

中华按蚊雌、雄蚊转录组比较分析^{*}

刘 杰, 李 勃, 陈 斌

(重庆师范大学 昆虫与分子生物学研究所 媒介昆虫重庆市重点实验室, 重庆 401331)

摘要:【目的】对中华按蚊(*Anopheles sinensis*)雌、雄蚊的转录组测序数据进行差异表达分析, 获得两者间差异表达基因, 对它们进行功能富集分析, 挖掘这些基因的功能表现和参与的通路。【方法】基于已知中华按蚊雌、雄蚊转录组数据进行了 RNA-Seq 数据相关性检查, 利用 R 语言 DESeq 包比较雌蚊和雄蚊中基因的差异表达情况, 并用 GOseq 和 KOBAS 对筛选出来的差异表达基因进行 GO 和 KEGG 富集分析。【结果】Pearson 相关系数的平方(R^2)大于 0.92, 说明样本在 3 个生物学重复间具有高度相关性, 可用于进一步分析; 中华按蚊雌、雄蚊差异表达基因共 3 298 个, 其中 1 589 个基因在雌蚊中表达上调, 1 709 个基因在雌蚊中表达下调; 差异表达基因明显富集在酰胺生物合成、有机酸代谢、肽代谢过程等功能类别上, 参与核糖体在真核生物中的生物发生、RNA 转运等 KEGG 通路。【结论】上述结果对深入研究不同性别的中华按蚊基因表达及差异具有潜在价值。

关键词: 中华按蚊; 转录组; GO 分析; KEGG 分析

中图分类号: Q966

文献标志码: A

文章编号: 1672-6693(2020)03-0054-06

中华按蚊(*Anopheles sinensis*)为蚊科(Culicidae)按蚊亚科(Anophelinae)按蚊属(*Anopheles*)昆虫。雌性中华按蚊兼吸人、畜血液, 因它具有嗜人和重复吸血的行为, 从而成为了中国和东南亚地区主要的疟疾传播媒介; 同时它也可以传播流行性乙型脑炎病毒(JEV)、猫立克次氏体等病原体^[1-2]。由中华按蚊参与传播的传染病严重危害人类健康。以疟疾为例, 2018 年全球新增约 2.28 亿疟疾病例, 同年有 40.5 万人死于疟疾^[3]。对蚊虫病媒的控制是预防和控制疟疾等疾病最有效的措施之一, 目前特别依赖于杀虫剂的使用^[4]。然而过度和持续使用杀虫剂导致了相关昆虫抗药性的增强。目前, 中华按蚊已对拟除虫菊酯等杀虫剂产生了高水平的抗药性^[5-7], 因此探索新的媒介控制策略势在必行。

有研究表明, 可采用将雌蚊转化为无害雄蚊的策略而实现蚊虫防控, 并提示调节蚊虫性别分化过程可能是当前蚊虫防控工作的一个重要突破口^[8]。当前为了揭示蚊虫的性别差异并进而开发新的病媒控制策略, 研究人员对不同性别的病媒蚊虫进行了大量研究, 例如: Koutsos 等人^[9]研究了冈比亚按蚊(*Anopheles gambiae*)在胚胎、5 个幼虫期、蛹、成蚊等 8 个连续生命周期中基因的表达谱, 发现在雌蚊中高表达的基因主要与免疫有关, 而雄蚊中高表达的基因则明显富集在代谢过程; Whittle 等人^[10]分别鉴定和研究了埃及伊蚊(*Aedes aegypti*)2 927 个卵巢偏好性表达、2 013 个精巢偏好性表达以及 4 449 个非性别偏好性表达的基因编码序列; Gamez 等人^[11]对白纹伊蚊(*Aedes albopictus*)在不同发育阶段的 mRNA 表达谱进行了特征描述, 其中包括雌蛹与雄蛹、雌蚊与雄蚊的转录组数据, 并着重分析和报道了性别偏倚基因。

