	14623002	14455098	2.17G	0.02	98.40	95.11	53.34
WT_3199_1	16149026	16056480	2.41G	0.02	98.42	95.19	53.46
WT_3199_2	15410044	15273880	2.29G	0.02	98.49	95.40	53.50
WT_3199_3	14385080	14252984	2.14G	0.02	98.50	95.43	53.36

2.2 样品与参考基因组比对情况统计

一般来说,如果参考基因组选择合适,且相关实验不存在污染,能够定位到基因组上的测序序列数量占比(Total mapped)

应大于 70%^[22],根据表 3 数据,所有样品成功比对到参考基因组的比例均大于 94%,表明本次实验相关数据可靠。

表 3 样品与参考基因组比对分析
Table 3 Comparison of samples and reference genome

样品名称 Sample name	过滤后 reads 数 Total reads	能定位到基因组上 的 Clean reads 数 Total mapped	占比 (%)	参考序列上有多个比对 位置的 clean reads 数 Multiple mapped	占比 (%)	参考序列上有唯一比对 位置的 Clean reads 数 Uniquely mapped	占比 (%)
WT_1	15676870	14999431	95.68	270620	1.73	14728811	93.95
WT_2	15087686	14458614	95.83	247293	1.64	14211321	94.19
WT_3	14455098	13729144	94.98	241582	1.67	13487562	93.31
WT_3199_1	16056480	15735132	98.00	237261	1.48	15497871	96.52
WT_3199_2	15273880	14938005	97.80	251372	1.65	14686633	96.16
WT_3199_3	14252984	13927260	97.71	409730	2.87	13517530	94.84

3 WT_3199 vs WT 转录组基因序列差异分析

差异表达基因(DEGs)的显著性分析是 RNA-seq 的核心,由于 read count 所表示的数值往往会受到测序深度和基因长度影响,一般采用计算校正后的 FPKM 代替 read count 表示各个样品全部基因的表达

量,通常认为 FPKM 大于 1 时,该基因表达^[23]。对于有生物学重复的实验,采用负二项分布法筛选出满足 $|\log_2(\text{FoldChange})| > 0$ 且 $\text{padj} < 0.05$ 的差异基因。

3.1 差异基因统计

对差异基因进行数量统计和数据筛选后,WT_3199vsWT 共检测出差异基因 1 906

个,其中上调 887 个,下调 1 019 个,火山图(图 3)可以直观地显示 WT_3199vsWT 差异基因的分布情况,其横坐标表示差异基因在 WT_3199 组与 WT 组中表达倍数的变化($\log_2\text{FoldChange}$),纵坐标表示差异基因在 WT_3199 组与 WT 组中差异的显著性水平($-\log_{10}\text{padj}$),红点表示上调的差异基因,绿点表示下调的差异基因。由于邻近的基因可能共同参与同一代谢通路或功能,因此通过对差异基因的聚类分析即可推测该未知基因的功能。聚类分析结果显示 WT_3199 部分基因的转录与 WT 相比存在差异。

