

## 骆驼斯氏副柔线虫雌虫和雄虫比较转录组学分析

冯陈晨 王文龙\* 红梅 刘宁 呼和巴特尔

(内蒙古农业大学 兽医学院/农业部动物临床诊疗技术重点实验室,呼和浩特 010018)

**摘要** 为探明骆驼斯氏副柔线虫雌虫和雄虫的转录组差异,挖掘出与性别决定相关的功能基因,采用 Illumina HiSeq™2000 分别对其雌虫和雄虫样本进行转录组测序,构建 cDNA 文库,通过建库质量评估和组装效率评估后对其进行功能注释及相关生物信息学分析。结果显示,比较雌虫和雄虫的转录本,获得 6 430 个差异表达基因,其中 2 131 个基因上调;4 299 个基因下调。对差异表达基因进行 GO 功能分类,有 341 328 和 714 个 Unigenes 分别注释到生物学过程、细胞组成和分子功能三大类 41 个分支中。KEGG 数据库分析,有 1 116 个差异表达基因参与到 198 条 KEGG 通路中,显著富集在机体免疫调节,卵母细胞成熟等通路。功能聚类分析中,雌虫样本在生殖发育、产卵、生殖行为中高表达;雄虫在主要精子蛋白、丝氨酸/苏氨酸特异性蛋白磷酸酶中高表达。此外,在雌虫和雄虫中筛选出 *Tab-3*, *Tab-5*, *Fem-1* 和 *Mog-3* 等 10 种线虫性别决定相关基因,以及 8 种雄性性别特异表达基因和 4 种雌性性别特异表达基因,这些基因在虫体性别决定和性别特异性表达中发挥重要作用。本研究利用生物信息技术预测出雌虫和雄虫差异表达基因在功能分类和代谢通路等方面的生物学特征,挖掘出性别决定基因和性别特异性表达相关基因,为进一步开展斯氏副柔线虫的功能基因组学研究、药物靶点预测和疫苗候选基因筛选提供有价值的数据库资源。

**关键词** 斯氏副柔线虫;转录组;差异表达基因;生物信息学分析

中图分类号 S858.24; S852.73+1

文章编号 1007-4333(2017)07-0061-10

文献标志码 A

## Comparative transcriptome analysis of female and male adults of *Parabronema skrjabini* in camel

FENG Chenchen, WANG Wenlong\*, HONG Mei, LIU Ning, Huhebateer

(College of Veterinary Medicine, Inner Mongolia Agricultural University/Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture, Hohhot 010018, China)

**Abstract** Aiming to find sex-determining genes, transcriptome profile of female and male adults of *Parabronema skrjabini* in camel were investigated. Samples of female and male were analyzed using Illumina HiSeq2000™ RNA sequencing platform and their transcriptome libraries were constructed. After quality filtering and de novo assembling, functional annotation and bioinformatics analysis were carried out. The results showed that, A total of 6 430 differentially expressed genes between female and male were obtained, of which 2 131 and 4 299 of genes were up- and down-regulated, respectively. And 341 328 and 714 differentially expressed genes were annotated into biological process, cellular component and molecular function and 41 functional categories. KEGG analysis showed that 1 116 differentially expressed genes were assigned to 198 KEGG pathways, and clustered significantly in immuno regulation, oocyte maturation and so on. Functional cluster analysis indicted that reproductive development, oviposition and reproductive behavior were high enriched in female adult stage, whilst the sperm main protein, serine/threonine specific protein phosphorylas etc. were high enriched in male adults stage. Moreover, we identified ten the significant sex-determining genes, such as *Tab-3*, *Tab-5* and *Fem-1*, eight gender-specific expression genes of male adult and four gender-specific expression genes of female adult. These genes were necessary to decide nematode gender. In this study, the biological

收稿日期: 2016-06-08

基金项目: 国家自然科学基金资助项目(31260603)

第一作者: 冯陈晨, 博士研究生, E-mail: 410580576@qq.com

通讯作者: 王文龙, 副教授, 主要从事寄生虫组学及分子免疫学研究, E-mail: wwl.imau@163.com

characteristics of differentially expressed genes were predicted by bioinformatics, which involved in functional genes and metabolic pathways. This study provides a foundation for the further whole genome sequence analysis, drug target prediction and vaccine candidate gene screening.

**Keywords** *Parabronemas krjabini*; transcriptome; differentially expressed genes; bioinformatics analysis

骆驼斯氏副柔线虫(*Parabronema skrjabini*)属于旋尾目、副柔属<sup>[1-3]</sup>,是一种寄生于骆驼真胃的吸血性线虫。大量感染斯氏副柔线虫后,可引起骆驼腹泻,贫血甚至死亡。据资料显示,我国内蒙古的双峰驼斯氏副柔线虫的感染率高达91.7%,感染强度最高可达1315条,严重威胁骆驼的健康<sup>[4-5]</sup>。斯氏副柔线虫病的研究仅限于传播媒介的确定和ITS序列的克隆<sup>[6]</sup>,缺乏有关斯氏副柔线虫生长、发育和繁殖相关的基础理论研究。因此,选用骆驼斯氏副柔线虫雌虫和雄虫为研究素材,从分子水平上寻找性别决定和性别特异表达相关基因。研究雌虫和雄虫转录组在生长繁殖代谢水平上的差异,挖掘差异表达基因的功能,为免疫抗原筛选和新药物靶点预测提供基础理论,对解决斯氏副柔线虫对骆驼的危害具有十分重要的意义。近些年来,随着高通量测序技术的不断发展,转录组测序技术已被广泛应用到许多未知动植物的研究中,而寄生虫转录组学研究也越来越受到重视。Ma等<sup>[7]</sup>利用高通量转录组测序技术对猪蛔虫性腺的生殖细胞系和体细胞进行比较转录组学研究,分别得到30137个和39654个表达基因,并以秀丽隐杆线虫基因组为参考进行蛋白质功能注释和KEGG代谢通路研究。巨艳等<sup>[8]</sup>利用Illumina的solexa测序平台对细粒棘球绦虫原头蚴转录组进行测序,得到9029条Unigenes,并对这些基因进行Nr蛋白数据库比对,KEGG代谢通路及GO功能分类研究。Liu等<sup>[9]</sup>对*S. turkestanicum*血吸虫成虫的雌虫和雄虫进行转录组测序,发现有3150个基因在雌虫中特异性表达,1014个基因在雄虫中特异性表达。前人在多种寄生虫性别转录组学研究中取得了阶段性成果,但是对斯氏副柔线虫的研究还很匮乏,尤其是对其雌虫和雄虫生殖发育相关的转录组基因表达研究仍属于未知。本研究通过开展高通量转录组测序分析,探讨斯氏副柔线虫雌虫和雄虫的转录组在基因功能和代谢水平上的差异,筛选关键的性别决定基因和性别特异性表达基因,旨在为开展斯氏副柔线虫病的预防和治疗研究提供参考。

