

“转基因稻谷储藏适宜性及机理”特约专栏文章之二

DOI: 10.16210/j.cnki.1007-7561.2020.06.012

转 *Cry1Ab/Cry1Ac* 基因大米粉饲喂 印度谷螟后的比较转录组分析

唐培安, 陶冶心, 王康旭, 吴学友, 陈二虎✉

(南京财经大学 食品科学与工程学院,
江苏高校粮油质量安全控制及深加工重点实验室, 江苏 南京 210023)

摘要: 转 *Cry1Ab/Cry1Ac* 基因作物可以有效防治害虫。以世界性的储粮害虫印度谷螟 *Plodia interpunctella* (Hübener) 为研究对象, 借助高通量测序法研究试虫中与 Bt 毒蛋白密切相关的基因, 以期初步揭示转 Bt 稻谷的杀虫作用机制。结果表明, 在对印度谷螟非胁迫种群和转 Bt 基因大米粉胁迫种群分别进行转录组测序, 经 *De novo* 组装后共得到 37 246 个 Unigene, 其中有 23 310 个 Unigene 获得注释。随后, 通过比较分析上述两种品系试虫基因表达量, 共筛选出 34 466 条显著差异表达的基因, 其中在胁迫品系中有 15 741 条基因表达量显著上调, 另有 18 725 条基因表达量显著下调。一方面大大扩充了针对该物种的基因信息, 另一方面为转基因稻谷抗印度谷螟的相关机理研究提供了新思路。

关键词: 转基因稻谷; 印度谷螟; 转录组分析; 高通量测序法

中图分类号: TS201.4 文献标识码: A 文章编号: 1007-7561(2020)06-0104-07

网络首发时间: 2020-10-27 10:05:58

网络首发地址: <https://kns.cnki.net/kcms/detail/11.3863.TS.20201027.0937.002.html>

Comparative Transcriptome Analysis of the Indian Meal Moth Fed with Transgenic *Cry1Ab/Cry1Ac* Rice Powder

TANG Pei-an, TAO Ye-xin, WANG Kang-xu, WU Xue-you, CHEN Er-hu✉

(Collaborative Innovation Center for Modern Grain Circulation and Safety, College of Food Science and Engineering, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210023, China)

Abstract: The *Cry1Ab/Cry1Ac* transgenic crops can effectively control pests. In the present study, the world-wide distribute pest, Indian meal moth *Plodia interpunctella* (Hübener), was taken as the research object. The genes closely related to Bt toxin protein were sequenced using a high-throughput sequencing to reveal the insecticidal mechanism of Bt transgenic rice. The results showed that the *de novo* assembly generated a total of 37,246 unigenes among of which 23,310 unigenes could be annotated after the transcriptome sequencing was performed on the non stressed population and Bt transgenic rice meal stressed

收稿日期: 2020-07-22

基金项目: 粮食公益性行业科研专项 (201413007-02, 201513002-05-03); 国家重点研发计划 (2016YFD0401004-4, 2017YFD0401003-5); 江苏省六大人才高峰高层次人才项目 (NY-057); 江苏高校优势学科建设工程资助项目

作者简介: 唐培安, 男, 1981 年出生, 博士, 教授, 研究方向为粮食储藏新技术、储粮害虫防治以及转基因稻谷的储藏安全性评估。
E-mail: tangpeian@163.com.

通讯作者: 陈二虎, 男, 1987 年出生, 博士, 讲师, 研究方向为储粮害虫绿色防控。E-mail: erhuchen1104@yahoo.com.

population of Indian rice borer, respectively. Then, 34,466 significantly differentially expressed genes including 15,741 up-regulated genes and 18,725 down-regulated genes were determined by the Reads Per kb per Million reads (RPKM) method. Our database provides powerful tools for transcriptome analysis of *P. interpunctella* and further understanding of the molecular basis of Bt related mechanisms in *P. interpunctella*.