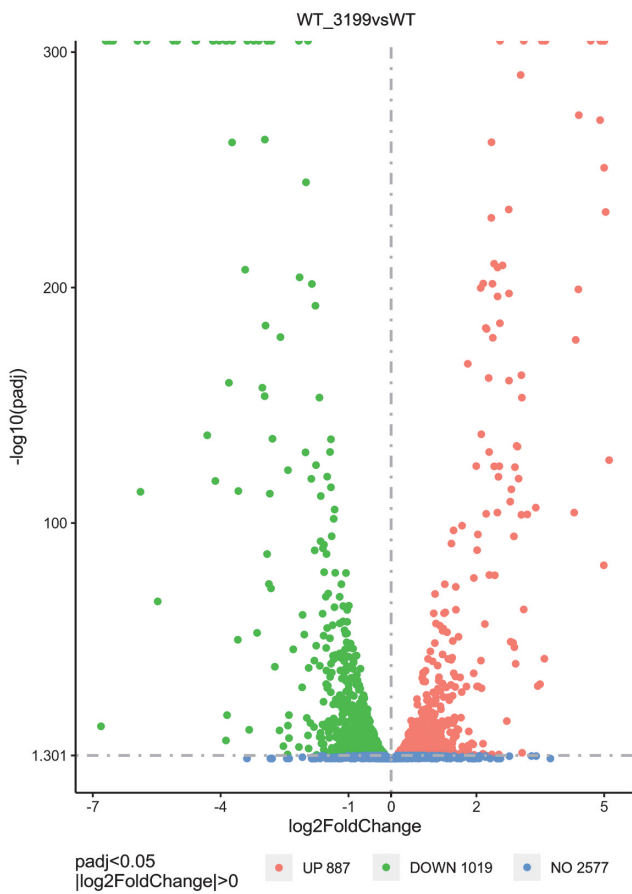


图 3 WT_3199vsWT 差异表达基因火山图
Fig. 3 Volcano plot analysis of WT_3199vsWT

3.2 差异基因聚类 差异基因聚类分析是指将表达模式相近或相同的基因聚在一起,这类基因可能参与共同的生物学过程及代谢通路或可以共同作用发挥一定的功能。热图(图 4)中表达模式相近或相同的基因聚集在一起,红色表示基因的高表达,绿色表示基因低表达,横坐标为样品名称,纵坐标表示差异基因 FP-KM 均一化处理后的数值。

将差异基因聚类分析后的结果按照功能分类(表 4)可以发现,上调的差异基因在乙醇胺利用、无氧呼吸及丙二酸利用等与碳源利用有关的途径差异较为明

显,而在下调的差异基因中与镍铁氢化酶、柠檬酸盐利用相关的基因变化较大,提示 STM14_3199 及其编码的蛋白有可能在碳源利用、精氨酸的降解及生物合成等过程中发挥一定的功能。

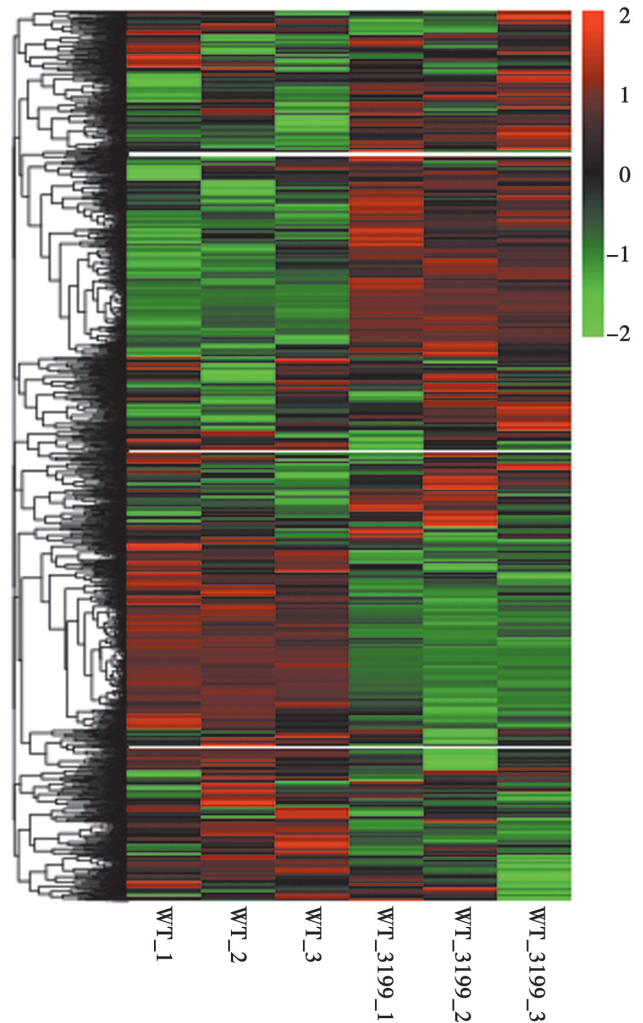


图 4 差异基因聚类分析图
Fig. 4 Cluster analysis of different genes

4 WT_3199 vs WT 转录组差异基因富集分析

利用软件对差异基因集进行 GO 富集分析(Gene Ontology)以及 KEGG 富集分析(Kyoto Encyclopedia of Genes and Genomes),即对差异基因进行显著性分析后,将相应的基因注释到 GO 或 KEGG 数据库中。富集分析的结果对探索目的基因可能在体内外发挥的功能及作用的通路有重要意义。

