

## 中华蜜蜂幼虫肠道参考转录组的 *de novo* 组装 及 SSR 分子标记鉴定

徐细建<sup>1</sup>, 郭睿<sup>1</sup>, 骆群<sup>2</sup>, 熊翠玲<sup>1</sup>, 梁勤<sup>1</sup>, 张串联<sup>2</sup>, 郑燕珍<sup>1</sup>, 张墨楠<sup>1</sup>,  
黄积健<sup>1</sup>, 张璐<sup>1</sup>, 李汶东<sup>1</sup>, 陈大福<sup>1</sup>

(<sup>1</sup>福建农林大学蜂学学院, 福州 350002; <sup>2</sup>江西省养蜂研究所, 南昌 330201)

**摘要:**【目的】利用 RNA seq 技术对中华蜜蜂 (*Apis cerana cerana*, 简称中蜂) 幼虫肠道参考转录组进行 *de novo* 组装, 并进行功能及代谢通路注释, 进而利用该转录组数据进行中蜂幼虫的 SSR 分子标记鉴定。【方法】实验室条件下饲养中蜂幼虫, 将纯化的蜜蜂球囊菌 (*Ascospaera apis*, 简称球囊菌) 孢子饲喂 3 日龄幼虫, 剖取 4、5 和 6 日龄幼虫肠道, 液氮速冻。将健康幼虫肠道与感染球囊菌的幼虫肠道同时进行 Illumina 测序。通过对 raw reads 的过滤得到 clean reads, 利用 Trinity 软件组装得到 unigenes。通过 BLASTx (E-value < 10<sup>-5</sup>) 比对 NCBI Nr、Swiss-Prot、KOG 和 KEGG 数据库, 对 unigenes 进行功能和代谢通路注释。利用 MISA 软件对所有 unigenes 进行 SSR 搜索, 并利用 Primer Premier 5 软件设计特异性 SSR 引物, 通过常规 PCR 对来源于北京、辽宁兴城和四川成都的中蜂幼虫肠道样本进行 SSR 位点鉴定。【结果】中蜂幼虫肠道的 RNA seq 共得到 35 670 000 条 reads, *de novo* 组装得到 43 557 个 unigenes, 平均长度为 898 nt。共有 18 225 个 unigenes 可被注释到上述公共蛋白数据库, 单独注释到 NCBI Nr、Swiss-Prot、KOG 和 KEGG 数据库的 unigenes 数分别为 3 899、443、37 和 10 个。KOG 注释结果显示, 11 442 条 unigenes 分布于 25 个基因家族, 其中注释到 RNA 加工和修饰家族的基因数最多, 达 1 249 个。9 679 个 unigenes 的 GO 分类结果显示, 在生物学进程分类中, 注释到细胞进程的基因最多, 达 4 201 个, 在分子功能和细胞组分类中, 注释到结合与细胞的基因数最多, 分别为 4 935 和 2 900 个。4 517 个 unigenes 可注释到 KEGG 数据库中的 216 个代谢通路, 注释到核糖体的基因数最多, 达 385 个。利用 MISA 软件, 可在 7 763 个 unigenes 搜索到 13 448 个 SSR 位点, 随机选取 20 对 SSR 引物对国内 3 个不同来源的中蜂幼虫肠道样本的 SSR 位点进行扩增, 有 6 对引物可鉴定出 SSR 分子标记。【结论】成功组装并注释了中蜂幼虫肠道参考转录组, 可为中蜂及其幼虫的分子生物学及组学研究提供重要的参考信息, 也可用于补充、丰富和检验东方蜜蜂的参考基因组, 基于此转录组数据开发出 6 个中蜂的 SSR 分子标记, 可应用于中蜂的基因图谱构建、基因多样性分析、基因定位等研究, 也说明利用转录组数据开发非模式生物 SSRs 的方法可行。

**关键词:** RNA seq; 参考转录组; 中华蜜蜂; unigene; SSR

## *De novo* Transcriptome Assembly for *Apis cerana cerana* Larval Gut and Identification of SSR Molecular Markers

XU XiJian<sup>1</sup>, GUO Rui<sup>1</sup>, LUO Qun<sup>2</sup>, XIONG CuiLing<sup>1</sup>, LIANG Qin<sup>1</sup>, ZHANG ChuanLian<sup>2</sup>, ZHENG YanZhen<sup>1</sup>,  
ZHANG ZhaoNan<sup>1</sup>, HUANG ZhiJian<sup>1</sup>, ZHANG Lu<sup>1</sup>, LI WenDong<sup>1</sup>, CHEN DaFu<sup>1</sup>

(<sup>1</sup>College of Bee Science, Fujian Agriculture and Forestry University, Fuzhou 350002; <sup>2</sup>Apiculture Institute of Jiangxi Province, Nanchang 330201)

收稿日期: 2016-09-29; 接受日期: 2016-12-12

基金项目: 国家现代农业产业技术体系(蜜蜂)建设专项(CARS-45-KXJ7)、福建省自然科学基金(2013J01070)

联系方式: 徐细建, Tel: 18020870542; E-mail: xxjlh2006@163.com。郭睿, Tel: 15205080780; E-mail: fafu\_ruiguo@126.com。徐细建和郭睿为同等贡献作者。通信作者陈大福, Tel: 0591-83726835; E-mail: dfchen826@163.com

**Abstract:** 【Objective】 The objective of this study is to *de novo* assemble a reference transcriptome for *Apis cerana cerana* larval gut, perform gene function and pathway annotation for this transcriptome, and to identify specific SSR molecular markers for *A. c. cerana* larvae. 【Method】 3-day-old instar *A. c. cerana* larvae were fed with the purified *Ascosphaera apis* spores, the guts of 4-, 5- or 6-day-old honeybee larvae were sampled and used as sequencing material for RNA seq. After filtration, clean reads were obtained, and unigenes were assembled using Trinity software. BLASTX tool ( $E\text{-value} < 10^{-5}$ ) was used to search the unigenes against NCBI Nr, Swiss-Prot protein, KOG as well as KEGG databases to perform gene function and pathway annotation. MISA software was used to search microsatellite markers in the larval gut's transcriptome. The specific primers of all SSRs were designed using Primer Premier 5 program and several pairs were used to amplify SSR loci in *A. c. cerana* larvae samples from 3 different regions (Beijing, Xingcheng, and Chengdu) in China by method of PCR. 【Result】 In this study, RNA seq produced 35 670 000 high quality reads, which were assembled into 43 557 unigenes with a mean length of 898 nt. 18 225 unigenes were annotated in the public protein databases. A total of 11 442 unigenes had a KOG classification and they distributed in 25 KOG categories, among them, RNA processing and modification was the largest group (1 249). 9 679 unigenes could be classified into three gene ontology (GO) categories, in which the mostly enriched ones were cellular process (4 201 unigenes), cell (2 900 unigenes) and binding (4 935 unigenes). 4 517 unigenes were annotated to 216 KEGG pathways, among them, ribosome (385 unigenes) was the largest. Finally, 13 448 SSRs were found in 7 763 unigenes, and 6 out 20 SSR loci could be successfully amplified in *A. c. cerana* larvae samples from 3 different regions in China using PCR. 【Conclusion】 This study assembled and annotated a reference transcriptome for *A. c. cerana* larval gut, which will provide a key information not only to studies on eastern honeybee and its larvae such as molecular biology and omics, but also to improve and validate the genome of *A. cerana*. SSR markers developed here could be applied to future investigation of *A. c. cerana* including gene map construction, genetic diversity analysis as well as gene location. Meanwhile, this study suggested that developing molecular markers using transcriptome data of non-model organism is a rapid and efficient method.

