研究论文

棉铃虫雌雄性腺转录组的比较分析*

刘香亚** 江 婷 肖海军 张万娜***

（江西农业大学昆虫研究所，南昌 330045）

摘 要 **【目的】** 棉铃虫 *Helicoverpa armigera*（Hübner）是我国重要的农业害虫，生殖力强是其发生为害的内在生物学基础，而已知的与棉铃虫生殖相关的基因信息相对较少，本研究旨在筛选棉铃虫性腺发育的相关基因。**【方法】** 通过高通量测序技术对棉铃虫成虫的卵巢和精巢进行转录组测序。**【结果】** 共获得100 603 条 unigenes，平均长度为 666.05 bp，其中 N50 为 1 114 bp。经同源性比对，52 071 条 unigenes 获得注释信息，其中注释到 Nr 数据库的序列最多。通过比较转录组分析发现，在棉铃虫精巢和卵巢中存在 7 714 个差异表达的基因（Differentially expressed genes，DEGs），其中 3 288 个 DEGs 表达水平在卵巢中上调，4 426 个 DEGs 在精巢中上调。差异基因富集结果涉及到生殖系统发育的相关通路有 mTOR 信号通路、卵母细胞减数分裂、促性腺激素释放激素信号通路和胰岛素信号通路等。同时筛选得到部分与性腺发育相关的基因，包括 *vitellogenin*（*Vg*）、*vitellogenin receptor*（*VgR*）、*chorion-related gene*、*testis-specific serine/threonine-protein kinase* 和 *spermatogenesis-associated protein* 等。选取其中 12 个 DEGs 用实时荧光定量检测其表达量，结果表明 RT-qPCR 与 RNA-Seq 的结果一致。**【结论】** 本研究获得了棉铃虫雌雄性腺的转录组数据及主要性腺发育相关基因，将有助于丰富棉铃虫繁殖相关的基因资源，为进一步研究棉铃虫生殖调控机理提供基础数据。

关键词 棉铃虫；高通量测序；繁殖；卵巢；精巢

Comparative transcriptome analysis of sex-biased genes in the ovary and testis of *Helicoverpa armigera*

LIU Xiang-Ya** JIANG Ting XIAO Hai-Jun ZHANG Wan-Na***

(Institute of Entomology, Jiangxi Agricultural University, Nanchang 330045, China)

Abstract **[Objectives]** The cotton bollworm, *Helicoverpa armigera*, is one of the most serious pests of cotton, corn and many other crops in China. Although high reproduction is the basis for outbreaks of this pest, information on reproduction-related genes of this species is lacking. **[Methods]** We sequenced ovary and testis transcriptomes of *H. armigera* on an Illumina HiSeq2500 platform and identified numerous genes that may be involved in gonadal development, gametogenesis and reproduction. [**Results**] Raw reads were assembled into 100 603 unigenes with a mean length of 666.05 bp and an N50 of 1 114 bp. A similarity search resulted in 52 071 unigenes being annotated, most of which with reference to the Nr database. Gene expression in the ovary and testis were then compared, and 7 714 differentially expressed genes (DEGs) were identified. Among these genes, 3 288 were up-regulated in the ovary and 4 426 were up-regulated in the testis. Based on KEGG enrichment analyses, the pathways involved in the regulation of gonadal development and gametogenesis were identified, including the “mTOR signaling”, the “Oocyte meiosis”, “Insulin signaling” and “GnRH signaling”, pathways. To further screen the transcriptome database, numerous DEGs involved in gonadal development and gametogenesis were identified, including *vitellogenin* (*Vg*), *vitellogenin receptor* (*VgR*), *chorion-related genes*, *testis-specific serine/threonine- protein kinase*

*资助项目 Supported projects：国家自然科学基金项目（32060642）；江西省自然科学基金项目（20202ACBL205004）；江西省自然科学基金项目（20212BAB205004）

**第一作者 First author，E-mail：liuxiangya715@163.com

***通讯作者 Corresponding author，E-mail：zhangwanna880210@yeah.net

收稿日期 Received：2021-03-22；接受日期 Accepted：2021-09-27

and *spermatogenesis-associated protein*. Finally, twelve DEGs were selected for RT-qPCR analysis, the expression patterns of which were consistent with the results obtained from transcriptome sequencing. **[Conclusion]** This study acquired the transcriptome data and identified a series of genes related to gonadal development, providing a basis for exploring the mechanisms regulating reproduction in *H. armigera*.

Key words *Helicoverpa armigera*; RNA-Seq; reproduction; ovary; testis

近年来高通量测序技术的迅速发展为昆虫分子生物学研究提供基因数据，转录组数据分析已经成为发掘非模式昆虫功能基因的有效手段，广泛应用于昆虫生殖调控、农药抗性、种群分化和生态适应等方向研究（Legeai and Derrien，2015；Oppenheim *et al.*，2015；Zhou *et al.*，2018）。性腺是昆虫繁殖的主要器官，精巢和卵巢作为生殖系统的重要组成部分，只有发育成熟并正常行使其生理功能，才能产生有活性的精子和卵子，保障种群的繁衍。深入发掘昆虫性腺发育的功能基因，有助于阐明害虫生殖调控的分子机制，为害虫综合治理提供支持。近年来转录组测序技术已广泛应用于昆虫性腺的研究，用以解析昆虫生殖调控的机制。例如，杨婧（2013）对短额负蝗 *Actractomorpha sinensis* 若虫和成虫进行转录组分析，鉴定了多个昆虫激素合成和调控的通路基因，推测组织蛋白酶可能参与卵子发生。在桔小实蝇 *Bactrocera dorsalis* 中，利用多组学技术鉴定出了参与其精子发生和卵子发生的关键基因，有助于了解卵巢发育和精子发生的分子机制（Wei *et al.*，2015，2017）。对意大利蝗 *Calliptamus italicus* 精巢和卵巢组织的转录组进测序，筛选得到性腺发育的有关基因，为后续研究其生殖调控机理提供数据支持（向敏等，2017）。Peng 等（2017）对不同发育级别的小菜蛾 *Plutella xylostella* 卵巢进行转录组测序分析，结果发现卵壳蛋白基因和卵膜蛋白基因在卵巢内大量表达，推测其与卵巢发育密切相关。