在有关中华按蚊的研究中, Chen 等人^[12]对该物种各个发育时期进行了转录组测序和从头组装, 然后分析了非重复序列(Unigenes)的基因组图谱、功能注释、同源性、密码子使用偏差和简单序列重复(SSR); 张晶晶等人^[13]则比较分析了化学感受基因在中华按蚊雌、雄蚊的触角、喙以及下颚须中的表达情况。然而, 上述关于中华按蚊的研究主要基于发育成熟的成蚊。目前, 通过对生物的转录组进行分析, 可以揭示生物体的基因表达, 研究基因的结构变异, 检测未知基因, 还可以发现新的转录本^[14]; 为了更好地理解中华按蚊的性别差异, 进而探索和

* 收稿日期: 2020-03-18 修回日期: 2020-04-09 网络出版时间: 2020-06-01 08:40

资助项目: 国家自然科学基金(No.31872262; No.31672363); 重庆市自然科学基金面上项目(No.CSTC2019JCYJ-MSXMX0527); 重庆市教育委员会科学研究项目(No.KJQN201800523); 重庆师范大学科研启动基金(No.17XLB017)

第一作者简介: 刘杰, 男, 研究方向为生物信息学, E-mail: 1471447118@qq.com; 通信作者: 陈斌, 男, 教授, 博士, E-mail: bin.chen@cqnu.edu.cn

网络出版地址: <http://kns.cnki.net/kcms/detail/50.1165.N.20200529.1101.010.html>

开发新的病媒控制策略,有必要对中华按蚊雌、雄蚊的转录组进行比较分析。因此,本研究通过比较分析笔者所在研究团队前期 RNA-Seq 测序所获中华按蚊雌、雄蚊的转录组数据,初步筛选出 3 298 个在中华按蚊雌、雄蚊中表达存在差异的基因,将获得的差异表达基因进行 GO 和 KEGG 富集分析,挖掘这些基因的功能表现以及参与的通路,以便为研究不同性别的中华按蚊基因表达及差异提供基础资料。

1 材料与方法

1.1 转录组数据来源

研究所用的羽化 72 h 的中华按蚊雌、雄蚊转录组数据来自重庆师范大学昆虫与分子生物学研究所,其中雌、雄蚊各取 3 次生物学重复。数据主要包括比对到基因组区域或基因外显子区的测序序列(Reads)的计数(Read-counts),以及用 FPKM 标准化后的基因表达水平值^[15]。

1.2 RNA-Seq 数据相关性检查

样品间基因表达水平相关性是检验实验可靠性和样本选择是否正确的重要指标,这一指标值越接近 1,表明样品之间表达模式的相似度越高。利用 R 语言,以 FPKM 值为基础计算样品间 Pearson 相关系数的平方(R^2)。通常要求生物学重复样品间的 R^2 需大于 0.92^[2]。

1.3 差异表达基因分析

采用 DESeq 软件^[16]筛选中华按蚊雌、雄蚊之间差异表达的基因,输入数据类型为 Readcounts 数据。差异表达基因根据矫正后的 p 值(Corrected p -Value, p_{adj})与样本之间表达量的差异倍数(Fold change, V_{FC})取以 2 为底数的对数的绝对值来判断,筛选标准为 $p_{adj} < 0.05$ 且 $|\log_2 V_{FC}| \geq 0.5$ 。

1.4 差异表达基因 GO 富集分析

采用 Goseq 软件^[17]对筛选到的差异表达基因进行 GO 富集分析,一般情况下满足 $p_{adj} < 0.05$ 的条目即为显著富集条目。

1.5 差异表达基因 KEGG 富集分析

采用 KOBAS 软件^[18]对筛选到的差异表达基因进行 KEGG 通路(Pathway)分析。同样地,满足 $p_{adj} < 0.05$ 的条目即为显著富集通路。