## 1 材料与方法

### 1.1 斯氏副柔线虫样本采集

斯氏副柔线虫成虫样本采集于每年11—12月间内蒙古巴盟乌拉特后旗的双峰驼真胃中,根据雌虫和雄虫的形态特征,在显微镜下将其鉴定,分装标记后在液氮中保存,用于RNA提取。

### 1.2 方法

#### 1.2.1 RNA提取

参照Invitrogen公司(美国)的Trizol Reagent说明书分别对斯氏副柔线虫雌虫和雄虫的总RNA进行提取。使用RQ1酶消解RNA中的DNA,纯化后将样本稀释进行UV检测及1.5%普通琼脂糖凝胶电泳质检合格后备用。

#### 1.2.2 cDNA文库构建及Illumina测序

利用oligo dT-磁珠富集捕获带有polyA尾巴的mRNA。在高温盐离子作用下,mRNA被随机打断,修复并连接5' adaptor,然后用带有3' adaptor和随机六聚体的RT引物反转录合成cDNA。对反转录后的cDNA进行PCR扩增,扩增时引入barcode序列,最终选取片段大小为300~500 bp的PCR产物,利用IlluminaHiSeq2000™测序平台进行文库的测序,由武汉生命之美科技有限公司完成。

#### 1.2.3 测序数据分析

将测序所得的原始数据进行质量评估和可信度分析,并删除测序过程中质量低的序列和不确定的序列( $Q < 20$ ),将得到的Clean reads合并,利用Trinity软件做转录组重头组装,对样品组装得到的Unigenes做进一步序列拼接、去冗余处理和同源聚类,最终得到转录本。将2样本Unigenes的表达量进行RPKM值归一化处理。

在差异表达基因(Differentially expressed genes, DEG)的筛选中,使用edgeR软件进行2样本间的差异表达分析。检测过程中,将差异倍数(Fold change, FC)  $\geq 2$  或  $\leq 0.5$  且  $P$  值  $\leq 0.01$  作为筛选标准,并利用logCMP模型对2个样本进行标准化,这样可以避免不明确的值和不明确的少数趋向于零的logFC,使2样本之间的比较更加详细和全面。

### 1.2.4 差异表达基因注释分析

斯氏副柔线虫雌虫和雄虫相比,将获得的差异表达基因分别进行 GO 功能注释和富集分析;同时将差异表达基因进行 KEGG pathway 分析,找到显著性差异变化的生物学调控通路;将差异表达基因与 *C. elegans* 基因编码的蛋白进行比对注释 (blastx)。然后,将注释出的差异基因所对应的 *C. elegans* 蛋白 GI 号在 DAVID 平台进行功能聚类分析,获得可信度高的功能通路和更细致全面的功能聚类。

## 2 结果

### 2.1 差异表达基因分析结果

#### 2.1.1 转录组测序结果

对骆驼斯氏副柔线虫雌虫和雄虫分别进行转录组测序,将得到的数据进行质量控制、de novo 拼接组装和聚类(Cluster)。结果显示,雌虫和雄虫分别获得 2 810 万以上的原始 Reads,且 Clean reads 比例均在 83%以上,说明建库质量较好。将 2 个样本

Unigenes 的表达量进行 RPKM 值归一化处理,如表 1 显示,雌虫转录本 RPKM 值>0 有 54 624 个 Unigenes,RPKM 值≥10 有 8 592 个 Unigenes;雄虫转录本中 RPKM 值>0 有 54 619 个 Unigenes,RPKM 值≥10 有 8 618 个 Unigenes;雌虫和雄虫中 RPKM 值≤10 的转录本均占 80%左右。

#### 2.1.2 差异表达基因筛选结果

选用 edgeR 软件对基因表达量做差异表达分析,以显著水平  $P\text{-value} \leq 0.01$  及差异倍数 fold change ≥ 2 或 ≤ 0.5 为标准筛选差异表达基因。根据火山图 1 所示,红色的点表示显著差异表达基因 (DEGs),纵坐标为  $\log_2 FC$  表示某一个基因在 2 样本中表达量差异倍数的对数值,且绝对值越大,表明基因表达量变化的倍数越大;横坐标为  $\log_2 CPM$  表示 2 样本之间比较同 1 个基因时的总聚集点,且坐标值越大,表明筛选的差异表达基因越可靠。将雌虫和雄虫转录本比较后得到的差异表达基因进行统计。结果表明,差异表达基因共有 6 430 个,其中表达上调基因有 2 131 个,表达下调基因有 4 299 个。