**Key words:** RNA seq; *de novo* assembly; *Apis cerana cerana*; unigenes; SSR

## 0 引言

【研究意义】蜜蜂是重要的授粉昆虫和社会学模式昆虫, 因其对研究神经生物学、发育、社会行为和表观基因组学的重要性而广受关注<sup>[1-4]</sup>, 西方蜜蜂 (*Apis mellifera*) 早在 2006 年就已完成基因组测序<sup>[5]</sup>, 为研究蜜蜂行为、遗传进化和基因功能提供了重要的信息和线索。与西方蜜蜂相比, 东方蜜蜂 (*Apis cerana*) 更易适应极端天气、飞行距离更长、具有更强的梳理行为和清洁行为以及群体防御能力<sup>[6-9]</sup>。中华蜜蜂 (*Apis cerana cerana*, 简称中蜂) 长期进化适应本土环境, 相比于西方蜜蜂具有抗螨害、耐寒、善利用零星蜜粉源等优点<sup>[8-12]</sup>。利用 RNA seq 技术对中蜂幼虫肠道进行深度测序, *de novo* 组装其参考转录组并进行功能及代谢通路注释, 可为中蜂幼虫的分子及组学研究提供重要参考信息, 在此基础上鉴定出的 SSR 分子标记可为在分子水平深入研究中蜂的重要性状、复杂行为及遗传进化提供宝贵信息。【前人研究进展】近年来, 以 RNA seq 为代表的高通量测序技术发展迅猛, 广泛应用于动植物及微生物研究<sup>[13-19]</sup>, 在蜜蜂研究方面也取得了一系列重要进展<sup>[20-21]</sup>。中国养蜂生产中的常见蜂种为意大利蜜蜂 (*Apis mellifera*

*ligustica*, 简称意蜂) 和中蜂。PARK 等<sup>[22]</sup>通过对 *A. cerana* 雄蜂的基因组测序和对 *A. mellifera* 及 *A. cerana* 工蜂多个组织的转录组测序, 获得 238 Mb 的基因组数据和 10 651 个基因信息, 并对 *A. cerana* 特有基因进行了分析, 但作者当时并未公布基因位置及功能注释信息。SSR 分子标记开发的传统方法是通过构建基因组 DNA 文库, 成本昂贵且费时费力, 而利用高通量测序技术的新一代 SSR 鉴定则更为经济、高效<sup>[13-15]</sup>。梁勤等<sup>[23]</sup>利用 6 对微卫星 DNA 标记对福建省 4 个中蜂群体进行遗传多样性分析, 评估群体内的遗传变异和群体间的遗传分化; 徐新建等<sup>[24]</sup>应用 10 个微卫星 DNA 标记对海南岛 11 个地点和大陆 2 个地点中蜂分析表明, 海南中蜂多样性丰富, 岛屿和邻近大陆种群发生了明显的遗传分化。目前, 已开发的中蜂 SSR 分子标记很少<sup>[25-26]</sup>, 制约了中蜂分子进化及种群遗传学的发展。【本研究切入点】东方蜜蜂的基因组已完成测序并公布, 但缺乏专一的中蜂幼虫肠道参考转录组, 严重制约中蜂幼虫的病原-宿主互作及免疫应答研究。【拟解决的关键问题】利用 RNA seq 数据组装并注释中蜂幼虫肠道参考转录组, 并鉴定出若干 SSR 分子标记, 解决中蜂幼虫参考转录组缺失以及 SSR 分子标记较少的问题。

## 1 材料与amp;方法

试验于 2015 年 12 月至 2016 年 8 月在福建农林大学蜂学学院蜜蜂保护学实验室完成。

### 1.1 供试材料

中蜂幼虫取自福建农林大学蜂学学院教学蜂场, 蜜蜂球囊菌 (*Ascosphaera apis*) 菌株保存于福建农林大学蜂学学院蜜蜂保护学实验室。

### 1.2 主要试剂及仪器

RNase-free 水购自中国上海生工生物公司; DNaseI 和 Oligotex mRNA Kits Midi 试剂盒购自德国 Qiagen 公司; Dynal M280 磁珠购自 Invitrogen 公司; 高碘酸钠购自美国 Sigma 公司; DNA ligase 购自美国 Thermo 公司; RNA Reagent 抽提试剂盒、Ex Taq polymerase 及 Superscript II reverse transcriptase 均购自日本 TaKaRa 公司; 纯化 cDNA 的 Ampure beads 为美国 Agencourt 产品; cDNA 文库构建试剂盒 TruSeq™ DNA Sample Prep Kit -Set A 为美国 Illumina 公司产品。其他试剂均为国产分析纯。

恒温恒湿气候箱购自中国宁波江南仪器厂; pH 计购自中国上海仪电科学股份有限公司; 超纯水仪购自中国四川沃特水设备有限公司; 高速冷冻离心机购自德国 Eppendorf 公司; 倒置显微镜为中国上海光学仪器五厂产品; 超净工作台为中国苏州安泰空气技术有限公司产品; PCR 仪为美国 Bio Rad 公司产品; 凝胶成像系统为中国上海培清科技有限公司产品; 超低温冰箱为中科美菱低温科技股份有限公司产品。

### 1.3 方法

**1.3.1 幼虫的人工饲养** 中蜂幼虫的人工饲料参照王倩等<sup>[26]</sup>的方法配制并改良, 将 D-果糖和 D-葡萄糖换为新鲜蜂蜜。预试验中对照组中蜂幼虫 7 日龄成活率达到 70% 以上。从福建农林大学蜂学学院蜂场选择经 PCR 检测为球囊菌阴性的健康蜂群。用灭菌的移虫针挑取 2 日龄幼虫, 放入无菌的 24 孔细胞培养板 (每孔对应 1 只幼虫, 孔内加有 35℃ 预温的幼虫饲料), 将 24 孔板放入恒温恒湿培养箱, 每隔 24 h 吸去旧饲料、加入新饲料。3 日龄时, 一组幼虫饲喂含有球囊菌孢子 ( $1 \times 10^7$  孢子/mL) 的人工饲料, 另一组幼虫饲喂以正常人工饲料。35℃, 90% RH 条件下饲喂幼虫至 7 日龄, 上述两组幼虫组均设置 3 个生物学重复。

**1.3.2 测序样品准备** 分别于 4、5、6 日龄剖取中蜂幼虫肠道组织, 为尽量减少肠道 RNA 的降解, 将从解剖取样到样品放入液氮速冻的时间控制在 30 s 以

内, 每剖取一组样品, 液氮速冻后迅速放入 -80℃ 超低温冰箱保存。

**1.3.3 cDNA 文库构建及 RNA seq** 利用 RNAiso Reagent 试剂盒抽提中蜂幼虫肠道组织的总 RNA, 然后用 RNase-free DNaseI 去除基因组 DNA 残留。RNA 的质量通过琼脂糖凝胶电泳和 NanoDrop ND-2000 (NanoDrop, Wilmington, DE, USA) 进行检测。利用 Oligotex mRNA Kits Midi 试剂盒说明书, 纯化各样品总 RNA 中的 mRNA。以 10 μg mRNA 作为模板, GsuI-oligo dT 作为反转录引物, 用 1 000 U Superscript II reverse transcriptase 在 42℃ 下孵育 1 h 合成第 1 链 cDNA; 随后利用高碘酸钠氧化 mRNA 的 5' 端帽子结构, 并连接生物素; 通过 Dynal M280 磁珠筛选连接了生物素的 mRNA/cDNA, 并通过碱裂解释放第 1 链 cDNA; 然后通过 DNA ligase 在第 1 链 cDNA 的 5' 末端加上接头, 利用 Ex Taq polymerase 合成第 2 链 cDNA。最后, 通过 GsuI 酶切去除 polyA 和 5' 端接头。利用 Ampure beads 对上述 cDNA 进行纯化, cDNA 文库通过 TruSeq™ DNA Sample Prep Kit-Set A 进行构建和 TruSeq PE Cluster Kit 进行扩增。委托广州基迪奥生物技术有限公司对上述 12 个肠道样品进行深度测序, 测序平台为 Illumina Hiseq 2500, 各肠道样品的 3 个生物学重复均同时进行测序。