Key words: transgenic rice; *Plodia interpunctella*; transcriptome analysis; high-throughput sequencing

印度谷螟 *Plodia interpunctella* (Hübener) 是一种全球范围内广泛发生的储粮害虫, 不仅能够通过直接取食对粮食造成损失, 其生命活动过程还可以提高储粮温度, 从而引起粮堆霉变和品质劣变。苏云金芽孢杆菌 *Bacillus thuringiensis* (Bt) 是一种革兰氏阳性细菌, 能够产生具有杀虫效果的晶状蛋白 (cry), 对鳞翅目和鞘翅目害虫具有明显的防治效果^[1]。随着转基因技术的不断发展, 利用转基因进行 cry 蛋白异源表达的农作物越来越多, 目前广泛用于田间生产的有转基因棉花, 主要用于防治棉铃虫。另外, 转基因水稻也可以用于二化螟、大螟和稻纵卷叶螟等害虫的防治, 同时, 研究人员发现转基因稻谷对印度谷螟也有不错的防效^[2]。在 Bt 抗虫研究领域还存在几个亟需回答的问题, 一是 cry 蛋白的毒性机理还未有定论, 目前存在孔道假说和信号传导假说两个潜在的致死机制。另外, 研究表明对印度谷螟进行室内筛选可以使试虫种群获得 Bt 抗性, 但是相关的抗性发生机理尚未得到阐明^[3]。

高通量测序技术, 是核酸测序中最常用的工具, 其标志就是能一次并行对几十万到几百万条基因序列进行测定。在高通量测序技术的迅猛发展下, 生物学问题可以通过基因表达差异分析、突变分析等手段进行解决。在昆虫学研究领域, 高通量测序可以用于进化发育、生理生化以及毒理学等方面的研究。但是, 印度谷螟应用高通量测序技术的报道很少, 相关基因数据有限。本研究通过对长期饲喂转基因大米粉和常规大米粉的印度谷螟种群进行转录组高通量测序 (Illumina HiSeq™ 2000), 组装后分别得到了 41 597 和 42 306 条潜在基因序列, 注释其中的关键基因, 最后比较分析差异表达基因。本研究一方面大大扩充了该物种的基因信息, 另一方面也为转基因稻谷对印度谷螟乃至其他害虫的抗虫机理提供新的见解。

1 材料与方法

1.1 供试印度谷螟

本试验所用的印度谷螟来源于江苏省吴江市, 后经实验室多代饲养。其中, 非胁迫品系 (SS) 在实验室条件下饲养 20 代以上 (30 ± 0.5) °C、相对湿度 70%~80%, 无光照), 饲料为正常米粉 (不含 Bt 毒素)。从非胁迫品系中随机选择部分试虫用含有 2% 转基因米粉的混合饲料饲养, 筛选培养 5 个月, 视该种群为胁迫种群 (RS)。

试虫非胁迫品系 (SS) 饲料为: 明恢 63 (普通稻谷) 大米粉; 试虫胁迫品系 (RS) 饲料为: 明恢 63 (普通稻谷) 和华恢 1 号 (*Cry1Ab/Cry1Ac* 转基因稻谷) 大米粉混合物。上述两种稻谷均来自华中农业大学作物遗传改良国家重点实验室。

1.2 RNA 制备以及 cDNA 文库构建

以 3 龄幼虫作为供试对象, 提取 RNA 后送测。总 RNA 提取按照 RNeasy Plus Universal Mini Kit (Qiagen) 试剂盒说明书进行, 用 DNase I (Takara) 去除基因组杂质, 测定 RNA 浓度, 并通过 1.5% 琼脂糖凝胶电泳检测 RNA 完整性。然后, 使用华大基因 HiSeq™ 2000 (Illumina) 构建 cDNA 文库并进行测序。

1.3 生物信息学分析

筛选剔除低质量序列后以高质量的 reads 进行生物信息学分析。首先, Trinity 组合某一重叠长度的 reads 以形成更长的 contigs 片段^[4]; 然后, 将 reads 进行末端读数配对, 确定 contigs 间的差异情况; 最后, 利用 Trinity 将 contigs 连接成没有 Ns 的 unigene 序列。将组装的 unigenes 与 NR、Swiss-Prot、KEGG 和 COG 数据库进行比对 (设置临界 E 值 $< 10^{-5}$)^[5]。Unigene 序列结果全部提交至 NCBI SRA (登录号分别为 SRP060836, SRP060680)。