4.1 GO 富集分析 对差异基因的 GO 富集分析通常用于描述基因功能,它可以分为生物学过程(Biological Process, BP)、细胞组分(Cellular Component, CC)和分子功能(Molecular Function, MF) 3 个部分,对差异基因进行进一步的 GO 注释分析,并选取差异显著性水平最高的前 30 个条目绘制散点图,散点图的纵坐标代表差异条目,横坐标代表该条目差异基因的

数量占差异基因总数的比值。由 GO 富集柱状图(图 5)看出,在生物过程中与小分子代谢途径、药物代谢途径、氨基酸活化以及 tRNA 氨基酰化有关的差异基因数量变化较多;在细胞组分中,与胞内组分以及细胞质相关的基因差异较为明显;而在分子功能方面,差异基因主要富集在氨酰 tRNA 连接酶活性、碳氧键的形成以及 tRNA 有关的催化活性等功能上。

表 4 差异基因功能分类
Table 4 The classification of differential genes

差异类型 Difference	基因名称 Gene name	基因功能描述 Description
上调	STM2456,STM2455,STM2454,STM2463,STM2464,STM2465,STM2466,STM2467,STM2468,STM2469,STM2470	乙醇胺利用
	STM2284,STM2285,STM2286,STM2287,STM2282,STM2283	无氧呼吸
	STM2039,STM2040,STM2041,STM2042,STM2043,STM2044,STM2045,STM2046,STM2047,STM2048,STM2049,STM2050,STM2051,STM2052,STM2053,STM2054,STM2055,STM2056,STM2057,STM2058	丙二醇利用
	STM4463,STM4464,STM4465,STM4466,STM4467	L-精氨酸降解 精氨酸生物合成
	STM4071,STM4072,STM4073,STM4074,STM4075,STM4076,STM4077,STM4078,STM4079,STM4080	输送 LsrABCD 的自体诱导物
	STM3240,STM3241,STM3242,STM3243,STM3244,STM3245	丙酸途径降解 L-苏氨酸
	STM1786, STM787, STM1788, STM789, STM1790, STM791, STM1792, STM793, STM1794	镍铁氢化酶
	STM0618,STM0619,STM0620,STM0621,STM0622,STM0623,STM0624	柠檬酸盐利用
	STM4296,STM4294	精氨酸脱羧酶 精氨酸分解代谢

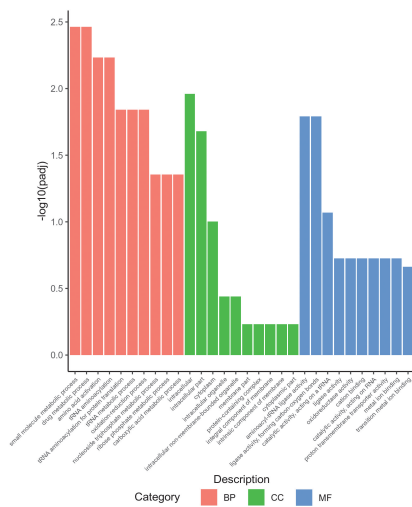


图 5 GO 富集柱状图
Fig. 5 GO enrichment column

4.2 KEGG 通路富集分析 对差异基因的 KEGG 富集分析主要分析的是基因组、化学和系统功能信息。

将 KEGG 富集分析结果中最显著的前 20 个 KEGG 通路绘制成柱状图(图 6),其中横坐标代表 KEGG 通路,纵坐标代表相应通路富集的显著性水平。从图中可以看出 WT_3199vsWT 的差异基因主要富集在碳代谢途径,此外还有多数基因富集到氧化磷酸化、柠檬酸循环、丙酮酸代谢等途径中。

