



大蜡螟触角转录组及嗅觉相关基因分析*

杨 爽^{1,2**} 赵慧婷³ 徐 凯² 郭丽娜² 杜亚丽² 李新宇²
苏文婷² 冯宇佳² 龙登隆² 姜玉锁^{2***}

(1. 云南省农业科学院蚕桑蜜蜂研究所, 蒙自 661101; 2. 山西农业大学动物科技学院, 太谷 030801; 3. 山西农业大学生命科学学院, 太谷 030801)

摘 要 【目的】大蜡螟 *Galleria mellonella* Linnaeus 是世界范围内蜂群中普遍存在的害虫, 给世界养蜂业带来了巨大的危害。本研究旨在建立大蜡螟触角转录组数据库, 挖掘大蜡螟嗅觉相关基因, 为今后利用嗅觉基因靶标防治大蜡螟提供分子生物学信息数据。【方法】利用高通量测序平台(Illumina HiSeq™ 2500)对大蜡螟触角进行转录组测序和组装, 通过筛选转录组数据库, 鉴定出嗅觉相关基因。【结果】成功构建了大蜡螟触角转录组, 经序列拼接获得 188 278 条 unigenes, 平均长度为 781 bp, N50 为 1 161 bp。使用 BLAST 软件将大蜡螟触角 unigenes 序列与 7 大公共数据库比对, 共注释到 108 047 条 unigenes, 其中 NR 数据库注释的 unigenes 最多, 为 78 746 条, 占 41.82%, 且 NR 注释的大蜡螟 unigenes 中与家蚕 *Bombyx mori* 类似基因最多, 为 18.4%。将 unigenes 与 GO 数据库比对发现, 共有 69 794 条 unigenes 注释到 GO 数据库中, 按照其所参与的生物过程、细胞组分和分子功能对其进行 GO 功能分类, 共含有 55 个亚类, 其中参与结合和催化功能及代谢过程的 unigenes 比例较大; KEGG 代谢途径分析表明, 14 699 条 unigenes 形成 231 条代谢通路, 其中被注释在信号传导通路中的 unigenes 最多, 为 2 401 条, 占 16.33%。进一步基因注释分析, 鉴定得到 226 个候选嗅觉相关基因, 包括 52 个气味结合蛋白 (Odorant binding proteins, OBPs) 基因, 36 个化学感受蛋白 (Chemosensory proteins, CSPs) 基因, 80 个气味受体 (Odorant receptors, ORs) 基因, 56 个离子型受体 (Ionotropic receptors, IRs) 基因和 2 个感觉神经元膜蛋白 (Sensory neuron membrane proteins, SNMPs) 基因; 通过比较雌、雄大蜡螟转录组鉴定出 114 个差异表达基因, 包括 5 个嗅觉相关基因 (*OBP8*、*OBP17*、*CSP3*、*CSP34* 和 *OR15*); 114 个差异表达基因中, 66 个基因在雌蛾触角中高表达, 48 个基因在雄蛾触角中高表达。【结论】本研究获得了大蜡螟触角转录组数据, 并鉴定出候选嗅觉相关基因, 结果为进一步研究大蜡螟的基因功能及嗅觉感受机制奠定分子基础。

关键词 大蜡螟; 触角转录组; 高通量测序; 基因注释; 嗅觉相关基因; 差异表达基因

Analysis of the antennal transcriptome and olfaction-related genes of the greater wax moth *Galleria mellonella* (Lepidoptera: Pyralidae)

YANG Shuang^{1,2**} ZHAO Hui-Ting³ XU Kai² GUO Li-Na² DU Ya-Li² LI Xin-Yu²
SU Wen-Ting² FENG Yu-Jia² LONG Deng-Long² JIANG Yu-Suo^{2***}

(1. Institute of Sericulture and Apiculture, Yunnan Academy of Agriculture Sciences, Mengzi 661101, China;

2. College of Animal Science and Technology, Shanxi Agricultural University, Taigu 030801, China;

3. College of Life Science, Shanxi Agricultural University, Taigu 030801, China)

Abstract [Objectives] The greater waxmoth (*Galleria mellonella*) is a ubiquitous pest of honeybee colonies and causes significant economic losses to the global apicultural industry. The objective of this study is to establish the antennal

*资助项目 Supported projects: 国家自然科学基金应急管理项目 (31640080); 国家蜂产业技术体系建设专项 (CARS-44-SYZ 16)

**第一作者 First author, E-mail: yangshuang19860724@163.com

***通讯作者 Corresponding author, E-mail: jiangys-001@163.com

收稿日期 Received: 2018-12-10; 接受日期 Accepted: 2019-03-25

transcriptome database of *G. mellonella* and explore olfaction-related gene data, which could provide information useful facilitating the prevention and control of *G. mellonella* using olfaction-related gene targeting. **[Methods]** The antennal transcriptome of *G. mellonella* adults was sequenced using an Illumina HiSeq™ 2500 platform, analyzed bioinformatically and successfully assembled. **[Results]** In total, 188 278 unigenes with a mean length of 781 bp and an N50 length of 1 161 bp were obtained. Of these unigenes, 108 047 unigenes were annotated using a BLAST search based on seven public databases. 78 746 unigenes (41.82%) were annotated to the NR database. *G. mellonella* unigenes had the highest similarity (18.4%) to those of *Bombyx mori*. 69 794 unigenes were divided into 55 branches and three categories (biological processes, cellular components and molecular function) using Gene Ontology (GO), and many were found to be related to binding, catalytic activity and metabolic processes. In the KEGG database, 14 699 unigenes were assigned to 231 known metabolic pathways. Among these, 2 401 unigenes (16.33%) are involved in the signal transduction pathway. By further analyzing transcriptome data, 226 olfaction-related genes of *G. mellonella* including 52 odorant binding protein (OBP) genes, 36 chemosensory protein (CSP) genes, 80 odorant receptor (OR) genes, 56 ionotropic receptor (IR) genes and 2 sensory neuron membrane protein (SNMP) genes, were identified. After comparison of the transcriptome of male and female antennae, 114 differentially expressed genes (DEGs) including 5 olfaction-related genes (*OBP8*, *OBP17*, *CSP3*, *CSP34* and *OR15*) were screened. Of these DEGs, 66 genes were highly expressed in the antennae of females and 48 in the antennae of males. **[Conclusion]** This study acquired the antennal transcriptome data of *G. mellonella* and identified olfaction-related genes. The results provide a foundation for the further study of gene functions and the molecular mechanisms of olfactory reception in *G. mellonella*.

Key words *Galleria mellonella*; antennal transcriptome; high-throughput sequencing; gene annotation; olfaction-related genes; differentially expressed genes

蜡螟属鳞翅目(Lepidoptera)螟蛾科(Pyralidae)蜡螟亚科(Galleriinae)蜡螟属 *Galleria* 昆虫(Kwadha *et al.*, 2017),危害养蜂生产的蜡螟主要有大蜡螟 *Galleria mellonella* 和小蜡螟 *Achroia grisella*。大蜡螟幼虫以蜜蜂巢脾为食,容易在长时间不清扫的蜂群和长时间存放不加以处理的巢脾上滋生、钻蛀隧道、吐丝作茧,破坏巢脾、蛀坏蜂具,在蜂群中造成蜜蜂的幼虫和蛹死亡,出现“白头蛹”,扰乱蜜蜂在蜂群中的正常活动,若不及时处理,整个蜂群会弃巢而逃(Ellis *et al.*, 2013)。大蜡螟是世界范围内普遍存在的害虫,给世界养蜂业带来了巨大的危害(Kwadha *et al.*, 2017)。目前防治大蜡螟主要靠化学防治(Svensson *et al.*, 2014; Kwadha *et al.*, 2017),但是化学药物具有高残留、污染环境、专一性不强等缺点,不适宜蜂场长效防控使用(Kwadha *et al.*, 2017)。另外,大蜡螟由于其幼虫钻蛀习性,不利于有效监测,给蜂群防治造成了极大阻碍,因此迫切需要寻求新的防治大蜡螟的方法。

嗅觉是昆虫取食、求偶、交配、寄主定位、寻找产卵场所、躲避敌害等行为活动的基础,对其种族的生存繁衍至关重要(Becher *et al.*, 2012;

Zhang *et al.*, 2015c)。对昆虫的嗅觉系统进行研究进而对嗅觉发生过程进行干扰是防治害虫的一条有效途径(胡颖颖等, 2013)。近年来,有关昆虫嗅觉基因及其功能的研究一直是人们关注的热点(Liu *et al.*, 2010; Silbering and Benton, 2010; Leal, 2013)。昆虫的嗅觉识别是一个非常复杂的过程,大致包括两个主要步骤:(1)在外周神经水平上,外周嗅觉系统对气味分子进行筛选和接收,并在神经细胞中将化学信号转化为电信号进行传导;(2)在中枢神经水平上,中枢神经系统对电信号进行整合和加工,引起昆虫对应的行为反应(Vogt and Riddiford, 1981; Kaissling, 2009)。位于昆虫触角感器中的多种蛋白参与这一过程,包括气味结合蛋白(Odorant binding proteins, OBPs),化学感受蛋白(Chemosensory proteins, CSPs)、气味受体(Odorant receptors, ORs)离子型受体(Ionotropic receptors, IRs)神经元膜蛋白(Sensory neuron membrane proteins, SNMPs)等(Vogt and Riddiford, 1981; Clyne *et al.*, 1999, 2000; Benton *et al.*, 2007, 2009; Pelosi *et al.*, 2014),这些蛋白都是未来害虫防治的潜在靶标。嗅觉系统在大蜡螟

寻偶、选择产卵场所等方面起着重要作用 (Kwadha *et al.*, 2017), 这一特性可用于大蜡螟的监测、诱杀和迷向防治。因此, 对大蜡螟嗅觉相关基因进行深入研究很有必要, 不仅可以从本质上揭示大蜡螟的嗅觉行为反应, 为研制出雌、雄大蜡螟均有良好效果的引诱剂或趋避剂提供理论依据, 还可以拓宽大蜡螟综合防控的思路。

