

云斑天牛成虫触角转录组及嗅觉相关基因分析*

胡佳萌^{1**} 徐丹萍² 卓志航^{1***} 杨伟³ 杨桦³ 郑奕然¹

(1. 海南大学林学院, 海口 570228; 2. 四川农业大学食品学院, 雅安 625014; 3. 四川农业大学林学院, 成都 611130)

摘要 【目的】建立云斑天牛 *Batocera horsfieldi* (Hope)成虫触角转录组数据库, 深入挖掘云斑天牛的基因数据信息。【方法】采用高通量测序平台 (Illumina HiSeq) 对云斑天牛成虫触角进行转录组测序、序列组装及生物信息学分析。【结果】云斑天牛成虫触角转录组共获得 137 485 条 Transcript 序列和 69 214 条 Unigene 序列; 其中, Transcript 序列平均长度 1 142 bp, Unigene 序列平均长度 1 983 bp。将 Unigene 分别比对到 NR、NT、SwissProt、KO、PFAM、GO、KOG 数据库进行基因功能注释, NR 注释 41 636 条, NT 注释 14 895 条, KO 注释 19 287 条, SwissProt 注释 33 442 条, PFAM 注释 34 687 条, GO 注释 35 321 条, KOG 注释 20 582 条。NR 注释表明, 72.0% 的云斑天牛 Unigene 与赤拟谷盗 *Tribolium castaneum* 和中欧山松大小蠹 *Dendroctonus ponderosae* 具有相似性。基因功能注释分类表明, 云斑天牛成虫触角转录组在 GO 数据库三大类中包含 5 个最主要功能, 分别是细胞过程、代谢过程、单有机体过程、结合和催化活性, 分别占 20 912、19 086、17 202、21 477 和 15 823 条 Unigenes; 云斑天牛转录组在 KOG 数据库 26 个功能目录共注释 20 585 条 Unigenes, 其中, 翻译后修饰、蛋白质转换和伴侣共有 1 977 条, 一般功能预测 3 285 条 (最多), 信号传导机制 3 053 条, 合计 8 315 条占全部 Unigenes 40.39%; 总共 19 287 条 Unigenes 分至 5 个 KEGG 功能类别, 其中细胞过程 6 793 条, 环境信息处理 6 255 条, 遗传信息处理 3 038 条, 代谢 3 852 条, 有机系统 3 508 条。进一步基因功能注释分析筛选得到 161 个嗅觉相关基因, 包含 96 个气味结合蛋白 (Odorant binding protein, OBP), 34 个化学感受蛋白 (Chemosensory protein, CSP) 和 31 个气味受体 (Odorant receptor, OR)。【结论】本研究获得了云斑天牛成虫触角转录组数据库, 为进一步研究云斑天牛的基因功能分析及嗅觉感受机制奠定了分子基础。

关键词 云斑天牛; 触角转录组; 嗅觉相关基因; 基因组注释; 高通量测序

Analysis of antennal transcriptome and olfaction-related genes of adult *Batocera horsfieldi* (Hope)

HU Jia-Meng^{1**} XU Dan-Ping² ZHUO Zhi-Hang^{1***} YANG Wei³ YANG Hua³ ZHENG Yi-Ran

(1. College of Forestry, Hainan University, Haikou 570228, China; 2. College of Food Science, Sichuan Agricultural University, Yaan 625014, China; 3. College of Forestry, Sichuan Agricultural University, Chengdu 611130, China)

Abstract [Objectives] To obtain genetic data on, and establish an antenna transcriptome database for, the adult white-striped longhorn beetle, *Batocera horsfieldi* (Hope). [Methods] The antenna transcriptome of adult *B. horsfieldi* was sequenced using an Illumina HiSeq 4000 platform and analyzed with bioinformatic software. [Results] The antennal transcriptome database of adult *B. horsfieldi* contained 137 485 transcripts and 69 214 Unigenes. The mean number of transcripts was 1 142 bp, and the mean number of Unigenes was 1 983 bp. Of the Unigenes found, 41 636 were annotated in the NR database, 14 895 were annotated in the NT database, 19 287 were annotated in the KO database, 33 442 were annotated in the SwissProt database, 34 687 were annotated in the PFAM database, 35 321 were annotated in the GO database, and 20 582 were annotated in the KOG database. The results of NR comment showed that 72.0% of *B. horsfieldi* Unigenes were

*资助项目 Supported projects: 国家自然科学基金项目 (31270694); 四川农业大学“211”工程双支计划资助项目 (00370101); 海南大学科研启动项目 [KYQD(ZR)1823]; 海南大学大学生创新创业训练计划项目 (201810589251)

**第一作者 First author, E-mail: jiameng.hu@foxmail.com

***通讯作者 Corresponding author, E-mail: zhuozhihang@foxmail.com

收稿日期 Received: 2018-08-04; 接受日期 Accepted: 2018-12-11

similar to those of *Tribolium castaneum* and *Dendroctonus ponderosae*. Transcriptome gene function classification indicated five main functional classifications annotated in the GO database; these were cellular process, metabolic process, single-organism process, binding and catalytic activity, each classification containing 20 912, 19 086, 17 202, 21 477 and 15 823 unigenes, respectively. 20 585 unigenes were annotated in 26 groups of the KOG database, 1 977 of which were involved in post-translational modification, protein turnover and chaperoning. Only 3 285 unigenes were involved in general function prediction, and 3 053 in signal transduction mechanisms, together comprising 40.39% of the total. 19 287 unigenes were annotated in 5 functional classifications in the KEGG database; cellular process (6 793), environmental information processing (6 255), genetic information processing (3 038), metabolism (3 852) and organismal system (3 508). By further screening and identification, 161 olfaction-related genes, including 96 odorant binding protein (OBP) genes, 34 chemosensory protein (CSP) genes and 31 odorant receptor (OR) genes, were identified. **[Conclusion]** The transcriptome database of the antenna of adult *B. horsfieldi* was successfully obtained. The results provide a foundation for the further study of gene function and the molecular mechanisms responsible for olfactory reception in *B. horsfieldi*.