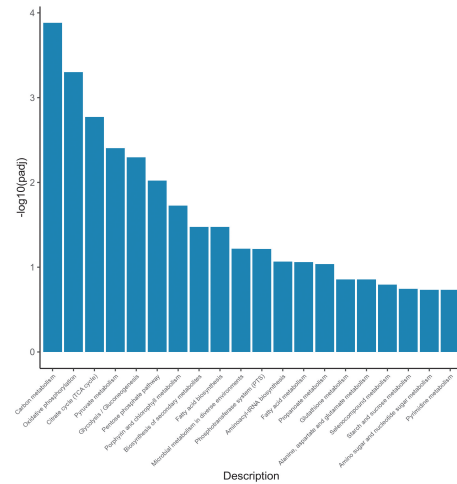


图 6 KEGG 富集柱状图
Fig. 6 KEGG enrichment column

讨论

鼠伤寒沙门菌是一种革兰阴性细长杆菌,属于沙门菌属 B 群,有周鞭毛,无芽孢,无荚膜;兼性厌氧,最适温度生长温度为 35~37 °C,最短 pH 6.8~7.8。该菌对营养要求不高,普通培养基即可生长出圆形、光滑、湿润、半透明边缘整齐的菌落。鼠伤寒沙门菌主要通过粪口途径进入宿主,黏附并定植于肠腔,突破肠粘膜上皮屏障后,被巨噬细胞吞噬。在巨噬细胞中鼠伤寒沙门菌以 SCV 为庇护场所抵抗种种不利因素对自身的杀伤,并随巨噬细胞播散到肠外组织引起更为严重的全身感染,因此对于鼠伤寒沙门菌的研究有着非常重要的临床意义^[24]。

本研究前期通过鼠伤寒沙门菌入侵巨噬细胞实验筛选出了一些入侵前后拷贝数变化较大的,尤其是入侵后拷贝数明显上调的,功能未知的十几种蛋白,这些蛋白可能在鼠伤寒沙门菌的致病过程中发挥重要作用^[12]。其中,STM14_3199 基因编码的蛋白颇受重视,其在入侵巨噬细胞后拷贝数上升高达 9.5 倍。然而,目前有关 STM14_3199 基因及其编码的蛋白质 STM14_3199 的相关研究甚少。为了进一步了解该基因及其编码的蛋白质,首先对其进行了生物信息学分析^[25],结果显示 STM14_3199 基因所编码的蛋白是一种未知蛋白(Uncharacterized protein),相对分子质量约为 7 ku,只有 62 个氨基酸,属于微小蛋白(micro-

protein)。虽然 microprotein 只是一种小分子蛋白,但其编码基因却在动植物基因组中大量存在。然而分布如此广泛的 microprotein 在动植物体内的功能至今尚不清楚^[26]。

由于 STM14_3199 蛋白无晶体衍射结构信息及同源结构,无法从结构方面入手推测其可能在体内外发挥的作用。同时,其在鼠伤寒沙门菌 ATCC 14028 株中氨基酸序列同源性高达 75% 的蛋白 STM14_1447 也是一种功能未知的蛋白,这为对 STM14_3199 功能的研究带来一定难度。

为了进一步探索 STM14_3199 蛋白的生物学功能,通过原核转录组测序技术分析了 STM14_3199 过表达后鼠伤寒沙门菌 ATCC 14028 株中基因表达的变化。在本研究中,构建了鼠伤寒沙门菌 STM14_3199 基因过表达菌株 WT_3199,将 WT 菌株及 WT_3199 菌株培养至对数生长期,提取 RNA 后建库测序。每一环节都进行了严格质控,保证所得测序数据可靠,与参考基因组比对后进行差异基因分析。

聚类分析显示 WT_3199 vs WT 上调的差异基因中有许多基因与乙醇胺的利用、无氧呼吸、丙二醇的利用、以及精氨酸的降解和合成有关,而一些在柠檬酸盐利用以及精氨酸的代谢途径中发挥功能的基因则发生了不同程度的下调,提示 STM14_3199 基因及其编码的蛋白质很可能在这些途径中发挥一定的功能。

GO 富集分析显示,与 WT 菌株相比 WT_3199 菌株在与小分子代谢途径、药物代谢途径、氨基酸活化以及 tRNA 氨基酰化等生物学过程中有明显的基因表达差异;而在细胞组分方面,与胞内组分以及细胞质相关的基因有显著差异;在分子功能上,差异基因在氨酰 tRNA 连接酶活性、碳氧键的形成以及与 tRNA 有关的酶催化活性等功能上有明显富集。KEGG 分析显示,差异基因主要富集在在碳代谢、氧化磷酸化、柠檬酸循环、丙酮酸代谢等途径,这些结果无疑为相关研究的展开以及研究方向提供了重要线索和思路。