表 1 样本有效 RNA-seq 序列及表达 Unigenes 数

Table 1 Number of high-quality clean reads and expressed unigenes of each sample

| 样本<br>Sample | 原始数据<br>Raw read | 有效数据<br>Clean read | 表达 Unigenes 数 (RPKM>0)<br>Expressed unigenes cluster | 表达 Unigenes 数 (RPKM≥10)<br>Expressed unigenes cluster |
|--------------|------------------|--------------------|------------------------------------------------------|-------------------------------------------------------|
| 雌虫           | 28 130 680       | 23 603 172(83.91%) | 54 624(54.91%)                                       | 8 592(15.73%)                                         |
| 雄虫           | 29 451 206       | 24 543 395(83.34%) | 54 619(54.90%)                                       | 8 618(15.78%)                                         |

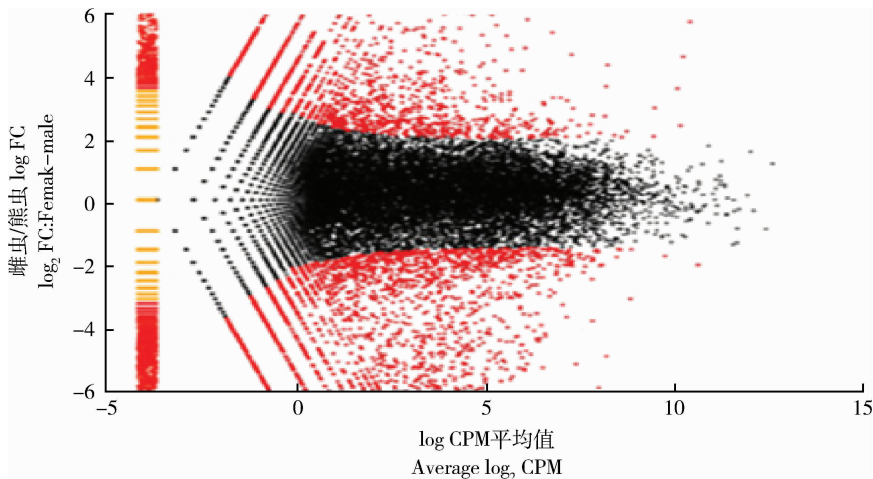


图 1 EdgeR 方法鉴定线虫差异表达基因

Fig. 1 Identification of the differential expressed unigenes between females and males by edgeR program

## 2.2 差异表达基因 Gene ontology 分析结果

### 2.2.1 差异表达基因 GO-生物过程富集分析

雌虫和雄虫的差异表达基因比对到 GO-生物过程数据库中,共注释出 341 个基因。表 2 中表明,显著富集在蛋白质氨基酸磷酸化过程中的差异表达基因(DEG)有 59 个,占该生物过程基因数(EG)的 25.65%,占比对到 BP 数据库所有差异表达基因数的 17.30%;富集在基于微管过程中基因数有 5 个,占该生物过程基因数的 26.32%,占比对到 BP 数据库所有差异表达基因数的 1.47%;富集在胚胎发育过程中的差异表达基因有 5 个,占该生物过程基因数的 23.8%,占比对到 BP 数据库所有差异表达基因数的 1.50%,具体分布见图 2(a)。

### 2.2.2 差异表达基因 GO-细胞组成富集分析

雌虫和雄虫所有差异表达基因比对到 GO-细

胞组成数据库中,注释出差异常基因 328 个,表 2 中表明,显著富集在细胞核中的差异表达基因有 46 个,占该细胞组成基因数的 10.16%,占比对到 CC 数据库所有差异表达基因数的 14.02%;富集在细胞骨架中的差异表达基因有 11 个,占该细胞组成基因数的 18.97%,占比对到 CC 数据库所有差异表达基因数的 3.35%;富集在微管中的基因有 5 个,占该细胞组成基因数的 14.71%,占比对到 CC 数据库所有差异表达基因数的 1.52%,具体分布见图 2(b)。

### 2.2.3 差异表达基因 GO-分子功能富集分析

雌虫和雄虫所有差异表达基因比对到 GO-分子功能数据库中,共注释出 714 个差异基因。表 2 中表明,显著富集在酪氨酸蛋白激酶活性功能的差异表达基因有 15 个,占该分子功能基因数的 45.45%,占比对到 MF 数据库所有差异表达基因数