2 结果

2.1 转录组测序结果概述

用从中华按蚊雌、雄蚊提取的总 RNA 制备了 6 个 RNA-Seq 文库,并通过 Illumina Hiseq 平台进行测序。6 个样本共产生了超过 3.23×10^8 条原始测序数据(Raw reads, RR),经过质量评估和过滤,在去除含有测序接头、未知碱基“N”以及低质量的 Reads 后,得到约 3.19×10^8 条过滤后的数据(Clean reads, CR),共含碱基个数为 47.79×10^9 。每个样本的 CR 范围为 $4.45 \times 10^7 \sim 5.90 \times 10^7$ 条,高质量的 Reads 碱基总数(Clean bases, CB)范围为 $6.68 \times 10^9 \sim 8.48 \times 10^9$ 个。各样本的测序过程碱基识别准确率达到 99% 以上的碱基占总碱基数的百分比(Q_{20})大于或等于 97.59%;而各样本的测序过程碱基识别准确率达到 99.9% 以上的碱基占总碱基数的百分比(Q_{30})大于或等于 93.17%。此外各样本的 GC 碱基对的比例在 46.74% 以上,说明各样本的基因相对较稳定。各样本实验所产生的测序序列比对个数(Total mapped, TM)范围为 $4.05 \times 10^7 \sim 5.43 \times 10^7$;它们占 CR 的平均值为 92.40%,大于 70%,满足有参分析要求(表 1)。测序共得到了 20 131 个基因,用 FPKM 值对基因表达水平进行标准化,各样本不同基因表达水平区间的基因数量和占样本基因总数的比例如表 2 所示。

2.2 RNA-Seq 整体质量评估

利用 Pearson 相关系数确定各条件下 3 个生物重复的相关性(图 1)。结果显示,生物学重复样品间相关系数的平方(R^2)均大于 0.92。因此,认为并行库的相关性是合理的,可以用于进一步分析。

2.3 差异表达基因的筛选结果

通过对中华按蚊雌、雄蚊基因的比较,发现两者的差异表达基因共 3 298 个,其中分别有 1 589 和 1 709 个基因在雌蚊中表达上调和下调(图 2)。在表达上调的基因中:1 500 个基因在雌、雄蚊间有超过 2 倍的表达变化($|\log_2 V_{FC}| \geq 1$);577 个基因在雌、雄蚊间有超过 4 倍的表达变化($|\log_2 V_{FC}| \geq 2$);186 个基因在雌、雄蚊间有超过 16 倍的表达变化($|\log_2 V_{FC}| \geq 4$)。在表达下调的基因中:有 1 568 个基因在雌、雄蚊间有超过 2 倍的表达变

化($|\log_2 V_{FC}| \geq 1$);575 个基因在雌、雄蚊间有超过 4 倍的表达变化($|\log_2 V_{FC}| \geq 2$);186 个基因在雌、雄蚊间有超过 16 倍的表达变化($|\log_2 V_{FC}| \geq 4$)。另外,有 56 个基因只在雌蚊中表达而在雄蚊中表达水平 FPKM 值为 0;有 122 个基因只在雄蚊中表达而在雌蚊中表达水平 FPKM 值为 0(表 3)。

表 1 中华按蚊雌、雄蚊转录组测序质量检测、参考基因组比对情况

Tab.1 Transcriptome sequencing quality detection and reference genome comparison of female and male *A. sinensis*

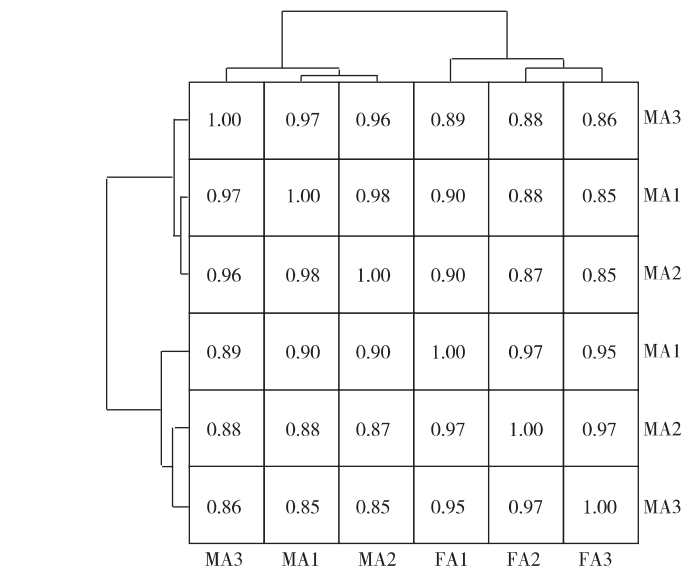
样本名称	RR/条	CR/条	CB/个	$Q_{20}/\%$	$Q_{30}/\%$	GC 碱基比例/%	TM	TM 占 CR 的比例/%
MA1	5.27×10^7	5.20×10^7	7.79×10^9	97.73	93.52	48.44	4.78×10^7	92.09
MA2	4.51×10^7	4.45×10^7	6.68×10^9	97.59	93.17	46.74	4.05×10^7	90.93
MA3	5.99×10^7	5.90×10^7	8.85×10^9	97.74	93.52	48.60	5.43×10^7	91.96
FA1	5.44×10^7	5.38×10^7	8.07×10^9	97.69	93.44	50.75	5.00×10^7	92.84
FA2	5.35×10^7	5.28×10^7	7.92×10^9	97.69	93.49	51.19	4.92×10^7	93.22
FA3	5.72×10^7	5.66×10^7	8.48×10^9	97.65	93.36	51.12	5.28×10^7	93.34