1.4 差异 Unigenes 分析

利用 RPKM 方法（每 kb 每百万读数）来鉴定两个样品（RS 和 SS）之间的基因表达差异^[6]，同时使用 FDR 方法来确定 P 值的阈值^[7]。基因显著差异表达的阈值为“FDR ≤ 0.001 和 log₂ 比 ≥ 1 的绝对值”。最后以 GO 富集和 KEGG 代谢通路深入分析转录组数据^[8]。

1.5 显著差异基因的 GO 分析

将 DEG 对应 GO 数据库，计算 GO 项基因数，分析 GO 功能差异，计算公式如下^[10]：

$$P = 1 - \frac{\sum_{i=0}^{m-1} \binom{m}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

N: KEGG 基因数；n: N 中 DEG 数；M: 特定途径基因数；m: M 中的 DEG 数。

1.6 显著差异基因的代谢通路分析

通过公共代谢路径相关数据库 KEGG 本体富集分析进一步了解显著差异基因的生物学功能^[9]，计算公式同上^[10]。

2 结果与分析

2.1 转录组数据 De novo 组装

我们从胁迫种群（RS）和非胁迫种群（SS）中均获得超过 51（百万）的 reads 数，且 Q20 > 98%，测序质量较好，深度和覆盖率结果如图 1 所示。之后，通过 Trinity 软件组装 reads，共分别获得 54 648 和 54 647 个 contigs，平均长度分别均在 600 bp 以上。之后再进行 unigenes 组装，RS-Unigenes 和 SS-unigenes 个数分别为 41 597 和 42 306 个，平均长度均超过 1 000 bp，另外还获得 37 246 个 All-unigenes（表 1）。

表 1 测序产量统计

	RS	SS	RS&SS
Total clean reads	52 038 128	51 170 436	-
Q20 percentage/%	98.28	98.31	-
GC percentage/%	44.79	44.93	-
Total number of contigs	54 648	54 647	-
Mean length of contigs (bp)	610	621	-
Total number of unigenes	41 597	42 306	-
Mean length of unigenes (bp)	1 037	1 093	-
Total number of all-unigenes	-	-	37 246
Mean length of all-unigenes (bp)	-	-	1 334

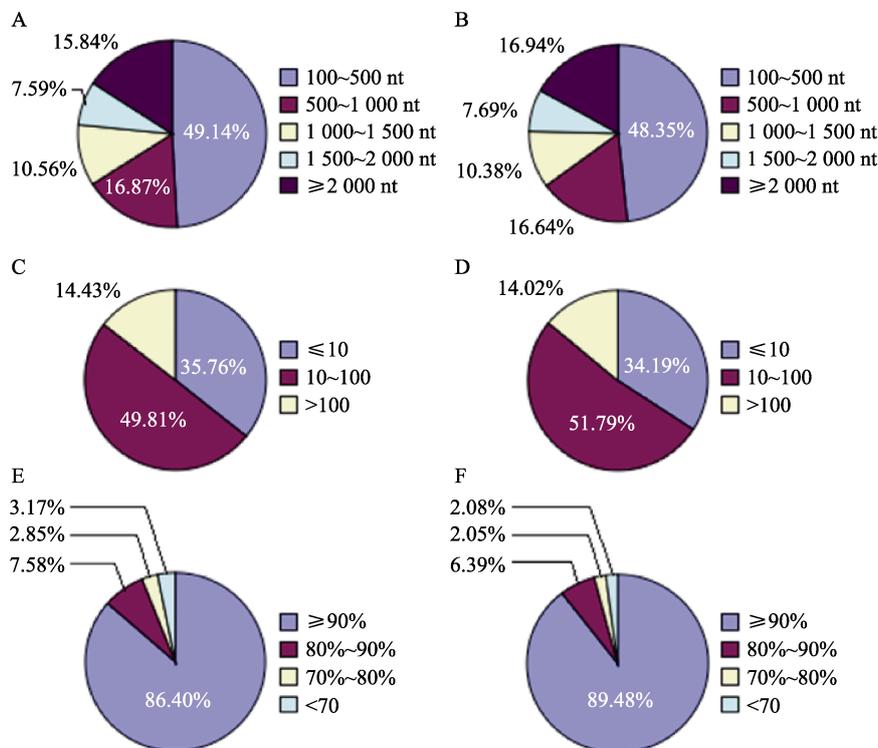


图 1 不同种群 Unigene 的分布情况

注：A: 胁迫品系 Unigene 长度分布，B: 非胁迫品系 Unigene 长度分布，C: 胁迫品系 Unigene 深度，D: 非胁迫品系 Unigene 深度分布，E: 胁迫品系 Unigene 覆盖度分布，F: 敏感品系 Unigene 覆盖度分布。