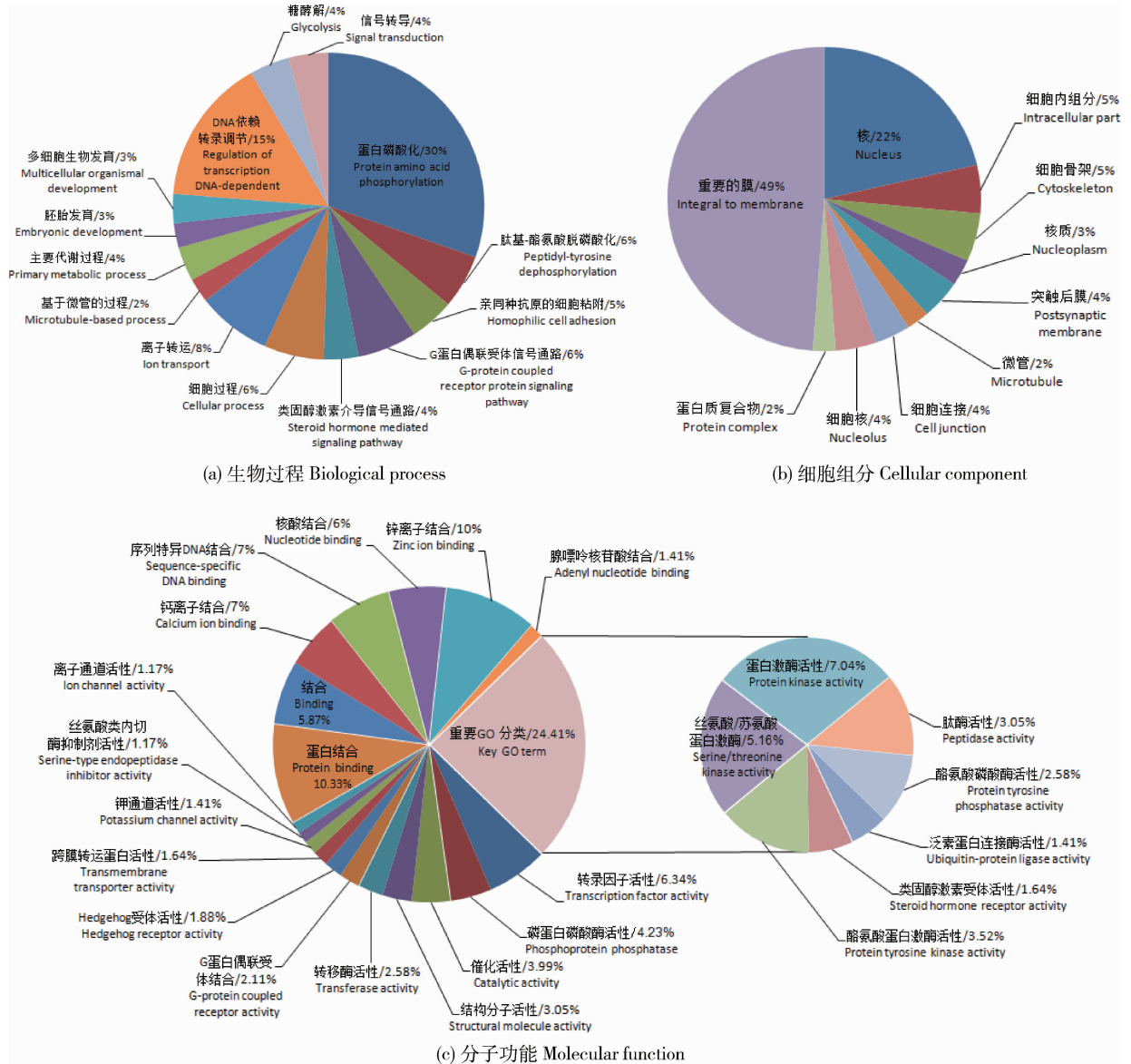
表 2 差异表达基因 GO-ontology 富集分析结果

Table 2 GO enrichment result of differentially expressed genes

| GO 分类<br>Gene ontology | 基因簇频率 <sup>①</sup>     |       | 基因组频率 <sup>②</sup>       |       | P 值               |
|------------------------|------------------------|-------|--------------------------|-------|-------------------|
|                        | Gene cluster frequency |       | Genome cluster frequency |       | P-value           |
|                        | DEG                    | CF/%  | EG                       | GF/%  | ( $P \leq 0.01$ ) |
| 生物过程                   |                        |       |                          |       |                   |
| 蛋白氨基酸磷酸化过程             | 59                     | 17.30 | 230                      | 2.73  | 0                 |
| 肽基酪氨酸磷酸化过程             | 11                     | 3.23  | 42                       | 0.50  | 3.29E-10          |
| 类固醇激素介导信号通路            | 7                      | 2.05  | 26                       | 0.31  | 4.85E-7           |
| 以微管为基础的生物过程            | 5                      | 1.47  | 19                       | 0.23  | 2.56E-5           |
| 胚胎发育过程                 | 5                      | 1.47  | 21                       | 0.25  | 4.34E-5           |
| 细胞组分                   |                        |       |                          |       |                   |
| 细胞核                    | 46                     | 14.02 | 453                      | 5.49  | 4.20E-13          |
| 细胞骨架                   | 11                     | 3.35  | 58                       | 0.70  | 1.29E-6           |
| 完整的膜                   | 105                    | 32.01 | 2 489                    | 30.16 | 0.000 1           |
| 蛋白质复合物                 | 5                      | 1.52  | 28                       | 0.34  | 0.001 5           |
| 微管                     | 5                      | 1.52  | 34                       | 0.41  | 0.003 0           |
| 分子功能                   |                        |       |                          |       |                   |
| 蛋白激酶活性                 | 30                     | 4.20  | 93                       | 0.56  | 0                 |
| 丝氨酸/苏氨酸蛋白激酶活性          | 22                     | 3.08  | 147                      | 0.89  | 1.53E-10          |
| 酪氨酸蛋白激酶活性              | 15                     | 2.10  | 33                       | 0.20  | 2.89E-15          |
| 类固醇激素受体活性              | 7                      | 0.98  | 26                       | 0.16  | 9.40E-8           |
| 泛素蛋白连接酶活性              | 6                      | 0.84  | 53                       | 0.32  | 0.003 6           |

注:① $CF = DEG / TDEG \times 100\%$ ,CF 为在该分类中差异表达基因簇的频率;DEG 为在该分类中差异表达基因的数量;TDEG 为比对到该分类中所有差异表达基因数。② $GF = EG / TEG \times 100\%$ ,GF 为比对到该分类中基因的频率;EG 为比对到该分类中所有基因数量;TEG 为比对到分类中所有基因数。

Note:① $CF = DEG / TDEG \times 100\%$ ,CF is the cluster frequency of differentially expressed gene annotated to each term;DEG is the number of differentially expressed gene annotated to each term;TDEG is the number of all differentially expressed genes annotated to GO ontology. ②  $GF = EG / TEG \times 100\%$ ;GF is the genome frequency of all genes annotated to the term;EG is the number of genes annotated to each term;TEG is the number of all genes annotated to GO ontology.



雌虫和雄虫的差异表达基因在 GO 三大分类((a)生物过程,(b)细胞组分,(c)分子功能)的分布情况。

Distribution of differentially expressed genes of female and male adults in three main categories (biological process, cellular component and molecular function) are shown in the subgraphs (a)-(c), respectively.