注:RR,CR,CB, Q_{20} 、 Q_{30} 和 TM 的含义见正文;MA 表示雄蚊,FA 表示雌蚊,下同

表 2 不同相对表达水平区间的基因数量

Tab.2 Number of genes at different relative expression levels

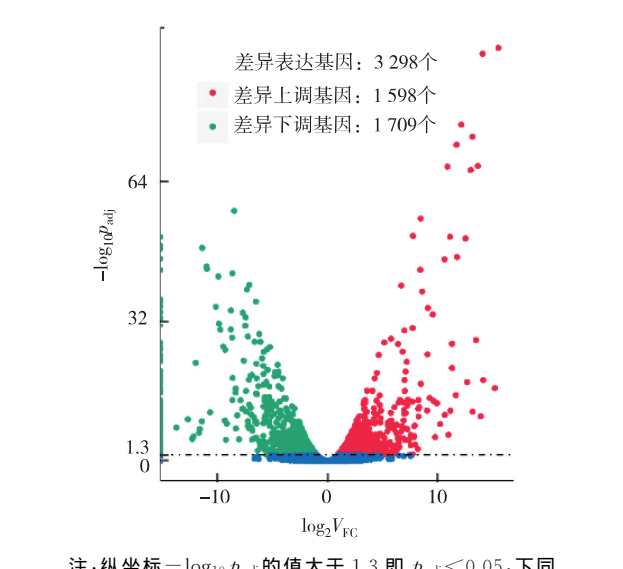
样本名称	基因相对表达水平									
	0~1		>1~3		>3~15		>15~60		>60	
	数量/个	占比/%	数量/个	占比/%	数量/个	占比/%	数量/个	占比/%	数量/个	占比/%
MA1	7 269	36.11	1 431	7.11	2 877	14.29	3 365	16.71	5 189	25.78
MA2	7 199	35.76	1 393	6.92	2 803	13.92	3 409	16.94	5 327	26.46
MA3	7 179	35.66	1 401	6.96	2 950	14.66	3 461	17.19	5 140	25.53
FA1	7 644	37.97	1 357	6.74	2 817	13.99	3 333	16.56	4 980	24.74
FA2	8 001	39.75	1 388	6.89	2 754	13.68	3 153	15.66	4 835	24.02
FA3	7 820	38.85	1 372	6.81	2 661	13.22	3 296	16.37	4 982	24.75



注:图中数据为 R^2

图 1 样本生物重复相关性检验

Fig.1 Biological repeat correlation test of samples



注:纵坐标 $-\log_{10} p_{adj}$ 的值大于 1.3 即 $p_{adj} < 0.05$,下同

图 2 中华按蚊雄蚊与雌蚊间差异表达基因的比较

Fig.2 Comparison of differential expressed genes between male and female *A. sinensis*

表 3 仅在中华按蚊雌蚊/雄蚊表达的部分基因
Tab.3 Partial genes expressed only in female or male *A. sinensis*