Key words *Batocera horsfieldi* (Hope); antennal transcriptome; olfaction-related genes; genome annotation; high-throughput sequencing

云斑天牛 *Batocera horsfieldi* (Hope) 又称云斑白条天牛 (李建庆等, 2009a), 隶属鞘翅目 (Coleoptera) 天牛科 (Cerambycidae), 该虫由卵孵化为幼虫后, 首先蛀食寄主韧皮部, 之后进入木质部并蛀食木质部 (王绍林等, 2004), 造成寄主植株树势衰弱, 或全株枯死, 易风折, 致使寄主植株材质损失, 果实品质下降 (梅爱华等, 1998), 是用材林、经济林和防护林的重要蛀干害虫 (陈京元和罗治建, 2001; 骆有庆, 2005)。云斑天牛寄主种类多样, 常见的有杨树 *Populus tomentosa*、核桃 *Juglans regia* 和桉树 *Eucalyptus robusta* 等 (刁志娥和丁福波, 2004), 补充营养寄主种类繁多, 常见的有小果蔷薇 *Rosa cymosa*、野蔷薇 *Rosa multiflora* Thunb., 核桃、白蜡树 *Fraxinus chinensis* Roxb., 桉树、桉木 *Alnus cremastogyne* Burkill、青冈 *Cylobalanopsis glauca*、梨树 *Pyrus spp.*、枣树 *Zizyphus jujuba* 等 (肖刚柔, 1992; 刁志娥和丁福波, 2004; 胡瑜珩, 2015), 该虫在国外主要分布于日本、越南、印度、缅甸等亚洲国家 (杨桦等, 2010), 国内广泛分布于中南、西南、华南和华北等地, 对生态、林业经济等都能造成一定破坏和损失 (高瑞桐等, 1995; 严敖金等, 1997)。此外, 由于云斑天牛生活隐蔽、分布广、寄主多、危害重、抗逆强, 且寄主及补充营养寄主种类多样, 并以幼虫在寄主树干内钻蛀危害, 幼虫蛀道向上

延伸较深, 造成树干内的蛀道常纵横交错, 此外该虫以虫粪木屑堵塞孔口, 导致该虫防控不佳 (李娟等, 2008; 李建庆等, 2009b)。

昆虫通过转录数据分析, 可以获取非编码及未知编码 RNA 种类、低表达量基因及 sRNA 分析等 (Etebari *et al.*, 2011; Xu *et al.*, 2012), 此外对不同种类、不同生境条件下及不同生长发育昆虫转录数据分析可以捕获转录发生位点、转录本表达丰度、功能基因信息、可变剪接和转录本单核苷酸多态性 (Single nucleotide polymorphisms, SNPs) 和差异表达基因等信息 (Anholt *et al.*, 2003; Xu *et al.*, 2012; Zhang *et al.*, 2012; 张棋麟和袁明龙, 2013), 如中华稻蝗 *Oxya chinensis* 生长发育相关基因挖掘 (吕红娟, 2012)、矮竹节虫 *Timema cristinae* 分子标记开发及遗传分化推论 (Comeault *et al.*, 2012)、白蜡蚧 *Ericerus pela* 蜡合成相关基因表达 (Yang *et al.*, 2012)、松褐天牛 *Monochamus alternatus* 气味结合蛋白基因挖掘 (周长祥, 2016)、灰飞虱 *Laodelphax striatellus* 与病毒互作相关基因表达 (Zhang *et al.*, 2010) 等都利用了转录文库分析。转录组文库构建已在天牛类蛀干害虫中广泛应用, 如黄斑星天牛及松褐天牛等都应用转录组文库对特定基因进行挖掘捕获及相关功能分析。昆虫的嗅觉感受器主要分布于触角 (Steinbrecht and Stankiewicz, 1999), 其嗅觉相关蛋白含量及

嗅觉相关基因表达均在触角中较高(李慧, 2012), 因此本研究选择云斑天牛成虫触角进行转录组及嗅觉相关基因分析。综上, 云斑天牛作为重要林业类蛀干害虫, 本研究拟采用高通量测序技术对其成虫触角转录组进行测序, 并通过基因功能注释及分析挖掘该虫的大量嗅觉相关基因, 为其相关基因互作、转录表达、功能注释及分子标记和开发提供理论依据, 为进一步从分子水平开展云斑天牛行为调控与防治奠定基础, 同时也丰富鞘翅目昆虫基因数据库。

1 材料与方法

1.1 供试昆虫

供试云斑天牛成虫采集: 于5月中上旬, 采集刚出木、未取食、未交配且健康良好的云斑天牛成虫(寄主杨树); 采集地点: 四川省德阳市罗江县调元镇; 成虫采集后, 放入60 cm×60 cm×60 cm 不锈钢养虫笼, 当天带回取样; 选取上述活动能力较强的云斑天牛成虫(雌雄成虫各6头), 切下触角, 迅速用液氮冷冻并保存于-80℃低温冰箱中备用。

1.2 Total RNA 样品提取与检测

提取上述供试云斑天牛雌雄成虫触角 Total RNA(雌:雄=1:1, 混合样); 琼脂糖凝胶电泳分析 RNA 降解程度以及是否有污染; Nanodrop 检测 RNA 的初度(OD_{260/280} 比值); Qubit 对 RNA 浓度进行精确定量; Agilent 2100 精确检测 RNA 的完整性。确保供试云斑天牛成虫触角 Total RNA 样品合格性, 进而确保样品能进行高质量的转录组测序。

1.3 文库构建及库检

样品检测合格后, 用带有 Oligo(dT) 的磁珠富集真核生物 mRNA。随后加入 Fragmentation buffer 将 mRNA 打断成短片段, 以 mRNA 为模板, 用六碱基随机引物(Random hexamers)合成一链 cDNA, 然后加入缓冲液、dNTPs 和 DNA polymerase I 和 RNase H 合成二链 cDNA, 再用

AMPure XP beads 纯化双链 cDNA。纯化的双链 cDNA 先进行末端修复、加 A 尾并连接测序接头, 再用 AMPure XP beads 进行片段大小选择。最后进行 PCR 扩增, 并用 AMPure XP beads 纯化 PCR 产物, 得到最终的文库。文库构建完成后, 先使用 Qubit 2.0 进行初步定量, 稀释文库至 1.5 ng·μL⁻¹, 随后使用 Agilent 2100 对文库的 Insert size 进行检测, Insert size 符合预期后, 使用 RT-qPCR 方法对文库的有效浓度进行准确定量(文库有效浓度 > 2 nmol·L⁻¹), 以保证文库质量。

1.4 上机测序及生物信息分析

库检合格后, 把不同文库按照有效浓度及目标下机数据量的需求 Pooling 后进行 Illumina HiSeq 测序。

1.5 转录组测序数据组装

对测序得到的原始数据(Raw data)进行过滤, 去除带接头(Adapter)的 Reads, 去除 N(N 表示无法确定碱基信息)的比例大于 10% 的 Reads, 去除低质量 Reads(质量值 Qphred ≤ 20 的碱基数占整个 Reads 的 50% 以上的 Reads)。鉴于云斑天牛无参考基因组, 获得 Clean reads 后, 需要对 Clean reads 进行拼接以获取后续分析的参考序列, 采用 Trinity(Grabherr *et al.*, 2011)对 Clean reads 进行拼接。将 Trinity 拼接得到的转录本序列, 作为后续分析的参考序列。以 Corset(Nadia M Davidson and Alicia Oshlack, 2014)层次聚类后得到最长 Cluster 序列进行后续的分析。