RNA-Seq 是近年发展起来的基于新一代测序技术进行转录组分析的技术, 该技术在推动功能基因组发展的同时, 极大促进了非模式昆虫转录组学方面的研究 (张麒麟和袁明龙, 2013)。目前, 该技术被成功的应用于多种昆虫, 如鳞翅目昆虫二化螟 (Cao *et al.*, 2014)、亚洲玉米螟 (Zhang *et al.*, 2015a)、棉铃虫和烟草夜蛾 (Zhang *et al.*, 2015b)、荔枝蒂蛀虫 (孟翔等, 2016)、鞘翅目昆虫松墨天牛 (Wang *et al.*, 2014)、黄粉虫 (Liu *et al.*, 2015)、绿豆象 (郑海霞等, 2018)、膜翅目昆虫天牛肿腿蜂 (Zhou *et al.*, 2015)、切叶蚁 (Koch *et al.*, 2013)、双翅目昆虫东方小实蝇 (Zhao *et al.*, 2016)、脉翅目昆虫中华草蛉 (Li *et al.*, 2015)、直翅目昆虫大垫尖翅蝗 (金永玲等, 2015)、同翅目昆虫扶桑绵粉蚧 (胡俊杰等, 2017)、半翅目昆虫点蜂缘蝽 (宋月芹等, 2017) 等, 筛选出昆虫中大量与嗅觉相关的基因, 为深入研究嗅觉相关基因功能奠定了基础。本研究以大蜡螟触角转录组测序为基础, 对大蜡螟嗅觉相关基因进行了发掘、筛选, 研究结果可为今后大蜡螟的生物防治提供分子方面的理论依据。

1 材料与方法

1.1 供试虫源

大蜡螟于山西农业大学动物科技学院养蜂学实验室人工饲养所得。将刚交配后雌蛾分装于 1.5 mL 的离心管中 (顶端开孔), 待其产卵后, 将同一批卵 (约 1 000 粒) 转入已配好的人工饲料中, 置于温度为 (30 ± 1) 、湿度为 $60\% \pm 5\%$ 、全黑暗环境的人工气候箱中饲养, 待其羽化出房后, 取第 1 天雌、雄蛾各 300 头, 用镊子将其

触角夹下, 雌、雄触角各 3 管, 每管约 100 头触角, -80°C 保存备用。

1.2 样品 cDNA 文库构建与测序

样品 RNA 的提取、质量检测、文库构建及测序由北京诺禾致源科技股份有限公司协助完成。取检测合格的 RNA 样品各 1.5 μg , 用带有 Oligo (dT) 的磁珠富集测序样品中的 mRNA, 随后加入 fragmentation buffer 将 mRNA 打断成短片段, 以打断后的 mRNA 为模板, 用六碱基随机引物 (Random hexamers) 合成第一链 cDNA, 然后加入缓冲液、dNTPs、DNA polymerase 和 RNase H 合成二链 cDNA, 再用 AMPure XP beads 纯化双链 cDNA。纯化双链 cDNA 先进行末端修复、cDNA 的 3' 末端加上碱基 "A", 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 $\text{ng}/\mu\text{L}$, 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 Q-PCR 对文库的有效浓度进行准确定量 (文库有效浓度 $>2 \text{ nmol}\cdot\text{L}^{-1}$), 以保证文库质量。最后将合成的 cDNA 样品在 Illumina HiSeqTM 2500 测序平台上测序, 测序深度 $10\times$ 。

1.3 转录组的拼接组装与注释

高通量测序得到的原始图像数据文件经 CASAVA 碱基识别 (Base calling) 分析转化为原始测序序列 raw reads, 对 raw reads 进行数据过滤、处理: (1) 去除带接头 (Adapter) 的 reads; (2) 去除 N (N 表示无法确定碱基信息) 的比例大于 0.1% 的 reads; (3) 去除低质量 reads (Qphred ≤ 20 的碱基数占整个 reads 的 50% 以上的 reads), 得到 clean reads。采用 Trinity 软件对 clean reads 进行组装拼接 (Grabherr *et al.*, 2011), 以获得高质量的大蜡螟触角的 unigenes 库。为使转录结果更完整, 本研究测序样品数据采用合并组装方式, 以利于后续分析。

使用 BLAST 软件, 将大蜡螟 unigenes 序列与 NCBI 蛋白数据库 (NCBI non-redundant protein sequences, NR), NCBI 核酸序列数据库 (NCBI non-redundant nucleotide sequences, NT), Swiss-Prot 数据库 (A manually annotated and reviewed protein sequence database), 基因本体论 (Gene ontology, GO), 京都基因与基因组百科全书 (KEGG ortholog database, KO), 蛋白家族 (Protein family, Pfam) 和直系同源基因簇 (Clusters of orthologous groups of proteins, KOG/COG) 七大数据库进行比对, 获得 unigenes 的注释信息 (郑海霞等, 2018)。

1.4 嗅觉相关基因的鉴定及序列分析

参考之前对昆虫不同种类嗅觉蛋白的划分报道 (Cao *et al.*, 2014), 对大蜡螟嗅觉相关蛋白进行分类, 同时对蛋白序列中保守氨基酸的特性、序列一致性进行分析。Classic OBPs 具有 6 个保守的半胱氨酸 Cys, 第 2 个与第 3 个 Cys 之间存在 3 个氨基酸, 第 5 个和第 6 个 Cys 之间存在 8 个氨基酸, 通式是: C1-X₁₅₋₃₉-C2-X₃-C3-

X₂₁₋₄₄-C4-X₇₋₁₂-C5-X₈-C6 或者 C1-X₂₀₋₆₆-C2-X₃-C3-X₂₁₋₄₃-C4-X₈₋₁₄-C5-X₈-C6。Minus-C OBPs 要比 Classic OBP 少两个 Cys, 一般缺少 C2 和 C5; Plus-C OBPs 含有 8 个保守的 Cys; CSP 只有 4 个保守的 Cys, 其判定通式为: C1-X₆₋₈-C2-X₁₈₋₁₉-C3-X₂-C4。

1.5 差异表达基因分析

为了比较嗅觉相关基因在雌、雄大蜡螟成虫触角之间表达量的差异, 利用 FPKM 值对基因表达量进行评估 (Li and Dewey, 2011); 利用 DESeq 软件来筛选差异表达基因, 筛选阈值为 $\text{padj} < 0.05$ (Anders and Huber, 2010)。

2 结果与分析

2.1 大蜡螟触角转录组测序统计与数据组装

本试验测序统计如表 1 所示, 从测序得到 reads 考察其碱基质量值 (Quality score, Q-Score) 可以看出, 雌、雄大蜡螟触角各 3 个样品碱基质量值超过 Q30 的碱基所占比例均高于 89%, 说明本次转录组测序质量良好, 可信度高。

表 1 大蜡螟触角转录组测序统计
Table 1 Statistics of RNA-Seq of antennae in *Galleria mellonella*

样品 Sample	原始读数 Raw reads	筛选后读数 Clean reads	总核酸数(G) Clean bases	Error (%)	Q20(%)	Q30(%)	GC(%)
C_1	40 713 720	39 655 732	5.95	0.02	97.30	93.07	41.60
C_2	51 470 294	50 320 026	7.55	0.02	95.87	89.89	42.09
C_3	43 438 618	42 286 760	6.34	0.02	97.27	92.98	41.37
X_1	44 502 844	43 312 838	6.50	0.02	97.23	92.90	42.28
X_2	43 207 642	42 083 112	6.31	0.02	97.33	93.14	42.04
X_3	46 581 712	45 445 076	6.82	0.02	97.17	92.84	41.32

C_1, C_2, C_3 为大蜡螟雌蛾触角生物学重复样品编号; X_1, X_2, X_3 为大蜡螟雄蛾触角生物学重复样品编号; Q20 和 Q30 分别表示质量值 ≥ 20 和 ≥ 30 的碱基所占百分比。

C_1, C_2, and C_3 indicate biological replicate samples of female *G. mellonella*; X_1, X_2, and X_3 indicate biological replicate samples of male *G. mellonella*; Q20 and Q30 indicate the percentages of bases with the quality value ≥ 20 and ≥ 30 , respectively.

Illumina HiSeqTM2500 对测序文库进行双末端上机测序后, 从大蜡螟触角测得 raw reads 数 269 914 830 条, 其中雌、雄蛾触角分别含 135 622 632 条和 134 292 198 条 raw reads, 去除

带头的 reads、无法确定碱基的 reads 和低质量的 reads, 得到 clean reads 数为 263 103 544 条, 总 clean reads 碱基数 39.47 Gb (NCBI GEO 数据库登录号: GSE107887), GC 碱基占总碱基数的

百分比为 41.78% (表 1); clean reads 经 Trinity 软件拼接组装后,得到 unigenes 数为 188 278 条,平均长度为 781 bp, N50 长度为 1 161 bp; 拼接出的 unigenes 与转录本长度分布显示,长度在 200-500 bp 的 unigenes 数为 107 679 条,占比 57.19%, 长度大于 1 000 bp 的 unigenes 数为 34 013 条,占比 18.1%, 组装完整性较好。

2.2 unigene 功能注释

组装完成后,利用 BLAST 软件将 unigenes 序列与 NR, NT, Swiss-Prot, GO, KO, Pfam 和 KOG 7 大数据库进行比对,获得 unigenes 的功能注释(表 2)。BLAST 参数 E-value 不大于 10^{-5} , HMMER 参数 E-value 不大于 10^{-10} 。通过 BLAST 比对,最终获得 188 278 条有注释信息的 unigenes,其中 NR 数据库中注释到的 unigenes 数目最多,为 78 746 条,占 41.82%; KO 数据库中注释到的 unigenes 数目最少,为 14 699 条,占 7.8%。

表 2 大蜡螟基因注释成功率统计
Table 2 Successful annotation rate of genes of *Galleria mellonella*

数据库 Annotated database	注释基因数量 Number of annotated unigenes	注释基因所占比例 Percentage of annotated unigenes
NR	78 746	41.82%
NT	69 581	36.95%
KO	14 699	7.8%
Swiss-Prot	72 709	38.61%
Pfam	69 018	36.65%
GO	69 794	37.06%
KOG	42 001	22.3%
All	7 388	3.92%
At least one	108 047	57.38%
Total unigenes	188 278	100.00%

根据 NR 数据库注释结果,进一步对这些注释的序列的同源物种进行统计,结果发现同源序列主要来自家蚕 *Bombyx mori*、大白斑蝶 *Danaus plexippus*、小菜蛾 *Plutella xylostella*、藏羚羊 *Pantholops hodgsonii*、智人 *Homo sapiens*, 类似基因分别占 18.4%、9.2%、8.2%、8.0%、2.4% (图 1)。