**1.3.4 中蜂幼虫肠道参考转录组的 *de novo* 组装** 首先, 利用 Perl 脚本去除含有 adaptor、未知核苷酸比例 > 5% 和低质量 reads, 获得 clean reads。利用软件 Trinity<sup>[27]</sup> 进行中蜂转录组的 *de novo* 组装 (缺省值 Kmer=25)。长度短于 200 bp 的 contigs 和 unigenes 将被舍弃。过滤和组装以后得到高质量的 unigenes。

**1.3.5 Unigenes 注释** 利用 BLASTx (E-value <  $10^{-5}$ ) 将测序序列比对 NCBI nr 数据库 (<http://www.ncbi.nlm.nih.gov>)、Swiss-Prot 数据库 (<http://www.expasy.ch/sprot>)、KOG (Clusters of orthologous groups for eukaryotic complete genomes) 数据库 (<ftp://ftp.ncbi.nih.gov/pub/COG/KOG/kyva>) 和 KEGG 代谢通路 (pathway) 数据库 (<http://www.genome.jp/kegg/>)。利用 BLASTX 将组装出来的 unigenes 序列与 Nr 数据库进行比对后, 取每个 unigenes 在 Nr 库中比对结果最好 (E 值最低) 的那一条序列为对应同源序列 (如有并列, 取第一条) 确定同源序列所属物种, 统计比对到各个物种的同源序列数量。基于 Nr database 注释结果, 利用 Blast2GO 进行 unigenes 的 GO 注释, 利用 WEGO 软件对每一个转录本进行 GO 分类。

1.3.6 SSR 分子标记开发 利用软件 MISA (<http://pgrc.ipk-gatersleben.de/misa/>) 搜索 unigenes 的微卫星标记, 按照以下标准从 unigenes 中查找 SSR 位点: 二核苷酸重复 $\geq 6$ 次, 三核苷酸重复 $\geq 5$ 次, 四核苷酸重复 $\geq 5$ 次, 五核苷酸重复 $\geq 5$ 次和六核苷酸重复 $\geq 5$ 次。根据 MISA 的输出结果, 利用 Primer Premier 5 (PREMIER Biosoft Int., Palo Alto, CA) 对每一个含有 16 bp 碱基重复的 SSR 设计引物。

选取北京 (B)、辽宁兴城 (L)、四川成都 (S) 3 个不同来源的中蜂幼虫肠道样本作为模板, 随机选取 20 对 SSR 引物进行 PCR 扩增, PCR 程序: 94℃预变性 5 min; 94℃变性 50 s, 55℃退火 30 s, 72℃延伸 30 s, 共 33 个循环, 72℃再延伸 10 min。PCR 产物经 1%琼脂糖凝胶电泳检测。

## 2 结果

### 2.1 中蜂幼虫肠道的 RNA seq 及参考转录组 *de novo* 组装

对上述 12 个肠道样品进行 Illumina 测序, 平均得到 30 584 420 条原始读段 (raw reads), 去除低质量和含有接头的 reads 后平均获得 29 726 139 条有效读段 (clean reads), 总测序长度为 3 715 767 396,

Q20 平均为 98.31%, 说明测序数据质量较好, 可用于下一步分析。各样品的测序详细信息如附表 1 所示。

对 clean reads 进行进一步序列拼接和去冗余处理, 组装得到 43 557 条 unigenes, 平均长度达 898 nt, N50 为 1 704 nt (表 1)。统计结果显示, unigenes 的数目随着序列长度的增加而减少, 在 200—299 nt 长度范围内数目最多, 符合生物体序列长度分布的基本规律。长度 $>1\ 000$  nt 的 unigenes 有 10 454 条, 占总 unigenes 的 24.00%。上述结果说明中蜂幼虫肠道的组装质量较好。转录组测序数据已上传 NCBI SRA 数据库, SRA 号: SRA456721。

### 2.2 Unigenes 注释

利用 BLASTx (E-value $<10^{-5}$ ) 将测序序列比对 NCBI Nr、Swiss-Prot、KOG 和 KEGG pathway 数据库, 结果显示分别有 17 456、12 830、11 442 和 9 045 个 unigenes 能够注释到上述数据库, 有功能或代谢通路注释的 unigenes 数目为 18 225, 占全部 unigenes 的 41.84%, 此外, 有 58.16%的 unigenes 无功能注释 (表 2)。有 29 个 unigenes 在 4 大数据库均有注释, 而仅能注释到 NCBI Nr、Swiss-Prot、KOG 和 KEGG pathway 数据库的 unigenes 分别为 3 899、443、37 和 10 个。

表 1 中蜂幼虫肠道参考转录组组装结果统计

Table 1 Summary of *A. c. cerana* larval gut's reference transcriptome assembled in this study

基因数目	GC 百分比	居中长度	最长长度	最短长度	平均长度	总长度
Genes number	GC percentage (%)	N50	Max length	Min length	Average length	Total assembled bases
43557	37.10	1704	22162	201	898	39117009

表 2 公共蛋白数据库注释统计表

Table 2 Summary of annotation information of all unigenes in public protein databases

Unigene 总数	Nr 数据库	Swiss-Prot 数据库	KOG 数据库	KEGG 数据库	注释基因数
Unigene number	NCBI Nr database	Swiss-Prot database	KOG database	KEGG database	Annotated gene
43557	17456	12830	11442	9045	18225

注释到 Nr 数据库中 unigenes 的 E-value 分布显示 (图 1), 比对到物种序列的 E-value 均 $<10^{-5}$ , 其中 E-value $<10^{-100}$  的有 49.76%, 说明比对结果可信度较高。注释基因同源序列的物种分布统计结果显示前 10 位的物种依次为 *Apis mellifera*、*Apis dorsata*、*Apis florea*、*Bombus impatiens*、*Bombus terrestris*、*Lasius niger*、*Megachile rotundata*、*Harpegnathos saltator*、*Capsaspora owczarzewski* ATCC 30864 和 *Cerapachys biroi*, 注释到 *A. mellifera* 的基

因数为 5 753 (31.57%), 注释到 *A. dorsata* 和 *A. florea* 的基因数分别为 3 695 (20.27%) 和 2 489 (13.66%) (表 3)。

KOG 注释结果显示, 11 442 个 unigenes 分布于 25 个基因家族 (图 2)。其中, 注释基因数最多的为信号转导机制, 其次为一般功能预测和翻译后修饰、蛋白翻转和分子伴侣。值得注意的是, 有 170 条 unigenes 注释到防御机制, 它们可能在中蜂幼虫抵御病原入侵过程发挥重要作用。