1.6 转录组数据分析

为获得全面的基因功能信息, 对云斑天牛成虫触角转录组进行了七大数据库的基因功能注释, 包括: Nr, Nt, Pfam, KOG/COG, Swiss-prot, KEGG, GO。通过与 Nr 库进行比对注释, 可以获取本物种基因序列与近缘物种基因序列的相似性以及本物种基因的功能信息; 对基因进行 GO 注释之后, 将注释成功的基因按照 GO 三个

大类的下一层级进行分类；KOG 分为 26 个 Group 将 KOG 注释成功的基因按 KOG 的 Group 进行分类；对基因做 KO 注释后，可根据它们参与的 KEGG 代谢通路进行分类。云斑天牛成虫触角转录文库主要由北京诺禾致源科技股份有限公司协助完成。

2 结果与分析

2.1 转录组序列信息

经高通量转录测序、清除和质量检测后，共得到 137 485 条 Transcript 序列和 69 214 条 Unigene 序列，其中 Q20 比例 96.74%，Q30 比例 91.8%，GC 比例 42.35%（表 1）。

经转录本拼接后，Transcript 合计 137 485 条，其中长度 200-500 bp 的 79 246 条，长度 500-1 000 bp 的 20 244 条，长度 1 000-2 000 bp 的 15 069 条，长度大于 2 000 bp 的 22 926 条；此外，Transcript 最短长度 201 bp，平均长度 1 142 bp，中位长度 399 bp，最长长度 42 567 bp，核苷酸合计 156 999 768 bp。Unigenes 合计 69 214 条，其中长度 200-500 bp 的 12 802 条，长度 500-1 000 bp 的 18 441 条，长度 1 000-2 000 bp 的 15 045

条，长度大于 2 000 bp 的 22 926 条；此外，Transcript 最短长度 201 bp，平均长度 1 983 bp，中位长度 1 166 bp，最长长度 42 567 bp，核苷酸合计 137 221 519 bp（图 1，表 2）。

2.2 转录组基因功能注释

为获取更全面的基因功能信息，将转录数据 Unigenes 分别比对到 NR、NT、SwissProt、KO、PFAM、GO、KOG 数据库进行基因功能注释，结果表明：NR 数据库注释基因共有 41 636 条，占 60.15%；NT 数据库注释基因共有 14 895 条，占 21.52%；PFAM 数据库注释基因共有 34 687 条，占 50.11%；KO 数据库注释基因共有 19 287 条，占 27.86%；SwissProt 数据库注释基因共有 33 442 条，占 48.31%；GO 数据库注释基因共有 35 321 条，占 51.03%；KOG 数据库注释基因共有 20 585 条，占 29.74%；其中，在以上 7 个数据库中都能注释成功的基因共有 6 359 条，占 9.18%；在以上 7 个数据库中至少 1 个数据库成功注释的基因共有 46 047 条，占 66.52%。此外，通过其中 5 种主要数据库（NT、NR、GO、KOG 和 PFAM）维恩图可以更清晰了解各数据库交叉注释情况（表 3，图 2，图 3）。

表 1 云斑天牛成虫触角转录本数据

Table 1 Overview of antennal transcriptome of adult *Batocera horsfieldi*

项目 Item	原始序列 Raw reads	干净序列 Clean reads	Q20 比例 Percentage of Q20	Q30 比例 Percentage of Q30	GC 比例 Percentage of GC
云斑天牛成虫 Adult <i>B. horsfieldi</i>	51 993 763	49 935 242	96.74	91.8	42.35

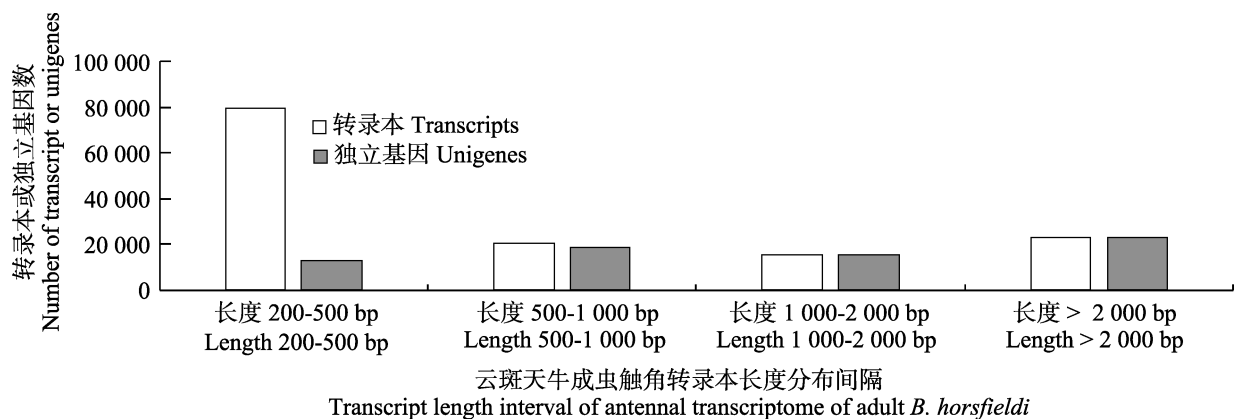


图 1 云斑天牛成虫触角转录本长度分布

Fig. 1 Length frequency distribution of antennal transcriptome of adult *Batocera horsfieldi*

表 2 云斑天牛成虫触角转录本长度分布
Table 2 Length distribution of antennal transcriptome of adult *Batocera horsfieldi*

项目 Item	云斑天牛成虫 Adult <i>B. horsfieldi</i>	
	转录本 Transcripts	独立基因 Unigenes
长度 200-500 bp Length 200-500 bp	79 246	12 802
长度 500-1 000 bp Length 500-1000 bp	20 244	18 441
长度 1 000-2 000 bp Length 1 000-2 000 bp	15 069	15 045
长度>2 000 bp Length>2 000 bp	22 926	22 926
合计 Total number	137 485	69 214
最短长度 Minimum length	201	201
平均长度 Mean length	1 142	1 983
中位长度 Median length	399	1 166
最长长度 Maximum length	42 567	42 567
N50	2 928	3 450
N90	350	828
核苷酸合计 Total nucleides	156 999 768	137 221 519

表 3 云斑天牛成虫触角转录组基因功能注释情况
Table 3 Annotation of antennal transcriptome of adult *Batocera horsfieldi*

项目 Item	独立基因数目 Number of unigenes	百分比 Percentage (%)
NR 数据库注释 Annotated in NR database	41 636	60.15
NT 数据库注释 Annotated in NT database	14 895	21.52
KO 数据库注释 Annotated in KO database	19 287	27.86
SwissProt 数据库注释 Annotated in SwissProt database	33 442	48.31
PFAM 数据库注释 Annotated in PFAM database	34 687	50.11
GO 数据库注释 Annotated in GO database	35 321	51.03
KOG 数据库注释 Annotated in KOG database	20 582	29.74
所有数据库均有注释 Annotated in all database	6 359	9.18
至少一个数据库注释 Annotated in at least one database	46 047	66.52