样本	基因
MA	AS10010045, AS10012766, AS10016379, AS10007731, AS10004323, AS10003959, AS10010178, AS10007056,
	AS10001142, AS10004936, AS10014591, AS10005391, AS10005378, AS10005298, AS10010837, AS10008888,
FA	AS10010201, AS10016370, AS10014622, AS10011814
	AS10012746, AS10001717, AS10013516, AS10012745, AS10013657, AS10002717, AS10013661, AS10013268,
	AS10012637, AS10003998, AS10007866, AS10010591, AS10004922, AS10016003, AS10006761, AS10003403,
	AS10016755, AS10002095, AS10000957, AS10012747

2.4 差异表达基因 GO 富集结果

中华按蚊雌、雄蚊的 3 298 个差异表达基因中,存在 GO 注释的有 1 964 个,共涉及 3 个类别:生物过程类、细胞组成类和分子功能类。差异表达基因显著富集在 11 个亚类,如:酰胺生物合成的过程、细胞酰胺代谢过程、肽生物合成的过程、羧酸代谢过程等(图 3)。从图 4 可知:表达上调的 1 589 个基因中,存在 GO 注释的有 1 017 个。在生物过程一类中,差异表达基因显著富集在 62 个亚类,最显著富集的亚类有细胞氮化合物代谢过程、氮化合物代谢过程、肽生物合成的过程等;其中富集到代谢过程、细胞过程和有机物代谢过程的基因数目较其他类别的基因数目更多,分别有 554,525 和 474 个。在细胞组成一类中,差异表达基因显著富集在 18 个亚类,最显著富集的亚类有细胞、细胞内部、核蛋白复合体等;其中富集到细胞和细胞部分的基因数目均为 325 个。在分子功能一类中,差异表达基因显著富集在 45 个亚类,最显著富集的亚类如杂环化合物结合、核酸结合、ATP 结合、RNA 结合等;其中富集到杂环化合物结合的基因数目最多,有 349 个。而表达下调的 1 709 个基因中,存在 GO 注释的有 947 个,且没有显著富集的亚类。

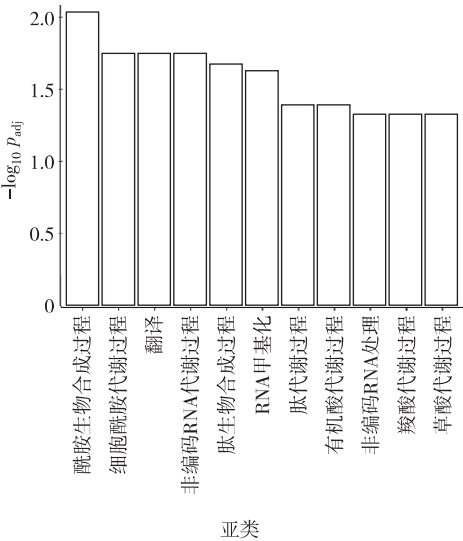


图 3 所有差异表达基因的 GO 功能分类
Fig.3 GO function classification of all differential expressed genes

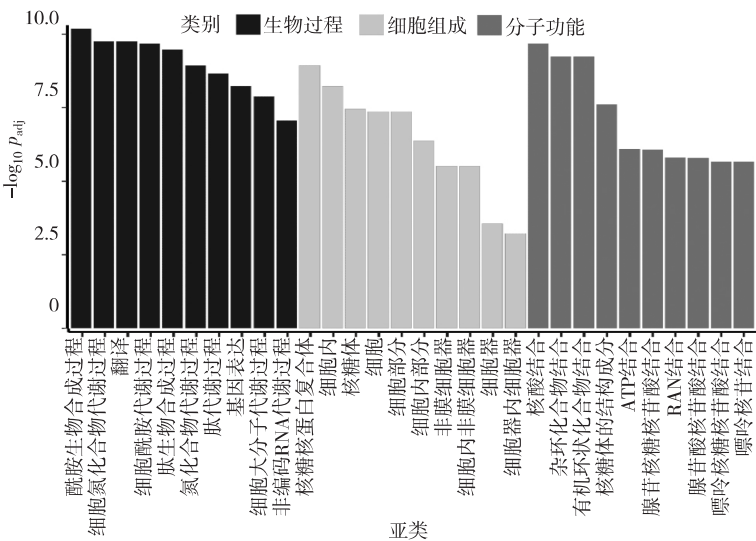


图 4 上调差异表达基因的 GO 功能分类
Fig.4 GO function classification of up-regulated differential expressed genes