此次测序共有 69 794 条 unigenes 注释到 GO 数据库中 (图 2), 按照其所参与的生物过程、细胞组分和分子功能对其进行 GO 功能分类,总共含有 55 个亚类。结果表明,在参与的生物过程 (Biological process) 本体中,参与细胞过程 (Cellular process) 的 unigenes 数量最多 (38 872 条),代谢过程 (Metabolic process) 次之 (36 716 条); 在细胞组分 (Cellular component) 本体中,参与细胞 (Cell) 的 unigenes 数量最多 (19 815 条), 其次是细胞部分 (Cell part) 为 19 814 条; 在分子功能 (Molecular function) 本体中,参与结合功能 (Binding) 和催化功能 (Catalytic activity) 的 unigenes 数量最多,分别为 35 705 条和 30 893 条。

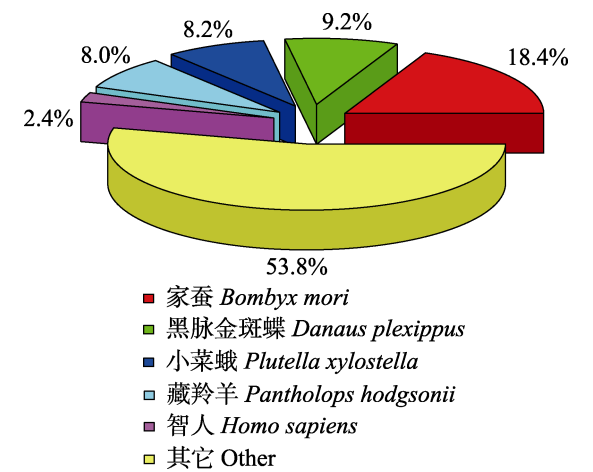


图 1 NR 数据库中注释的大蜡螟触角 unigenes
Fig. 1 Annotated unigenes in antennae of *Galleria mellonella* in NR database

将大蜡螟转录组的 unigenes 和 KOG 数据库进行比对,共 42 001 条 unigenes 在 KOG 数据库中得到注释,分为 26 类,其中注释最多的是 R 类 (基本功能预测),含有 6 580 条;其余依次是 T 类 (信号传导机制) 5 600 条;O 类 (翻译后修饰、蛋白转运及蛋白伴侣) 4 777 条, J 类 (参与翻译、核糖体结构及生物合成) 3 464 条, C 类 (能量生产与转换) 2 690 条, E 类 (氨基酸运输与代谢) 2 558 条, 其它类别均低于 2 500 条, 分布如图 3。

结合 KEGG 数据库对大蜡螟的 unigenes 可能参与或涉及的代谢途径进行分析, 结果表明,

14 699 条 unigenes 参与了 5 大类代谢通路分支，
即细胞过程 (A, Cellular processes) , 环境信息处

理 (B, Environmental information processing) , 遗传
信息处理 (C, Genetic information processing) ,