图 2 Gene ontology (GO)分类图

Fig. 2 Pie charts showing gene ontology (GO) classification

的 2.10%;富集在丝氨酸/苏氨酸蛋白激酶活性功能中的差异表达基因有 22 个,占该分子功能基因数的 14.97%,占比对到 MF 数据库所有差异表达基因数的 3.08%;富集在泛素蛋白连接酶活性功能中的基因有 6 个,占该分子功能基因数的 11.32%,占比对到 MF 数据库所有差异表达基因数的 0.84%,具体分部见图 2(c)。

### 2.3 差异表达基因 KEGG Pathway 分析

将雌虫和雄虫所有差异表达基因比对到

KEGG 数据库,结果显示,2 778 个差异表达基因注释到该数据库中,其中 1 116 个差异表达基因有具体定义,并显著富集到 198 条通路。下表 3 中列出差异表达基因显著富集的前 10 个通路,其中与机体免疫调节相关的通路有 3 个;与机体组织器官形态形成、mRNA 调节相关的通路有 2 个;与机体所需物质代谢合成有关的通路有 2 个;与卵母细胞减数分裂、成熟等方面相关的通路有 3 个。

表3 差异表达基因显著性富集的通路

Table 3 Enriched pathways of differentially expressed genes

| 通路<br>Pathway | 差异表达基因数 <sup>①</sup><br>DEGs | KEGG <sup>②</sup> | P 值<br>P-value | Ko ID   |
|---------------|------------------------------|-------------------|----------------|---------|
| 吞噬体           | 35                           | 214               | 2.89E-7        | Ko04145 |
| 肾素血管紧张素系统     | 7                            | 14                | 1.07E-5        | Ko04614 |
| 蛋白酶体          | 17                           | 86                | 2.95E-5        | Ko03050 |
| Hh 信号通路       | 12                           | 157               | 0.00046        | Ko04340 |
| 卵母细胞减数分裂      | 13                           | 77                | 0.00123        | Ko04114 |
| mRNA 监测通路     | 22                           | 170               | 0.00141        | Ko03015 |
| 类固醇生物合成       | 3                            | 8                 | 0.00118        | Ko00100 |
| 泛素介导蛋白水解      | 21                           | 199               | 0.00183        | Ko04120 |
| 黄体酮调节卵母细胞成熟过程 | 8                            | 56                | 0.00269        | Ko04914 |
| 糖酵解/糖异生过程     | 25                           | 263               | 0.00343        | Ko00010 |

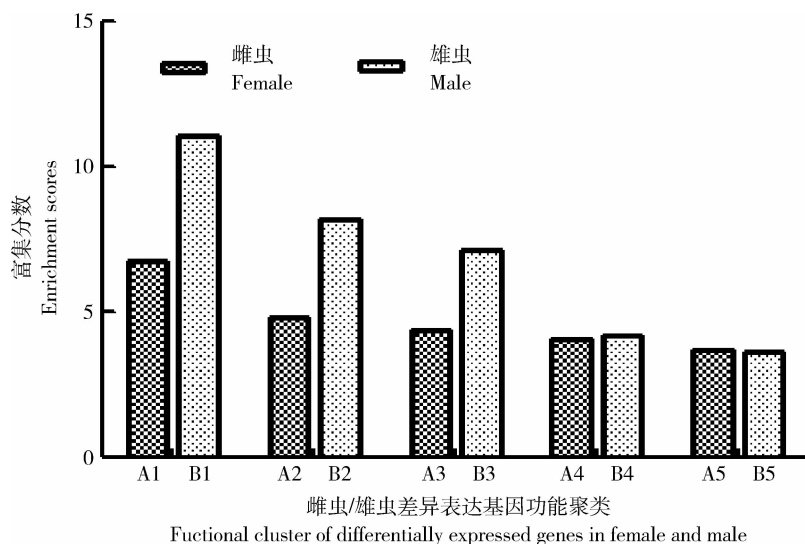
注:①DEGs,在每个 KEGG 代谢通路中差异表达基因数量;②KEGG,在 KEGG 数据库中涉及到该通路的基因数。

Note:①DEGs,the number of differentially expressed genes in each KEGG pathway;②KEGG,the number of genes in the pathway in KEGG database.

## 2.4 差异表达基因功能聚类分析

斯氏副柔线虫雌虫和雄虫相比,利用 DAVID 平台对差异表达基因进行功能聚类分析并计算功能聚类富集分数。结果显示,雌虫样本中在质膜、跨膜;生殖发育、产卵、生殖行为;EGF、EGF 样钙离子

结合等相关功能中富集性高表达。雄虫样本在无膜细胞器、核糖体结构成分、细胞骨架;核酸、嘌呤结合、ATP 结合、磷酸代谢过程、丝氨酸/苏氨酸蛋白激酶活性;主要精子蛋白;丝氨酸/苏氨酸特异性蛋白磷酸化酶等相关功能中富集性高表达(图 3)。



A1,跨膜区 Transmembrane region,细胞膜 Membrane;A2,产卵 Oviposition,生殖发育 Reproductive development,生殖行为 Reproductive behavior;A3,表皮生长因子 Epidermal growth factor,EGF 样钙离子结合 EGF-like calcium-binding;A4,信号肽 Signal peptide,二硫键 Disulfide bond;A5,刺猬状受活性 Hedgehog receptor activity;B1,无膜细胞器 Non-membrane-bounded organell,细胞骨架 Cytoskeleton;B2,丝氨酸/苏氨酸蛋白激酶 Serine/threonine protein kinase,ATP 结合 ATP binding,核苷酸结合 nucleotide binding;B3,主要精子蛋白 Major sperm protein;B4,丝氨酸/苏氨酸蛋白磷酸酶 Serine/threonine-protein phosphatase;B5,糖酵解/糖异生 Glycolysis/Gluconeogenesis.

