

DOI: 10.3724/SP.J.1005.2011.00636

## 家蚕核型多角体病毒水平转移基因分析

段海蓉<sup>1,3</sup>, 丘德彬<sup>2</sup>, 贡成良<sup>3</sup>, 黄茉莉<sup>1</sup>

1. 苏州大学系统生物研究中心, 苏州 215006;
2. 苏州大学数学科学学院, 苏州 215006;
3. 苏州大学医学部, 苏州 215123

**摘要:** 为了探讨杆状病毒基因组的遗传进化模式, 文章利用家蚕核型多角体病毒(BmNPV)和其宿主家蚕全基因组数据, 进行了全基因组的同源性搜索和系统进化分析, 结果显示, BmNPV 的几丁质酶(Chi)基因、凋亡抑制蛋白 3(IAP3)基因和尿苷二磷酸葡萄糖转移酶(UGT)基因为水平转移基因。这 3 个基因都来源于其宿主昆虫。通过核苷酸组成、密码子偏好性、选择压力等基因特征分析, 发现 BmNPV 水平转移基因与其基因组序列存在明显差异, 进一步验证水平转移基因的外源性。对 3 个水平转移基因的功能分析发现它们有利于杆状病毒在宿主昆虫中的侵染与繁殖, 并提高杆状病毒在昆虫中的生存能力。

**关键词:** 家蚕核型多角体病毒; 家蚕; 杆状病毒; 水平转移基因

## Analysis of horizontal transfer gene of *Bombyx mori* NPV

DUAN Hai-Rong<sup>1,3</sup>, QIU De-Bin<sup>2</sup>, GONG Cheng-Liang<sup>3</sup>, HUANG Mo-Li<sup>1</sup>

1. Center for Systems Biology of Soochow University, Suzhou 215006, China;
2. School of Mathematical Sciences of Soochow University, Suzhou 215006, China;
3. Medical College of Soochow University, Suzhou 215123, China

**Abstract:** For research on genetic characters and evolutionary origin of the genome of baculoviruses, a comprehensive homology search and phylogenetic analysis of the complete genomes of *Bombyx mori* NPV and *Bombyx mori* were used. Three horizontally transferred genes (inhibitor of apoptosis, chitinase, and UDP-glucosyltransferase) were identified, and there was evidence that all of these genes were derived from the insect host. The results of analysis showed lots of differences between the features of horizontal transferred genes and the ones of whole genomic genes, such as nucleotide composition, codon usage bias and selection pressure. These results reconfirmed that the horizontally transferred genes are exogenous. The analysis of gene function suggested that horizontally transferred genes acquired from an ancestral host insect can increase the efficiency of baculoviruses transmission.

**Keywords:** BmNPV; *Bombyx mori*; baculoviruses; horizontal gene transfer

收稿日期: 2010-09-29; 修回日期: 2010-12-14

基金项目: 国家重点基础研究发展计划“973”项目(编号: 2005CB121000), 江苏省高校自然科学基金项目(编号: SZ107917)和 2009 年度大学生创新性实验计划项目(编号: 57315969)资助

作者简介: 段海蓉, 硕士研究生, 专业方向: 生化与分子生物学。Tel: 0512-65112418; E-mail: melissa.duan@hotmail.com