2.2 Unigene 功能注释与 COG 分类

在 37 246 个 unigenes 中有 23 310 个 unigenes 可以被注释, 其中有 22 211 个 unigenes 在 NR 数据库中可以被注释。通过 NR 数据库中的 E 值分析, 20.5% 的注释序列 E 值小于 $1.0E-100$, 15.5% 注释序列 E 值范围在 $1.0E-100$ 到 $1.0E-60$ 之间, 表明二者之间具有强同源性 (表 2)。序列相似性分布表明, 26.1% 的序列之间具有较高的同源性, 而另外 73.9% 的序列同源性介于 17%~80%。另外, 与印度谷螟相似度最高的物种是帝王蝶

表 2 注释结果统计

Sequence File	NR	NT	Swiss-Prot	KEGG	COG	GO	ALL
All-Unigene	22 211	14 161	17 755	15 750	9 348	10 694	23 310

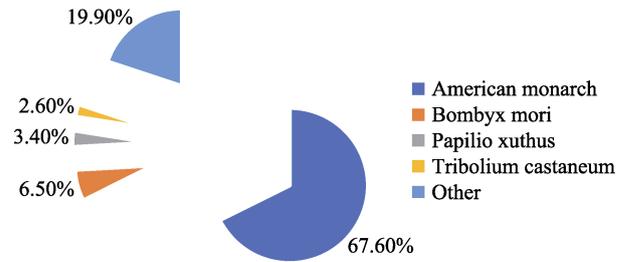


图 2 Nr 注释的物种分布

(*American monarch*, 67.60%), 其次是家蚕 (*Bombyx mori*, 6.5%) 和赤拟谷盗 (*Tribolium castaneum*, 3.4%) (图 2)。

2.3 注释基因功能研究

10 694 个序列可以注释到 GO 的 58 个功能组, 主要包括: 生物过程, 细胞组分和分子功能 (图 3)。同时, 细胞加工 (cellular process) (6 665,