棉铃虫 *Helicoverpa armigera*（Hübner）属于鳞翅目夜蛾科，是我国重大的农业害虫，为害棉花、小麦、玉米、大豆和花生等多种作物。高通量测序技术已广泛用于棉铃虫功能基因和代谢调控的研究。例如，Zhao 等（2013）对棉铃虫幼虫取食致病菌后的血淋巴和脂肪体组织进行转录组测序，发现差异表达的基因主要涉及免疫、解毒代谢和个体发育，结果有助于丰富昆虫抗菌免疫的分子机制。Zhang 等（2015）通过构建棉铃虫雌雄成虫触角转录组文库，鉴定出 133 个嗅觉基因，为全面阐释棉铃虫的嗅觉机制奠定了基础。刘云泽等（2019）通过对单个棉铃虫抗性品系与敏感品系进行转录组测序，筛选得到与 Bt 抗性相关的靶标基因。Jin 等（2019）通过对 5 种不同杀虫剂处理的棉铃虫幼虫进行转录组测序，阐明了 54 个与解毒代谢相关的 ATP 结合盒式转运蛋白（ATP-binding cassette transporters，ABC transporters）的差异表达。

成虫繁殖力强是引起棉铃虫种群危害与区域暴发成灾的关键所在，通过干扰棉铃虫成虫的繁殖潜力可以减少其种群数量，而棉铃虫生殖相关信息的匮乏限制了相关策略的发展。因此，本文利用高通量测序技术对棉铃虫成虫精巢和卵巢组织进行转录组测序，以期发掘调控棉铃虫生殖的重要功能基因，为基于生殖调控的策略防治害虫提供支持。

1 材料与方法

1.1 供试昆虫饲养与材料收集

供试虫源为室内用人工饲料长期饲养的棉铃虫种群，人工饲料配方参照梁革梅等（1999）。饲养条件为温度（27 ± 2）℃，相对湿度 70% ± 5%，光周期为 14 L：10 D。待幼虫化蛹后，将雌蛹、雄蛹分别放入养虫笼（40 cm×26 cm×21 cm）内，待成虫羽化后饲喂 10%的蜂蜜水。随机选取羽化后 3 d 未交配的棉铃虫雌雄蛾，置于体视显微镜下解剖，将获取卵巢和精巢迅速投入液氮中冷冻，然后置于 - 80 ℃冰箱保存。将 10 头雌蛾的卵巢组织作为一个检测样品，共设置 3 个生物学重复。用同样的方法准备棉铃虫雄蛾精巢的检测样品。

1.2　RNA 提取与 cDNA 文库构建

采用 Trizol 法提取棉铃虫总 RNA，然后用 1%的琼脂糖凝胶电泳检测获取总 RNA 的完整性，通过微量分光光度计 Nanodrop2000 测定样品的 RNA 浓度，最后使用 Agilent 2100 bioanalyzer 检测 RNA 完整值。选取 1.8< OD_{260}/OD_{280}<2.0，RNA 完整度值≥8.0 的样品，委托北京百迈客生物科技有限公司使用 Illumina HiSeq2500 平台进行转录组测序。

1.3　转录组数据组装和功能注释

为了保证数据分析的质量及可靠性，对测序获得的原始数据进行质控分析，经过去除低质量序列及去接头等质控过程，最终获得高质量的质控数据（clean reads）。同时，计算 Q20 和 Q30，GC 含量和重复序列水平，并利用 Trinity 软件对获得的 clean data 数据进行拼接组装，去除冗余序列后最终获得 unigene。随后对转录本进行功能注释，将 unigene 序列分别与 NR、Swiss-Prot、Pfam、CO、GOG 和 KEGG 数据库进行 BLAST 比对，获取转录本的功能注释信息。

1.4　差异基因的表达与分析

以拼接好的转录组数据作为参考数据库，使用 Bowtie 将各样品测序所得的 reads 与数据库进行比对。根据比对结果，结合 RSEM 进行表达量水平估计。利用 FPKM（Fragments per kilobase per million）值量化转录本的表达丰度，然后使用 DEGSeq 软件分析样品间的基因差异表达情况。样品间差异表达基因的筛选阈值为$|\log_2(\text{fold change})|>1$ 且 FDR<0.01。对差异基因进行 KEGG 通路富集分析、GO 功能及 Pfam 结构域分析。

1.5　棉铃虫参与性腺发育基因的筛选

结合已有报道，本研究根据雌雄性腺差异表达基因的功能注释信息，筛选得到了大量棉铃虫卵巢和精巢发育的相关功能基因（Nie *et al.*，2011；Peng *et al.*，2017；Wei *et al.*，2017；向敏等，2017）。另外，我们从转录组数据中筛选出雌雄性腺特异表达的基因，筛选阈值为$|\log_2(\text{fold change})|>10$ 且 unigene 在雌雄性腺的 FPKM 值（其中较小值）<10。然后利用 RT-qPCR 对 12 个 DEGs 在棉铃虫雌雄性腺的表达水平进行检测（表 1）。以 *β-actin* 作为内参基因，RT-qPCR 的反应体系和操作参考作者之前的报导（Zhang *et al.*，2021）。采用 $2^{-\Delta\Delta Ct}$ 法分析基因相对表达水平，使用 SPSS17.0 软件对结果进行统计分析。

表 1　本实验所用的引物序列

Table 1　Primers used in this experiment

引物 Primer	基因名称 Gene name	引物序列 Primer sequence（5′-3′）	
		正向引物 Forward primer	反向引物 Reverse primer
c40012	*keratin*	GCTTCAACAGCTACCCGAAC	CGTATCCTCCTCCACAGCAT
c37343	*leucine-rich repeat extension-like protein 1*	TCCATACCCCTCTCATACGC	TAAGCTGGAGTGGCTGGTCT
c40761	*lipase 3*	TCCCTGTGGTATTGGGACAT	TCTGCAGTCTCACGACATCC
c40153	*cortex*	TATGCTGCTTTTCCCGAGTT	CTTGCAGCAGTGGATGTTGT
c30351	*G2/mitotic-specific cyclin*	AGACACCGAACGAAGCAGAT	AGTGTTGATGTCCCCCAAAG
c38734	*cell division cycle protein 20*	ACCTCAACCTCGTTGACTGG	GAGCAAACAGTCTCGGAACC
c32244	*transcription factor Sox 14*	GGTGCACAGTGGAAAGACCT	TGTAATCCGGGTGTTCCTTC
c31141	*elastin-like protein*	CTTAGGAGGCCTAGGCAGTG	CACCTCCATAGCCAATACCG
c34667	*spidroin*	AGGTTACAGCGATGCAAGTG	GTAGCCTCCGTATCCACCAT
c36498	*threonine-protein kinase 2*	TAGCGGCCTGCAATACCTAC	AATCCGCCAGTTTTACGTTG
c35887	*trypsin3A*	AGTATCAGCCTGCGGTCTTC	ATCTCAGGGTGGATCTCCTG
c44218	*chitinase*	TGCGGTATCGTCTGTGAGAG	GGGCAACCTTCACTAATCCA
β-actin	*β-actin*	GCGACATCAAGGAGAAGCTG	CGTCGCACTTCATGATGGAG