通讯作者: 贡成良, 博士, 教授, 研究方向: 家蚕生化与分子生物学。E-mail: gongcl@suda.edu.cn

黄茉莉, 博士, 副教授, 研究方向: 数据的整合分析、基因组水平的数学模型构建。E-mail: huangml@suda.edu.cn

网络出版时间: 2011-4-19 9:20:32

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20110419.0920.005.html>

水平基因转移(Horizontal gene transfer, HGT)<sup>[1,2]</sup>, 又称侧向基因转移(Lateral gene transfer, LGT), 是指生物个体之间、生物与环境之间、或单个细胞内部细胞器之间及细胞器与细胞核之间进行的遗传物质的横向交流。随着高等真核生物基因组测序的完成, 许多真核生物的基因水平转移现象通过生物信息学方法的研究被发现<sup>[3,4]</sup>。目前, 一般可以根据密码子偏爱性、BLAST 分析、基因分布和系统发育分析鉴定水平转移基因<sup>[5-8]</sup>。另外, 选择压力分析法、内含子序列的比较和特殊的插入序列等方法也常用于预测基因水平转移事件<sup>[9-13]</sup>。但最常用的方法是 BLAST 和系统发育分析法。BLAST 方法提供了一种最快最简单的发现一个数据库中与研究蛋白最相似蛋白的方法。系统发育分析法是基于用单个基因构建的进化树与传统物种进化树相比较来发现水平转移事件。与其他方法不同的是, 系统发育分析方法不仅能预测水平转移事件, 还能指出该基因的来源。

杆状病毒是专门寄生于节肢动物, 特别是鳞翅目(Lepidoptera)昆虫的一个很大的 DNA 病毒家族, 该家族可以分为两类: 颗粒病毒(Granulosis virus, GV)和核型多角体病毒(Nucleopolyhedrovirus, NPV)。家蚕核型多角体病毒(*Bombyx mori* NPV, BmNPV)<sup>[14]</sup>主要感染家蚕, 并引起家蚕因血液型脓病而死亡, 该疾病目前还没有有效的治疗方法。但同时通过重组 BmNPV, 在利用家蚕表达外源基因等方面有广泛应用。2003 年, Hughes 等<sup>[15]</sup>用 13 种不同的杆状病毒基因组数据探测了宿主细胞与杆状病毒之间的基因水平转移, 发现在杆状病毒中有 6 个基因与宿主细胞之间有水平转移的迹象, 而且有证据显示其中两个基因是从宿主鳞翅目昆虫水平转移而来。但由于当时真核生物基因组数据的缺乏, 研究中并没有分析这些基因水平转移的位点和转移方式等问题。2008 年, Susum 等<sup>[16]</sup>通过比较家蚕与 BmNPV 的基因组, 发现在 BmNPV 编码的 136 个蛋白质中有 15 个蛋白与家蚕的蛋白有较大的相似性, 这些基因可能是鳞翅目杆状病毒与鳞翅目昆虫之间的水平转移基因, 但是它们到底是不是水平转移基因还需通过其他检测方法来进行进一步验证。

为了深入理解 NPV 的基因组特征, 探讨病毒基因组的遗传进化模式, 本文基于 BmNPV 的基因组数据, 利用序列相似性搜索及系统进化分析相结合

的方法分析鉴定了 BmNPV 基因组中的水平转移基因, 特别是与宿主家蚕之间的可能水平转移基因。同时根据相似性搜索结果及系统进化树的拓扑结构判断水平转移基因的来源物种及转移方向。利用多种软件分析了水平转移基因的核苷酸组成、密码子偏好性、选择压力及基因排布, 发现它们不同于基因组的普遍特性, 从而进一步确定它们的外源性。

## 1 材料和方法

### 1.1 材料

从 GenBank 中下载了 BmNPV 的全基因组序列(NC\_001962.1), 同时获取了全基因组中已预测的 143 条蛋白质序列(GI: 9630817-9630959)。从家蚕基因组数据库(Silkworm genome database, SilkDB)(<http://silkworm.genomics.org.cn/>)<sup>[17]</sup>下载了家蚕的全基因组序列, 同时获取下载了已预测的家蚕的 14 623 条蛋白质序列(BGIBMGA000001-BGIBMGA 014623)。