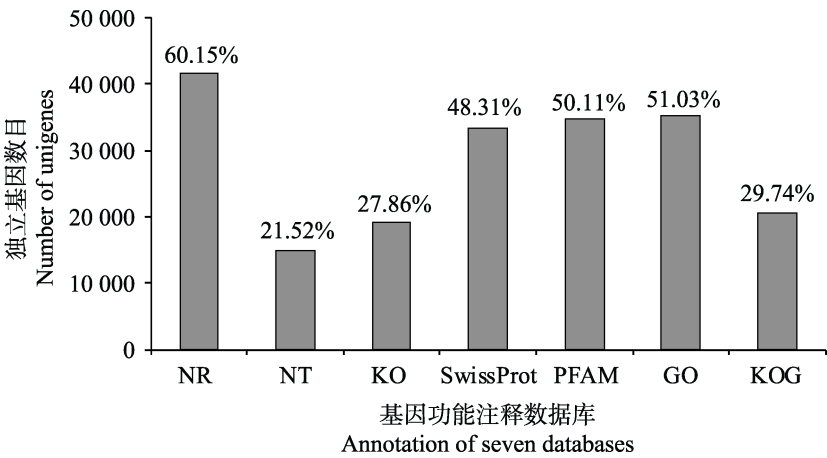


图 2 云斑天牛成虫触角转录组基因功能注释情况
Fig. 2 Annotation of antennal transcriptome of adult *Batocera horsfieldi*

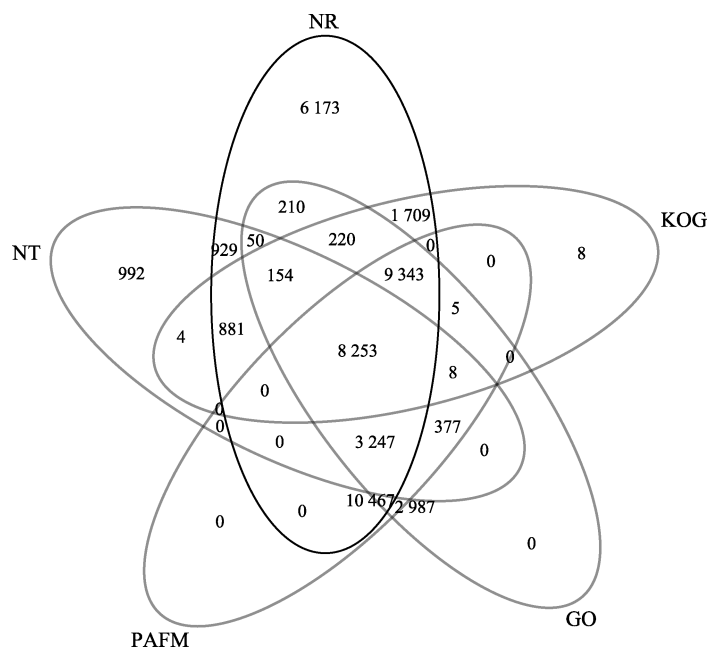


图3 云斑天牛成虫触角转录组5种数据库维恩图

Fig. 3 Venn figure for five databases of antennal transcriptome of adult *Batocera horsfieldi*

通过 BLASTX 运算法则 ($E\text{-values} \leq 10^{-5}$) 与 NR 数据库进行比对注释, 结果显示, 云斑天牛与赤拟谷盗 *Tribolium castaneum* 相似性最高, 57.4%, 其次为中欧山松大小蠹 *Dendroctonus ponderosae*, 14.6%; 云斑天牛与豌豆长管蚜 *Acyrtosiphon pisum*, 马铃薯甲虫 *Leptinotarsa decemlineata*, 内华达古白蚁 *Zootermopsis nevadensis* 相似度较低, 分别为 0.9%, 1.2% 和 1.5%, 其它物种占 24.4% (图 4)。

2.3 转录组基因功能分类

GO 分类共有 55 个功能类别, 分为三大类 Ontology: 生物过程 (Biological process)、分子功能 (Molecular function) 和细胞组分 (Cellular

component), 分别用来描述基因编码的产物所参与的生物过程、所具有的分子功能及所处的细胞环境。云斑天牛成虫触角转录组在 GO 数据库三大类 (Ontology) 包含 5 个最主要功能, 分别是: 细胞过程 (Cellular process)、代谢过程 (Metabolic process)、单有机体过程 (Single-organism process)、结合 (Binding) 和催化活性 (Catalytic activity), 云斑天牛成虫触角转录组 5 个最主要功能分别占有 20 912、19 086、17 202、21 477 和 15 823 条 Unigenes (图 5)。

KOG 是 NCBI 基于基因直系同源关系, 针对真核生物建立的数据库。云斑天牛成虫触角转录组在 KOG 数据库 26 个功能目录共注释 20 585 条 Unigenes, 其中翻译后修饰 (Posttranslational

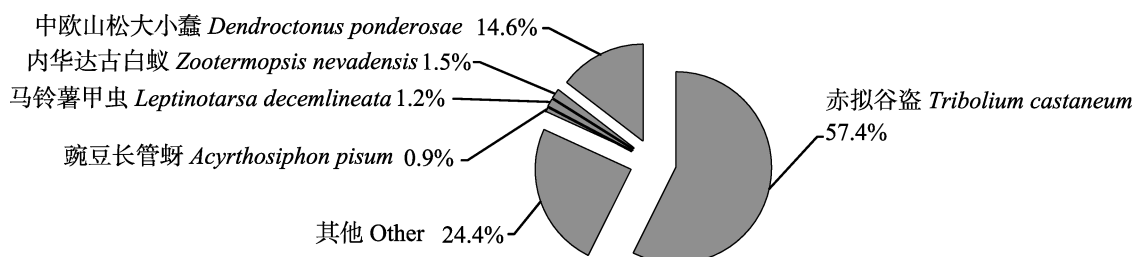
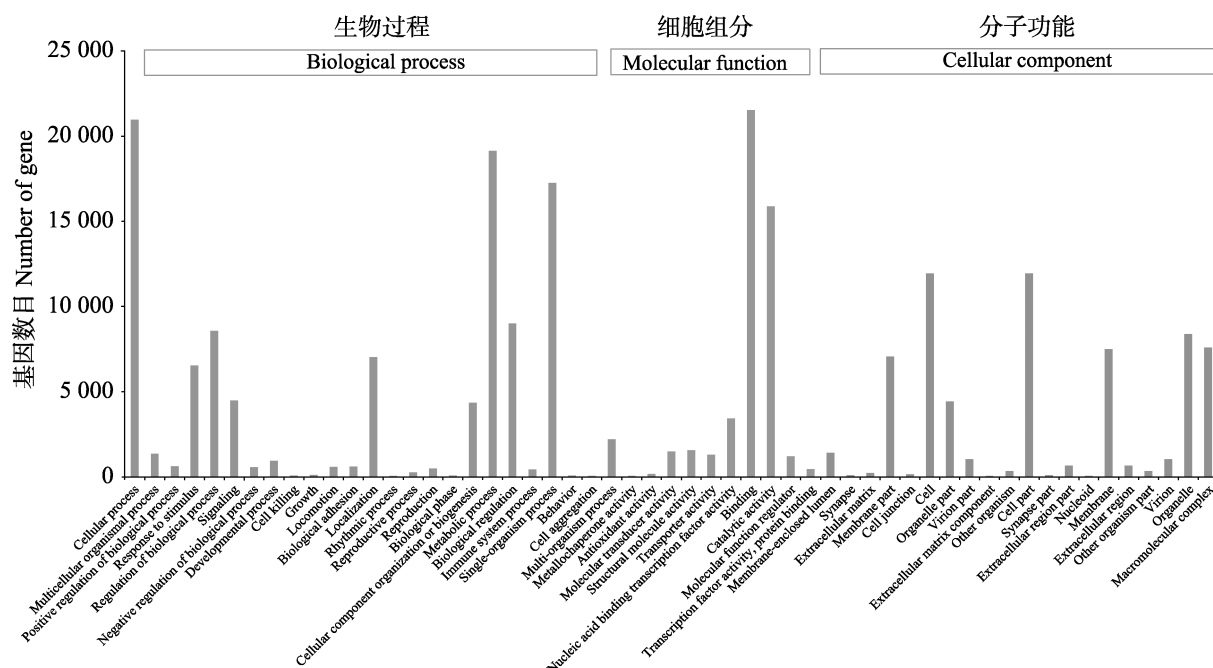