2 结果与分析

2.1 棉铃虫卵巢和精巢转录组测序与组装

对棉铃虫卵巢和精巢样本进行转库组测序，每个样本均产生 8.13 Gb 以上的高质量数据（表 2）。GC 含量在 48.18%- 51.81%之间。每个样本 Q20 均大于 97%，Q3 均大于 93%（表 2）。经 Trinity 组装，共得到 145 345 个转录本，其中 41 778 个长度大于 1 000 bp，N50（覆盖 50%所有核苷酸的最大 unigene 长度）长度为 2 355 bp。对转录本进行组装分析，获得 100 603 个 unigenes，其中长度在 1 kb 以上的 unigene 有 13 573 条，平均长度为 666 bp（表 3）。

表 2 棉铃虫转录组测序产量统计

Table 2 Statistics of transcriptome assembly for *Helicoverpa armigera*

样品名 Sample	质控数据（个） Clean data (*n*)	数据大小（Gb） Clean bases (Gb)	GC 含量（%） GC content (%)	Q20 (%)	Q30 (%)
精巢 1 Testis 1	30 206 422	9.00	48.97	97.43	93.30
精巢 2 Testis 2	27 371 235	8.17	49.24	97.36	93.11
精巢 3 Testis 3	29 929 395	8.93	48.18	97.38	93.18
卵巢 1 Ovary 1	27 327 955	8.13	51.81	97.65	93.72
卵巢 2 Ovary 2	29 899 952	8.91	51.44	97.58	93.58
卵巢 3 Ovary 3	33 339 151	9.94	51.55	97.54	93.47

表 3 棉铃虫转录组测序数据的组装质量分析

Table 3 Quality of the assembly results of *Helicoverpa armigera* transcriptome

长度范围 Length range (bp)	转录本 Transcripts		unigenes	
	数量 Number	百分比 Percentage	数量 Number	百分比 Percentage
<500	79 905	54.97	72 518	72.08%
500-1 000	23 662	16.28	14 512	14.43%
1 000-2 000	18 758	12.91	6 895	6.85%
>2 000	23 020	15.84	6 678	6.64%
总数 Total number	145 345		100 603	
总长度（bp） Total length (bp)	154 402 239		67 006 574	
N50 长度（bp） N50 length (bp)	2 355		1 114	
平均长度（bp） Mean length (bp)	1 062.32		666.05	

2.2 转录本功能注释和分类

将获得转录本序列与 Nr、Swiss-Prot、GO、COG、Pfam 和 KEGG 等数据库进行比对，共有 52 071 条 unigenes 获得注释结果。与 Nr 数据库比对发现，其中与棉铃虫相匹配的基因序列最多（34.21%），其次为为烟芽夜蛾 *Heliothis virescens*（7.76%），柑桔凤蝶 *Papilio xuthus*（4.99%），斜纹夜蛾 *Spodoptera litura*（4.23%）等（图 1）。根据 GO 数据库的注释结果，注释较多的功能分属于“代谢过程”（Metabolic process）、“催化活性”（Catalytic activity）和“细胞”（Cell），分别包含 9 100、9 386 和 6 272 条 unigenes，分属“生物学过程”、“分子功能”和“细胞成分”大类（图 2）。此外，经过 KOG 功能分析划分为 25 类（图 3），其中 19.04%的 unigenes 划归为“一般功能”（General function prediction only），然后依次为“翻译后修饰”（9.89%），“信号转导机制”（8.66%），“翻译、

核糖体结构和生物合成”（6.99%）。根据 KEGG 数据库的注释结果，15 810 个 unigenes 可分为 6 个分支 44 个二级通路，其中与“翻译 translation”过程相关的 unigene 最多（图 4）。

2.3　差异表达基因分析

对雌雄性腺差异表达的基因进行分析，发现有 2 077 个 unigenes 在卵巢特异性表达，2 655 个在精巢特异性表达（图 1）。以精巢为基准，分析棉铃虫性腺基因的差异表达情况，发现 3 288 个表达量上调，4 426 个下调（图 1）。差异基因经 GO 功能富集分析发现，在“生物过程”大类中“代谢过程”和“细胞过程”所含 unigene 最多；在“细胞组分”亚类中，“细胞”和“细胞部分”包含注释 unigene 的最多；在“分子功能”下属分类中，“结合”和“催化活性”所占比例最大（图 2）。

通过 KEGG 功能富集发现，DEGs 映射到一级通路中涉及“遗传信息处理”所占比例最多，其次为“代谢通路”。所涉及的二级分支代谢通路中，“翻译”、“折叠、分类和降解”及“信号转导”占据前 3 位。此外，涉及“mTOR 信号通路”的基因在卵巢中表达量显著上调（图 3：A），部分卵巢上调表达的基因富集到卵母细胞减数分裂（Oocyte meiosis），胰岛素信号通路（Insulin signaling pathway），甾体激素生物合成（Steroid hormone biosynthesis）和 TGF-β 信号通路。而在卵巢下调表达基因中，涉及寿命调节通路（Longevity regulation pathway）和吞噬体（Phagosome）的通路转录本显著富集（图 3：B）。

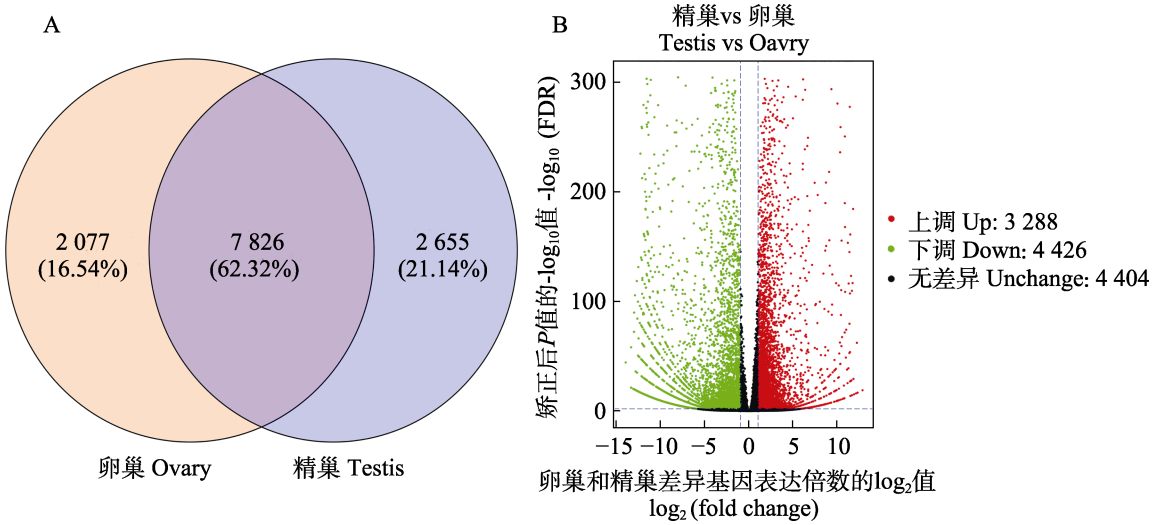

图 1　棉铃虫卵巢和精巢差异表达基因的比较

Fig. 1　Comparison of differentially expressed genes (DEGs) in ovary and testis of *Helicoverpa armigera*