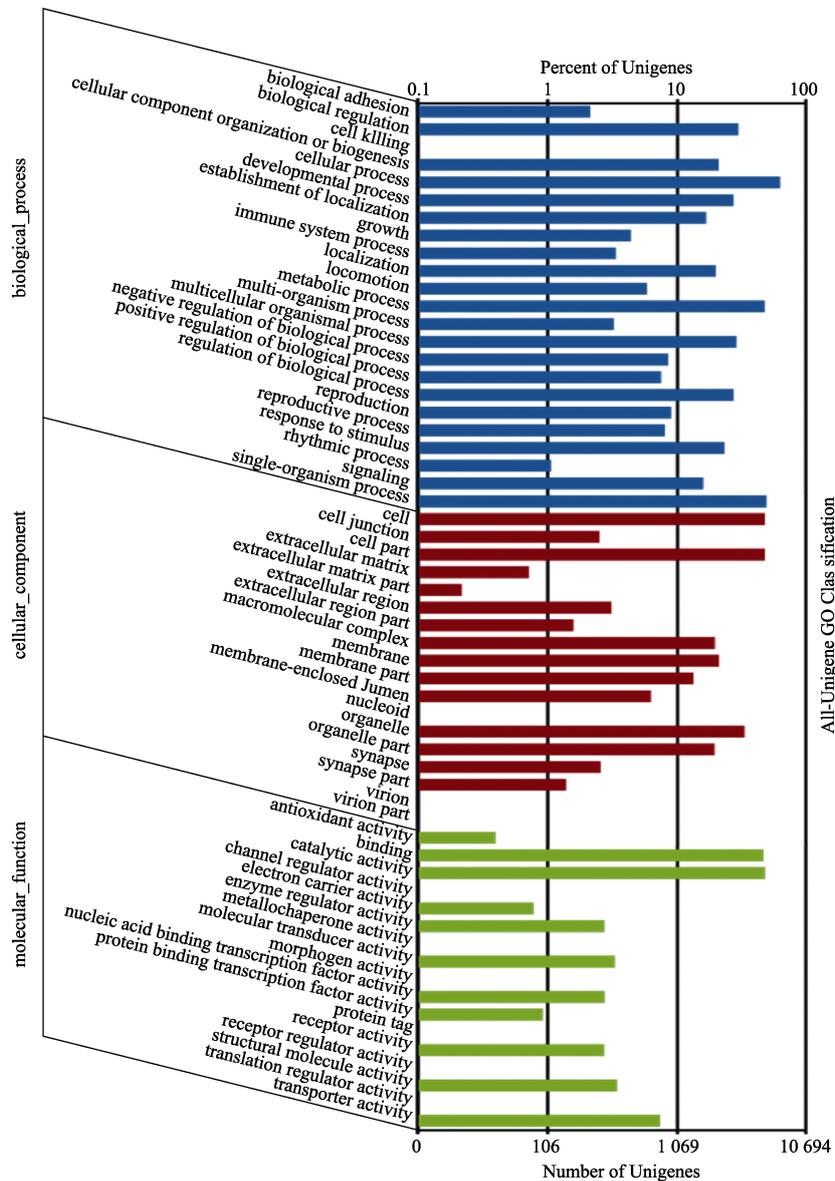


图 3 Unigene 的 GO 分类图

62.32%)、细胞组分 (cell) (5 063, 47.34%) 和催化活性 (catalytic activity) (5 101, 47.7%) 为主。

共 9 348 条 unigenes 注释到 COG 并分为 25 个功能类别 (图 4)。其中, 最大的聚类组是“一般功能基因” (3 662, 39.17%), 其次是“复制、重组

和修复基因” (1 575, 16.85%) 以及“翻译、核糖体结构和生物发生” (1 519, 16.25%)。而其它类别基因较少, 如“细胞外结构” (92, 0.98%), “RNA 加工和修饰” (88, 9.41%) 和“核结构” (5, 0.05%)。