### 1.2 方法

#### 1.2.1 水平转移基因的鉴定

利用本地 Blast 软件中 Formatdb 程序将家蚕的 14 623 条蛋白质序列文件格式化为 BLAST 可识别的检索数据库文件, 然后用 Blastall 程序对 BmNPV 的 143 条蛋白质序列与家蚕蛋白库中序列进行相似性搜索, 筛选出 E 值小于  $e^{-30}$ , 序列一致性高于 30%(Identity>30%)的结果作为候选水平转移基因。对于这些候选水平转移基因, 我们又在 NCBI 的非冗余蛋白质序列数据库中进行了新一轮的相似性搜索, 以期在其他物种中找到相似性序列。利用比对结果中的物种信息, 将其通过 Taxonomy 数据库对应成谱系数据。再根据以下标准进行水平转移基因的来源物种分析: 前 3 个比对结果的非同物种系谱都属于某一域(如: 昆虫), 水平转移基因即可能来源于该域。同时从比对结果中挑选出代表性物种的相似性序列进行进化分析, 用于构建进化树的序列数据来源见附表 1。利用 ClustalW 软件<sup>[5]</sup>进行蛋白质家族的多序列比对, 比对结果通过手工校正, 并去除了所有包含空格和失意的位点。为了查清每个候选基因的系统发生关系, 分别用软件 MEGA 4<sup>[18]</sup>和 PhyML v3.0 aLRT<sup>[19]</sup>中 WAG 模型构建了邻接(Neighbor-joining, NJ)进化树和最大似然法(Maximum

likelihood, ML)进化树。系统发生进化树的可靠性通过运行 1 000 次重复的自引导值进行检验。综合考虑水平转移基因的系统进化树拓扑结构及序列比对的同源性,分析水平转移基因的来源物种和转移方向。运用软件 PILER-DF<sup>[20]</sup>、LRT\_Finder, 鉴定 BmNPV 基因组中的可能转座元件,同时比较水平转移基因上下游 1 000 bp 不同物种间的基因组序列。

### 1.2.2 水平转移基因的基因特征分析

运用 MEGA 4 软件批量分析 BmNPV 基因组中所有 143 个编码基因的核苷酸组成,从中抽提出水平转移基因的数据。根据这些数据绘制了 BmNPV 基因组中各基因的 GC 含量分布图。运用 Codon W 软件计算了 BmNPV 基因组中 143 个基因的密码子指数,包括密码子适应指数(Codon adaptation index, CAI)、密码子偏好性指数(Codon bias index, CBI)、有效密码子数(Effective number of codon)和最优密码子使用频率(Frequency of optimal codons, FOP)。并同时计算了如第三密码子核苷酸组成的 GC<sub>3S</sub> 值等一些其他常用数值和各基因的密码子使用情况。BmNPV 水平转移基因的相关数据按上述方法计算获得。运用 KaKs\_Calculator 软件<sup>[13]</sup>中的所有模型计算部分物种的水平转移基因(数据来源见附表 1~3)的非同义替换率(The number of non-synonymous substitutions per non-synonymous site, Ka)和同义替换率(The number of synonymous substitutions per synonymous site, Ks),并通过 Fisher 检验检测各个模型得出的结果,最后选取 *P* 值最低的结果作为该基因的选择压力数据。根据 GenBank 中的基因组数据及位于水平转移基因上下游的编码基因的注释信息,分析这水平转移基因在不同核型多角体病毒中的分布并作图。

## 2 结果与分析

### 2.1 水平转移基因的鉴定

家蚕有 28 条染色体,目前已经预测出 14 623 个蛋白, BmNPV 基因组全长共 128 413 bp, 包含 143 个阅读框。从基因组水平上,用 BmNPV 的全部蛋白序列对家蚕全部蛋白进行序列相似性搜索( $E < e^{-30}$ , Identity>30%),发现 BmNPV 的 143 个蛋白中有 6 个蛋白与家蚕中蛋白有很高的相似性(表 1),它们分

别为蜕皮甾体尿苷二磷酸葡糖转移酶(Ecdysteroid UDP-Glucosyl Transferase, *EGT*)、凋亡抑制蛋白(Inhibitor of apoptosis, *IAP*)<sup>[21, 22]</sup>、超氧化物歧化酶(Superoxide dismutase, *SOD*)、泛素(Ubiquitin, *Ubi*)、几丁质酶(Chitinase, *Chi*)和组织蛋白酶(Cathepsin, *Cat*)。