A. 卵巢和精巢基因表达差异的维恩图；B. 卵巢和精巢表达基因的散点图。

A. Venn diagram showing the DEGs between ovary and testis; B. Scatter plot showing the DEGs between ovary and testis.

2.4　棉铃虫生殖相关基因分析

根据转录组分析结果，结合相关文献报道，我们鉴定出与雌雄性腺发育相关的一些基因（表 4）。其中 *chorion*、*Vg*、*VgR* 及 *lipase 3-like* 等基因在卵巢中显著上调，相关基因涉及卵子发生（Oogenesis）、卵泡细胞发育（Ovarian follicle cell development）、卵壳蛋白形成（Chorion-containing eggshell formation）、卵母细胞发育（Oocyte development）及卵母细胞分化（Oocyte differentiation）等雌性生殖器官发育过程。而 *glycine-rich protein*、*spidroin-like* 和 *testis-specific serine/threonine-protein kinase* 等基因在精巢中显著高表达。选取其中 12 个差异基因 DEGs，利用 RT-qPCR 对其表达量进行验证。结果表明，RT-qPCR 检测和 RNA-Seq 测序结果的一致（图 4），说明 RNA-Seq 测序结果可信。另外，我们筛选

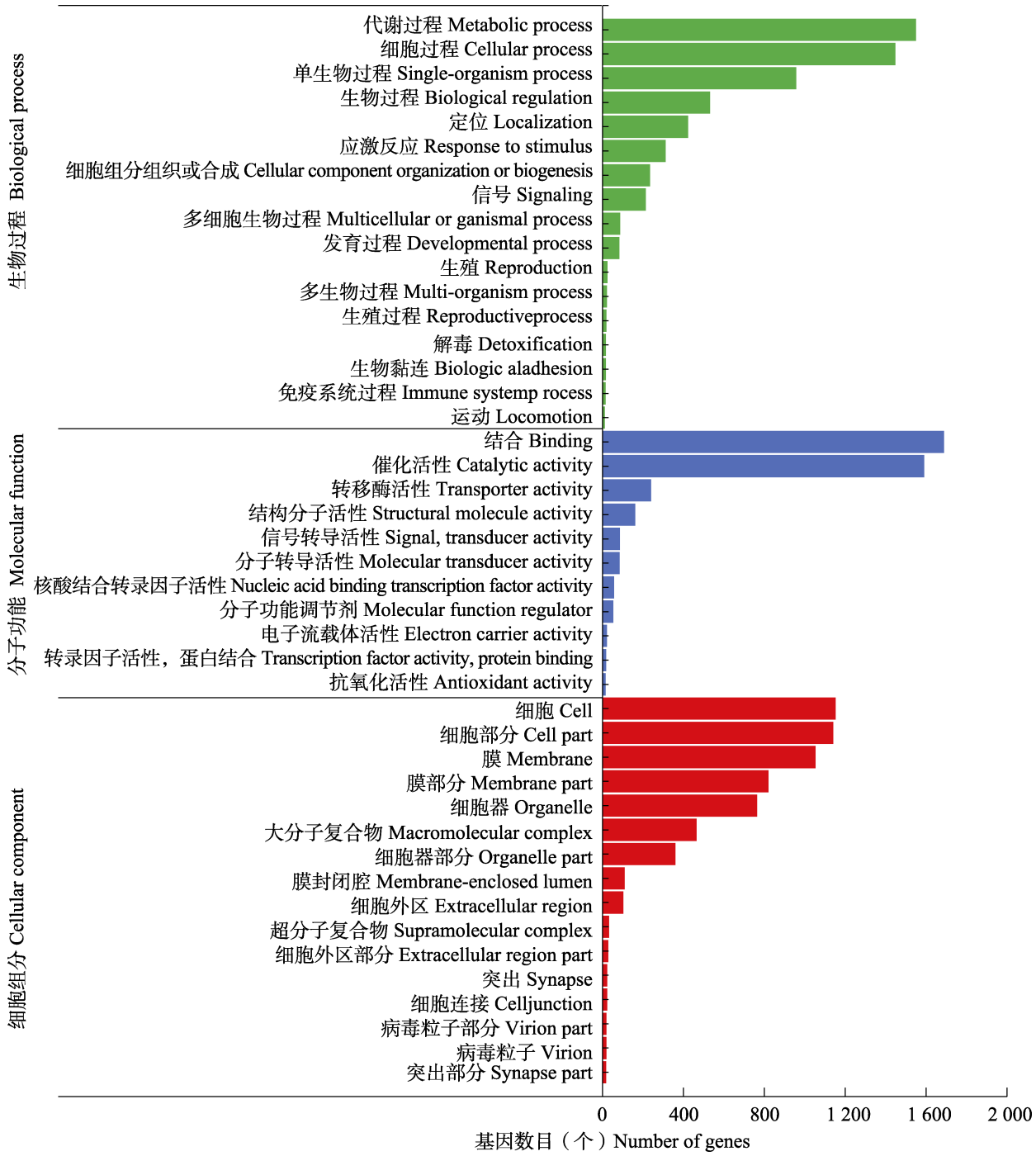

图 2　棉铃虫卵巢和精巢差异表达基因的 GO 功能分类

Fig. 2　GO classification of the differentially expressed genes (DEGs) in ovary and testis of *Helicoverpa armigera*

到了在卵巢或精巢特异表达的基因（表 5），其中部分基因缺少功能注释信息。

3　讨论

棉铃虫是我国一种重大的农作物害虫，繁殖力强是其发生为害的内在生物学基础（郭予元，1998），而已知的棉铃虫性腺发育相关的基因信息较少。为了探索棉铃虫性腺发育及繁殖机制，阐明性腺基因表达模式，本研究选取性成熟期的棉铃虫卵巢和精巢组织进行转录组测序，以期获