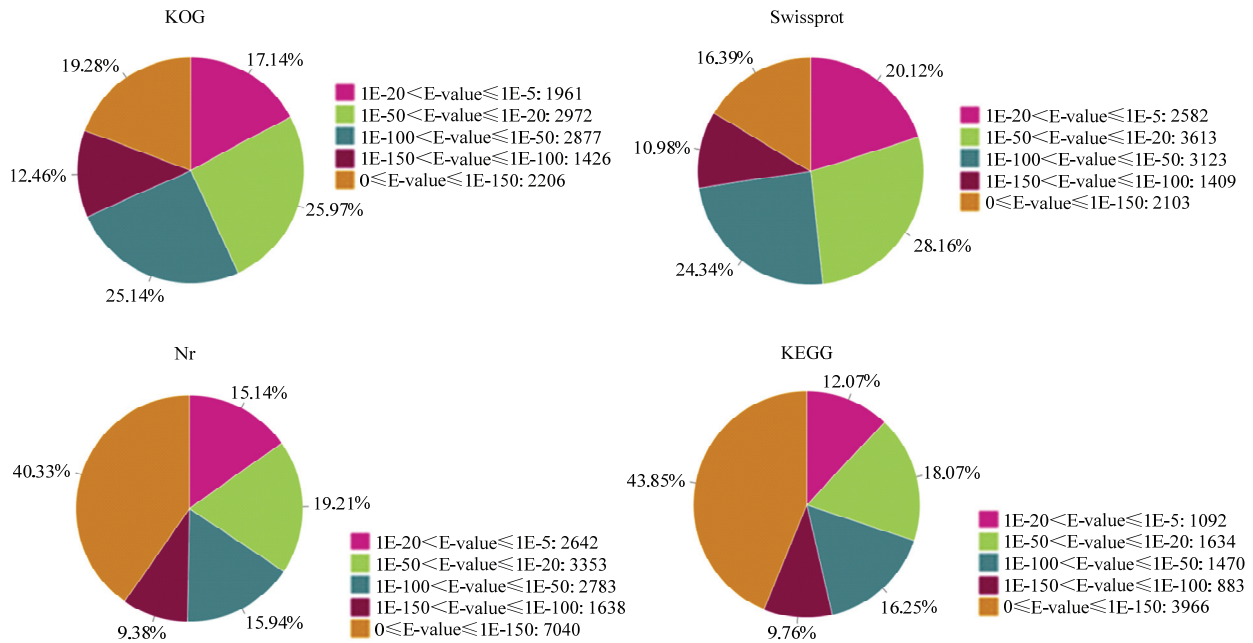


图 1 E 值分布

Fig.1 Distribution of E-value in four databases

表 3 Unigenes 的物种分布统计表 (前 10 位)

Table 3 Unigenes distribution in different species (top 10 species)

物种 Species	Unigenes 数量 Unigene number
西方蜜蜂 <i>Apis mellifera</i>	5753
大蜜蜂 <i>Apis dorsata</i>	3695
小蜜蜂 <i>Apis florea</i>	2489
凤仙花熊蜂 <i>Bombus impatiens</i>	319
欧洲熊蜂 <i>Bombus terrestris</i>	313
黑褐草蚁 <i>Lasius niger</i>	309
苜蓿切叶蜂 <i>Megachile rotundata</i>	291
印度跳蚁 <i>Harpegnathos saltator</i>	173
变形虫 ATCC 30864	163
<i>Capsaspora owczarzaki</i> ATCC 30864	163
毕氏粗角猛蚁 <i>Cerapachys biroi</i>	156

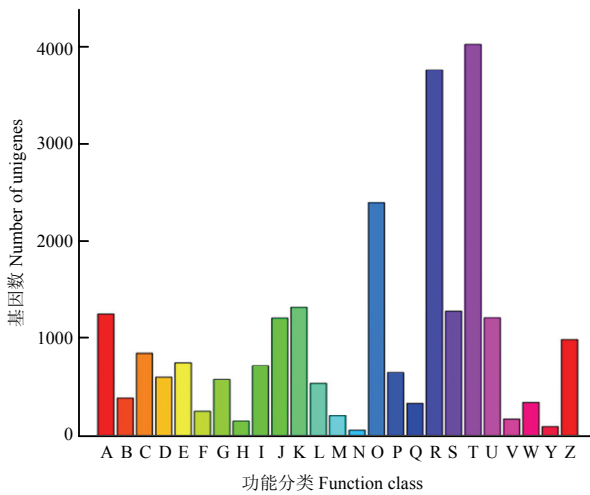
### 2.3 Unigenes 的 Gene Ontology (GO) 分类

对所有 unigenes 进行 GO 分类, 共有 9 679 个 unigenes 具有 GO 功能注释, 这些基因的功能分为生物学过程、细胞组分和分子功能 3 类。如图 3 所示, 生物学进程中, 注释到行为、生物黏附、生物调控、细胞杀伤、细胞成分组织或生物合成、细胞进程、生长、免疫系统进程、定位、运动、代谢进程多组织进程、多细胞组织进程、生殖、生殖进程、应激、信号、

单一有机体进程的 unigenes 数目分别为 22、92、1 655、2、519、4 156、7、16、1 132、36、4 146、52、220、25、31、819、593 和 3 263 个; 细胞组分中, 注释到细胞、细胞连接、细胞零、细胞外基质、细胞外基质组分、胞外区、胞外区零件、大分子复合物、细胞膜、细胞膜零件、细胞膜内腔、细胞器、细胞器零件、突触、突触零件、病毒、病毒零件的 unigenes 数目分别为 2 900、15、2 900、39、6、27、26、1 287、1 702、1 511、65、1 893、700、41、37、49 和 49; 分子功能中, 注释到抗氧化活性、结合、催化活性、通道调节子活性、电子转运活性、酶调节活性、脘基核苷酸交换因子活性、分子功能调节因子、分子转换器活性、核酸结合转录因子活性、蛋白结合转录因子活性、结构分子活性、转运子活性的 unigenes 数目分别为 48、4 935、2、34、89、61、150、316、177、30、402 和 521。

### 2.4 Unigenes 的 KEGG 代谢通路注释

对所有 unigenes 进行 KEGG 代谢通路注释, 共有 4 517 个 unigenes 注释到 KEGG 数据库中, 这些 unigenes 的通路信息如图 4 所示。这些 unigenes 分布于 216 个已知的代谢通路中, 其中富集数量最多的 10 个代谢通路是核糖体、碳代谢以及内质网蛋白加工、内吞、RNA 转运、嘌呤代谢、氧化磷酸化、剪接体、



A: RNA 加工和修饰 RNA processing and modification; B: 染色质结构与动力学 Chromatin structure and dynamics; C: 能量产生和转换 Energy production and conversion; D: 细胞周期调控、细胞分裂、染色体 Cell cycle control, cell division, chromosome partitioning; E: 氨基酸运输和代谢 Amino acid transport and metabolism; F: 核苷酸运输和代谢 Nucleotide transport and metabolism; G: 碳水化合物的运输和代谢 Carbohydrate transport and metabolism; H: 辅酶运输和代谢 Coenzyme transport and metabolism; I: 脂质运输和代谢 Lipid transport and metabolism; J: 翻译、核糖体结构和生物合成 Translation, ribosomal structure and biogenesis; K: 转录 Transcription; L: 复制、重建和修复 Replication, recombination and repair; M: 细胞壁/细胞膜/膜结构的生物合成 Cell wall/membrane/envelope biogenesis; N: 细胞运动 Cell motility; O: 翻译后修饰、蛋白质周转、伴侣 Posttranslational modification, protein turnover, chaperones; P: 无机离子转运与代谢 Inorganic ion transport and metabolism; Q: 次生代谢产物的合成、转运和代谢 Secondary metabolites biosynthesis, transport and catabolism; R: 一般功能预测 General function prediction only; S: 未知功能 Function unknown; T: 信号转导机制 Signal transduction mechanisms; U: 胞内运输、分泌和囊泡运输 Intracellular trafficking, secretion, and vesicular transport; V: 防御机制 Defense mechanisms; W: 细胞外结构 Extracellular structures; Y: 核结构 Nuclear structure; Z: 细胞骨架 Cytoskeleton