图3 差异基因功能聚类表达量最高的前5类

Fig. 3 Top 5 functional clusters of differentially expressed genes

## 2.5 性别决定及性别特异性相关基因分析

### 2.5.1 性别决定相关基因分析

通过对斯氏副柔线虫雌虫和雄虫差异表达基因的研究,筛选出与性别决定相关的重要基因并对其进行富集分析。在雌虫和雄虫中发现 10 种重要的性别决定相关基因,分别为 3 种 *Tra* 基因 (*Transformer gene*): *Tra-1*, *Tra-3*, *Tra-4*; 2 种 *Mab* 基因 (*Male abnormal*): *Mab-3*, *Mab-23*; 2 种 *Fem* 基因 (*Feminization gene*): *Fem-1*, *Fem-2*;

*Fog-3* 基因 (*Feminization of germ line*); *Mog-3* 基因和 *Sex-1* 基因。其中,有关促进雄性个体发育的基因 *Tra-1*、*Mab-3*、*Fog-3* 在雄虫中较高表达;有关抑制雌性个体发育的基因 *Tra-3* 和 *Tra-4* 在雌虫和雄虫中都有表达,但在雌虫中相对较高表达;有关雌性个体选择的基因 *Sex-1* 在雌虫中高表达。促进雄性发育和调节精子形成的基因 *Fem-1* 和 *Fem-2* 以及影响雄虫生殖能力的基因 *Mab-23* 和 *Mog-3* 在雌虫和雄虫中都有表达(表 4)。

表 4 性别决定相关基因  
Table 4 Sex determination genes

| 基因<br>Gene    | 基因表达量                       |         | GO 功能注释<br>Functional annotation |
|---------------|-----------------------------|---------|----------------------------------|
|               | Expression quantity of gene |         |                                  |
|               | 雌虫 Female                   | 雄虫 Male |                                  |
| <i>Tra-1</i>  | 4.43                        | 3.05    | 核酸结合模板 DNA 转录调节                  |
| <i>Tra-3</i>  | 20.43                       | 10.41   | 雌性生殖细胞性别决定<br>雌雄同体细胞受精           |
| <i>Tra-4</i>  | 28.12                       | 13.15   | 未知                               |
| <i>Mab-3</i>  | 0.43                        | 8.29    | 转录因子活性<br>序列特异性 DNA 结合           |
| <i>Mab-23</i> | 8.51                        | 10.35   | 转录因子活性<br>模板 DNA 转录调节            |
| <i>Fem-1</i>  | 15.14                       | 8.62    | 未知                               |
| <i>Fem-2</i>  | 12.75                       | 5.09    | 未知                               |
| <i>Fog-3</i>  | 0.16                        | 39.38   | 未知                               |
| <i>Mog-3</i>  | 95.57                       | 86.43   | 蛋白结合                             |
| <i>Sex-1</i>  | 32.87                       | 2.21    | 主要性别决定<br>生殖过程<br>雌雄同体生殖期发育      |

### 2.5.2 性别特异表达基因分析

在斯氏副柔线虫雌虫和雄虫差异表达基因研究中,探索出性别特异性表达基因并对其进行富集分析。结果显示,在雄虫个体中精子蛋白(*MSPs*)家族和精子特异性(*SSP*)家族的基因表达量明显升高。其中主要 *MSPs* 为:*MSP-4*、*MSP-2*、*MSP-77* 和 *MSP-78*,这些基因是雄虫特异性细胞骨架蛋白,它与精子的运动有关;主要 *SSPs* 为:*SSP-34*、

*SSP-35*和 *SSP-36*,这些基因是精子特异性的表达基因。在雌虫个体中 *EGGs* 相关基因表大量明显上调,主要有:*EGG-2*, *EGG-3*, *EGG-5* 和 *EGG-6*。其中,*EGG-2* 在卵母细胞发育过程中被表达,在受精期间与配体相互作用,来调节虫体繁殖能力;*EGG-3* 在卵子活化期间起关键作用,并且直接参与到虫卵向胚胎转变的过程中。*EGG-5* 在虫卵向胚胎转变和胚胎发育过程中起到关键作用(表 5)。

表5 性别特异性表达基因  
Table 5 Gender-specific expression genes

| 性别特异性表达基因<br>Gender-specific<br>expression gene | 基因富集分数<br>Expression quantity of gene |          | 蛋白描述<br>Protein description |
|-------------------------------------------------|---------------------------------------|----------|-----------------------------|
|                                                 | 雌虫 Female                             | 雄虫 Male  |                             |
|                                                 | 雄虫特异性表达基因                             |          |                             |
| <i>F52F12.8</i>                                 | 0                                     | 58.03    | 主要精子结构域蛋白 2                 |
| <i>ZK354.7</i>                                  | 0.30                                  | 2 582.50 | 主要精子结构域蛋白 2                 |
| <i>F53B6.4</i>                                  | 0.21                                  | 105.37   | 主要精子结构域蛋白 4                 |
| <i>MSP-77</i>                                   | 0                                     | 227.24   | 主要精子蛋白                      |
| <i>MSP-78</i>                                   | 3.73                                  | 958.02   | 主要精子蛋白                      |
| <i>SSP-34</i>                                   | 0.13                                  | 526.90   | 精子特异性蛋白                     |
| <i>SSP-35</i>                                   | 0.11                                  | 77.95    | 精子蛋白结构域包含蛋白                 |
| <i>SSP-36</i>                                   | 0.06                                  | 35.23    | 精子特异性蛋白                     |
| 雌性特异性表达基因                                       |                                       |          |                             |
| <i>EGG-2</i>                                    | 3.01                                  | 0        | 低密度脂蛋白受体结构域 A               |
| <i>EGG-3</i>                                    | 14.21                                 | 0.80     | 蛋白酪氨酸磷酸酶                    |
| <i>EGG-5</i>                                    | 91.46                                 | 2.53     | 细胞核激素受体家族成员                 |
| <i>EGG-6</i>                                    | 2.70                                  | 0        | 富集亮氨酸重复蛋白家族                 |