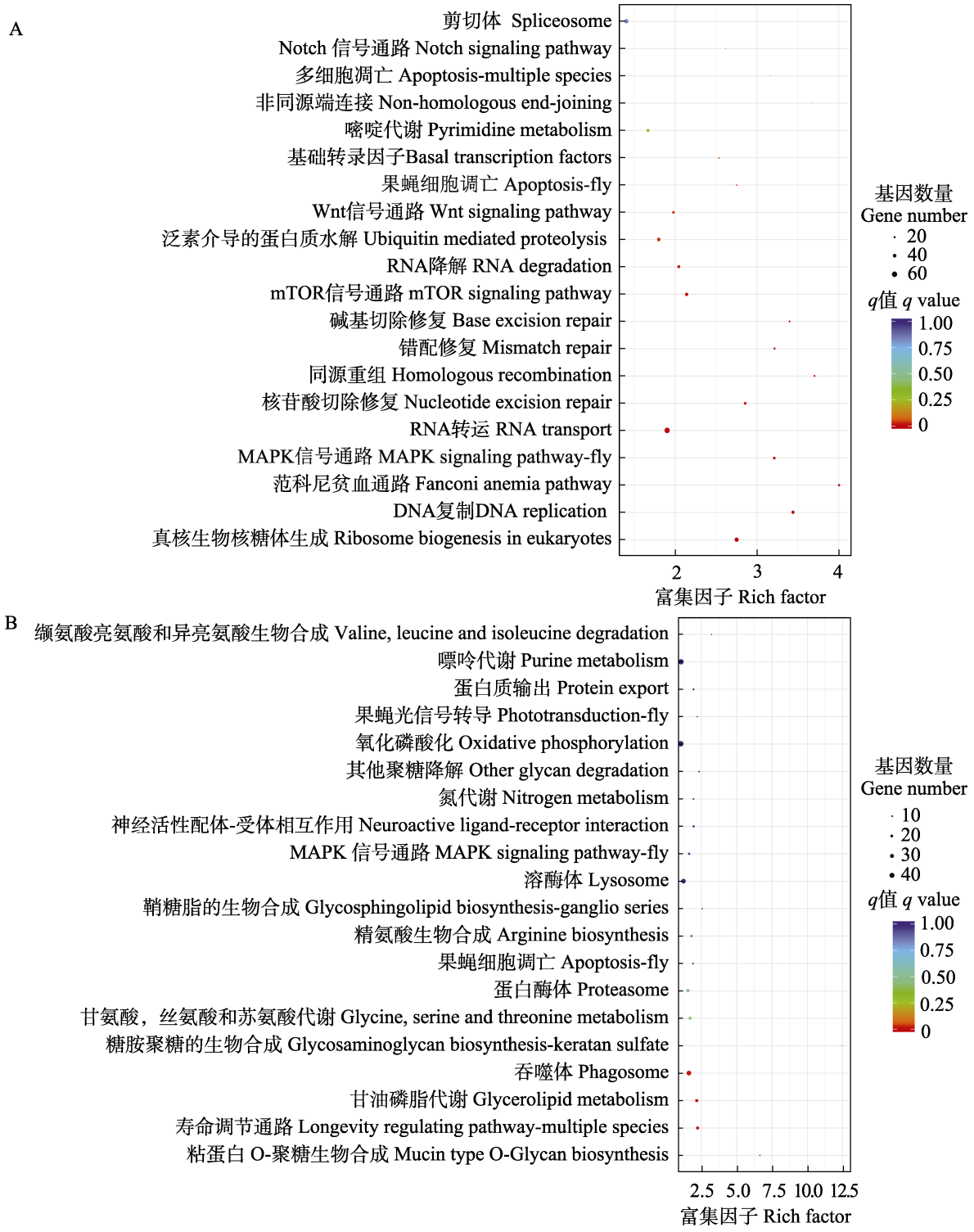

图 3　棉铃虫卵巢与精巢差异表达基因的 KEGG 代谢通路富集分析

Fig. 3　KEGG pathway enrichment of the differentially expressed genes (DEGs) in ovary and testis of *Helicoverpa armigera*

A. 卵巢中上调表达基因的 KEGG 代谢通路富集分析；B. 卵巢中下调表达基因的 KEGG 代谢通路富集分析。

A. KEGG pathway enrichment analysis of the up-regulated DEGs in ovary relative to testis;
B. KEGG pathway enrichment analysis of the down-regulated DEGs in ovary relative to testis.

表 4 棉铃虫性腺发育相关基因的鉴定以及表达量分析

Table 4 Identification and the expression analysis of genes associated with ovary and testis development of *Helicoverpa armigera*

基因序列号 unigene ID	功能注释 Functional annotation	$\log_2$FC	*q* 校验值 *q* value
卵巢发育相关基因 Genes associated with ovary development			
c32631.graph_c0	chorion class B protein Ld34-like	11.85	5.31×10^{-16}
c36551.graph_c0	chorion class B protein PC10-like	11.71	0
c34767.graph_c0	chorion class A protein Ld2/Ld41-like	11.69	1.51×10^{-117}
c40012.graph_c0	keratin-associated protein 5-5-like	11.58	0
c35052.graph_c0	chorion class A protein Ld12-like	11.41	8.71×10^{-28}
c37343.graph_c0	leucine-rich repeat extensin-like protein 1	11.40	4.74×10^{-108}
c40761.graph_c0	lipase 3-like	11.15	0
c40153.graph_c0	cortex-like isoform X3	10.96	5.93×10^{-75}
c34529.graph_c0	chorion class A protein Ld19-like	10.84	1.26×10^{-168}
c43147.graph_c1	VgR	10.77	0
c34331.graph_c0	chorion class A proteins Ld24	10.44	2.82×10^{-156}
c39566.graph_c0	chemosensory protein 26	10.43	5.33×10^{-60}
c43279.graph_c0	Vg	9.89	0
c34478.graph_c0	chorion protein E1-like	8.67	3.09×10^{-209}
c30351.graph_c1	G2/mitotic-specific cyclin-B	8.50	0
c39326.graph_c0	ubiquitin-conjugating enzyme E2-24 kDa	8.15	0
c32244.graph_c0	transcription factor Sox-14	7.37	0
c36427.graph_c0	acetylcholinesterase-like	7.25	0
c38734.graph_c0	cell division cycle protein 20-like protein	6.03	0
c41524.graph_c0	tudor domain-containing protein 7A	5.02	0
c29245.graph_c0	protein lethal (2) essential for life	4.90	0
精巢发育相关 Genes associated with testis development			
c31141.graph_c0	elastin-like	- 12.32	0
c34667.graph_c0	spidroin-1-like isoform X3	- 12.13	0
c35626.graph_c0	seminal fluid protein	- 11.92	1.14×10^{-119}
c36498.graph_c0	testis-specific serine/threonine-protein kinase 2	- 11.92	3.99×10^{-104}
c33283.graph_c0	seminal metalloprotease 1-like	- 11.69	2.31×10^{-143}
c35887.graph_c0	trypsin 3A1-like	- 11.66	3.66×10^{-276}
c33878.graph_c0	small proline-rich protein 2H-like	- 11.64	1.98×10^{-192}
c36588.graph_c0	trypsin beta-like	- 11.61	0
c36609.graph_c0	trypsin-7-like	- 11.34	0
c34548.graph_c0	testis-specific serine/threonine-protein kinase 3	- 11.05	8.52×10^{-26}
c44218.graph_c0	probable chitinase 2	- 11.00	0
c38005.graph_c0	testis-specific serine/threonine-protein kinase 6-like	- 2.78	5.39×10^{-179}