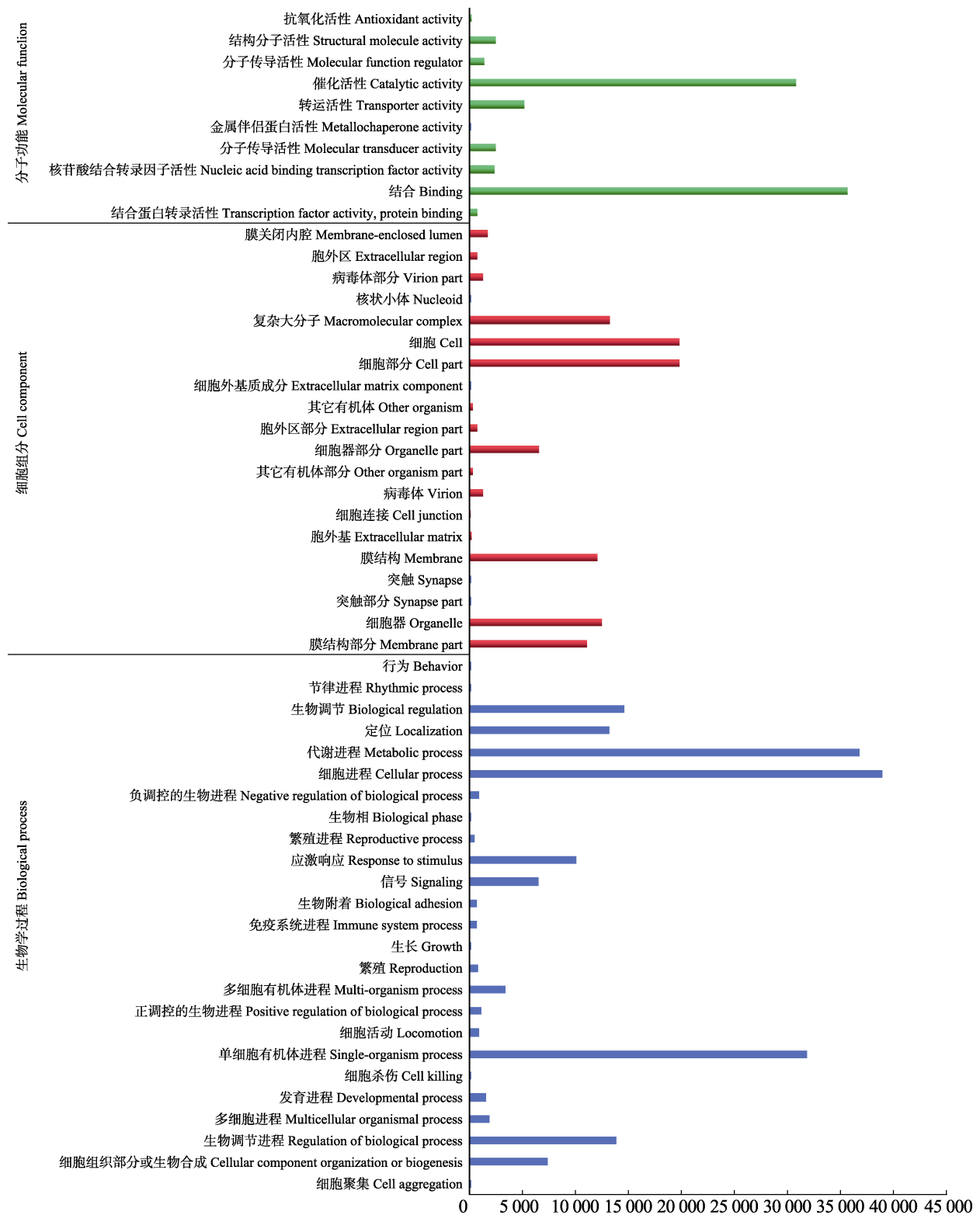


图 2 大蜡螟触角 unigenes GO 功能注释

Fig. 2 Annotated unigenes in antennae of *Galleria mellonella* in GO database

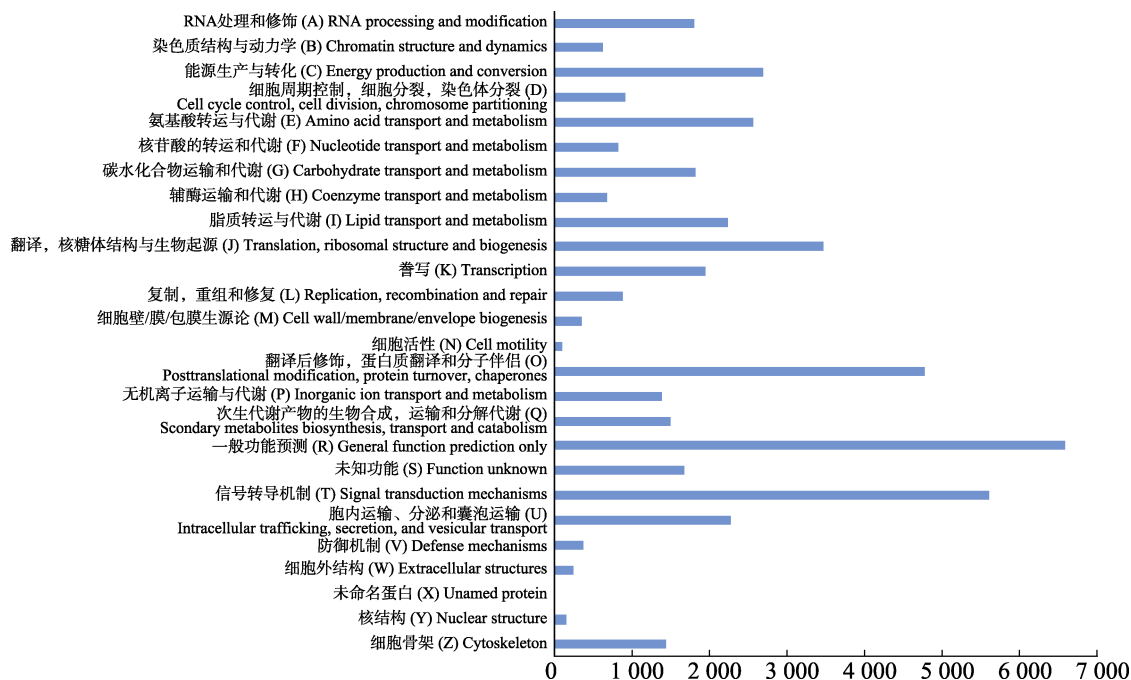


图 3 大蜡螟触角 unigenes KOG 分类图

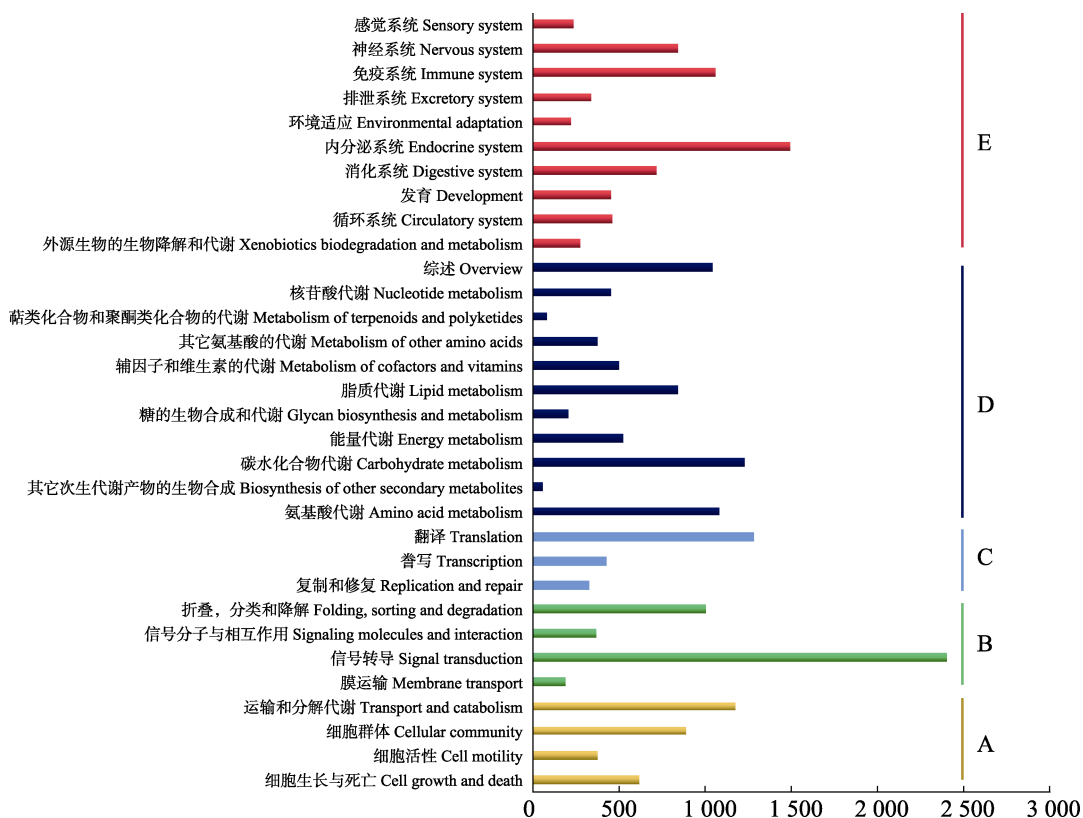
Fig. 3 KOG classification of the annotated unigenes in antennae of *Galleria mellonella*

图 4 大蜡螟触角 unigenes KEGG 分类图

Fig. 4 KEGG classification of the annotated unigenes in antennae of *Galleria mellonella*

A: 细胞过程; B: 环境信息处理; C: 遗传信息处理; D: 代谢; E: 有机系统。

A: Cellular processes; B: Environmental information processing;

C: Genetic information processing; D: Metabolism; E: Organismal systems.

代谢(D, Metabolism)和有机系统(E, Organismal systems), 分布在 231 个代谢通路中, 其中被注释在信号传导(Signal transduction)通路中的基因最多有 2 401 条(图 4), 占全部被注释基因的 16.33% ;其余注释较多的代谢通路依次是内分泌系统(Endocrine system) 1 495 条, 翻译(Translation) 1 284 条, 碳水化合物代谢(Carbohydrate metabolism) 1 226 条, 运输分解代谢(Transport and catabolism) 1 173 条, 氨基酸代谢(Amino acid metabolism) 1 079 条, 免疫系统(Immune system) 1 054 条。

2.3 嗅觉相关蛋白基因的鉴定与序列分析

通过关键词搜索和注释基因分析, 从大蜡螟触角转录组数据中, 鉴定出 226 个候选嗅觉相关基因, 包括 52 个气味结合蛋白、36 个化学感受蛋白、80 个气味受体(包括 79 个典型气味受体基因和 1 个非典型气味受体基因 *OR2*)、56 个离子型受体(包含共受体 *IR8a* 和 *IR25a*)和 2 个神经元膜蛋白基因。

序列分析结果显示, 52 个 OBPs 中 35 个具有完整开放阅读框(ORF), 它们的序列一致性为 23.27% ;其中 OBP1, OBP13, OBP21, OBP29 和 OBP40 含有 4 个保守的 Cys, 属于 Minus-C OBPs ; OBP13 和 OBP35 含有 5 个 Cys, OBP5, OBP20, OBP32, OBP42, OBP44 含有 7 个 Cys ; 剩余 24 个 OBPs 含有 6 个保守的 Cys, 属于 Classic OBP (图 5)。36 个 CSPs 中 33 个具有完整 ORF, 序列中均具有 4 个保守的 Cys, 序列一致性为 39.43% (图 6)。获得的 2 个 SNMPs, 即 SNMP1 和 SNMP2, 它们的序列一致性为 25.88% (图 7)。

2.4 差异表达基因分析

基于大蜡螟触角转录组测序 FPKM 值, 我们对雌、雄大蜡螟成虫触角嗅觉差异表达基因进行了分析, 共鉴定出 114 个差异表达基因(DEGs), 其中 66 个基因在雌性触角中高表达, 48 个基因在雄性触角中高表达(图 8); 在 114 个 DEGs, 鉴定出 5 个嗅觉差异表达基因, 包括

2 个 OBPs 基因(*OBP8* 和 *OBP17*)、2 个 CSPs 基因(*CSP3* 和 *CSP34*)和 1 个 OR 基因(*OR15*)。

3 讨论

随着新一代高通量测序技术的发展, 转录组测序逐渐被运用于非模式昆虫的功能基因发掘(Vogel *et al.*, 2011), 并趋向于解决有害生物防治、近缘种遗传关系与进化等更加复杂的实际问题(杨帆等, 2014)。本研究利用 RNA-Seq 技术对大蜡螟的触角进行了转录组测序, 共获得 188 278 条高质量的 unigenes, 所有样本的 Q30 碱基比例不低于 89.89%, N50 长度为 1 161 bp, 表明本研究获得的测序数据质量良好。对 unigenes 进行功能注释发现, 108 047 条 unigenes 成功注释, 在 NR 数据库中注释到的 unigenes 最多, 为 78 746 条, 占 41.82%, 这与大多数研究结果相似(金永玲等, 2015; 孟翔等, 2016); 其中大蜡螟触角 unigenes 序列与鳞翅目昆虫家蚕基因序列相似度最高, 达 18.4%, 与鳞翅目小菜蛾基因序列相似度为 8.2%, 表明大蜡螟触角基因与鳞翅目昆虫基因功能相似。

GO 数据库的基因功能分类体系有助于全面认识昆虫的基因和基因产物的属性, 此次测序共有 69 794 条 unigene 注释到 GO 数据库中, 按照其所参与的生物过程、细胞组分和分子功能对其进行 GO 功能分类, 总共含有 55 个亚类, 其中参与细胞、代谢、催化和结合活性的基因功能注释较多; KOG 数据库比对结果显示, 共 42 001 条 unigenes 在 KOG 数据库中得到功能注释, 除基本功能预测外, 注释最多的是信号传导机制, 含有 5 600 条; 结合 KEGG 数据库对大蜡螟的 unigenes 可能参与或涉及的代谢途径进行分析, 结果表明, 14 699 条 unigenes 参与了 5 大类代谢通路分支, 分布在 231 个代谢通路中, 其中被注释在信号传导通路中的基因最多, 有 2 401 条, 占全部被注释基因的 16.33%, 充分说明了触角在大蜡螟嗅觉识别行为活动中的重要性。