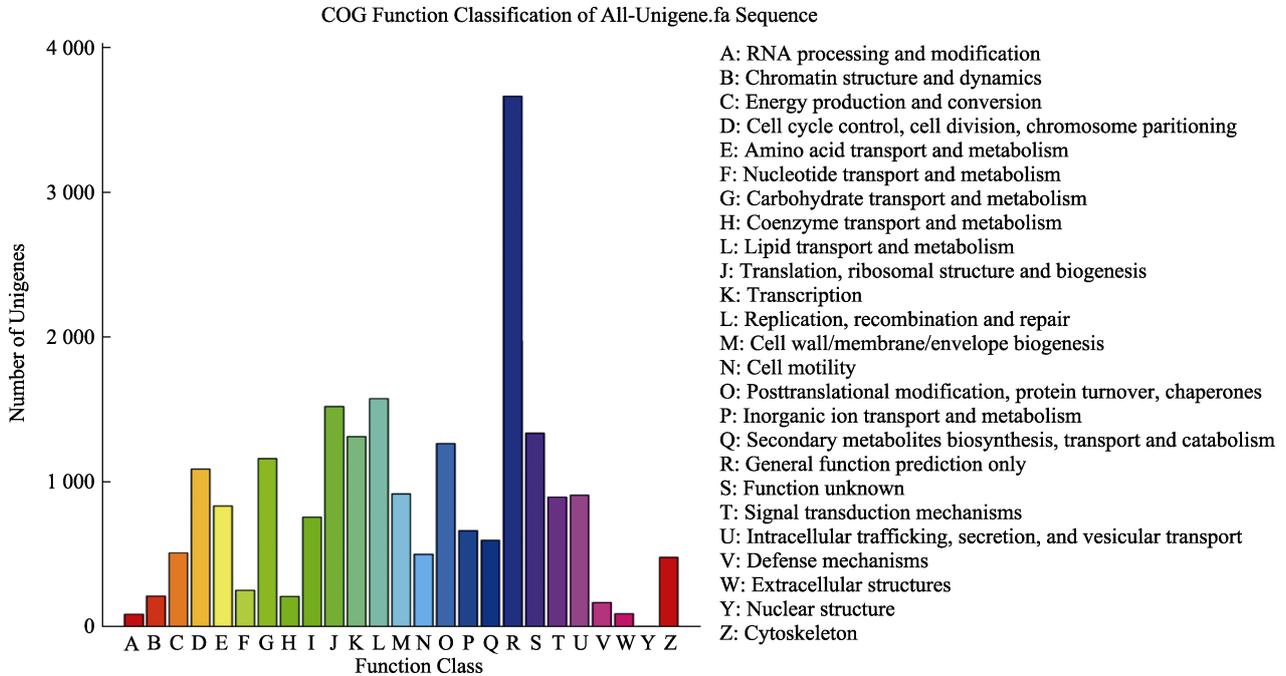


图 4 COG 功能注释

KEGG 注释结果显示, 15 750 条 unigenes 被注释到 258 个 KEGG 通路。其中, 以新陈代谢通路、肌动蛋白细胞骨架调节通路为主, 另外有少数的 unigene 被注释到了生物素代谢和赖氨酸生物合成通路。

2.4 不同品系印度谷螟基因表达分析

对印度谷螟胁迫品系与非胁迫品系的差异基因进行比较分析发现, 有 15 741 条基因上调, 18 725 条基因下调 (图 5)。进一步利用阈值 ($FDR \leq 0.01$) 和 $\log_2\text{Ratio} (\geq 1)$ 分析发现, 两个种群中差异极显著的基因共有 10 224 条, 其中 4 520 条基因上调, 5 704 条基因下调; 除此之外, 有 1 245

条基因在胁迫品系中表达, 同时, 有 503 条基因只在非胁迫种群中表达 (图 6)。

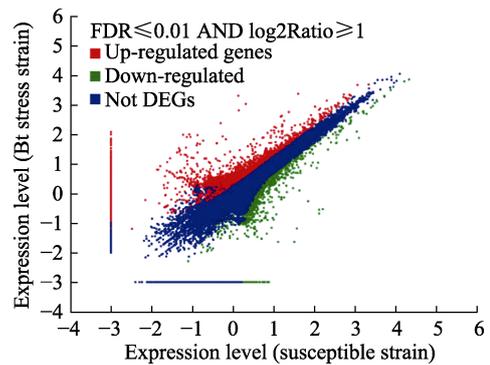


图 6 印度谷螟不同品系显著差异表达基因

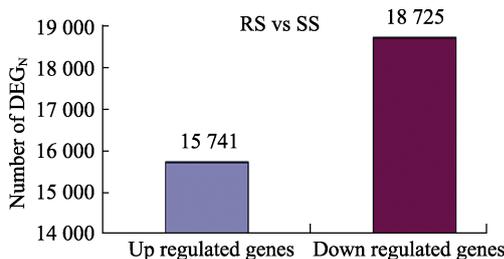


图 5 印度谷螟不同品系差异表达的基因

2.5 印度谷螟解毒和 Bt 毒素受体相关基因分析

结果表明, 有 87 条羧酸酯酶和谷胱甘肽-S-转移酶的相关基因在胁迫和非胁迫种群中具有表达差异, 其中 60.91% 的基因在胁迫品系中表达量上调 (表 3)。

另外, 我们还对不同种群试虫的 Bt 毒素受体相关的基因进行了比较分析。结果表明, 有 78 条

表 3 印度谷螟胁迫和敏感品系中差异显著的羧酸酯酶和谷胱甘肽-S-转移酶基因

Unigene ID	Gene length	Log2Ratio	Annotation
CL2151.Contig1_All	1 961	-2.038 732	carboxylesterase CarE-11 precursor [<i>Bombyx mori</i>]
CL2624.Contig2_All	1 886	-2.155 650	carboxylesterase ae2 [<i>B. mori</i>]
CL3154.Contig2_All	1 318	2.005 294	Carboxylesterase [<i>H. armigera</i>]
CL3736.Contig2_All	1 699	2.074 580	carboxylesterase CXE26 [<i>S. littoralis</i>]
Unigene1347_All	1 797	-3.236 728	carboxylesterase [<i>B. mori</i>]
Unigene3304_All	1 956	-2.193 912	PREDICTED: venom carboxylesterase-6-like [<i>Megachile rotundata</i>]
CL174.Contig2_All	802	-7.791 160	carboxylesterase [<i>Melitta cinxia</i>]
Unigene16047_All	446	1.520 618	glutathione-S-transferase [<i>Bemisia tabaci</i>]