FC：倍数变化；$\log_2$FC：卵巢和精巢的 RPKM 值之比取 $\log_2$。

FC: Fold change; $\log_2$FC: The ratio of the expression fold change between ovary and testis with a logarithm of base 2.

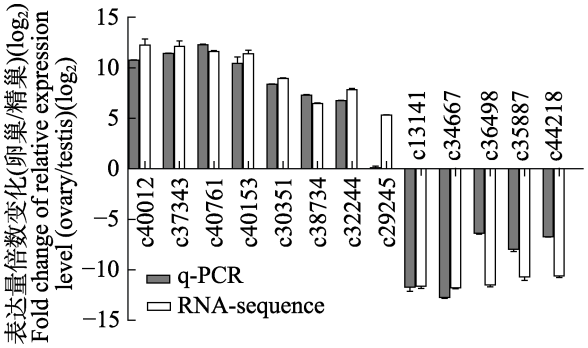

图 4　RT-qPCR 分析验证 RNA-Seq 的结果

Fig. 4　Validation of the RNA-seq data by RT-qPCR detection

得参与棉铃虫生殖的功能基因。本研究共生成 100 603 条转录本,其中 52 071 条获得功能注释,但仍有 48.24%（48 532 条）未获得功能注释,出现这种情况的原因可能是缺乏完整的棉铃虫基因组信息,或因为棉铃虫相近物种的基因注释信息及收录较少，或由于这些 unigene 包含 3′或 5′非翻译区及非编码 RNA，或属于新基因，或包含未知的蛋白质结构域。此外通过 KEGG 分析表明,11 433 条 unigenes 分属于 264 条代谢通路,其中包含与性腺发育和成熟相关的信号通路（Meng *et al.*，2015），例如：内分泌系统的胰岛

表 5　棉铃虫雌雄性腺特异表达基因的鉴定

Table 5　Identification of the specific expressed genes in ovary or testis of *Helicoverpa armigera*

基因序列号 unigene	精巢 FPKM 值 Testis FPKM	卵巢 FPKM 值 Ovary FPKM	$\log_2$FC	伪发现率 FDR	注释信息 Annotation
c35711.graph_c0	2 481.60±31.72	0.57±0.07	- 12.70	0	uncharacterized
c32808.graph_c0	3 002.33±37.23	0.83±0.03	- 12.46	0	myrosinase 1-like
c20522.graph_c0	1 496.48±13.71	0.43±0.13	- 12.42	0	uncharacterized
c40246.graph_c1	5 198.65±219.12	1.52±0.14	- 12.41	0	GATA zinc finger domain-containing protein 14-like
c42593.graph_c2	11 017.89±429.86	3.21±0.17	- 12.40	0	uncharacterized
c33306.graph_c0	36 520.73±124.13	10.68±0.19	- 12.39	0	uncharacterized
c37810.graph_c1	7 660.74±102.13	2.16±0.05	- 12.37	0	glycine-rich cell wall structural protein-like
c35315.graph_c0	7 536.49±175.21	2.23±0.25	- 12.37	0	uncharacterized
c38349.graph_c1	11 488.87±313.87	3.43±0.06	- 12.36	0	vegetative cell wall protein gp1-like
c26767.graph_c0	11 338.29±93.75	3.46±0.05	- 12.32	0	elastin-like
c34698.graph_c0	1.59±0.26	7 125.12±374.23	11.40	1.61×10^{-157}	chorion class A protein Ld2/Ld41-like
c36973.graph_c0	0.32±0.03	1 356.45±37.05	11.36	4.51×10^{-275}	glycine-rich RNA-binding protein 1
c39948.graph_c1	6.26±0.08	24 784.50±322.85	11.30	0	uncharacterized
c40733.graph_c0	0.29±0.07	1 078.93±15.03	11.29	0	keratin-associated protein 5-5-like
c36991.graph_c0	2.00±0.46	7 612.69±153.20	11.24	2.99×10^{-287}	chorion class B protein PC10-like
c43398.graph_c0	4.38±0.23	18 117.90±295.60	11.23	0	chorion class A protein Ld12-like
c42029.graph_c0	0.35±0.08	1 200.47±28.15	11.04	5.22×10^{-114}	uncharacterized
c40907.graph_c1	1.35±0.20	4 492.06±10.84	11.00	0	uncharacterized
c34228.graph_c0	0.56±0.09	1 866.79±32.69	10.97	3.42×10^{-174}	uncharacterized
c39734.graph_c0	4.76±0.33	14 232.85±264.64	10.88	0	lipase 3-like

FDR：伪发现率；FC：倍数变化；$\log_2$FC：卵巢和精巢的 RPKM 值之比取 $\log_2$。

FDR: False discovery rates; FC: Fold change; $\log_2$FC: The ratio of the expression fold change between ovary and testis with a logarithm of base 2.

素信号通路、雌激素信号通路、黄体酮介导的卵母细胞成熟途径、卵巢类固醇形成途径、FOXO 信号通路、PI3K-Akt 信号途径和 MAPK 信号通路，与能量代谢相关的氧化磷酸化途径、与信号转导（Signal transduction）相关的 M-TOR 信号途径等。