图 2 Unigenes 的 KOG 功能分类  
Fig.2 KOG classification of unigenes

氨基酸生物合成和泛素介导的蛋白水解 (表 4)。此外, 注释到溶酶体、MAPK 信号通路、Jak-STAT 信号通路、昆虫激素生物合成、黑化作用、Ras 信号通路、凋亡和嗅觉转化上的 unigenes 分别为 119、27、25、16、10、7、4 和 4 个, 其中富集在免疫通路上的 unigenes 有可能在中蜂幼虫响应病原微生物入侵的免疫应答过程中发挥关键作用。

2.5 SSR 分子标记鉴定

利用 MISA 软件从 43 557 条 unigenes 中共鉴定出 13 448 个 SSR 位点。其中二核苷酸重复最多, 数目达到 7 804 (58.03%), 其次依次为三核苷酸、四核苷酸、五核苷酸和六核苷酸重复, 数目分别为 3 797 (28.23%)、1 307 (9.72%)、339 (2.52%) 和 201

(1.49%) (表 5)。通过对 SSR 基元进行分析, 发现 AT/AT 出现的频率最高 (30.4%), 其次为 AG/CT (22%), 不同类型的 SSR 在总 SSR 中所占的比例如图 5 所示。

在上述的 13 448 个 SSR 位点中, 利用 Primer Primer 5 软件在随机挑选的 20 个 SSRs 序列两侧设计特异性引物, 引物序列信息如附表 2 所示。提取 4、5、

表 4 注释到 KEGG 数据库前 10 位代谢通路

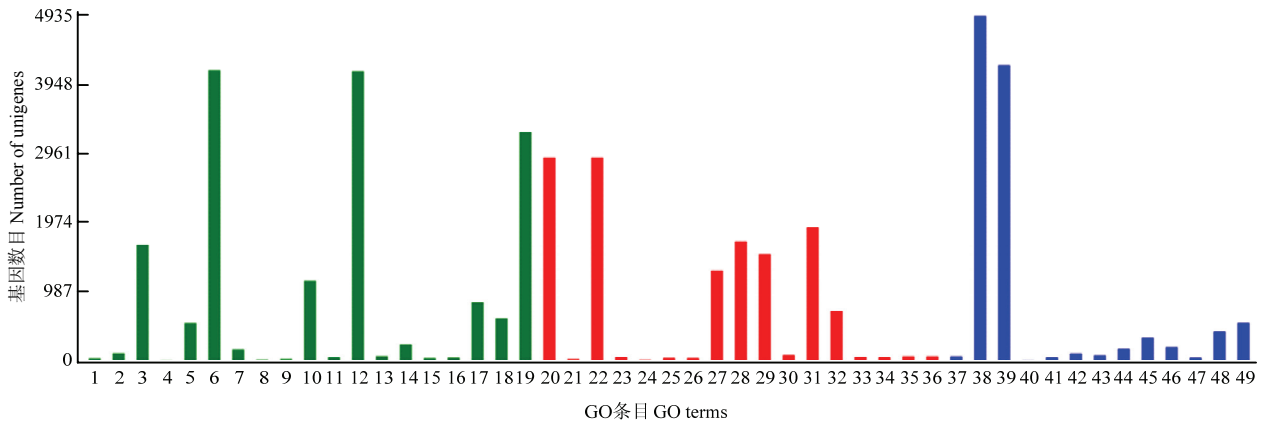
Table 4 Top 10 pathways of unigenes annotated in KEGG pathway database

代谢通路 Pathway	基因数目 Gene number (4517)	代谢通路 ID Pathway ID
核糖体 Ribosome	385 (8.41%)	ko03010
碳代谢 Carbon metabolism	270 (5.9%)	ko01200
内质网蛋白加工 Protein processing in endoplasmic reticulum	234 (5.11%)	ko04141
内吞 Endocytosis	217 (4.74%)	ko04144
RNA 转运 RNA transport	210 (4.59%)	ko03013
嘌呤代谢 Purine metabolism	207 (4.52%)	ko00230
氧化磷酸化 Oxidative phosphorylation	205 (4.48%)	ko00190
剪接体 Spliceosome	201 (4.39%)	ko03040
氨基酸生物合成 Biosynthesis of amino acids	170 (3.71%)	ko01230
泛素介导蛋白水解 Ubiquitin mediated proteolysis	146 (3.19%)	ko04120

表 5 中蜂幼虫肠道 SSR 位点统计

Table 5 Characteristics of SSRs in *A. c. cerana* larval gut

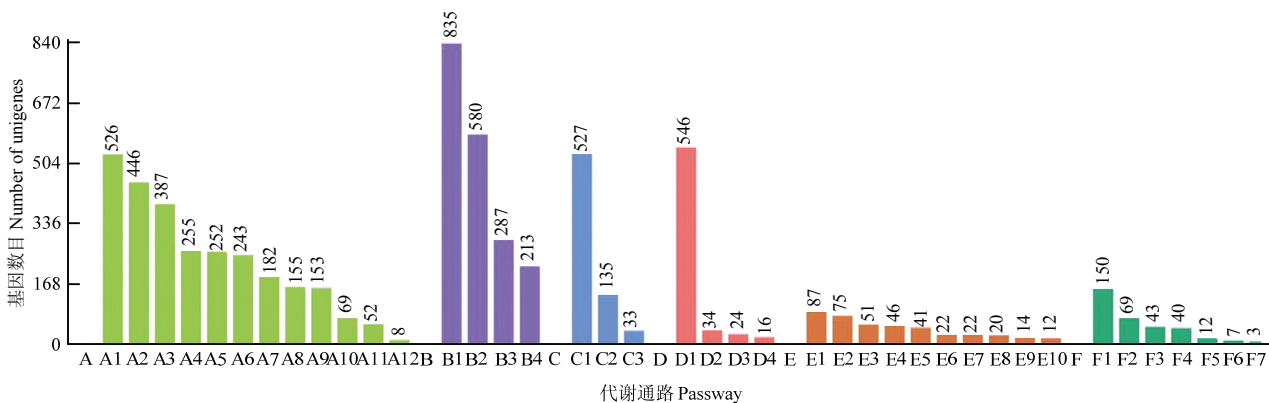
统计项目 Statistical item	数量 Number
序列总数 Total number of sequences examined	43557
序列的总长 Total size of examined sequences	39117009
SSR 位点总数 Total number of identified SSRs	13448
含 SSR 位点的序列数 Number of SSR containing sequences	7763
含 1 个以上 SSR 序列数 Number of sequences containing more than one SSR	2793
复合型 SSR 位点数 Number of SSRs present in compound formation	2347
二核苷酸重复 SSR 位点数 Di-nucleotide	7804
三核苷酸重复 SSR 位点数 Tri-nucleotide	3797
四核苷酸重复 SSR 位点数 Tetra-nucleotide	1307
五核苷酸重复 SSR 位点数 Penta-nucleotide	339
六核苷酸重复 SSR 位点数 Hexa-nucleotide	201