2.5 差异表达基因 KEGG 富集结果

对中华按蚊雌、雄蚊比较所得的 3 298 个差异表达基因进行 KEGG 注释,结果表明,有 968 个基因被注释到 102 条 KEGG 通路上,其中显著富集的通路有 4 条,分别是 72 个差异表达基因参与的核糖体(Ribosome)通路、47 个差异表达基因参与的核糖体在真核生物中的生物发生(Ribosome biogenesis in eukaryotes)通路、47 个差异表达基因参与的 RNA 转运(RNA transport)通路以及 26 个差异表达基因参与的 RNA 降解(RNA degradation)通路(表 4)。

3 讨论

中华按蚊雌、雄蚊表现出高度的性别二态性,例如在形态上,雄蚊触角的轮生毛较雌蚊长而密,并且梗节也

明显比雌蚊大;雌蚊顶鳞较密,有明显的沟区,头部最后方的鳞片呈黄褐色,前部鳞片颜色淡,而雄蚊较雌蚊色淡,且触角梗节无鳞片^[19]。又如,在行为上,成年中华按蚊雌、雄蚊主要以花蜜和蜜露中的糖分为食,以满足基础代谢和飞行的能量需求。但除此之外雌蚊还需吸食血液以完成卵的发育^[20]。目前有不少研究涉及不同性别的蚊虫转录组,包括不同性别的组织^[21]以及不同性别发育阶段(如雌蛹和雄蛹)^[11]。通过转录组分析,

可研究具有性别偏好性表达的基因,如 Hahn 等人^[22]报道了在冈比亚按蚊雌蚊中偏好性表达的基因。也可比较分析在蚊虫性别间差异表达的基因,为研究蚊虫性别差异的分子机制提供有用的信息。

笔者所在研究团队对中华按蚊的分类学及分子生物学开展了大量的研究。本研究基于该团队前期通过 RNA-Seq 测序所得的中华按蚊雌、雄蚊的转录组数据,在经过数据相关性检查后,对这些数据进行了差异比较分析。共筛选出 3 298 个差异表达基因,其中 1 589 个基因在雌蚊中表达上调,1 709 个基因在雌蚊中表达下调。表达上调的基因中,有 56 个基因只在雌蚊中表达,这些基因主要与先天免疫反应的激活、防卫正调节有关。这与前人的报道相符^[9],说明雌蚊因吸食血液,故需激活免疫以增加生存的机会^[23];表达下调的基因中,有 122 个基因只在雄蚊中表达。所有差异表达基因 GO 和 KEGG 富集分析结果表明,这些基因主要与酰胺生物合成、细胞酰胺代谢、RNA 甲基化、有机酸代谢、肽生物合成、羧酸代谢等生物学过程有关,参与核糖体、核糖体在真核生物中的生物发生、RNA 转运及降解等通路。推测其中原因,可能是在蚊虫的生活史中,羽化 72 h 后的雌蚊开始经历最重要的环节,即吸血并完成交配,为繁殖后代做准备。在这期间,雌蚊集中于合成与吸血和血液消化相关的蛋白;而羽化 72 h 后的雄蚊将完成精子的储备工作以备交配。在这期间,雄蚊集中于大量合成与精子形成有关的蛋白。故差异表达基因主要参与肽生物合成等过程。由于生物体中蛋白质的合成与所行使的功能是密切相关的,因此蛋白质的成功合成是至关重要的。这个历程与核糖体、RNA 转运等活动密不可分。RNA 从细胞核到细胞质的转运是基因表达的基础。核糖体是负责制造蛋白质的细胞工厂。在真核生物中,核糖体生物发生涉及 4 种 rRNA 和大约 80 种核糖体蛋白的产生和正确装配。在缺乏这些蛋白的情况下,即使在最佳生长条件下,也会造成核糖体的生物发生停滞,细胞生长终止。因此核糖体在真核生物中的生物发生通路、RNA 转运通路对蛋白质的成功合成尤为重要。同样地, RNA 降解通路在细胞 RNA 分子的正确加工、质量控制和更新等方面起着重要作用,会影响基因的表达。因此,中华按蚊雌、雄蚊的基因表达围绕核糖体和 RNA 出现差异,并且差异表达基因集中在核糖体通路、核糖体在真核生物中的生物发生通路、RNA 转运通路以及 RNA 降解通路。总之,本研究结果对深入研究不同性别的中华按蚊基因表达及差异具有潜在的价值。