为了筛选与棉铃虫性腺发育相关的基因，我们利用 FPKM 值校正 unigene 的表达量，通过比较分析棉铃虫卵巢和精巢的转录组数据，共获得了 7 714 个 DEGs。将差异表达基因富集到 KEGG 数据库，发现一些与棉铃虫卵巢发育相关的代谢通路，如 mTOR 信号通路和 Wnt 信号通路和 MAPK 信号通路。卵子发生是昆虫卵巢发育的关键过程，是昆虫种群繁衍的基础，这一过程包括卵黄发生前期、卵黄发生期和卵黄成熟期（卵壳发生期）。卵黄原蛋白（Vitellogenin，Vg）合成及其受体（Vitellogenin receptor，VgR）介导的 Vg 摄取是昆虫卵黄发生的基础，也是卵子发生的关键时期，显著影响昆虫种群数量，是昆虫生殖的关键环节（Muhammad *et al.*，2014），因此这两个基因可以作为害虫防控的重要分子靶标（Roy-Zokan *et al.*，2015；Upadhyay *et al.*，2016；赵静等，2019）。本文研究发现 Vg 和 VgR 在卵巢组织中显著上调，这表明了它们对卵巢发育有促进作用，这与我们前期的 VgR 在棉铃虫卵巢中功能研究结果相吻合（Zhang *et al.*，2016）。此外，我们还发现卵壳蛋白基因在棉铃虫卵巢中的表达量显著高于精巢，说明了卵壳蛋白的合成对卵巢发育及卵子发生至关重要。同样在家蚕 *Bombyx mori* 中，Chen 等（2015）结合家蚕基因组和转录组数据鉴定出了 126 个 Chorion 相关基因。Peng 等（2017）发现 Chorion 卵壳蛋白相关基因在小菜蛾卵巢转录组文库中大量表达，说明了 Chorion 卵壳蛋白在昆虫卵子发生中的功能较为保守，此外有研究表明干扰掉褐飞虱 *Nilaparvata lugens* 的 NlChP 基因能显著降低其产卵量和胚胎的存活率（Lou *et al.*，2018），但鳞翅目昆虫中卵壳蛋白的功能研究相对较少，卵壳蛋白对生殖的调控作用还有待发掘。在昆虫卵巢发育过程中，卵原细胞通过有丝分裂进行增殖，然后历经减数分裂最终产生卵母细胞。在这一过程中，调控细胞分裂周期的各种因子显示与卵巢发育和成熟密切相关（Wu *et al.*，2016）。本文中筛选到参与细胞分裂周期（CDC）蛋白 cdc 20 基因在卵巢内高表达（表 4），且 cdc B3 和 cdc 22 基因以及细胞周期蛋白 A（cyclin A）和细胞周期蛋白 H（cyclin H）等细胞周期调控因子基因在棉铃虫卵巢的表达量是精巢表达量的 1.72-2.64 倍，推测细胞分裂周期相关因子参与棉铃虫卵巢发育，这与其他昆虫中的报道相一致（Jacobs *et al.*，1998；Peng *et al.*，2017）。此外，Peng 等（2017）发现小分子的热激蛋白基因 （Small heat shock protein，sHSP）在小菜蛾卵巢内高表达，推测其可能参与卵子发生，然而我们在棉铃虫卵巢高表达的基因中并未发现 *sHSP*。

已有研究表明睾丸特异性蛋白激酶（Testis-specific serine/threonine-protein kinase，TSSK）、精子发生相关蛋白（Spermatogenesis-associated protein）及 seminal fluid protein 是参与昆虫精子发生的重要功能基因（Wei *et al.*，2015；Sun *et al.*，2019），我们研究发现这些基因在棉铃虫雄蛾精巢都高表达，说明它们可能参与棉铃虫精巢分化和精子发生过程。本文还发现 Trypsin-like 蛋白酶基因在精巢内高表达，在埃及伊蚊 *Aedes aegypti* 中证实 trypsin-like 蛋白酶在体外能够激活精子活力（Sirot *et al.*，2008），这说明 trypsin-like 蛋白酶对激发精子活性具有重要作用。另外， 通过 KEGG 功能富集证实参与泛素介导的蛋白水解相关通路基因在棉铃虫卵巢与精巢都大量表达，已有研究表明这些信号通路涉及性腺发育和配子发生等生殖调控（Meng *et al.*，2015；董忠典等，2021）。

本文对棉铃虫成虫精巢和卵巢组织进行转录组测序，获得了大量参与棉铃虫性腺发育和配子发生的功能基因。但由于缺乏棉铃虫及其近源种的基因信息，仍有大量性腺特异表达的基因未获得功能注释，其中可能存在性腺发育和生殖调控的相关基因，这些未知基因的功能还有待深入研究。

参考文献 (References)

Chen Z, Nohata J, Guo H, Li S, Liu J, Guo Y, Yamamoto K, Kadono-Okuda K, Liu C, Arunkumar KP, Nagaraju J, Zhang Y, Liu S, Labropoulou V, Swevers L, Tsitoura P, Iatrou K, Gopinathan KP, Goldsmith MR, Xia Q, Mita K, 2015. A comprehensive analysis of the chorion locus in silkmoth. *Scientific Reports*, 5: 16424.

Dong ZD, Li XY, Huang CQ, Zhang HR, Huang SK, Zhang N, Guo YS, Wang ZD, 2021. Comparative transcriptome analysis of the gonad of the lacustrine goby (*Gobiopterus lacustris*). *Journal of Fisheries of China*, 45(3): 365–380. [董忠典, 黎学友, 黄承勤, 张海瑞, 黄顺楷, 张宁, 郭昱嵩, 王中铎, 2021. 湖栖鳍虾虎鱼性腺转录组比较分析. 水产学报, 45(3): 365–380.]

Guo YY, 1998. Researches on the Cotton Bollworm, *Helicoverpa armigera*. Beijing: China Agricultural Press. 134–138. [郭予元, 1998. 棉铃虫的研究. 北京: 中国农业出版社. 134–138.]

Jacobs HW, Knoblich JA, Lehner CF, 1998. Drosophila cyclin B3 is required for female fertility and is dispensable for mitosis like cyclin B. *Gene Development*, 12(23): 3741–3751.

Jin MH, Liao CY, Chakrabarty S, Zheng WG, Wu KM, Xiao YT, 2019. Transcriptional response of ATP-binding cassette (ABC) transporters to insecticides in the cotton bollworm, *Helicoverpa armigera*. *Pesticide Biochemistry and Physiology*, 154: 46–59.

Legeai F, Derrien T, 2015. Identification of long non-coding RNAs in insects genomes. *Current Opinion in Insect Science*, 7: 37–44.

Liang GM, Tan WJ, Guo YY, 1999. An improvement in the technique of artificial rearing cotton bollworm. *Plant Protection*, 25(2):15–17. [梁革梅, 谭维嘉, 郭予元, 1999. 人工饲养棉铃虫技术的改进. 植物保护, 25(2): 15–17.]

Liu YZ, Zhang DD, Qu AJ, Zheng FQ, Xiao YT, 2019. Transcriptome analysis of differentially expressed genes involved in Bt resistance in cotton bollworm (*Helicoverpa armigera*). *Shandong Agricultural Sciences*, 51(7): 5–9. [刘运泽, 张丹丹, 曲爱军, 郑方强, 萧玉涛, 2019. 棉铃虫Bt抗感品系的转录组差异表达分析. 山东农业科学, 51(7): 5–9.]