通过基因功能注释, 本研究从大蜡螟触角中鉴定获得 226 个候选嗅觉相关基因, 包括 52 个

GmelOBP1	MIMVSVLLLAIAITVGYAEQRTALYPAEKVGQFLG.....	IFTNCSISEAGATEDSLIKLKEVKLSDDKN....	64
GmelOBP3	MLAMRRFSIALFSLCAVCGVFAEDIENPITAVVRKILIS.....	TARSCMDHVKATEADIEYLRGDPPEKAA...	69
GmelOBP5	MTATLTLFLVLIAFGNCGKDKPVSSEIEKTIQIVH.....	DACVGTGASEEDITNCENGI..FKEDPK..	64
GmelOBP7	MLLIRIFELVLIAASEAMIMKQIKNTGKMR.....	KSCQPKNNVEDKIDPINDGV..FIEEKE..	59
GmelOBP8	MLLIAAAKFLLLIAFCEAMIMKQIKNTGKMR.....	KSCQPKNNVEDKIDPINDGV..FIEEKE..	59
GmelOBP9	MYAVIISVLLIFISVGGGKDKPVFGDEINEMNIH.....	NEVSKTKVAEEDITNCENGI..FKEDVK..	64
GmelOBP13	MSRFLFLLIQVGVFVSCVGVPKKTIVHITPELAG.....	KILLIALQCPETGIDSDAIVMIFYGKYRDDQ...	67
GmelOBP15	MNSFLSLAVLTIALSYVT..ALTQBELDKVREPMIN.....	FTKQSKDYGVSDDELKAIKESGKVDGVEP..	64
GmelOBP16	MYNVSVSGVNIIFKMNSFLCLVILGLAVSNVT..ALTNDDELDKIHESMVV.....	FTKQSKDEGVSDDELKAIAKESGKVDGVEP..	79
GmelOBP17	MEARAARVCAVLGLLLVGLVTAIDDDMAELIKMVH.....	DSQGEETGVDFGLVDRKNAGADIMPDEK..	63
GmelOBP19	MAKVHRVFLCVVAAAFSSVS..ALSDEKETIKGMVMP.....	LLTTCQKEHDITMEDVIAAKKANDGDAINP..	67
GmelOBP20	MYAVIISVLLIFISVGGGKDKPVFGDEINEMNIH.....	NEVSKTKVAEEDITNCENGI..FKEDVK..	64
GmelOBP21	MNPSTIVCLIFSVAGINAGTVHLSGAQKEKATQAL.....	TCIKETGVKPEVIAELIKKGLHTEDEG...	62
GmelOBP22	MSTVLKVLCAPIFIRFFVQGDPIDELKQTMN.....	IVKGLTKYPTVISEEMHMKNGTLKSGQAN..	64
GmelOBP29	MPSVKLFVSVMTGLSAAQKRVKFPPEVATEIIQA.....	ASQCVIETGVNPELLIKFANYDGKIIEDEQ..	66
GmelOBP30	MHTKKLHLHYCAFLIQLIIFDYSYGMIRQQLKNSGKLM.....	KSCMPKNDVTEEEIGEIEQK..FIEDRK...	65
GmelOBP31	MVASYCFLFTAIVLASISNPVGTAEVMSHVSAAFGK.....	TLEROREESGLTAEIMEEFKFWSDDFDIHRE	70
GmelOBP32	MAARKLILVIVISLPGIMA..DMVMKVDTVLGFE.....	ALEHOREETALTTEKMEQFFHWRDDDFKFEHRE	66
GmelOBP34	MKYVLVAVLAAGLGLVAGTAEVMSHVTAEFGK.....	ALTTCRDESGLSAEVLIEEFQHFHWRDDDFKFEHRE	66
GmelOBP35	MAAPPMDMEDMAELAAIVR.....	SSQSGESGVDLGLVERVNAGTAIMEADAK...	47
GmelOBP37	MEKYWFILIVAVITVGDCCDMTKEQLKKTGMRL.....	KOCVTKVGVVEEKISQIEKQK..FIEEDK...	61
GmelOBP38	MKVLYTVSVIVLKLTAATLTHEHITTIQPKLTN.....	HADDLEEYPLSTEDIAAFKNGSEFPDGDSD...	63
GmelOBP40	MHFIETITVVISIALACVKGSPSNINMGGAQFLG.....	HVMSCNEEFKIPRIKLESEM..TDDIGDIDP..	64
GmelOBP41	MYRSTILLFCLLENFMHSTIATIDEQRAQIHQHEFEMIG.....	IBQIKDHTVIEDDITNLRTRKLPKSEYG...	66
GmelOBP42	MSRVLLLVIAVQSFVPTELASVSKTGSVVKETDRDDYKMMSTGLIT.....	VFTESRS..SETELDQAMNEONETFRVMSYLESINESGSPFDETRT...	94
GmelOBP43	MITRQQLKNSGKLM.....	KTCMPKNDVTEEEVGDIKQK..FVEDRR...	41
GmelOBP44	MINFAVILCLLLTILSVS..ILVTCMKSTEMSTAPQDKTIDNTKVMNSKMIANDSNAKDEIDLMSAMT...	ONESFRIEMSEFLIALNESGSPFDETRT...	103
GmelOBP45	MHFIETITVVISIALACVKGSPSNINMGGAQFLG.....	HVMSCNEEFKIPRIKLESEM..TDDIGDIDP..	64
GmelOBP46	MSKRIVIVVTIYLAIT..EPSTDSQDINMKELMNVK.....	AFDIQCKELDENTVHADRYNFWKREDFITHRQ...	68
GmelOBP47	MGRSKLGVIAFCLTCLLLH..VRRANSSAEVWKLVSITFAK.....	AMDACKKELNVGDHILQEMQNFVREYDVLVHRD	73
GmelOBP48	MNMIMKGFAR.....	VLEDCKKQENVDGHIMQDFANFWHEEYDLVNR	43
GmelOBP49	MAG..AVFVGLLLAALISVRDVEPSREVLYITITGFLK.....	TLDDCKHELNLSDHITLIDLYHYWKQBYDMLDKD	68
GmelOBP50	MIEPKLLILVLLVSV..ISKVELSQDVMKNSINFAK.....	PMEIICKKEMDLPDSAKKEFGNFKWDGYVLSDRQ	68
GmelOBP51	MENKHVLLVVVLVAV..VRRSE..AADAMKLASGELIT.....	VLEDCKKELNIDNITIDLYHYWKQBYFALLHRE	67
GmelOBP52	MENKLDNNEERKVFESFVITVLMGL..NCKINASQIVMKMNVNFGK.....	ALEACKKELSLPDTTFEDFYNFWKVDYDLKHRT	79
Consensus			
GmelOBP1	FNKYLYOVLTSGYSKDKGHVTEEEKIVE.....LFPE..SDRK.....IVKEVIEECNKNQGLEA.....ADAALAFYKCTIEKSPFWLVIRQ.....		140
GmelOBP3	CIILCKLEKGVIGVKNKYSKIGFMATVSPVLFKN..KKLE.....HMKIVSENCDEKVNHAT..ATPQQLGNEVTSCLFYAPELHFKS.....		150
GmelOBP5	LKCYMECLLEASLADGVDYDMLVS.....LIPE..EYLD.....RTSKMIFGCKHLDTDPK..D..KQSAFTVHKCSYEKDPDLYFLF.....		141
GmelOBP7	VKCYIACIMKMANITK..NGKINFDAAIKQADLLLPD..EIKQ.....PAKEAIGACRKIADGHK..D..ICDASFVYTKIHDHNSVFFTF.....		139
GmelOBP8	LKCYIACIMKMANITK..NGKINFDAAIKQADLLLPD..EIKQ.....PAKEAIGACRKIADGHK..D..ICDASFVYTKIHDHNSVFFTF.....		139
GmelOBP9	LKCYMECVVEEINLVDDNNEVDYEMLIS.....LFPE..QYQK.....RAISLIVESCKHLGN.....CFS.....		117
GmelOBP13	FKDFVECCYKITGFLRNDGFLNEDKATK.....AFRN.....EP.....LIEEIQRCGLRQKDP.....KESLFKFKYCFVDTPITITL.....		139
GmelOBP15	CLIGCVFKAGFIDGAKFLPKIKELSKYLSN..EEDIAKFGTIVDCAANDEQVTGG..DPAICIRSKLVLECYAKHREELIKVTN.....		149
GmelOBP16	CLIGCVFKAGFIDGAKFLPKIKELSKYLSN..EEDQTLFGSVVDTCAANDEEVDTG..DAGVRSKLVFECEIKHKGQ..FKVSR.....		163
GmelOBP17	LKCYIACIMVITGGMS..DGEVDIDAVLT.....LLPE..NIGK.....KNEPFLRCGCTTKGADD.....CDTAFITQVQWQANKADYFLI.....		137
GmelOBP19	CFYACFFKKIGVITEEGRYAPDVAKEGHKKFIHD..DELVSKLDSVTDKCSSINDQIVITDQ..AEGCEISKLLVNCLKENKDEFSLFDD.....		152
GmelOBP20	LKCYMECVVEEINLVDDNNEVDYEMLIS.....LFPE..QYQK.....RAISLIVESCKHLDMKEK..E..SQRVFDIHKCSYAVDPDPFYFIF.....		141
GmelOBP21	LKFTFLQFQKAGIIGHDGKINAVGAALS.....KLPPGVDKA.....EAAKVLCECKNKKGKA.....ADTAYQVFKYHAGAKTHIVL.....		137
GmelOBP22	CLIAAYKRTGMDEAGMLSLGVNKAATGMVFSNPNKMKKAEFEIACKGVNEEVDNDGDKGCTRAALIFRCITIEKAPGFDLI.....		149
GmelOBP29	FNKFAYCIVLVTKYGYKGRNGRVKLDKATA.....IMPKDVDDK.....AVRKIMECNEETGKDA.....POTAFKIFKCLNEKSPVIMIMI.....		142
GmelOBP30	VKCYIACVYMSQVIK..NNKLYEAVLKQVDMFPA..EMKD.....AVKASATHCKDISKRYK..D..ICEASYWTAKQYDFDPSNFWFP.....		145
GmelOBP31	LGCAIICMNSKSLMK..DDSRILHVNMDHYKVSFPN..GEVLAA.....KMVELIHNCEKKYDDIE..D..DCTRVMKSAACFKQDAKKEGIAP..EVAMIEAVVEKY.....		164
GmelOBP32	VGCALICVMSRYNLLIT..DSHRMHENIEKFKSEFPN..GEVLAK.....QMVELIHNCEKKYDDIE..D..DCTRVMKSAACFKQDAKKEGIAP..EVAMIEAVVEKY.....		164
GmelOBP34	LGCAIICMNSKSLMK..EDSRMHVNMDHYKVSFPN..QGVLSA.....KLVELIHNCEKKYDDIT..D..DCTRVMKSAACFKQDAKKEGIAP..EVAMIEAVVEKY.....		160
GmelOBP35	LKCYIACIMETAGMS..DEEVDVEAVLA.....LLDD..NTRS.....RNEKSLRCGCTQKQADK.....CDTAWKTQRCQWQANPNDGYLLI.....		121
GmelOBP37	VKCYIACVYMSQVIK..NNKLYEAVLKQVDMFPA..EMKA.....SIRKTSIENCKDKVAKRYK..D..VCEASYPTAKCTYDDNPKDFIFA.....		141
GmelOBP38	AACLGAVLKTIGLFDLGLSVKDAALDKAKEIFTG..EGEMDAITDIITCAKINDEVTEDG..EKGGERARLLFNCLNEIK.....		142
GmelOBP40	CLFACVFRNTGLLDDKGMNPEVNDKLNAFET..DITYIKKFDGIIKGCYSYGN.....		115
GmelOBP41	APCFIACVMTKGLVLDQGLMKENALELARKVFED..EEVQKISDYMHSCHAVNTASVSDG..EKGGERARFLAYKCMENAPQGFEDI.....		152
GmelOBP42	PKCYVROVLEKTVGASED..GNYPARTA.....VIFAG..KRAGRPMNDIEEATGCAANKESCK.....CERAYQYIKCMETIEIRKYE.....		171
GmelOBP43	VKCYIACVYMSQVIK..NNKLYEAVLKQVDMFPP..EMRT.....AVKAAAEKNDIAKRYK..DDICEASYRTAKCTYDDNPFDAENFWFP.....		122
GmelOBP44	PKCFIACVMTKGLVLDQGLMKENALELARKVFED..EEVQKISDYMHSCHAVNTASVSDG..EKGGERARFLAYKCMENAPQGFEDI.....		183
GmelOBP45	CLFACVFRNTGLLDDKGMNPEVNDKLNAFET..DITYIKKFDGIIKGCYSYGN.....		141
GmelOBP46	TGCATKCLSAKDLID..PEGKLRHGNACQKFAEYGA..DENMAN.....QVVDLIHACEKSFPSND..D..EOMRTIANTNCFKTEIHLKWLAPDMILLVIGELLLEY.....		163
GmelOBP47	TGCMITCMAKHDLIVITEDRMHENAHAFAKAHA..DDALAK.....QIVIMLHDCESDNPQGE..D..FCLRALAVAKCFKTKIHALKWAPSMIEELLSEVMSDM.....		169
GmelOBP48	MGCVIMQWASKLELID..EDMKMHGNAQEFKAHAHA..DDALAK.....QIVIMLHDCESDNPQGE..D..FCLRALAVAKCFKTKIHALKWAPSMIEELLSEVMSDM.....		138
GmelOBP49	VGCALICVMSKLNLDV..TSGRLHGHSAKDFAVQHAHA..AESVAD.....KLVELIHNCEKKYDDIT..D..DCTRVMKSAACFKQDAKKEGIAP..EVAMIEAVVEKY.....		163
GmelOBP50	TGCATKCLSAKDLID..QEVNLHGHNAQEFKAHAHA..DEAMAK.....QVVDLIHACEKSFPSND..D..EOMRTIANTNCFKTEIHLKWLAPDMILLVIGELLLEY.....		163
GmelOBP51	TGCATKCLSAKLELLS..EDGNLHGHNAQEFKAHAHA..ADDIAT..QMISIIHACEKSGEGIE..D..DOWRILEVAKCFRVIHIEHWEKPDVIVTEVITEI.....		162
GmelOBP52	TGCATKCLSSKLDLVD..PBGNLHGHNAHEFAKTHGA..DDAMAQ.....KIVVDLIHNCCKSTPDNA..D..ICLILILEVAKCFKVEIHLKWLAPDMILLVIGELLLEY.....		174
Consensus			