图4 基于 BLAST 的云斑天牛转录组同源性物种分布图

Fig. 4 Proportional homology distribution among other species based on the BLAST hits against the NR database of antennal transcriptome of adult *Batocera horsfieldi*



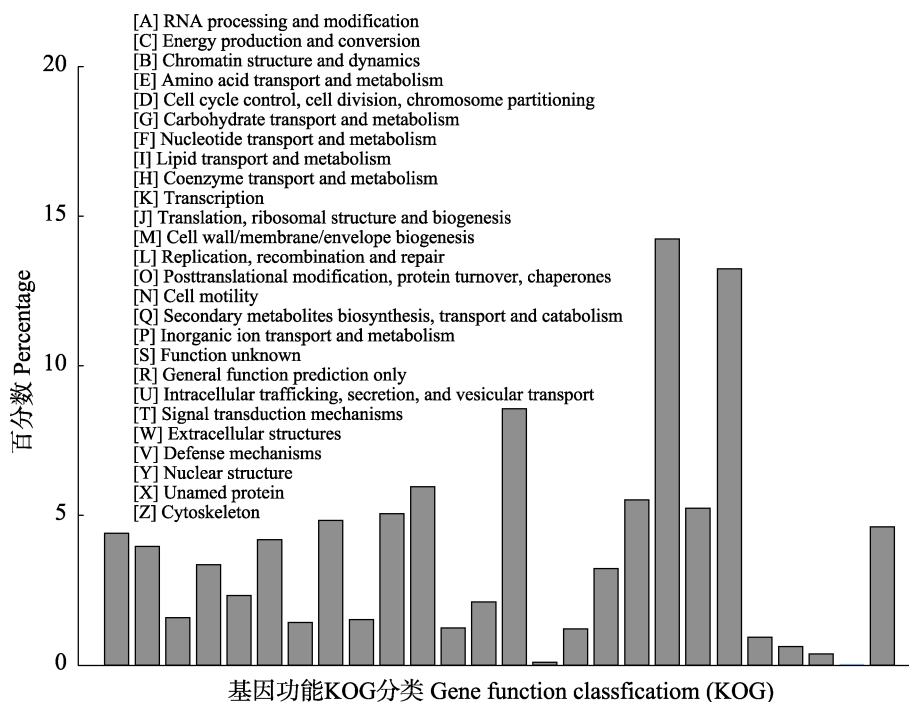
基因功能GO分类 Gene function classification (GO)

图 5 云斑天牛成虫触角转录组 GO 分类图

Fig. 5 GO classification of antennal transcriptome of adult *Batocera horsfieldi*

modification), 蛋白质转换 (Protein turnover), 伴侣 (Chaperones) 共有 1 977 条, 一般功能预测 (General function prediction only) 有 3 285 条 (最多), 信号传导机制 (Signal

transduction mechanisms) 有 3 053 条, 合计 8 315 条占全部 Unigenes 的 40.39%, 而未命名蛋白 (Unnamed protein) 和细胞运动 (Cell motility) 功能目录中 Unigenes 数目较少 (图 6)。



基因功能KOG分类 Gene function classification (KOG)

图 6 云斑天牛成虫触角转录组 KOG 分类图

Fig. 6 KOG Classification of antennal transcriptome of adult *Batocera horsfieldi*

KEGG 数据库是系统分析基因产物和化合物在细胞中代谢途径以及这些基因产物功能的数据库。它整合了基因组、化学分子和生化系统等方面的数据,包括代谢通路(KEGG pathway)、药物(KEGG drug)、疾病(KEGG disease)、功能模型(KEGG module)、基因序列(KEGG genes)及基因组(KEGG genome)等。KO(KEGG ORTHOLOG)系统将各个 KEGG 注释系统联系在一起,且 KEGG 已建立了一套完整 KO 注释的系统,可完成新测序物种的基因组或转录组的功能注释。云斑天牛成虫触角转录表明:总共 19 287 条 Unigenes 被分至五个 KEGG 功能类别,其中细胞过程(Cellular process)6 793 条,环境信息处理(Environmental information processing)

6 255 条,遗传信息处理(Genetic information processing)3 038 条,代谢(Metabolism)3 852 条,有机系统(Organismal system)3 508 条;其中,其它次生代谢产物生物合成(Biosynthesis of other secondary metabolites)和萜类及酮类化合物代谢(Metabolism of terpenoids and polyketides)占有 Unigenes 最少(图 7)。

2.4 转录组嗅觉相关基因

云斑天牛成虫触角转录组共注释得到 161 个嗅觉相关基因,包含 96 个气味结合蛋白(Odorant binding protein, OBP),34 个化学感受蛋白(Chemosensory protein, CSP)和 31 个气味受体(Odorant receptor, OR)(表 4)。