参考文献:

- [1] FENG X Y, ZHANG S S, HUANG F, et al. Biology, bio-nomics and molecular biology of *Anopheles sinensis* Wiedemann 1828 (Diptera: Culicidae), main malaria vector in China[J]. *Frontiers in Microbiology*, 2017, 8: 1473.
- [2] CHEN Q, PEI D, LI J Y, et al. The antenna transcriptome changes in mosquito *Anopheles sinensis*, pre- and post-blood meal[J]. *PLoS ONE*, 2017, 12(7): e0181399.
- [3] World Health Organization. World malaria report 2019[EB/OL]. (2019-12-04) [2020-03-18]. <https://www.who.int/publications-detail/world-malaria-report-2019>.
- [4] PLUESS B, TANSER F C, LENGELER C, et al. Indoor residual spraying for preventing malaria[EB/OL]. (2010-04-14) [2020-03-18]. <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD006657.pub2/full>.
- [5] CUI F, RAYMOND M, QIAO C L. Insecticide resistance in vector mosquitoes in China[J]. *Pest Management Science*, 2006, 62(11): 1013-1022.
- [6] KIM H, BAEK J H, LEE W J, et al. Frequency detection of pyrethroid resistance allele in *Anopheles sinensis* populations by real-time PCR amplification of specific allele (rt-PASA)[J]. *Pesticide Biochemistry and Physiology*, 2007, 87(1): 54-61.
- [7] ZHONG D, CHANG X, ZHOU G, et al. Relationship between knockdown resistance, metabolic detoxification and organismal resistance to pyrethroids in *Anopheles sinensis* [J]. *PLoS One*, 2013, 8(2): e55475.
- [8] HALL A B, BASU S, JIANG X, et al. A male-determining factor in the mosquito *Aedes aegypti*[J]. *Science*, 2015, 348(6240): 1268-1270.
- [9] KOUTSOS A C, BLASE C, MEISTER S, et al. Life cycle

表 4 差异表达基因的 KEGG 通路显著富集分析结果
Tab.4 Significant enrichment results of the KEGG pathway of the different expressed genes

KEGG 条目	描述	矫正后的 p 值
aga03010	核糖体	5.18×10^{-7}
aga03008	核糖体在真核生物中的生物发生	1.73×10^{-6}
aga03013	RNA 转运	4.08×10^{-2}
aga03018	RNA 降解	4.08×10^{-2}