图 5 大蜡螟 35 个气味结合蛋白的序列比对

Fig. 5 Sequence alignment of thirty-five odorant binding proteins (OBPs) of *Galleria mellonella*

阴影部分 C 表示 6 个保守的半胱氨酸残基。Shaded part C indicates six conserved cysteine residues.

OBPs 基因、36 个 CSPs 基因、80 个 ORs 基因、56 个 IRs 基因和 2 个 SNMPs 基因。OBPs 被认为是参与嗅觉信号传导过程的第一类蛋白质 (Cao et al., 2014), 大多数昆虫 OBPs 都属于 Classic

OBPs 家族, 具有 6 个保守的 Cys 位点, OBPs 保守的氨基酸基序是形成功能的重要元素 (Vogt and Riddiford, 1981)。本研究通过序列比对及基因注释发现, 大蜡螟 35 个全长 OBPs

GmelCSP1	...MIANKYLLVLACAFG..VFAEDS.....KYTDKFDINIQEILGNRLIMNYHECIMSRRGKCTNEGKELKANMO	68
GmelCSP2	MKLLLVSVVVVASV.YAKPGS.....TYTDKWDHINVDIEILESQRLLRGYVDCILLDKGRCTPDGKALKETIP	67
GmelCSP3	MKFTTII.CIMSLAAVVMAYPR..E.....TYTDREDFNVNLEILENRRLLVPYIKCILDQGGCTSPDGKELKLIHIK	67
GmelCSP4	MNNFEITMKTAFAFALCVLVAVAVCAAD.....TYNPQFDNENAEVAGNIRLLKNYKGCFLDQGGCTAEGSDFKRVIP	73
GmelCSP5	MKAIVCVCFIAMVAVALARPNDDE.....KYTDKYDNVNLEILLVNDRLRGYPYKCMLESFCAPDGKELKSHIK	70
GmelCSP6	MKLLLVIALACMAAVAYGKPAT.....KYTDKWDNINVDIEILESRLMKGYVDCILLDRGCTPDGKALKDITLP	68
GmelCSP7	MNLILFTVVATIG.YATAQSE.....QYTDYDSLRIIDDLANKRLTAAYTKCMLEKGCCTAEGRELKSHIT	66
GmelCSP8	MKIFLLSFILLYLTGKHFSEES.....KYTNKYDGVNLEILLTNERLLTGYIKCLLEGGCTPDGKELKNNLP	68
GmelCSP9	MRSFIILCCLVAAVFAADK.....YNSKYDNFDVDTLITNDRLLKAYINCFILDKGRCTPEGSDFKKTLP	64
GmelCSP10	MKLIITVL.ALVAALGRPDE.....TYEGKEFNINVDIEISNPTLLKNYINCLLGKGSCTPDGKQIKGLIP	65
GmelCSP11	MKS.FCIFVALVVCVSAE.....QYSAENDDFDIDKAVEDDATTLSMVDCTFDKDSCTPLAALKEIMP	64
GmelCSP12	MQNLLLVLIICGIMMNLPTA.....RAVPQMIDAQLERTLTDRTIMORHLKCALGEGFCDLVGRRLITLAP	66
GmelCSP13	MQ.LITQAMVICAVLYACAAQ.....NDRPQVTDALDDALNDKRFIQRLKCALGEGFCDPIGKRLKITLAP	65
GmelCSP14	MNVVILSAVLAMAAFAAADS.....QYTDYDELINVEQIMANRRLLVPMNVCVLDKGCCTAEGKELKAHIQ	68
GmelCSP16	MKFTIILFCTVLAVALVGE.....DILKEEDALFVQDETAFTQLYACIMDKGCEGKMATYKDMIP	61
GmelCSP17	MRAFFILAIWVA.....VYAVDYAAILQDETRLKAADVCLLDGRGCV.EVQELKDKFP	53
GmelCSP18	MKITILVLCALVAVVCRPEE.....TYSTQYDNFDEAHLAGNVRLLQNYAKCFLDKGCTAEGTDFKKTIP	66
GmelCSP19	MKLLLVSVVVVASV.YAKPGS.....TYTDKWDHINVDIEILESQRLLRGYVDCILLDKGRCTPDGKALKETIP	67
GmelCSP20	MK.VVFLVCVLAADVVAHPHE.S.....HYTDKWDNIDLEILLNNKILASVVKCCLDQGGCTPDGKELKSHIK	67
GmelCSP21	MKSFTVLCFLGLAAVALAKPNG.S.....TYTDYDNVNLEILLGNRRLLTPYIKCILEEGKCTADGKELKSHIR	69
GmelCSP23	MKAIVCICFVGLIAVALARPNE.E.....KYTTKYDDVNDEIISNERLRKPYVDCVLDHGRCAPDAKELKSHIQ	69
GmelCSP25	MKLAIVLCLVAALVADDK.....YKDISDDFDITEVLDNERLINSYAKCLLNKGCCTPEVKQVQDKLP	64
GmelCSP26	MKSATALLVLTIAFAAAQNE.....QYTDKFDNVNLEILLNNKRLVPYIKCALDQGGCTADGKELKSHIK	67
GmelCSP27	MK..FWILLAFATAGAR.....LYSSKDNMDIDTLVRNPRYMKSVGCYLDNRNSCTTKKTATILKQAIP	63
GmelCSP28	MK..FVIALSFLLALAAAE.....FYSNVNKFEDVALAKNEPELKTIFIGCFLDATPCTDVPAYTIEDIP	63
GmelCSP29	MKIFLLSFILLYLTGKHFSEES.....KYTNKYDGVNLEILLTNERLLTGYIKCLLEGGCTPDGKELKNNLP	68
GmelCSP30	MKLLVISMILLAVATLALAEES.....KYTDYDNVNLEILLNNRRLLVAVYKCIETGRCCTAEGKELKSHIK	67
GmelCSP31	MRVVFLELCVSVYVVMQNDMANMPKYDQRYDYLVDALFNKRLVRNVDCLISASGRCTPEGKALKRLIP	73
GmelCSP32	MKSILVLCILVAAVSCRPE.....SYDTRYDNFDEALVGNVRLITAYGHCFLNGGCTPEGSADFKKTIP	65
GmelCSP33	MLSLYLSTIFILV...HVCYEE.....EYTNKYDGVDLLEEILANERLLSGYVDCILLDLGCTPDGKELKNNLP	65
GmelCSP34	MKTICIVLCILAFVAVVGRPG..G.....QYTDYDNINLEILLGNRRLLVPYIHCVLDEGRCTPDGRELKSHIK	68
GmelCSP35	MKWLVLCVAVV..AYAQE.....KYTDKYDNVNDEIMNNARLINGYASCLLDKGCCTPEGTTELKAMI	64
GmelCSP36	MKYLLVLCIAFFV..AFAGE.....QYSDKYDNLVDIEIMGNRI FAGYSDCILLDKGCTAEGKELKDHQ	64
Consensus		
GmelCSP1	EAIETGCAKCTEKQKGCYEVIEYLIKNDLMKELCEKYDPEGKWRKKYEDRAR.ANGIVIPDN.....	132
GmelCSP2	DALEFDCKCTAKQKTSSDKVIRFLVNRPELWELKLVRYDPPDSIYEQRYKDKLNALSE.....	126
GmelCSP3	EALSDNCAKCTEKQKSGTRKVLTHIISNEKDYWNLKAKYDPPNNEFKKYEKEYGLV.....	124
GmelCSP4	EALQITGCKCTPKQVRLIRITVNAFKTKLPDIWEALVQKHDPQGIYKASFDEFLSSOD.....	131
GmelCSP5	EALETHCAKCNKQEDGVRKVIKYLINQKDEWHLVAKYDPPNNSKKYETELKEAEA.....	129
GmelCSP6	DALENESCKCTEKQKSGSKVLKHLINKRPDWFKEALVAKYDPPNNSKQARYKDKTESVKGKA.....	129
GmelCSP7	EALQNGCKCTQKQLAGVKKVILKHLINKENAYWVELTDMYDPPKKEYRHKYERELKSL.....	123
GmelCSP8	DAIQNDCKCTINRQQGADQVMGYIIDHRPDDWNKLEEK.....	107
GmelCSP9	EAVETVCAKCTEKQKLNIRKVIKATQKHPKQWELVKNNDPSGKHRADFDFKFIQGS.....	121
GmelCSP10	EGINTSCAKCTDKQKELVAKSILAAKEKFLKEYNELVAMHDPGKHITVELEAFVTKYSS.....	124
GmelCSP11	DAIERGCKCTDAQKHMFKVYLDALKRFPANENYRQKFDPPNNSKYFATLEAAVSKY.....	121
GmelCSP12	LVLRGSCPCQSVQETIQIRRTLAFTQRYNPEWESRIVRYG.....	107
GmelCSP13	LVLRGACPCQCTPQETKQIORTLSYVQRNPPQWAKIVRYAG.....	107
GmelCSP14	DAMQTGCKCTEKQKQKARKAVKQIKENEKDFWEQIKSKYDPCNKEKETYEAFIAAEN.....	126
GmelCSP16	KLLKINCESCSPKDKELFEKFSKIAAEKHPOEFKELMDKLTPKAQ.....	106
GmelCSP17	QLVATDGGCTTPBQKIRFEIAKKI IQDNYPADFEALKAKYLPSS.....	97
GmelCSP18	DAVKTKCAKCNPKQRELIRIVVKALQELPTYWDELVKKEDPEGLYKDDFNATKATD.....	124
GmelCSP19	DALEFDCKCTAKQKTSSDKVIRFLVNRPELWELKLVRYDPPDSIYEQRYKDKLNALSE.....	126
GmelCSP20	EALENRCKCTPAQKDGTRKVLTHLINHEPEMANQLCEKYDAEGKYRMYEYKSVKH.....	126
GmelCSP21	EALQNCCKCTDAQKSGTRRVLGHIINNEEESWNLKAKYDPPESKYTVKYELELRKLKQ.....	128
GmelCSP23	EALENNCCKCNEKQCGAVRKVIKYLINKEDEWQLVAKYDPPENKYSKKYHELRDVK.....	128
GmelCSP25	EALETRCCKCTDKQKQMGKQAREVKKHHPGLWKDLVAMYDPEGKHQAQFDFLASQ.....	121
GmelCSP26	EALENCAKCTEAQKDGTRRVLGHLINNEPQYWEELINRYDKDRKYVEKYESELKTIQ.....	126
GmelCSP27	DIVQTACAKCTHTQKYLIRRYLEELKKNLPSDYQAFRRKYDPEGITTFDALESATANN.....	120
GmelCSP28	DAVRTICARCSPIQKVLKFLINHAQAQLPEETAFAFNKKYDPEGKYPALIAVLSA.....	119
GmelCSP29	DAIQNDCKCTINRQQGADQVMGYIIDHRPDDWNKLEEKYKSDGSYKRYLEKKQSVKN.....	127
GmelCSP30	DGMQTACSKCTQWQKQARKVVKHIRENESEYWOQMLTKYDPPNGEFKSIYEPFLESDD.....	125
GmelCSP31	EALRTKVRCTERQKRTAVKVIKRLKYEYPEEWAELSSRWDPDTPGDFTRYFEFL.....	127
GmelCSP32	DALRTGCKCSPKQRLIRVVVQGFQNKTPALWQDLVKKQDPNGQYKEIFTRFLNGRD.....	123
GmelCSP33	DAIENDCKCTKQKREGADQVMHYIIDHRPDDWLKLEEKYKSDGSYKRYKIASQPEDNKSAEKPEDNTEKSEENK.....	142
GmelCSP34	EALENDCAKCTENQRNGSRVVISYLINHEQEYWNKLAAYDPTGKYKSKYENYRAAQ.....	127
GmelCSP35	EAIETGCAKCTEKQKTAAKI IKKILATEHLDIWEKFTAKYDPTGKWQNEYRERAK.ENGFEPL.....	126
GmelCSP36	EALETGCKSNKQKDTIVRVTKKILISDHLVDWNEVSAYDPTGKWRNEYRERAKAEYGDVPAK.....	129
Consensus		