本研究成功构建了沙门菌 STM14_3199 基因过表达菌株 WT_3199,并根据转录组学结果分析了 STM14_3199 基因可能发挥作用的生物学过程及通路,为进一步研究 STM14_3199 基因及其编码的蛋白质的功能奠定了良好的基础。然而,关于 STM14_3199 在碳源利用、精氨酸的合成及分解代谢以及其他代谢途径的具体作用目前尚不清楚,还需要进一步的探索及研究。

【参考文献】

[1] Crump JA, Mintz ED. Global trends in typhoid and paratyphoid fever[J]. Clin Infect Dis, 2010, 50(2): 241-246.
[2] Yuan J, Guo W. Mechanisms of resistance of *Salmonella typhi-*

murium to quinolones in infectious diarrhea [J]. Microbiol Immunol, 2017, 61(3-4): 138-143.
[3] Crump JA. Updating and refining estimates of typhoid fever burden for public health action[J]. Lancet Global Health, 2014, 2(10): e551-e553.
[4] Malaeb M, Bizri AR, Ghosn N, et al. Salmonella burden in Lebanon[J]. Epidemiol Infect, 2016, 144(8): 1761-1769.
[5] Aderem A, Underhill DM. Mechanisms of phagocytosis in macrophages [J]. Annu Rev Immunol, 1999, 17(1): 593-623.
[6] Belarbi K, Cuvelier E, A Destee, et al. NADPH oxidases in Parkinson's disease: a systematic review[J]. Mol Neurodegener, 2017, 12(1): 84.
[7] Halici S, Zenk SF, Jantsch J, et al. Functional analysis of the *Salmonella* pathogenicity island 2-mediated inhibition of antigen presentation in dendritic cells [J]. Infect Immun, 2008, 76(11): 4924-4933.
[8] Figueira R, Holden DW. Functions of the *Salmonella* pathogenicity island 2 (SPI-2) type III secretion system effectors[J]. Microbiology, 2012, 158(Pt 5): 1147-1161.
[9] Hensel M, Shea JE, Waterman SR, et al. Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages [J]. Mol Microbiol, 1998, 30(1): 163-174.
[10] Robinson N, McComb S, Muligan R, et al. Type I interferon induces necroptosis in macrophages during infection with *Salmonella enterica* serovar *Typhimurium* [J]. Nat Immunol, 2012, 13(10): 954-962.
[11] Vishal C, Siddharth M, Praveen R. Tumour necrosis factor alpha mediated apoptosis in murine macrophages by *Salmonella enterica* serovar Typhi under oxidative stress [J]. Fems Immunol Med Microbiol, 2010, 47(2): 278-286.
[12] Yang Y, Yue Y, Song N, et al. The YdiU Domain Modulates Bacterial Stress Signaling through Mn²⁺-Dependent UMPylation [J]. Cell Reports, 2020, 32(12): 108161.
[13] Lago M, Monteil V, Douche T, et al. Proteome remodelling by the stress sigma factor RpoS/σS in *Salmonella*: identification of small proteins and evidence for post-transcriptional regulation [J]. Scientific Reports, 2017, 7(1): 2127.
[14] 张春兰, 秦孜娟, 王桂芝, 等. 转录组与 RNA-Seq 技术[J]. 生物技术通报, 2012(12): 51-56.
[15] 祁云霞, 刘永斌, 荣威恒. 转录组研究新技术: RNA-Seq 及其应用[J]. 遗传, 2011, 33(11): 1191-1202.
[16] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics [J]. Nat Rev Gene, 2009, 10(1): 57.
[17] Shang X, Geng L, Yang J, et al. Transcriptome analysis reveals the mechanism of alkalinity exposure on spleen oxidative stress, inflammation and immune function of *Luciobarbus capito*[J]. Ecotoxicol Environ Safety, 2021(225): 112748.
[18] Parkhomchuk D, Borodina T, Amstislavskiy V, et al. Transcriptome analysis by strand-specific sequencing of complementary DNA[J]. Nucleic Acids Res, 2009, 37(18): e123-e123.
[19] Costa V, Angelini C, Feis I D, et al. Uncovering the Complexity of Transcriptomes with RNA-Seq[J]. J Biomed Biotechnol, 2015, 2010(5757): 853916.