### 3 讨论

新一代高通量 RNA-Seq 技术被应用到许多未知的研究领域,在科学探索中发挥了关键性的作用<sup>[10]</sup>。本研究首次针对骆驼斯氏副柔线虫雌虫和雄虫阶段进行转录组测序,探索出雌虫和雄虫中基因表达的差异和代谢通路的差异,挖掘性别决定相关基因和性别特异性表达基因。

1)在 GO 功能分析中,差异基因主要富集在类固醇激素介导通路、细胞骨架、微管、丝氨酸/苏氨酸蛋白激酶活性等功能中。在线虫体内类固醇激素的作用与不同哺乳动物体内类固醇复合物有所不同,其包含雌激素、雄激素、前体胆固醇和黄体酮等,这些在不同性别的虫体中对配子的形成发挥作用<sup>[11]</sup>。细胞骨架是构成微管的重要成分,而微管在性腺形成过程中也发挥着重要的作用,这与 Ma 等<sup>[7]</sup>对猪蛔虫性腺转录组测序结果一致。根据 Yi 等<sup>[12]</sup>对猪蛔虫的研究发现丝氨酸/苏氨酸激酶对雄虫精子蛋白运动性器官膜的相关组成发挥重要作用。在 KEGG pathway 分析中,差异基因显著富集在蛋白

酶体通路,蛋白酶体在虫体免疫调节中起到重要作用,根据报道在病原体入侵时蛋白酶体和泛素分子结合共同降解蛋白产生多肽抗原来保护机体<sup>[13]</sup>,这可能与成虫与终末宿主之间相互作用,而产生免疫逃避有关。类固醇合成在雌虫和雄虫第二性征发育过程中必不可少,而且可以调节雌虫子宫内膜黄体酮的改变,使其受精<sup>[11]</sup>。

2)在功能聚类分析中,雌虫在性腺发育、生殖发育以及表皮生长因子 EGF 等有关的功能中聚类,而 EGF 可能与线虫在宿主内需要表皮的生长来获得自身保护有关。雄虫在与精子形成过程中发挥主要作用的主要精子蛋白、丝氨酸/苏氨酸特异性蛋白磷酸化酶中聚类。根据捻转血矛线虫的转录组研究发现,丝氨酸/苏氨酸蛋白激酶鉴定为分泌性蛋白,在寄生虫与宿主互作中有重要作用<sup>[14]</sup>。

3)本研究以秀丽隐杆线虫为参考,发现在雌虫和雄虫中存在不同表达量的性别决定基因。经研究分析表明,*Tra-1* 基因在斯氏副柔线虫雌虫中的表达量稍高于雄虫,这与 Berkseth M 等<sup>[15]</sup>对 *C. elegans* 研究一致。*Tra-1* 基因在性别特异性发育



过程中调节雄虫体细胞性腺和精子发生,正调控雌虫体细胞性别分化,是性别决定过程中的最终调节基因<sup>[16]</sup>。*Tra-4* 基因通过与 *Tra-1* 相互作用来抑制雌虫中雄虫特异性基因表达,进而在性别决定中发挥作用<sup>[17]</sup>。Barnes T M 等<sup>[18]</sup> 研究发现 *C. elegans* 中 *Tra-3* 基因调节雌雄同体中体细胞和生殖细胞的正确性别发育,是雌雄同体性别决定基因,在雄虫中没有相关性别决定的作用,而在斯氏副柔线虫中 *Tra-3* 基因雌虫中的表达量高于雄虫,可能说明 *Tra-3* 基因更趋向于在具有雌性生殖能力的虫体中起性别决定作用。*Mab-3* 基因是 *Tra-1* 的靶位基因,能与 *Lin-32* 和 *Mab-5* 基因相互协调雄虫感觉器官分化,在调节雄性性别发育和行为等方面有重要作用<sup>[19]</sup>。*Mab-23* 作用于雄性虫体尾部的真皮、交合刺和肛道的形态形成,在雄性交配过程中尾部正常翻转起重要作用<sup>[20]</sup>。*Fog-3* 基因是生物体所必需的,主要控制决定生殖细胞发起精子形成的命令。*Sex-1* 基因在 XO 或 XX 胚胎中的表达量改变可使性染色体数量改变从而改变虫体性别<sup>[21]</sup>。在性别特异性表达基因的研究中,精子蛋白(MSP) 家族是第一个被鉴定的线虫性别特异表达蛋白,不同线虫中存在的 MSP 基因数不同,在营自由生活的线虫中 MSP 的数量为 15~50 个,而寄生性线虫仅为 5~13 个<sup>[22]</sup>,在斯氏副柔线虫中鉴定出 5 个 MSP 基因。MSPs 在精子移行运动中有重要作用,同时也是线虫中唯一能合成精母细胞的蛋白。雌虫中性别特异性表达基因 *Egg-3* 在卵活化期间,调节细胞表面和皮质重组来响应精子的进入<sup>[23]</sup>。*EGG-5* 基因的缺乏会引起卵母细胞减数分裂、极体形成以及阻止多精子进入的功能出现缺陷,而使虫卵无法正常受精<sup>[24-25]</sup>,所以 *EGG-5* 在虫卵向胚胎转变过程中起重要作用。

## 参考文献 References

[1] 卢俊杰. 人和动物寄生线虫图谱[M]. 北京:中国农业科学技术出版社,2003:405-459  
Lu J J. *The Map of Human and Animal Parasitic Nematodes* [M]. Beijing: China Agricultural Science and Technology Publishing House,2003:405-459 (in Chinese)

[2] 沈杰,黄兵. 中国家畜家禽寄生虫名录[M]. 北京:中国农业科学技术出版社,2004:105-111  
Shen J, Huang B. *The List of Parasites of Livestock and Poultry in China* [M]. Beijing: China Agricultural Science and

Technology Publishing House,2004:105-111 (in Chinese)

[3] Mehlhorn H. *Encyclopedic Reference of Parasitology* [M]. 2nd ed. Berlin:Springer,2001,