图 6 大蜡螟 33 个化学感受蛋白的序列比对

Fig. 6 Sequence alignment of thirty-three chemosensory proteins (CSPs) of *Galleria mellonella*

阴影部分 C 表示 4 个保守的半胱氨酸残基。Shaded part C indicates four conserved cysteine residues.

中仅有 24 个含有 6 个 Cys, 表明大蜡螟 OBPs 种类和功能的多样性。目前, 对于 OBPs 功能的假说和模式很多, 然而, 很少有实验能明确证实

OBPs 是否以及如何影响昆虫的生理和行为反应, 对于大蜡螟 OBPs 的功能还有待进一步分析验证。CSPs 在进化上高度保守, 不同物种间序

Gme1SNMP1	MKIPGHKLLAASIGVTAFSILFGWYLFVILRSQIRKEMALSKKTDVRKMEKIPFALDFKVLNFYTNAAEVQKGAIP	80
Gme1SNMP2	MLGKRSRLFFGISVAGLIISIILAVWGFPKIRSQIQNIQLNDSMMYEKWRVLPMLFKITVFNVTNTEENNGKPK	80
Consensus	m s g si fp i rsq k l w p l fk fn tn ee g p	
Gme1SNMP1	IVKEVGPYYFEWKEKVEIEDHEEDDTITYKKLDTFYFRPDLSGRDLTGEEVITMHPHYILSMLTFVARDKPSMINMVVK	160
Gme1SNMP2	KLQEIQPFVYKEYRERIVLG.YGNDTIKYMLKKTIFDAEESAP.FSEDDMTVINFSYMAALITINDMPALVSTLNM	158
Consensus	e qp e e dti y tf f s t l p	
Gme1SNMP1	AVNGIFDEP..PDVFLKIKVLDLLFRGLTINCARTFAPKAVCTALKKEG.ADLIFEPNNQYKFSFGKRNATADPHV	237
Gme1SNMP2	GFQEMYSQYGLDDPFTTVKARDLFFDGLFNCRGNHSSIALICSQISTTSPPTMRPTEDGSGFYFSMFSLNRTAS.GPY	237
Consensus	d f k dl f gl nc c e fs f n ta	
Gme1SNMP1	TVQRGIRNMDVQGVIAIDGKPKQMKWRDS..CNEYQGTGTIFPPFLTEND..RIMSFSGDLCRSFKPWYVKKTSYKSI	313
Gme1SNMP2	SMVRGIENVKELGHIVSYKEKTFMNNWSDPYOGQLNGSDSIYPPIDENNVPKRLYTFQPDICRSIYVDFMSTKINIFNM	317
Consensus	r gi nv g k mn w c g d i pp n rl f d crs kt	
Gme1SNMP1	PINHYTANIGDLAND...PEINCFCKA.....PDNCPKGLMDTKCIGAPMYASMPHFLDSDPEILLKNVVG.LHPNAEA	384
Gme1SNMP2	SAYYYELSESASFASKSANPNKCYCKKNWISGNDGCLLMGMNLFPOQQAFAIASLPHFYLAEEELLEFFDGGIKPKDEK	397
Consensus	y a p c ck d c g m c ap as phf ell g p e	
Gme1SNMP1	HSTAIDFESISGTFMVAQRVQFNIQLKHKDIPLCCKDLPTFAPLEWIEEGLAINKTVMKLRHQLFLPMLRVNVLRWV	464
Gme1SNMP2	HKSIMLEPNTIGAVLKGVRQLQFVVELR.QMKIPQLMKVPTGLFPMWLWIEESAEPVPSIQEELQQLKLG.YANTARW	475
Consensus	h i e g qr qfn l kip p p wiee l l l n rvw	
Gme1SNMP1	LLVFASMSILGCILFHEFRERILKFAVITTKTINTKINPEDALKSNNVVEQTHQEAQVD	524
Gme1SNMP2	LLALALIAAVVSTIVVRATLPGWGRNSVFLRPITNVNGIDLNGK.....	521
Consensus	l a r n	

图 7 大蜡螟两个神经元膜蛋白的序列比对

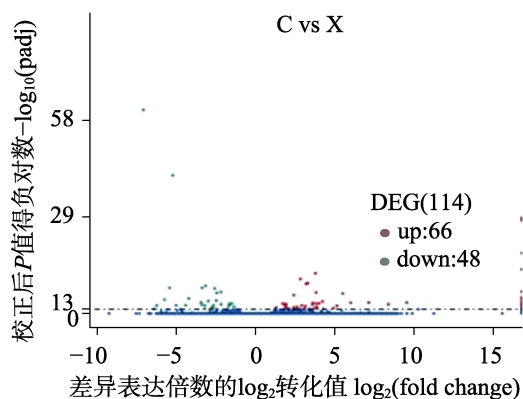
Fig. 7 Sequence alignment of two sensory neuron membrane proteins (SNMPs) of *Galleria mellonella*

图 8 雌、雄大蜡螟触角差异表达基因分析

Fig. 8 Analysis of differential expression genes of female and male *Galleria mellonella*

C 代表雌性大蜡螟, X 代表雄性大蜡螟。

C represents female *Galleria mellonella*,X represents male *Galleria mellonella*.

列相似性较高。本研究在大蜡螟触角转录组数据中鉴定出 36 个 CSPs, 其中 33 个具有完整 ORF, 序列中均含有 4 个保守 Cys, 并且存在明显的高度保守结构域, 说明这些 CSPs 在不同昆虫中具有相似的功能。SNMPs 是一种膜蛋白, 在果蝇中已被证实在性信息素的感受中必不可少 (Benton *et al.*, 2007)。最新研究表明, SNMP1 参与了嗅觉神经元对性信息素的识别过程, 而 SNMP2 没

有参与 (Cao *et al.*, 2014)。本研究从大蜡螟触角转录组中鉴定获得 2 个 SNMPs, 即 SNMP1 和 SNMP2, 它们的序列一致性为 25.88%, SNMPs 在大蜡螟中的发现可以为蛾科昆虫 SNMPs 的研究打开新的视野。

自然界中的昆虫为了感知复杂的外界环境, 进化出庞大的受体蛋白群 (Hansson and Stensmyr, 2011), 如 ORs 和 IRs, 它们确保嗅觉神经元反应的特异性。昆虫气味受体分为两大类: 即非典型型气味受体 Orco 和传统气味受体 ORs, Orco 可加速 ORs 与气味分子间的相互作用, 并能辅助 ORs 正确定位到嗅觉神经元树突膜上 (Cao *et al.*, 2014)。不同昆虫间 ORs 同源性较低, 在昆虫中的数量差异较大 (胡颖颖等, 2013), 本文从大蜡螟触角中鉴定出 80 个 ORs, 其数量相对其它昆虫偏少, 这可能与大蜡螟的栖境范围窄、生存环境的信息化合物单一有关。IRs 是近年来发现的一种离子型谷氨酸受体 (iGluR) 类蛋白家族, 已证实参与化学感受 (Cao *et al.*, 2014)。IRs 比 ORs 更古老, 与 ORs 不同的是, 在一个 IR 表达神经元中, 两个或三个 IRs 总是与离子型共受体 *IR25a* 和 *IR8a* 中的一个或两个