1: 行为 Behavior; 2: 生物黏附 Biological adhesion; 3: 生物调控 Biological regulation; 4: 细胞杀伤 Cell killing; 5: 细胞成分组织或生物合成 Cellular component organization or biogenesis; 6: 细胞进程 Cellular process; 7: 发育进程 Developmental process; 8: 生长 Growth; 9: 免疫系统进程 Immune system process; 10: 定位 Localization; 11: 运动 Locomotion; 12: 代谢进程 Metabolic process; 13: 多组织进程 Multi-organism process; 14: 多细胞组织进程 Multicellular organismal process; 15: 生殖 Reproduction; 16: 生殖进程 Reproductive process; 17: 应激 Response to stimulus; 18: 信号 Signaling; 19: 单一有机体进程 Single-organism process; 20: 细胞 Cell; 21: 细胞连接 Cell junction; 22: 细胞零件 Cell part; 23: 细胞外基质 Extracellular matrix; 24: 细胞外基质组分 Extracellular matrix component; 25: 胞外区 Extracellular region; 26: 胞外区零件 Extracellular region part; 27: 大分子复合物 Macromolecular complex; 28: 细胞膜 Membrane; 29: 细胞膜零件 Membrane part; 30: 细胞膜内腔 Membrane-enclosed lumen; 31: 细胞器 Organelle; 32: 细胞器零件 Organelle part; 33: 突触 Synapse; 34: 突触零件 Synapse part; 35: 病毒 Virion; 36: 病毒零件 Virion part; 37: 抗氧化活性 Antioxidant activity; 38: 结合 Binding; 39: 催化活性 Catalytic activity; 40: 通道调节子活性 Channel regulator activity; 41: 电子转运活性 Electron carrier activity; 42: 酶调节活性 Enzyme regulator activity; 43: 腺基核苷酸交换因子活性 Guanyl-nucleotide exchange factor activity; 44: 分子功能调节因子 Molecular function regulator; 45: 分子转换器活性 Molecular transducer activity; 46: 核酸结合转录因子活性 Nucleic acid binding transcription factor activity; 47: 蛋白结合转录因子活性 Protein binding transcription factor activity; 48: 结构分子活性 Structural molecule activity; 49: 转运子活性 Transporter activity

图 3 Unigenes 的 GO 分类

Fig.3 GO classification of all unigenes



A: 代谢 Metabolism; A1: 碳水化合物代谢 Carbohydrate metabolism; A2: 氨基酸代谢 Amino acid metabolism; A3: 总览 Global and overview; A4: 脂质代谢 Lipid metabolism; A5: 能量代谢 Energy metabolism; A6: 核苷酸代谢 Nucleotide metabolism; A7: 聚糖生物合成和代谢 Glycan biosynthesis and metabolism; A8: 辅助因子和维生素代谢 Metabolism of cofactors and vitamins; A9: 其他氨基酸代谢 Metabolism of other amino acids; A10: 外源性物质降解和代谢 Xenobiotics biodegradation and metabolism; A11: 萜类化合物和聚酮化合物代谢 Metabolism of terpenoids and polyketides; A12: 其他二级代谢的生物合成 Biosynthesis of other secondary metabolites; B: 遗传信息处理 Genetic information processing; B1: 翻译 Translation; B2: 折叠、分选和降解 Fold, sorting and degradation; B3: 转录 Transcription; B4: 复制和修复 Replication and repair; C: 环境信息处理 Environmental information processing; C1: 信号转导 Signal transduction; C2: 信号分子及互作 Signal molecules and interaction; C3: 膜运输 Membrane transport; D: 细胞进程 Cellular processes; D1: 运输和分解代谢 Transport and catabolism; D2: 细胞群体 Cellular community; D3: 细胞生长和死亡 Cell growth and death; D4: 细胞运动 Cell motility; E: 生物系统 Organism systems; E1: 衰老 Aging; E2: 内分泌系统 Endocrine system; E3: 发育 Development; E4: 感觉系统 Sensory system; E5: 免疫系统 Immune system; E6: 消化系统 Digestive system; E7: 循环系统 Circulatory system; E8: 神经系统 Nervous system; E9: 环境适应性 Environment adaptation; E10: 排泄系统 Excretory system; F: 人类疾病 Human diseases; F1: 内分泌代谢疾病 Endocrine and metabolic diseases; F2: 传染性疾病 Infectious diseases; F3: 癌症 Cancers; F4: 神经退行性疾病 Neurodegenerative diseases; F5: 物质依赖 Substance dependence; F6: 免疫疾病 Immune diseases; F7: 心血管疾病 Cardiovascular diseases

图 4 Unigenes 的 KEGG 代谢通路注释

Fig.4 KEGG pathway annotation of all unigenes

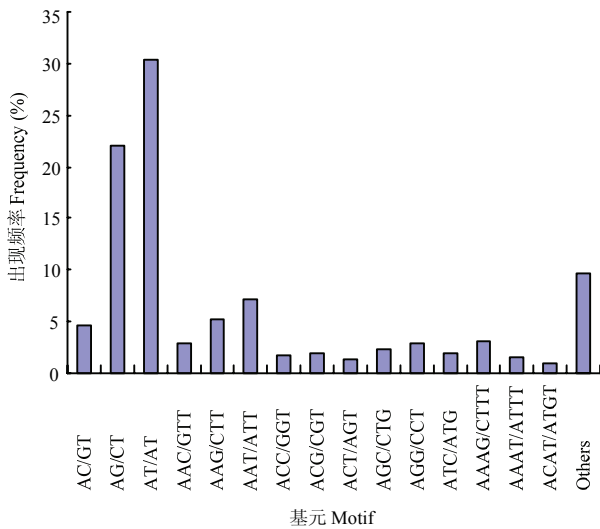
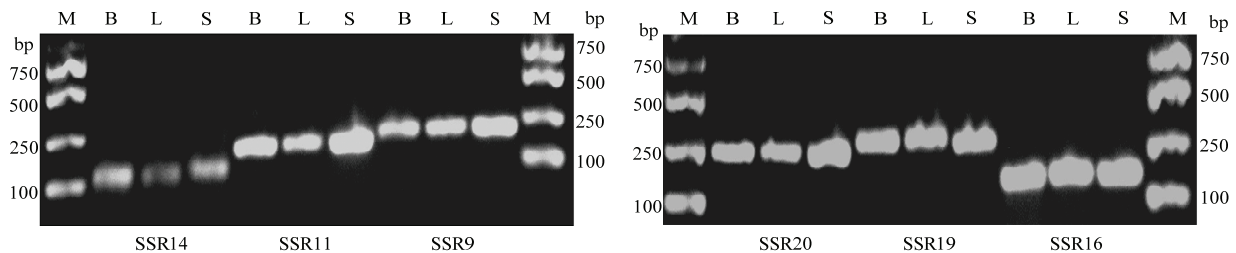


图5 不同串联重复单元类型的 SSR 在总 SSR 中所占比例

Fig.5 Frequency of SSR motif in total SSRs



M: 2000 bp ladder; B: 北京来源的中蜂幼虫样品 Larvae sample from Beijing; L: 来源于辽宁兴城的幼虫样品 Larvae sample from Xingcheng, Liaoning; S: 来源于四川成都的幼虫样品 Larvae sample from Chengdu, Sichuan

图6 国内3个来源中蜂幼虫肠道 SSR 位点鉴定

Fig.6 SSR loci identification of *A. c. cerana* larval gut samples from three different regions in China

较为复杂、不够专一。肠道是中蜂幼虫的主要免疫器官，在抵御病原微生物入侵过程中扮演着重要角色。本研究利用 RNA seq 技术对中蜂肠道进行深度测序，成功组装并注释了专一的中蜂幼虫肠道参考转录组，将有力推动中蜂及其幼虫的分子及组学研究，如中蜂幼虫响应球囊菌或东方蜜蜂微孢子虫 (*Nosema ceranae*) 侵染过程中的免疫应答及分子调控研究。

养蜂生产中，意蜂幼虫易被球囊菌感染而罹患白垩病<sup>[28]</sup>，而中蜂幼虫具有较强的球囊菌抗性，但偶尔可见患病幼虫。通常认为中蜂具有较强的清理行为，表现出更强的群体防御<sup>[7]</sup>，但中蜂幼虫个体水平的免疫防御却鲜有研究，其中蜂幼虫球囊菌抗性方面所发挥的作用值得深入探讨。未来笔者课题组将在本研