[4] 赵治国. 我国骆驼斯氏副柔线虫病传播媒介的研究[D]. 呼和浩特:内蒙古农业大学,2010  
Zhao Z G. Study on the vector of camel Parabronemosis in China [D]. Hohhot: Inner Mongolia Agricultural University, 2010 (in Chinese)

[5] 郑文青. 骆驼斯氏副柔线虫病流行病学调查及生前 PCR 诊断方法建立[D]. 呼和浩特:内蒙古农业大学,2010  
Zheng W Q. Study on epidemiological investigation and established the PCR antemortem diagnosis method of camel Parabronemosis [D]. Hohhot: Inner Mongolia Agricultural University,2015 (in Chinese)

[6] Seyed S. Molecular characterization of the first internal transcribed spacer of rDNA of *Parabronema skrjabini* for the first time in sheep[J]. *Annals of Parasitology*,2015,61(4): 241-246

[7] Ma X, Zhu Y, Li C, Shang Y, Meng F, Chen S, Miao L. Comparative transcriptome sequencing of germline and somatic tissues of the *Ascaris suum* gonad [J]. *BioMed Central Genomics*,2011,12:481

[8] 巨艳,李子华,王娅娜. 细粒棘球绦虫原头蚴 mRNA 测序及表达谱分析[J]. 中国人兽共患病学报,2015(1):21-25  
Ju Y, Li Z H, Wang Y N. mRNA sequencing and transcriptome characteristic of *Echinococcus granulosus* [J]. *Chinese Journal of Zoonoses*,2015(1):21-25 (in Chinese)

[9] Liu GH, Xu MJ, Chang QC, Gao J F, Wang C R. De novo transcriptomic analysis of the female and male adults of the blood fluke *Schistosoma turkestanicum* [J]. *Parasites & Vectors*,2016,9:143

[10] Reuter J A, Spacek D V, Snyder M P. High-throughput sequencing technologies[J]. *Cell Press*,2015,58(4):586-597

[11] David J C. *Biochemistry and Function of Nematode Steroids* [J]. *Critical Reviews in Biochemistry and Molecular Biology*, 1999,34(4):273-284

[12] Yi K, Buttery S, Stewart M, Roberts T. A Ser/Thr kinase required for membrane-associated assembly of the major sperm protein motility apparatus in the amoeboid sperm of *Ascaris* [J]. *Molecular Biology of the Cell*,2007,18(5):1816-1825

[13] Wang X, Zhao Y, Wong K, Ehlers P, Jones J, Yuji K, Marra A, Holt A, Hansen D, Moerman G. Identification of genes expressed in the hermaphrodite germ line of *C. elegans* using SAGE[J]. *BioMed Central Genomics*,2009,10:213

[14] Schwarz E, Korhonen P, Campbell B, Young N, Jex R. The genome and developmental transcriptome of the stronglylid nematode *Haemonchus contortus* [J]. *Genome Biology*,2013,14(8):R89.

[15] Berkseth M, Ikegami K, Arur S, Lieb J, Zarkower D. TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination [J].

- Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(40):16033-8
- [16] Li T, Kelly WG. A role for WDR5 in TRA-1/Gli mediated transcriptional control of the sperm/oocyte switch in *C elegans* [J]. *Nucleic Acids Res*, 2014, 42(9):5567-81
- [17] Grote P, Conradt B. The PLZF-like protein TRA-4 cooperates with the Gli-like transcription factor TRA-1 to promote female development in *C elegans* [J]. *Development Cell*, 2006, 11(4):561-573
- [18] Barnes T, Hodgkin J. The tra-3 sex determination gene of *Caenorhabditis elegans* encodes a member of the calpain regulatory protease family [J]. *The EMBO Journal*, 1996, 15(17):4477-4484
- [19] Inoue H, Nishida E. The DM domain transcription factor MAB-3 regulates male hypersensitivity to oxidative stress in *Caenorhabditis elegans* [J]. *Molecular and cellular biology*, 2010, 30(14):3453-3459
- [20] Lints R, Emmons S. Regulation of sex-specific differentiation and mating behavior in *C elegans* by a new member of the DM domain transcription factor family [J]. *Genes & Development*, 2002, 16(18):2390-2402
- [21] Carmi I, Kopczynski J, Meyer B. The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex [J]. *Nature*, 1998, 396(6707):168-173
- [22] 邹丰才, 吴绍强, 黄翠琴, 朱兴全. 寄生线虫性别特异表达基因研究进展 [J]. *中国寄生虫学与寄生虫病杂志*, 2005, 23(4):246-249
- Zou F C, Wu S Q, Huang C Q, Zhu X Q. The research on evolution of gender-specific expressed genes in parasitic nematodes [J]. *Chinese Journal of Parasitology and Parasitic Diseases*, 2005, 23(4):246-249 (in chinese)
- [23] Maruyama R, Velarde N, Klancer R, Gordon S, Kadandale P, Parry J, Hang J, Rubin J, Stewart-Michaelis A, Schweinsberg P, Grant BD, Piano F, Sugimoto A, Singson A. EGG-3 regulates cell-surface and cortex rearrangements during egg activation in *Caenorhabditis elegans* [J]. *Current Biology*, 2007, 17(18):1555-1560
- [24] Parry J, Velarde N, Lefkovith A, Zegarek M, Hang J, Klancer R, Maruyama R, Druzhinina M, Grant B, Piano F, Singson A. EGG-4 and EGG-5 link events of the oocyte-to-embryo transition with meiotic progression in *C elegans* [J]. *Current Biology*, 2009, 19(20):1752-1757
- [25] Stoeckius M, Grün D, Kirchner M, Ayoub S, Torti F, Piano F, Herzog M, Selbach M, Rajewsky N. Global characterization of the oocyte-to-embryo transition in *Caenorhabditis elegans* uncovers a novel mRNA clearance mechanism [J]. *The EMBO Journal*, 2014, 33(16):1751-1766

责任编辑: 苏燕