Lou YH, Pan PL, Ye YX, Cheng C, Xu HJ, Zhang CX, 2018. Identification and functional analysis of a novel chorion protein essential for egg maturation in the brown planthopper. *Insect Molecular Biology*, 27(3): 393–403.

Meng XL, Liu P, Jia FL, Li J, Gao BQ, 2015. De novo transcriptome analysis of *Portunus trituberculatus* ovary and testis by RNA-Seq: Identification of genes involved in gonadal development. *PLoS ONE*, 10(6): e0128659

Nie HY, Zhong XW, Yi QY, Zhang LP, Zou Y, Zhao P, 2011. Proteome analysis of small heat shock proteins in *Bombyx mori* testis at different developmental stages. *Scientia Agricultura Sinica*, 44(9): 1923–1930. [聂红毅, 钟晓武, 衣启营, 张利平, 邹勇, 赵萍, 2011. 家蚕精巢中小热激蛋白在不同发育时期的蛋白质组学分析. 中国农业科学, 44(9): 1923–1930.]

Oppenheim SJ, Baker RH, Simon S, DeSalle R, 2015. We can't all be supermodels: The value of comparative transcriptomics to the study of non-model insects. *Insect Molecular Biology*, 24(2): 139–154.

Peng L, Wang L, Yang YF, Zou MM, He MY, Wang Y, Wang Q, Vasseur L, You MS, 2017. Transcriptome profiling of the *Plutella xylostella* (Lepidoptera: Plutellidae) ovary reveals genes involved in oogenesis. *Gene*, 637: 90–99.

Roy-zokan EM, Cunningham CB, Hebb LE, Mckinney EC, Moore AJ, 2015. Vitellogenin and vitellogenin receptor gene expression is associated with male and female parenting in a subsocial insect. *Proceedings of the Royal Society B Biological Sciences*, 282(1809): 20150787.

Sirot LK, Poulson RL, Caitlin McKenna M, Girnary H, Wolfner MF, Harrington LC, 2008. Identity and transfer of male reproductive gland proteins of the dengue vector mosquito, *Aedes aegypti*: Potential tools for control of female feeding and reproduction. *Insect Biochemistry and Molecular Biology*, 38(2): 176–189.

Sun RR, Sun ZP, Chen YY, Zhu FY, Li Y, Zhong GH, Yi X, 2019. Comparative proteomic analysis of sex-biased proteins in ovary and testis at different stages of *Spodoptera litura*. *Journal of Proteomics*, 206: 103439.

Upadhyay SK, Singh H, Dixit S, Mendu V, Verma PC, 2016. Molecular characterization of vitellogenin and vitellogenin receptor of *Bemisia tabaci*. *PLoS ONE*, 11(5): e0155306.

Wei D, Li HM, Yang WJ, Wei DD, Dou W, Huang Y, Wang JJ, 2015. Transcriptome profiling of the testis reveals genes involved in spermatogenesis and marker discovery in the oriental fruit fly, *Bactrocera dorsalis*. *Insect Molecular Biology*, 24(1): 41–57.

Wei D, Li R, Zhang MY, Liu YW, Zhang Z, Smagghe G, Wang JJ, 2017. Comparative proteomic profiling reveals molecular characteristics associated with oogenesis and oocyte maturation during ovarian development of *Bactrocera dorsalis* (Hendel). *International Journal of Molecular Sciences*, 18(7): 1379.

Wu ZX, Guo W, Xie YT, Zhou ST, 2016. Juvenile hormone activates the transcription of cell-division-cycle 6 (Cdc6) for polyploidy-dependent insect vitellogenesis and oogenesis. *Journal of Biological Chemistry*, 291(10): 5418–5427.

Xiang M, Ye XF, Hu HX, Wang WL, Yu F, Xiao HW, Ji R, Wang H, 2017. Analysis of the transcriptome and gonadal development related genes of *Calliptamus italicus* (Orthopera: Acrididae). *Acta Entomologica Sinica*, 60(11): 1235–1246. [向敏, 叶小芳, 扈鸿霞, 王伟亮, 于非, 肖宏伟, 季荣, 王晗, 2017. 意大利蝗转录组及性腺发育相关基因分析. 昆虫学报, 60(11): 1235–1246.]

Yang J, 2013. Comparative transcriptome of *Atractomorpha sinensis* at three developmental stages and mitochondrial transcript mapping. Master dissertation. Xi'an: Shaanxi Normal University. [杨婧, 2013. 短额负蝗三种虫态的比较转录组及线粒体转录组作图研究. 硕士学位论文. 西安: 陕西师范大学.]

Zhang J, Wang B, Dong S, Cao D, Dong J, Walker WB, Liu Y, Wang G, 2015. Antennal transcriptome analysis and comparison of chemosensory gene families in two closely related noctuidae moths, *Helicoverpa armigera* and *H. assulta*. *PLoS ONE*, 10(2): e0117054.

Zhang WN, Ma L, Liu XY, Peng Y, Liang GM, Xiao HJ, 2021. Dissecting the roles of FTZ-F1 in larval molting and pupation, and the sublethal effects of methoxyfenozide on *Helicoverpa armigera*. *Pest Management Science*, 77(3): 1328–1338.

Zhang WN, Ma Long, Xiao HJ, Xie BT, Smagghe G, Guo YY, Liang GM, 2016. Molecular characterization and function analysis of the vitellogenin receptor from the cotton bollworm, *Helicoverpa armigera* (Hübner) (Lepidoptera, Noctuidae). *PLoS ONE*, 11(5): e0155785.

Zhao J, Tao R, Hao DJ, Xiao LB, Tan YA, 2019. RNA interference of vitellogenin receptor gene in beet armyworm (*Spodoptera exigua*). *Scientia Agricultura Sinica*, 52(1): 56–64. [赵静, 陶蓉, 郝德君, 肖留斌, 谭永安, 2019. 甜菜夜蛾卵黄原蛋白受体基因的 RNA 干扰. 中国农业科学, 52(1): 56–64.]

Zhao ZY, Wu GQ, Wang J, Qiu CL, Qiu LH, 2013. Next-generation sequencing-based transcriptome analysis of *Helicoverpa armigera* larvae immune-primed with *Photorhabdus luminescens*. *PLoS ONE*, 8(11): e80146.

Zhou C, Yang H, Wang Z, Long GY, Jin DC, 2018. Comparative transcriptome analysis of *Sogatella furcifera* (Horváth) exposed to different insecticides. *Scientific Reports*, 8(1): 8773–8785.