- Bacteriol, 1981, 145(2):796-802.
- [7] Murooka Y, Azakami H, Yamashita M. The monoamine regulon including syntheses of arylsulfatase and monoamine oxidase in bacteria[J]. Biosci Biotechnol Biochem, 1996, 60(6):935-941.
- [8] Ganguly A, Joerger RD. Sugar sulfates are not hydrolyzed by the acid-inducible sulfatase AslA from *Salmonella enterica* Enteritidis NalR and Kentucky 3795 at pH 5.5[J]. Can J Microbiol, 2017, 63(8):739-744.
- [9] Chang DE, Smalley DJ, Tucker DL, et al. Carbon nutrition of *Escherichia coli* in the mouse intestine[J]. Proc Natl Acad Sci USA, 2004, 101(19):7427-7432.
- [10] Robbe C, Capon C, Coddeville B, et al. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract[J]. Biochem J, 2004, 384(Pt 2):307-316.
- [11] Bell A, Juge N. Mucosal glycan degradation of the host by the gut microbiota[J]. Glycobiology, 2020, 31(6):691-696.
- [12] Hanson SR, Best MD, Wong CH. Sulfatases: Structure, mechanism, biological activity, inhibition, and synthetic utility[J]. Angew Chem Int Ed Engl, 2004, 43(27):3526-3548.
- [13] Vishal C, Siddharth M, Praveen R. Tumour necrosis factor alpha mediated apoptosis in murine macrophages by *Salmonella enterica* serovar Typhi under oxidative stress[J]. Fems Immunol Med Microbiol, 2010, 47(2):278-286.
- [14] M naco A, Canales-Huerta N, Jara-Wilde J, et al. *Salmonella typhimurium* triggers extracellular traps release in murine macrophages[J]. Front Cell Infect Microbiol, 2021(11):639768.
- [15] Halici S, Zenk SF, Jantsch J, et al. Functional analysis of the *Salmonella* pathogenicity island 2-mediated inhibition of antigen presentation in dendritic cells[J]. Infect Immun, 2008, 76(11):4924-4933.
- [16] Bueno SM, Riquelme S, Riedel CA, et al. Mechanisms used by virulent *Salmonella* to impair dendritic cell function and evade adaptive immunity[J]. Immunology, 2012, 137(1):28-36.
- [17] Zha L, Garrett S, Sun J. *Salmonella* infection in chronic inflammation and gastrointestinal cancer[J]. Diseases, 2019, 7(1):28.
- 【收稿日期】 2022-01-02 【修回日期】 2022-03-15

(上接 399 页)

- [20] 张晓明, 焦新安, 唐丽华, 等. 鼠伤寒沙门氏菌 SL7207 SifA⁻突变株的构建和鉴定[J]. 微生物学报, 2005, 45(3):349-354.
- [21] Goldstein LD, Cao Y, Pau G, et al. Prediction and quantification of splice events from RNA-seq data [J]. Plos One, 2016, 11(5):e0156132.
- [22] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2[J]. Nature Methods, 2012, 9(4):357-359.
- [23] Trapnell C, Hendrickson DG, Sauvageau M, et al. Differential analysis of gene regulation at transcript resolution with RNA-seq [J]. Nature Biotechnol, 2013, 31(1):46.
- [24] 宋楠楠, 杜忠君, 岳盈盈, 等. 鼠伤寒沙门氏菌 ydiU 基因敲除株的构建及其在压力培养条件下的生长特性研究[J]. 中国病原生物学杂志, 2019, 156(12):57-61, 70.
- [25] Galperin M, Koonin EV. 'Conserved hypothetical' proteins: prioritization of targets for experimental study[J]. Nucleic Acids Res, 2004, 32(18):5452-5463.
- [26] 何丹妮, 张琪丽, 倪欢欢, 等. microProtein 调控拟南芥开花时间的机理研究[C]// 第三届全国植物开花·衰老与采后生物学大会.
- 【收稿日期】 2021-12-11 【修回日期】 2022-03-05