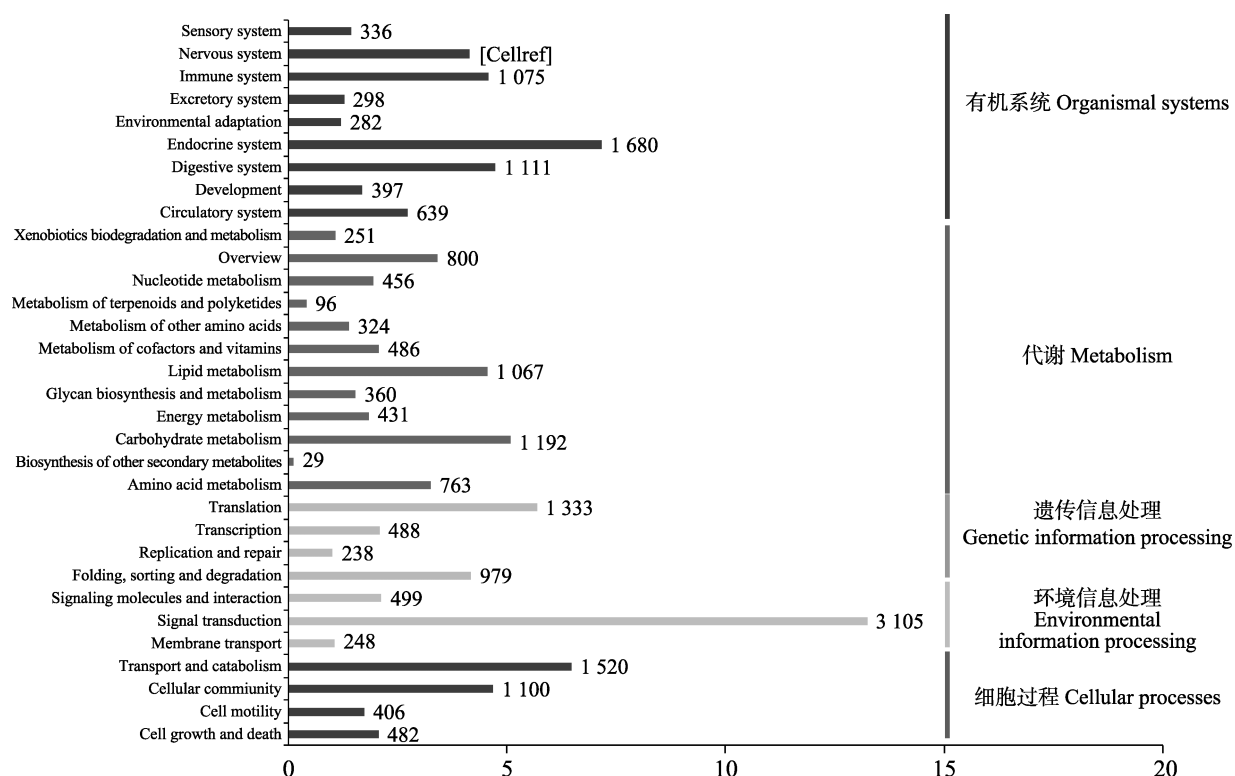


图 7 斑天牛成虫触角转录组 KEGG 分类图

Fig. 7 KEGG classification of antennal transcriptome of adult *Batocera horsfieldi*

表 4 云斑天牛成虫触角转录组嗅觉相关基因

Table 4 Olfaction-related genes in the antennal transcriptome of *Batocera horsfieldi*

项目 Item	气味结合蛋白 OBP Odorant binding protein	气味受体 CSP Chemosensory protein	化学感受蛋白 OR Odorant receptor
注释基因数 Number of annotated unigenes	96	31	34

3 讨论

利用高通量测序构建昆虫转录文库早已成为昆虫分子生物学研究的重要手段,它不仅可以捕获新基因、转录发生位点、差异表达基因、非编码及未知编码 RNA 种类,也能获取功能基因信息、转录本表达丰度、可变剪接和转录本 SNP,还能进行分子标记开发、基因表达与分析及 sRNA 分析等,因此对昆虫不同种类、不同生境条件及不同生长发育等研究通常采用该方法 (Anholt *et al.*, 2003; Etebari *et al.*, 2011; Xu *et al.*, 2012; Zhang *et al.*, 2012; 张麒麟和袁明龙, 2013); 通过转录文库构建对昆虫相关基因研究涉及众多昆虫,如直翅目进行了中华稻蝗生长发育相关基因挖掘 (吕红娟, 2012), 半翅目研究了白蜡蚧蜡合成相关基因表达 (Yang *et al.*, 2012), 鞘翅目进行了松褐天牛气味结合蛋白基挖掘 (周长祥, 2016), 竹节虫目研究了矮竹节虫分子标记开发及遗传分化推论 (Comeault *et al.*, 2012), 双翅目研究了黑腹果蝇生长发育因子的调控机制 (Hou *et al.*, 2012), 鳞翅目揭示了家蚕选择性剪接和反式剪接事件 (Shao *et al.*, 2012), 膜翅目研究了授花粉器不同发育阶段下欧洲熊蜂 *Bombus terrestris* 基因表达情况 (Colgan *et al.*, 2011) 等。

目前,利用转录文库研究昆虫分子生物学信息主要表现在以下几个方面:首先,分子标记开发,利用转录文库能获得较多昆虫分子标记,主要表现为微卫星 (Simple sequence repeats, SSRs) 及 SNPs (张麒麟和袁明龙, 2013), 例如盐沼甲虫 *Pogonus chalceus* 通过转录组文库构建捕获 38 141 个 SNPs 位点, B1 型和 B2 型大豆蚜虫分别捕获 635 个 SNPs 和 1 382 个 SSRs, 此外矮竹节虫、甜菜夜蛾 *Spodoptera exigua*、珠弄蝶 *Erynnis propertius* 和红珠灰蝶 *Lycaeides argyrogomon* 也都捕获大量的 SNPs 和 SSRs 位 (Bai *et al.*, 2010; Gompert *et al.*, 2010; O'Neil *et al.*, 2010; Comeault *et al.*, 2012; Pascual *et al.*, 2012)。其次,对昆虫行为生态学研究,如对飞蝗散居型和群居型的研究大量采用转录测序数据 (Chen *et al.*, 2010;

Jiang *et al.*, 2012; Wang and Kang, 2014)。另外,捕获昆虫生长发育相关基因,并进行表达水平分析,例如中华稻蝗生长发育相关基因挖掘 (吕红娟, 2012), 白蜡蚧蜡合成相关基因表达 (Yang *et al.*, 2012), 松褐天牛气味结合蛋白基挖掘 (周长祥, 2016), 黑腹果蝇生长发育因子的调控机制 (Hou *et al.*, 2012), 家蚕选择性剪接和反式剪接事件 (Shao *et al.*, 2012) 等;此外,昆虫对非生物逆境的响应,主要是昆虫对低温和农药的响应 (张麒麟和袁明龙, 2013), 如白纹伊蚊 *Aedes albopictus* 利用转录文库研究低温响应模式 (Poelchau *et al.*, 2011), 白杨叶甲 *Chrysomela tremulae* 利用转录文库研究了抗虫机理 (Pauchet *et al.*, 2009)。最后,昆虫与其它生物互作研究,如灰飞虱 *Laodelphax striatellus* 与病毒互作相关基因表达 (Zhang *et al.*, 2010), 大蜡螟 *Galleria mellonella* 对共生菌的免疫反应 (Vogel *et al.*, 2011) 等。