究组装并注释的参考转录组的基础上，对病原胁迫过程中中蜂幼虫的病原-宿主互作机制、免疫应答机制及分子调控机制进行深入系统的研究。

PCR 产物经 1%琼脂糖凝胶电泳检测，结果显示，有 6 对 SSRs 特异性引物 (SSR9、SSR11、SSR14、SSR16、SSR19、SSR20) 对 3 个不同来源的中蜂幼虫肠道样品都扩增出了具有多态性的特异性条带 (图 6)，说明这些 SSR 位点有望作为中蜂幼虫特有的分子标记，基于转录组数据大规模开发 SSR 分子标记具有良好的前景。

### 3 讨论

2015 年，韩国的研究人员公布了东方蜜蜂雄蜂的基因组信息，但当时并没有公布基因的位置及功能注释信息<sup>[22]</sup>。WANG 等<sup>[16]</sup>曾对中蜂进行过转录组测序，因测序组织包括 3 日龄工蜂幼虫、1 日龄工蜂蛹、1 日龄成年工蜂、采集蜂以及哺育蜂，故该转录组信息

究组装并注释的参考转录组的基础上，对病原胁迫过程中中蜂幼虫的病原-宿主互作机制、免疫应答机制及分子调控机制进行深入系统的研究。

SSR 分子标记的传统开发方法是通过构建 DNA 文库进行筛选，成本高且效率低，而高通量测序技术的应用为大规模筛选 SSR 分子标记带来曙光<sup>[15]</sup>。目前，已报道的中蜂 SSR 分子标记非常少<sup>[24-25]</sup>，严重阻碍中蜂的品种鉴定及遗传进化等研究。本研究基于中蜂幼虫肠道的转录组数据预测潜在的 SSR 分子标记，随机选取的 20 对特异性 SSR 引物中有 6 对可在北京、辽宁兴城和四川成都 3 个不同来源的中蜂幼虫样品中扩增出具有多态性的片段，这些新开发的 SSR 分子标记有助于中蜂的基因图谱构建、基因多样性分析、基



因定位等<sup>[29-30]</sup>研究的深入开展, 说明基于转录组测序数据大规模开发 SSR 分子标记具有良好的应用前景。

## 4 结论

成功组装中蜂幼虫肠道参考转录组并对其进行了功能及代谢通路注释, 可为中蜂幼虫的分子及组学研究提供重要的参考信息, 也可用于补充、丰富和检验已公布的东方蜜蜂基因组, 基于该转录组数据开发出的 6 个中蜂的 SSR 分子标记可应用于中蜂的基因图谱构建、基因多样性分析、基因定位等研究, 同时也说明利用转录组数据开发非模式生物 SSRs 的方法可行。

## References

- [1] ZAYED A, ROBINSON G E. Understanding the relationship between brain gene expression and social behavior: lessons from the honey bee. *Annual Review of Genetics*, 2012, 46: 591-615.
- [2] BEGNA D, HAN B, FENG M, FANG Y, LI J K. Differential expressions of nuclear proteomes between honeybee (*Apis mellifera* L.) queen and worker larvae: a deep insight into caste pathway decisions. *Journal of Proteome Research*, 2012, 11(2): 1317-1329.
- [3] FORET S, KUCHARSKI R, PELLEGRINI M, FENG S, JACOBSEN S E, ROBINSON G E, MALESZKA R. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(13): 4968-4973.
- [4] GALIZIA G E D, GIU M. *Honeybee Neurobiology and Behavior: A Tribute to Randolph Menzel*. Springer, 2012: 521.
- [5] WEINSTOCK G M, ROBINSON G E, GIBBS R A, WEINSTOCK G M. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, 2006, 443(7114): 931-949.
- [6] PENG Y S, NASR M E, LOCKE S J. Geographical races of *Apis cerana* Fabricius in China and their distribution. Review of recent Chinese publications and a preliminary statistical analysis. *Apidologie*, 1989, 20(1): 9-20.
- [7] LIN Z, PAGE P, LI L, QIN Y, ZHANG Y, HU F, NEUMANN P, ZHENG H, DIETEMANN V. Go east for better honey bee health: *Apis cerana* is faster at hygienic behavior than *A. mellifera*. *PLoS ONE*, 2016, 11(9): e0162647.
- [8] FRIES I, WEI H, SHI W, CHEN S J. Grooming behavior and damaged mites (*Varroa jacobsoni*) in *Apis cerana cerana* and *Apis mellifera ligustica*. *Apidologie*, 1996, 27(1): 3-11.
- [9] 周冰峰, 朱翔杰. 论中华蜜蜂种质资源的保护//中国蜂业科技可持续发展学术研讨会暨蜂业科技与生态研讨会, 2008: 110-117.
- ZHOU B F, ZHU X J. The protection of *Apis cerana cerana*//China Apicultural Conference on Sustainable Development & Apicultural Technology and Ecology, 2008: 110-117. (in Chinese)
- [10] RATH W. Co-adaptation of *Apis cerana* Fabr. and *Varroa jacobsoni* Oud. *Apidologie*, 1999, 30(2/3): 97-110.
- [11] PENG Y S, FANG Y, XU S, GE L. The resistance mechanism of the Asian honey bee, *Apis cerana* Fabr. to an ectoparasitic mite, *Varroa jacobsoni* Oudemans. *Journal of Invertebrate Pathology*, 1987, 49(1): 54-60.
- [12] 龚一飞, 张其康. 蜜蜂分类与进化. 福州: 福建科学技术出版社, 2000: 21-26.
- GONG Y F, ZHANG Q K. *Classification and Evolution of Apis*. Fuzhou: Fujian Science and Technology Press, 2000: 21-26. (in Chinese)
- [13] 李红亮. 中华蜜蜂头部 ESTs 文库构建和主要触角特异蛋白基因克隆、定位及其表达鉴定[D]. 杭州: 浙江大学, 2007.
- LI H L. Construction of cDNA libraries from head of bees, cloning, expression and subcellular localization of the antenna specific protein gene in the Chinese honeybee, *Apis cerana cerana*[D]. Hangzhou: Zhejiang University, 2007. (in Chinese)
- [14] NOVAES E, DROST D R, FARMERIE W G, PAPPAS G J, GRATTAPAGLIA D, SEDEROFF R R, KIRST M. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics*, 2008, 9(1): 312.
- [15] GUO R, WANG S, XUE R, CAO G, HU X, HUANG M, ZHANG Y, LU Y, ZHU L, CHEN F, LIANG Z, KUANG S, GONG C. The gene expression profile of resistant and susceptible *Bombyx mori* strains reveals cypovirus-associated variations in host gene transcript levels. *Applied Microbiology and Biotechnology*, 2015, 99(12): 5175-5187.
- [16] WANG Z, GERSTEIN M, SNYDER M. RNA-seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 2009, 10(1): 57-63.
- [17] HAAS B J, PAPANICOLAOU A, YASSOUR M, GRABHERR M, BLOOD P D, BOWDEN J, COUGER M B, ECCLES D, LI B, LIEBER M. *De novo* transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. *Nature Protocols*, 2013, 8(8): 1494-1512.
- [18] VAN DIJK E L, AUGER H, JASZCZYCYN Y, THERMES C. Ten years of next-generation sequencing technology. *Trends in Genetics*, 2014, 30(9): 418-426.
- [19] TRAPNELL C, WILLIAMS B A, PERTEA G, MORTAZAVI A, KWAN G, VAN BAREN M J, SALZBERG S L, WOLD B J, PACHTER L. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell

- differentiation. *Nature Biotechnology*, 2010, 28(5): 511-515.
- [20] WANG Z L, LIU T T, HUANG Z Y, WU X B, YAN W Y, ZENG Z J. Transcriptome analysis of the Asian honey bee *Apis cerana cerana*. *PLoS ONE*, 2012, 7(10): e47954.
- [21] CORNMANN R S, BENNETT A K, MURRAY K D, EVANS J D, ELSIK C G, ARONSTEIN K. Transcriptome analysis of the honey bee fungal pathogen, *Ascosphaera apis*: implications for host pathogenesis. *BMC Genomics*, 2012, 13: 285.
- [22] PARK D, JUNG J W, CHOI B S, JAYAKODI M, LEE J, LIM J, YU Y, CHOI Y S, LEE M L, PARK Y. Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. *BMC Genomics*, 2015, 16: 1.
- [23] 梁勤, 张立卿. 利用微卫星标记分析福建4个中华蜜蜂群体的遗传多样性. 福建农林大学学报 (自然科学版), 2009, 38(4): 388-392.  
LIANG Q, ZHANG L Q. Analysis of genetic diversity of four *Apis cerana cerana* populations in Fujian Province with microsatellite markers. *Journal of Fujian Agriculture and Forestry University (Natural Science Edition)*, 2009, 38(4): 388-392. (in Chinese)
- [24] 徐新建, 周姝婧, 朱翔杰, 周冰峰. 海南岛中华蜜蜂遗传多样性的微卫星 DNA 分析. 昆虫学报, 2013, 56(5): 554-560.  
XU X J, ZHOU S J, ZHU X J, ZHOU B F. Microsatellite DNA analysis of genetic diversity of *Apis cerana cerana* in Hainan Island, southern China. *Acta Entomologica Sinica*, 2013, 56(5): 554-560. (in Chinese)
- [25] 于藏龙, 周姝婧, 徐新建, 朱翔杰, 周冰峰. 长白山中华蜜蜂 (*Apis cerana cerana*) 遗传多样性分析. 福建农林大学学报 (自然科学版), 2013, 42(6): 643-647.  
YU Y L, ZHOU S J, XU X J, ZHU X J, ZHOU B F. Analysis on genetic diversity of *Apis cerana cerana* in Changbai Mountains. *Journal of Fujian Agriculture and Forestry University (Natural Science Edition)*, 2013, 42(6): 643-647. (in Chinese)
- [26] 王倩, 孙亮先, 肖培新, 刘锋, 康明江, 胥保华. 室内人工培育中华蜜蜂幼虫技术研究. 山东农业科学, 2009(11): 113-116.  
WANG Q, SUN L X, XIAO P X, LIU F, KANG M J, XU B H. Study on technology for indoor artificial feeding of *Apis cerana cerana* larvae. *Shandong Agricultural Sciences*, 2009(11): 113-116. (in Chinese)
- [27] GRABHERR M G, HAAS B J, YASSOUR M, LEVIN J Z, THOMPSON D A, AMIT I, XIAN A, FAN L, RAYCHOWDHURY R, ZENG Q. Trinity: reconstructing a full-length transcriptome without a genome from RNA-seq data. *Nature Biotechnology*, 2011, 29(7): 644-652.
- [28] GUPTA R K, REYBROECK W. *Honeybee Pathogens and Their Management*, 2014: Springer, Netherlands.
- [29] SOMRIDHIVEJ B, WANG S, SHA Z, LIU H, QUILANG J, XU P, LI P, HU Z, LIU Z. Characterization, polymorphism assessment, and database construction for microsatellites from bac end sequences of channel catfish (*Ictalurus punctatus*): A resource for integration of linkage and physical maps. *Aquaculture*, 2008, 275(1/4): 76-80.
- [30] GAO X, HAN J, LU Z, LI Y, HE C. Retracted: characterization of the spotted seal *phoca largha* transcriptome using Illumina paired-end sequencing and development of SSR markers. *Comparative Biochemistry and Physiology-Part D: Genomics and Proteomics*, 2012, 7(3): 277-284.

(责任编辑 岳梅)

附表 1 过滤前后碱基信息统计表

Table S1 Characteristics of bases before and after filtration

样本 Sample	过滤前 Before filtration			过滤后 After filtration		
	过滤前数据 Data before filtration (bp)	质量值为 20 的 数据百分比 Q20 (%)	质量值为 30 的 数据百分比 Q30 (%)	过滤后数据 Data after filtration (bp)	质量值为 20 的 数据百分比 Q20 (%)	质量值为 30 的 数据百分比 Q30 (%)
AcCK-1	4191778750	4122004862 (98.34%)	4030429418 (96.15%)	4120582000	4064246549 (98.63%)	3979833851 (96.58%)
AcCK-2	4334195500	4213996428 (97.23%)	4077228759 (94.07%)	4160134000	4083143848 (98.15%)	3965993864 (95.33%)
AcCK-3	3586255750	3521454328 (98.19%)	3438794473 (95.89%)	3519954500	3468760601 (98.55%)	3393319631 (96.40%)
AcT1-1	3553351500	3451094311 (97.12%)	3335859989 (93.88%)	3405554750	3340447479 (98.09%)	3242098346 (95.20%)
AcT1-2	3893773250	3787923084 (97.28%)	3667212827 (94.18%)	3741024750	3672494334 (98.17%)	3568820941 (95.40%)
AcT1-3	3997812500	3872456664 (96.86%)	3736549940 (93.46%)	3813806000	3736163317 (97.96%)	3622084861 (94.97%)
AcT2-1	3712367500	3645843238 (98.21%)	3556475695 (95.80%)	3643512000	3588525599 (98.49%)	3505311512 (96.21%)
AcT2-2	3367244250	3299992256 (98.00%)	3211236465 (95.37%)	3305180000	3250516762 (98.35%)	3168526789 (95.87%)
AcT2-3	3456114000	3391514363 (98.13%)	3305319508 (95.64%)	3394594250	3341561175 (98.44%)	3261515432 (96.08%)
AcT3-1	3955549750	3852320596 (97.39%)	3730784980 (94.32%)	3813542000	3745784753 (98.22%)	3640301441 (95.46%)
AcT3-2	3382756500	3312883983 (97.93%)	3223515743 (95.29%)	3313704750	3258324213 (98.33%)	3176565510 (95.86%)
AcT3-3	4445430000	4358157898 (98.04%)	4244135651 (95.47%)	4357619750	4287478600 (98.39%)	4182386252 (95.98%)

附表 2 PCR 扩增有条带的 SSR 位点引物序列

Table S2 Primers of SSR loci that can be amplified by PCR

引物名称 Primer name	引物序列 Primer sequence
Uni9-F	5'-GCGAGCTCGAATTGGATAA-3'
Uni9-R	5'-GAACGTCTCTCGTCTCCAG-3'
Uni11-F	5'-GAGTGCCCTGTGAGAGCAG-3'
Uni11-R	5'-CAAACACGTCACACTTGGG-3'
Uni14-F	5'-ATTCAACAGCAGCAGCAACA-3'
Uni14-R	5'-CGTTCAGCTGTTGGAACGTA-3'
Uni16-F	5'-GTTTCGCGACGAGTTTCTCTC-3'
Uni16-R	5'-GAGATGGGGATTGGAATTGA-3'
Uni19-F	5'-TCATCCCTCGATAAGTTGGG-3'
Uni19-R	5'-CAAATACGCATTTCCCGAAT-3'
Uni20-F	5'-GAAACGTAATCCGACCGAGA-3'
Uni20-R	5'-CATGATGAACGTATCAGCGG-3'