注释为氨肽酶-N, 碱性磷酸酶, 钙黏蛋白, 糖脂类的显著差异基因在两个品系中被发现, 其中有 55.13% 的基因在胁迫品系中上调。值得注意的是, 所有注释为 APN 的基因在胁迫品系中都显著上调。之后, 我们对筛选出来的 APN 基因进行了系统发育树构建, 6 个 APN 基因能够划到 4 个相关家族, 与家蚕、小菜蛾、烟草天蛾等鳞翅目昆虫的 APN 蛋白序列相似度较高 (图 7)。

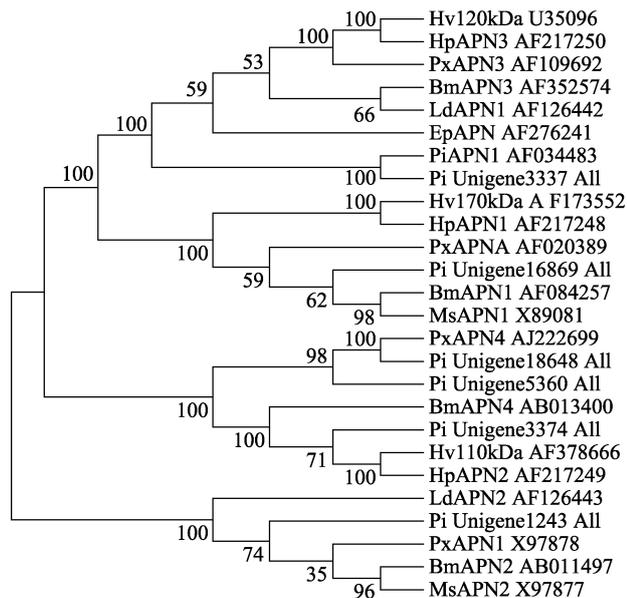


图 7 印度谷螟和其他鳞翅目昆虫中 APN 基因系统发育分析
 注: Bm: 家蚕; Hv: 烟芽夜蛾; Px: 小菜蛾; Hp: 棉铃虫;
 Ms: 烟草天蛾; Ld: 舞毒蛾; Ep: 苹淡褐卷蛾; Pi: 印度谷螟。

3 讨论

印度谷螟危害严重且全球性分布, 是转基因水稻的主要靶标害虫之一。过去几年, 和 Bt 抗性机制相关的解毒代谢机制以及毒素受体被很多人研究^[11-12]。然而印度谷螟对 Bt 毒素的抗性机理尚不清晰。本研究中, 我们用二代测序技术对印度谷螟 3 龄敏感及经转基因大米粉饲喂过的胁迫品

系进行转录组比较分析, 以研究试虫对 Bt 的相关毒理学机制。通过测序, 总计产出超过 100 万 nt 数据, 这些转录数据经过过滤被组装为 Unigenes, 并注释到 NR, NT, KEGG, COG 和 GO 数据库, 其中, 注释到 NR 中的 Unigenes 最多, 有 22 211 个, 同时发现和印度谷螟最相似的物种是帝王蝶 (67.6%), 其次是家蚕 (6.5%)。差异表达基因分析结果发现, 约有 46% 的差异表达基因在胁迫品系中显著上调。大量研究表明, 氨肽酶-N、碱性磷酸酶、钙黏蛋白、糖脂类可能是 Bt 毒素的识别受体^[13-14]。在本研究, 我们发现 9 个编码 APN 的基因, 这些基因在胁迫品系中的表达量明显升高, 另外还发现 55 个编码 Cadherin 的基因, 10 个编码 ALP 和 4 条编码 Glycolipid 的基因, 但这些受体基因在不同品系中的表达量无显著差异。由此可以推测, 印度谷螟对转基因大米粉的毒理学机制和 APN 受体密切相关。另外, 我们的另一项针对取食转基因大米粉后的印度谷螟体内抗氧化酶的活力动态进行了深入探索, 这将与本研究一起为后续的相关研究提供了坚实的理论依据^[15]。