云斑天牛作为我国用材林、经济林和防护林的重要蛀干害虫,目前对其空间格局分布、防控措施、植物源诱剂研发、天敌防控、蛋白纯化以及分子生物学等研究已较为广泛,如李建庆等 (2018) 研究了不同林分杨树云斑天牛种群空间格局地统计学分析,张冬勇等 (2016) 研究了两种对云斑天牛有林间诱捕效果的植物源物质,李慧 (2012) 研究了云斑天牛触角 cDNA 文库的构建及相关嗅觉蛋白的表达及功能分析,李建庆等 (2009c) 还研究了云斑天牛的风险分析及其防控对策;本文对云斑天牛成虫触角转录组及其嗅觉相关基因进行了相关分析,并成功注释了云斑天牛成虫触角转录组相关基因功能信息,不仅丰富了鞘翅目昆虫基因数据库,也将为下一步云斑天牛的分子标记开发、行为生态学研究、捕获其生长发育相关基因、研究其对非生物逆境的响应及与它生物互作研究提供数据支撑。相关研究表明,很多鞘翅目昆虫嗅觉基因在不同虫态表达还存在显著差,本文分析了云斑天牛成虫触角转录组及相关嗅觉基因信息,不同虫态的转录组信息及其基因差异表达仍有待进一步分析验证。

参考文献 (References)

- Anholt RR, Dilda CL, Chang S, Fanara JJ, Kulkarni NH, Ganguly I, Rollmann SM, Kamdar KP, Mackay TF, 2003. The genetic architecture of odor-guided behavior in *Drosophila*: Epistasis and the transcriptome. *Nature Genetics*, 35(2): 180–184.
- Bai XD, Zhang W, Orantes L, Jun TH, Mittapalli O, Mian MAR, Michel AP, 2010. Combining next-generation sequencing strategies for rapid molecular resource development from an invasive aphid species, *Aphis glycines*. *PLoS ONE*, 5(6): e11370
- Chen S, Yang P, Jiang F, Wei Y, Ma Z, Kang L, 2010. *De novo* analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PLoS ONE*, 5(12): e15633.
- Chen JY, Luo ZJ, 2001. Damage characteristics and its control of *Batocera horsfieldi* in Hanjiang plain. *Journal of Forestry Engineering*, 15(6): 46–48. [陈京元, 罗治建, 2001. 江汉平原杨树天牛的危害特点与防治对策. 林业科技开发, 15(6): 46–48.]
- Colgan TJ, Carolan JC, Bridgett SJ, Sumner S, Blaxter ML, Brown MJF, 2011. Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. *BMC Genomics*, 12: 623.
- Comeault AA, Sommers M, Schwander T, Buerkle CA, Farkas TE, Nosil P, Parchman TL, 2012. *De novo* characterization of the *Timema cristinae* transcriptome facilitates marker discovery and inference of genetic divergence. *Molecular Ecology Resources*, 12(3): 549–561.
- Diao ZE, Ding FB, 2004. Occurrence and control of white-stiped longhorn, *Batocera horsfieldi* (Hope) (Coleoptera: Cerambycidae) on Chinese ash, *Fraxinus chinensis* Roxb. *Entomological Journal of East China*, 13(2): 49–52. [刁志娥, 丁福波, 2004. 云斑天牛在白蜡树上的发生与防治研究. 华东昆虫学报, 13(2): 49–52.]
- Etebari K, Palfreyman RW, Schlipalius D, Asgari S, 2011. Deep sequencing-based transcriptome analysis of *Plutella xylostella* larvae parasitized by *Diadegma semiclausum*. *BMC Genomics*, 12(1): 446.
- Gompert Z, Forister ML, Fordyce JA, Nice CC, Williamson RJ, Buerkle CA, 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butterflies. *Molecular Ecology*, 19(12): 2455–2473.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7): 644–652.
- Hou D, Ruiz M, Andrulis ED, 2012. The ribonuclease Dis3 is an essential regulator of the developmental transcriptome. *BMC Genomics*, 13(1): 359.
- Hu YH, 2015. Spatial distribution analysis of *Batocera horsfieldi* (Hope) and influence of the replenishing nutrition plants. Master dissertation. Beijing: Beijing Forestry University. [胡瑜珩, 2015. 云斑天牛空间格局分析及补充营养植物对其的影响. 硕士学位论文. 北京: 北京林业大学.]
- Jiang F, Yang M, Guo W, Wang X, Kang L, 2012. Large-scale transcriptome analysis of retroelements in the migratory locust, *Locusta migratoria*. *PLoS ONE*, 7(7): e40532.
- Li H, 2015. Construction and analysis of cDNA libraries from the antenna of the *Batocera horsfieldi* (Hope) and cloning, expression and functional analysis of olfactory related proteins. Doctoral dissertation. Wuhan: Huazhong Agricultural University. [李慧, 2012. 云斑天牛触角 cDNA 文库的构建及相关嗅觉蛋白的表达及功能分析. 博士学位论文. 武汉: 华中农业大学.]
- Li J, Wang MT, Zhang ZC, Chen JY, Zhang GA, 2008. Behavioral response of *Batocera horsfieldi* adults to plant volatiles. *Scientia Silvae Sinicae*, 44(6): 168–170. [李娟, 王满园, 张志春, 陈京元, 张国安, 2008. 云斑天牛成虫对植物气味的行为反应. 林业科学, 44(6): 168–170.]
- Li JQ, Yang ZQ, Zhang YL, Mei ZX, Zhang YR, Wang XY, 2009a. Biological control of *Batocera horsfieldi* (Coleoptera: Cerambycidae) by releasing its parasitoid *Dastarcus helophoroides* (Coleoptera: Bothrideridae). *Scientia Silvae Sinicae*, 45(9): 94–100. [李建庆, 杨忠岐, 张雅林, 梅增霞, 张玉荣, 王小艺, 2009a. 利用花绒寄甲防治杨树云斑天牛的研究. 林业科学, 45(9): 94–100.]
- Li JQ, Yang ZQ, Zhang YL, Mei ZX, 2009b. Spatial distribution pattern and sampling technique for *Batocera horsfieldi* (Coleoptera: Cerambycidae) in poplar. *Acta Entomologica Sinica*, 52(8): 860–866. [李建庆, 杨忠岐, 张雅林, 梅增霞, 2009b. 杨树上云斑天牛种群的空间格局及抽样技术. 昆虫学报, 52(8): 860–866.]
- Li JQ, Yang ZQ, Mei ZX, Zhang YL, 2009c. Pest risk analysis and control counter measure of *Batocera horsfieldi*. *Forest Research*, 22(1): 148–153. [李建庆, 杨忠岐, 梅增霞, 张雅林, 2009c. 云斑天牛的风险分析及其防控对策. 林业科学研究, 22(1): 148–153.]
- Li JQ, Mei ZX, Yang ZQ, 2018. Geostatistical analysis on spatial distribution pattern of *Batocera horsfieldi* (Coleoptera: Cerambycidae) populations in different poplar forest types. *Scientia Silvae Sinicae*, 54(3): 83–90. [李建庆, 梅增霞, 杨忠岐, 2018. 不同林分杨树云斑天牛种群空间格局地统计学分析. 林业科学, 54(3): 83–90.]
- Luo YQ, 2005. Theory and techniques of ecological regulation of