共表达 (Cao *et al.*, 2014)。已有对昆虫 IRs 的研究表明, 嗅觉神经上的 IRs 起到了检测胺类、酸类的作用 (Kain *et al.*, 2013); 最近研究报道, 有些在味觉神经上表达的少数 IRs 参与了信息素感知 (Koh *et al.*, 2014)。在大蜡螟转录组数据中, 我们筛选出 56 个 IRs (包括 *IR25a* 和 *IR8a*), 这些 IRs 在大蜡螟嗅觉识别过程中的作用还有待后续深入研究。

通过比较雌、雄大蜡螟触角转录组分析发现 114 个差异表达基因, 包括 5 个嗅觉相关基因 (*OBP8*、*OBP17*、*CSP3*、*CSP34* 和 *OR15*), 这些嗅觉差异表达基因在大蜡螟性信息素识别过程中可能发挥着重要作用, 可以作为潜在的干扰大蜡螟嗅觉识别的靶标基因。

参考文献 (References)

- Anders S, Huber W, 2010. Differential expression analysis for sequence count data. *Genome Biology*, 11(10): R106.
- Becher PG, Flick G, Rozpedowska E, Schmidt A, Hagman A, Lebreton S, Larsson MC, Hansson BS, Piškur J, Witzgall P, Bengtsson M, 2012. Yeast, not fruit volatiles mediate *Drosophila melanogaster* attraction, oviposition and development. *Functional Ecology*, 26(4): 822–828.
- Benton R, Vannice KS, Gomez-Diaz C, Vossall LB, 2009. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell*, 136(1): 149–162.
- Benton R, Vannice KS, Vossall LB, 2007. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature*, 450(7167): 289–293.
- Cao D, Liu Y, Wei J, Liao X, Walker WB, Li J, Wang G, 2014. Identification of candidate olfactory genes in *Chilo suppressalis* by antennal transcriptome analysis. *International Journal of Biological Sciences*, 10(8): 846–860.
- Clyne PJ, Warr CG, Freeman MR, Lessing D, Kim J, Carlson JR, 1999. A novel family of divergent seven transmembrane proteins: Candidate odorant receptors in *Drosophila*. *Neuron*, 22(2): 327–338.
- Clyne PJ, Warr CG, Carlson JR, 2000. Candidate taste receptors in *Drosophila*. *Science*, 287(5459): 1830–1834.
- Ellis JD, Graham JR, Mortensen A, 2013. Standard methods for wax moth research. *Journal of Apicultural Research*, 52(1): 1–17.
- 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, di 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.
- Hansson B, Stensmyr M, 2011. Evolution of insect olfaction. *Neuron*, 72(5): 698–711.
- Hu JJ, Meng X, Zhou JB, Yang LX, Liu SH, Li RZ, 2017. Transcriptome analysis of the cotton mealybug, *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae). *Acta Entomologica Sinica*, 60(1): 9–17. [胡俊杰, 孟翔, 周佳滨, 杨陆兴, 刘善海, 李润钊, 2017. 扶桑绵粉蚧转录组分析. 昆虫学报, 60(1): 9–17.]
- Hu YY, Xu SF, Abebe JW, Li W, Guo ZB, Zhou T, 2013. Research advance of olfactory proteins and olfactory mechanism in insects. *Genomics and Applied Biology*, 32 (5): 667–676. [胡颖颖, 徐书法, Abebe JW, 李薇, 国占宝, 周婷, 2013. 昆虫嗅觉相关蛋白及嗅觉识别机理研究概述. 基因组学与应用生物学, 32(5): 667–676.]
- Jin YL, Cong B, Wang LY, Dong H, 2015. An analysis of the transcriptome of *Epacromius coerulipes* (Orthoptera: Acrididae). *Acta Entomologica Sinica*, 58(8): 817–825. [金永玲, 丛斌, 王丽艳, 张海燕, 董辉, 2015. 大垫尖翅蝗转录组分析. 昆虫学报, 58(8): 817–825.]
- Kain P, Boyle SM, Tharadra SK, Guda T, Pham C, Dahanukar A, Ray A, 2013. Odour receptors and neurons for detection of DEET and new insect repellents. *Nature*, 502 (7472): 507–512.
- Kaissling KE, 2009. Olfactory perireceptor and receptor events in moths: A kinetic model revised. *Journal of Comparative Physiology a Sensory, Neural, and Behavioral Physiology*, 195 (10): 895–922.
- Koch SI, Groh K, Vogel H, Hansson BS, Kleineidam CJ, Grosse-Wilde E, 2013. Caste-specific expression patterns of immune response and chemosensory related genes in the leaf-cutting ant, *Atta vollenweideri*. *PLoS ONE*, 8(11): e81518.
- Koh TW, He Z, Gorur-shandilya S, Menuz K, Larter NK, Stewart S, Carison JR, 2014. The *Drosophila IR20a* clade of ionotropic receptors are candidate taste and pheromone receptors. *Neuron*, 83 (4): 850–865.
- Kwadha CA, Ong'Amo GO, Ndegwa PN, Raina SK, Fombong AT, 2017. The biology and control of the greater wax moth, *Galleria mellonella*. *Insects*, 8 (2): 61, 1–17.
- Leal WS, 2013. Odorant reception in insects: Roles of receptors, binding proteins, and degrading enzymes. *Annual Review Entomology*, 58(1): 373–391.
- Li B, Dewey C, 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12: 323.
- Liu C, Pitts RJ, Bohbot JD, Jones PL, Wang G, Zwiebel LJ, 2010. Distinct olfactory signaling mechanisms in the malaria vector

- mosquito *Anopheles gambiae*. *PLoS Biology*, 8 (8): e1000467.
- Liu S, Rao XJ, Li MY, Feng MF, He MZ, Li SG, 2015. Identification of candidate chemosensory genes in the antennal transcriptome of *Tenebrio molitor* (Coleoptera: Tenebrionidae). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 13: 44–51.
- Li ZQ, Zhang S, Luo JY, Wang SB, Wang CY, Lv LM, Dong SL, Cui JJ, 2015. Identification and expression pattern of candidate olfactory genes in *Chrysoperla sinica* by antennal transcriptome analysis. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 15: 28–38.
- Meng X, Hu JJ, Liu H, Ouyang GC, Guo MF, 2016. Analysis of the transcriptome and olfaction-related genes of *Conopomorpha sinensis* Bradley (Lepidoptera: Gracilariidae). *Acta Entomologica Sinica*, 59(8): 823–830. [孟翔, 胡俊杰, 刘慧 欧阳革成, 郭明昉, 2016. 荔枝蒂蛀虫转录组及嗅觉相关基因分析. *昆虫学报*, 59 (8): 823–830.]
- Pelosi P, Iovinella I, Felicioli A, Dani FR, 2014. Soluble proteins of chemical communication: an overview across arthropods. *Frontiers in Physiology*, 5: 320.
- Silbering AF, Benton R, 2010. Ionotropic and metabotropic mechanisms in chemoreception: ‘chance or design’? *EMBO Reports*, 11(3): 173–179.
- Song YQ, Dong JF, Chen QX, Hu ZJ, Sun HZ, 2017. Analysis of the antennal transcriptome and chemoreception-related genes of the bean bug, *Riptortus pedestris* (Hemiptera: Alydidae). *Acta Entomologica Sinica*, 60(10): 1120–1128. [宋月芹, 董钧锋, 陈庆霄, 胡振杰, 孙会忠, 2017. 点蜂缘蝽触角转录组及化学感受相关基因的分析. *昆虫学报*, 60(10): 1120–1128.]
- Svensson GP, Gündüz EA, Sjöberg N, Hedenström E, Lassance JM, Wang HL, Löfstedt C, Anderbrant O, 2014. Identification, synthesis, and behavioral activity of 5,11-dimethylpentacosane, a novel sex pheromone component of the greater wax moth, *Galleria mellonella* (L.). *Journal of Chemical Ecology*, 40(4): 387–395.
- Vogel H, Altincicek B, Glöckner G, Vilcinskis A, 2011. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. *BMC Genom.*, 12(1): 308.
- Vogt RG, Riddiford LM, 1981. Pheromone binding and inactivation by moth antennae. *Nature*, 293(5828): 161–163.
- Wang J, Li DZ, Min SF, Mi F, Zhou SS, Wang MQ, 2014. Analysis of chemosensory gene families in the beetle *Monochamus alternatus* and its parasitoid *Dastarcus helophoroides*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 11(9): 1–8.
- Yang F, Huang LH, Zhang AB, 2014. High-throughput transcriptome sequencing technology and its applications in Lepidoptera. *Acta Entomologica Sinica*, 57(8): 991–1000. [杨帆, 黄立华, 张爱兵, 2014. 高通量转录组测序技术及其在鳞翅目昆虫上的应用. *昆虫学报*, 57(8): 991–1000.]
- Zhou CX, Min SF, Yan-Long T, Wang MQ, 2015. Analysis of antennal transcriptome and odorant binding protein expression profiles of the recently identified parasitoid wasp, *Sclerodermus* sp. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 16: 10–19.
- Zhang J, Wang B, Dong S, Cao D, Dong J, Walker WB, Liu Y, Wang G, 2015b. 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 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 T, Coates BS, Ge X, Bai S, He K, Wang Z, 2015a. Male-and female-biased gene expression of olfactory-related genes in the antennae of asian corn borer, *Ostrinia furnacalis* (Guenee) (Lepidoptera: Crambidae). *PLoS ONE*, 10(6): e0128550.
- Zhang S, Pang BP, Zhang L, 2015c. Novel odorant-binding proteins and their expression patterns in grasshopper, *Oedaleus asiaticus*. *Biochemical and Biophysical Research Communications*, 460(2): 274–280.
- Zheng HX, Zhang YW, Zhang XH, Yang MH, 2018. Analysis of the antennal transcriptome and olfaction-related genes of *Callosobruchus chinensis* (Coleoptera: Bruchidae). *Acta Entomologica Sinica*, 61(2): 168–177. [郑海霞, 张耀文, 张仙红, 杨美红, 2018. 绿豆象触角转录组及嗅觉相关基因的分析. *昆虫学报*, 61(2): 168–177.]
- Zhao Y, Wang F, Zhang X, Zhang S, Guo S, Zhu G, Liu Q, Li M, 2016. Transcriptome and expression patterns of chemosensory genes in antennae of the parasitoid wasp *Chouioia cunea*. *PLoS ONE*, 11(2): e0148159.