- transcriptome of the malaria mosquito *Anopheles gambiae* and comparison with the fruitfly *Drosophila melanogaster* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(27):11304-11309.
- [10] WHITTLE C A, EXTAVOUR C G. Rapid evolution of ovarian-biased genes in the yellow fever mosquito (*Aedes aegypti*) [J]. *Genetics*, 2017, 206(4):2119-2137.
- [11] GAMEZ S, ANTOSHECHKIN I, MENDEZ-SANCHEZ S C, et al. The developmental transcriptome of *Aedes albopictus*, a major worldwide human disease vector [J]. *G3: Genes, Genomes, Genetics*, 2020, 10(3):1051-1062.
- [12] CHEN B, ZHANG Y J, HE Z B, et al. *De novo* transcriptome sequencing and sequence analysis of the malaria vector *Anopheles sinensis* (Diptera: Culicidae) [EB/OL]. (2014-07-07) [2020-03-18]. <https://parasitesandvectors.biomedcentral.com/articles/10.1186/1756-3305-7-314>.
- [13] 张晶晶. 中华按蚊化感组织的转录组测序及化感基因的表达谱分析[D]. 重庆: 重庆师范大学, 2019.
- ZHANG J J. Transcriptome sequencing and expression profiling of allelopathic genes of *Anopheles sinensis* [D]. Chongqing: Chongqing Normal University, 2019.
- [14] WANG Z, GERSTEIN M, SNYDER M. RNA-Seq: a revolutionary tool for transcriptomics [J]. *Nature Reviews Genetics*, 2009, 10(1):57-63.
- [15] TRAPNELL C, WILLIAMS B A, PERTEA G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation [J]. *Nature Biotechnology*, 2010, 28(5):511-516.
- [16] ANDERS S, HUBER W. Differential expression analysis for sequence count data [EB/OL]. (2010-10-27) [2020-03-18] <https://link.springer.com/article/10.1186/gb-2010-11-10-r106>.
- [17] YOUNG M D, WAKEFIELD M J, SMYTH G K, et al. Gene ontology analysis for RNA-seq: accounting for selection bias [EB/OL]. (2010-02-04) [2020-03-18]. <https://link.springer.com/article/10.1186/gb-2010-11-2-r14>.
- [18] MAO X, CAI T, OLYARCHUK J G, et al. Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary [J]. *Bioinformatics*, 2005, 21(19):3787-3793.
- [19] 张莉娟. 中华按蚊不同龄期发育历期、有效积温和发育形态研究[D]. 重庆: 重庆师范大学, 2019.
- ZHANG L J. Studies on the developmental duration, effective accumulated temperature and developmental morphology of *Anopheles sinensis* in different ages [D]. Chongqing: Chongqing Normal University, 2019.
- [20] DISSANAYAKE S N, RIBEIRO J M C, WANG M H, et al. aeGEPUCI: a database of gene expression in the dengue vector mosquito, *Aedes aegypti* [EB/OL]. (2010-10-04) [2020-03-18] <https://bmcresnotes.biomedcentral.com/articles/10.1186/1756-0500-3-248>.
- [21] RIBEIRO J M C, MARTIN-MARTIN I, MOREIRA F R, et al. A deep insight into the male and female salivary transcriptome of adult *Culex tarsalis* mosquitoes [J]. *Insect Biochemistry and Molecular Biology*, 2018, 95:1-9.
- [22] HAHN M W, LANZARO G C. Female-biased gene expression in the malaria mosquito *Anopheles gambiae* [J]. *Current Biology*, 2005, 15(6):R192-R193.
- [23] KOKOZA V, AHMED A, CHO W L, et al. Engineering blood meal-activated systemic immunity in the yellow fever mosquito, *Aedes aegypti* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97(16):9144-9149.

Animal Sciences

Comparative Transcriptome Analysis of Female and Male *Anopheles sinensis*

LIU Jie, LI Bo, CHEN Bin

(Chongqing Key Laboratory of Vector Insects, Institute of Insect and Molecular Biology, Chongqing Normal University, Chongqing 401331, China)

Abstract: [Purposes] Differential expression analysis was carried out on transcriptome sequencing data of female and male *Anopheles sinensis* to obtain differentially expressed genes, carry out functional enrichment analysis on them, and dig out the functional expression and involved pathway of these genes. [Methods] Based on the transcriptome data of female and male *A. sinensis* obtained in the laboratory, the correlation test of RNA-Seq data was carried out. The differential expression of genes in female and male mosquitoes was compared by using the DESeq package of R language, and the GO and KEGG enrichment analysis of the screened differential genes was conducted by Goseq and KOBAS. [Findings] Pearson correlation coefficient squared (R^2) was greater than 0.92, indicating that the samples were highly correlated among the three biological replicates and could be used for further analysis. There were 3 298 different genes in female and male mosquitoes, of which 1 589 were up-regulated and 1 709 were down-regulated. The differentially expressed genes were significantly enriched in functional categories such as amide biosynthesis, organic acid metabolism and peptide metabolism, and were involved in the KEGG pathways such as ribosome biogenesis in eukaryotes and RNA transport. [Conclusions] It is of potential value to study gene expression and difference of different sex in *A. sinensis*.

Keywords: *Anopheles sinensis*; transcriptome; GO analysis; KEGG analysis

(责任编辑 方 兴)