- poplar longhorned beetle disaster in shelter-forest. Doctoral dissertation. Beijing: Beijing Forestry University. [骆有庆, 2005. 防护林杨树天牛灾害的生态调控理论与技术研究. 博士学位论文. 北京: 北京林业大学.]
- Lv HJ, 2012. Transcriptome compare between adult and nymph of *Oxya chinensis* and the construction of their transcript map. Master dissertation. Xian: Shaanxi Normal University. [吕红娟, 2012. 中华稻蝗若虫和成虫转录组的比较研究及线粒体转录组作图. 硕士学位论文. 西安: 陕西师范大学.]
- Mei AH, Chen JY, Wu GY, Du XS, Luo FS, 1998. The species investigation, occurrence and control of poplar pests in Hanjiang plain. *Forest Pest and Disease*, (2): 36–38. [梅爱华, 陈京元, 吴高云, 杜夕森, 罗福世, 1998. 江汉平原杨树害虫种类调查、发生原因及主要害虫防治对策. 森林病虫通讯, (2): 36–38.]
- O'Neil ST, Dzurisin JDK, Carmichael RD, Lobo NF, Emrich SJ, Hellmann JJ, 2010. Population-level transcriptome sequencing of nonmodel organisms *Erynnis propertius* and *Papilio zelicaon*. *BMC Genomics*, 11(1): 310.
- Pascual L, Jakubowska AK, Blanca JM, Cañizares J, Ferré J, Gloeckner G, Vogel H, Herrero S, 2012. The transcriptome of *Spodoptera exigua* larvae exposed to different types of microbes. *Insect Biochemistry and Molecular Biology*, 42(8): 557–570.
- Pauchet Y, Wilkinson P, van Munster M, Augustin S, Pauron D, ffrench-Constant RH, 2009. Pyrosequencing of the midgut transcriptome of the poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. *Insect Biochemistry and Molecular Biology*, 39(5/6): 403–413.
- Poelchau MF, Reynolds JA, Denlinger DL, Elisk CG, Armbruster PA, 2011. A *de novo* transcriptome of the Asian tiger mosquito, *Aedes albopictus*, to identify candidate transcripts for diapause preparation. *BMC Genomics*, 12: 619.
- Shao W, Zhao QY, Wang XY, Xu XY, Tang Q, Li M, Li X, Xu YZ, 2012. Alternative splicing and trans-splicing events revealed by analysis of the *Bombyx mori* transcriptome. *RNA*, 18(7): 1395–1407.
- Steinbrecht RA, Stankiewicz BA, 1999. Molecular composition of the wall of the insect olfactory sensilla-the chitin question. *Journal of Insect Physiology*, 45(8):785–790.
- Wang SL, Wang HQ, Xia MH, Dong SH, 2004. The occurrence and control of *Batocera horsfieldi* of walnut. *China Fruits*, (2): 12–14. [王绍林, 王宏琦, 夏明辉, 董士恒, 2004. 核桃树云斑天牛的发生规律与防治技术. 中国果树, (2): 12–14.]
- Wang X, Kang L, 2014. Molecular mechanisms of phase change in locusts. *Annual Review of Entomology*, 59(1): 225–243.
- Xiao GR, 1992. Forest Insects of China. Beijing: China Forestry Publishing House. 472–473. [肖刚柔, 1992. 中国森林昆虫. 北京: 中国林业出版社. 472–473.]
- Xu Y, Zhou WW, Zhou YJ, Wu JX, Zhou XP, 2012. Transcriptome and comparative gene expression analysis of *Sogatella furcifera* (Horváth) in response to southern rice black-streaked dwarf virus. *PLoS ONE*, 7(4): e36238.
- Yan AJ, Ji BZ, Qian FJ, Chen JY, Ye ZY, 1997. A study on *Batocera horsfieldi* (Hope). *Journal of Nanjing Forestry University*, 21(1): 3–8. [严敖金, 嵇保中, 钱范俊, 陈京元, 叶中亚, 1997. 云斑天牛 *Batocera horsfieldi* (Hope)的研究. 南京林业大学学报, 21(1): 3–8.]
- Yang H, Yang MF, Yang W, Yang CP, Dong JC, Shen YH, 2010. A study on the spatial distribution pattern and the living-inhabiting tunnel of the larvae of *Batocera horsfieldi* (Hope). *Journal of Sichuan Agricultural University*, 28(2): 148–152. [杨桦, 杨茂发, 杨伟, 杨春平, 董杰超, 沈远华, 2010. 杨树云斑天牛幼虫空间分布格局及生息坑道. 四川农业大学学报, 28(2): 148–152.]
- Yang P, Zhu JY, Gong ZJ, Xu DL, Chen XM, Liu WW, Lin XD, Li YF, 2012. Transcriptome analysis of the Chinese white wax scale *Ericerus pela* with focus on genes involved in wax biosynthesis. *PLoS ONE*, 7(4): e35719.
- Zhang DY, Liu JD, Wang JY, Zhu N, Xu HC, Fan JT, 2016. Two plant compound lures for trapping *Batocera horsfieldi* (Hope) in the field. *Chinese Journal of Applied Entomology*, 53(4): 856–863. [张冬勇, 柳建定, 王菊英, 朱宁, 徐华潮, 樊建庭, 2016. 两种对云斑天牛有林间诱捕效果的植物源物质. 应用昆虫学报, 53(4): 856–863.]
- Zhang F, Guo H, Zheng H, Zhou T, Zhou Y, Wang S, Fang R, Qian W, Chen X, 2010. Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). *BMC Genomics*, 11(1): 303.
- Zhang QL, Yuan ML, 2013. Progress in insect transcriptomics based on the next-generation sequencing technique. *Acta Entomologica Sinica*, 56(12): 1489–1508. [张麒麟, 袁明龙, 2013. 基于新一代测序技术的昆虫转录组学研究进展. 昆虫学报, 56(12): 1489–1508.]
- Zhang Y, Jiang R, Wu H, Liu P, Xie J, He Y, Pang H, 2012. Next-generation sequencing-based transcriptome analysis of *Cryptolaemus montrouzieri* under insecticide stress reveals resistance-relevant genes in ladybirds. *Genomics*, 100(1): 35–41.
- Zhou CX, 2016. Host location mechanism of *Scleroderma* sp., a parasitoid of *Monochamus alternatus* Hope. Master dissertation. Wuhan: Huazhong Agricultural University. [周长祥, 2016. 松褐天牛肿腿蜂的寄主定位机制. 硕士学位论文. 武汉: 华中农业大学.]