本研究通过二代测序技术对印度谷螟转录组进行了解析, 同时利用生物信息分析方法对不同转基因大米粉饲喂种群的表达差异基因进行了比较研究, 研究结果将有利于后续基因功能验证和印度谷螟的生物学相关研究, 为储粮害虫防治提供了重要的理论依据和参考资料。

参考文献:

- [1] LI Y, HALLERMAN E M, LIU Q, et al. The development and status of Bt rice in China[J]. Plant Biotechnol J, 2016, 14: 839-848.
- [2] LIU Q, HALLERMAN E, PENG Y, et al. Development of Bt rice and Bt maize in China and their efficacy in target pest

- control[J]. *Int J Mol Sci*, 2016, 17: 1-15.
- [3] BATES S L, ZHAO J Z, ROUSH R T, et al. Insect resistance management in GM crops: Past, present and future[J]. *Nat Biotechnol*, 2005, 23: 57-62.
- [4] GRABHERR M G, HAAS B J, YASSOUR M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. *Nature biotechnology*, 2011, 29(7): 644-652.
- [5] KIM Y K, SON J W, KIM H Y, et al. New occupational allergen in citrus farmers: citrus red mite (*Panonychus citri*)[J]. *Annals of Allergy, Asthma & Immunology*, 1999, 82(2): 223-228.
- [6] TORRE S, TATTINI M, BRUNETTI C, et al. RNA-seq analysis of *Quercus pubescens* Leaves: de novo transcriptome assembly, annotation and functional markers development[J]. 2014.
- [7] AUDIC S, CLAVERIE J M. The significance of digital gene expression profiles[J]. *Genome research*, 1997, 7(10): 986-995.
- [8] FINK G, McCONNELL M, VOLLMER S. Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures[J]. *Journal of Development Effectiveness*, 2014, 6(1): 44-57.
- [9] RAJAGOPAL R, SIVAKUMAR S, AGRAWAL N, et al. Silencing of midgut aminopeptidase N of *Spodoptera litura* by double-stranded RNA establishes its role as *Bacillus thuringiensis* toxin receptor[J]. *Journal of Biological Chemistry*, 2002, 277(49): 46849-46851.
- [10] YE J, FANG L, ZHENG H, et al. WEGO: a web tool for plotting GO annotations[J]. *Nucleic acids research*, 2006, 34: W293-W297.
- [11] LIU N. Insecticide resistance in mosquitoes: impact, mechanisms, and research directions[J]. *Annual review of entomology*, 2015, 60: 537-559.
- [12] ZHANG H, TIAN W, ZHAO J, et al. Diverse genetic basis of field-evolved resistance to Bt cotton in cotton bollworm from China[J]. *Proceedings of the National Academy of Sciences*, 2012, 109(26): 10275-10280.
- [13] PIGOTT C R, ELLAR D J. Role of receptors in *Bacillus thuringiensis* crystal toxin activity[J]. *Microbiology and Molecular Biology Reviews*, 2007, 71(2): 255-281.
- [14] 李光涛, 曹阳, 叶恭银, 等. 转 Bt 基因抗虫稻谷对麦蛾的抗性评价[J]. *植物保护学报*, 2008, 35(3): 205-208.
- [15] 王康旭, 吴学友, 陶冶心, 等. 转基因稻谷 Bt 蛋白降解及长期取食对印度谷螟体内 3 种抗氧化酶活性的影响[J/OL]. *粮油食品科技*: 1-8. <https://doi.org/10.16210/j.cnki.1007-7561.2020.06.013>. 
- 备注: 本文的彩色图表可从本刊官网 (<http://lspkj.ijournal.cn/ch/index.aspx>), 中国知网、万方、维普、超星等数据库下载获取。