表 1 家蚕核型多角体病毒与家蚕相似程度高的蛋白

| 基因名称             | BmNPV<br>基因编号 | 家蚕基因编号(蛋白序列一致性%, <i>E</i> 值)            |
|------------------|---------------|---|
| <i>UGT</i>       | 9630823       | BGIBMGA013834(35.43, 6E-84),<br>4 个其他蛋白 |
| <i>IAP1</i>      | 9630835       | BGIBMGA011524(30.91, 1E-35)             |
| <i>SOD</i>       | 9630841       | BGIBMGA001307(54, 9E-49),<br>2 个其他蛋白    |
| <i>Ubiquitin</i> | 9630844       | BGIBMGA001549(76.32, 5E-31),<br>2 个其他蛋白 |
| <i>Chitinase</i> | 9630926       | BGIBMGA008709(62.34, 0)                 |
| <i>Cathepsin</i> | 9630927       | BGIBMGA005131(38.17, 2E-60),<br>3 个其他蛋白 |

随后,用 BmNPV 的这 6 个蛋白在 NCBI 的非冗余蛋白质序列数据库中进行了新一轮的相似性搜索,发现这 6 个蛋白家族并不是病毒中所特有,它们同时存在于病毒、细菌及动植物等各种不同物种中。根据搜索结果筛选了杆状病毒类、细菌类、鳞翅目昆虫及其他真核动植物中的一些有代表性物种的同源蛋白,对上述 6 个基因进行系统发育分析。进化树的拓扑结构显示其中有 3 个基因(*SOD*、*Ubi* 和 *Cat*), BmNPV 与其他杆状病毒形成一个独立的分支,家蚕与其他鳞翅目昆虫聚集,与传统物种进化树没有明显差异,推测没有基因水平转移的迹象。在基于 *IAP* 蛋白序列的进化树中, NPV 并没有形成一个独立的分支,有部分 NPV 与真核生物中的鳞翅目聚集在一起形成一个整体的分支。在基于 *Chi* 和 *UGT* 序列的进化树中, NPV 组成独立分支,并且与家蚕所在的鳞翅目的分支相近,结果显示这 3 个基因有水平转移的迹象。

### 2.2 水平转移基因的来源物种及转移方向和方式

通过对 BmNPV 的 *IAP* 蛋白的相似性搜索,发现 *IAP* 同系物出现在各种不同的杆状病毒和绝大部分动物中。在比对结果中,除杆状病毒外,相似程度较高的前 3 个非杆状病毒物种分别为 *B. mori*、*S. frugiperda* 和 *S. exigua*, 它们都属于鳞翅目昆虫,故推测杆状病毒中 *IAP* 基因与鳞翅目昆虫的 *IAP* 基因

具有共同的祖先。一个杆状病毒基因组中通常有一个或多个 *IAP* 基因, BmNPV 中有 *IAP1* 和 *IAP2* 两个基因。从 *IAP* 基因进化树拓扑结构看(图 1), 脊椎动物中非鳞翅目昆虫的 *IAP* 基因形成了独立分支, 而杆状病毒中所有 *IAP* 基因却没有聚集在一起。其中, 杆状病毒的 *IAP1* 分支和 *IAP2* 分支相互邻接, 远离真核生物的 *IAP* 分支。杆状病毒的 *IAP3* 基因形成单独分支, 与鳞翅目昆虫的 *IAP* 基因邻近。据此, 可以推测某个杆状病毒的祖先从一个远祖的鳞翅目昆虫中获取并保留, 形成了 *IAP3* 基因。杆状病毒基因组中有一个或多个 *IAP* 基因, 且这些不同的 *IAP* 在氨基酸水平上有较高的相似性, 这一现象指出杆状病毒基因组在进化过程中, *IAP* 基因可能发生了基因复制事件。通过分析不同物种中 *IAP* 基因的两侧序列, 发现 *IAP3* 基因的上下游序列在杆状病毒之间缺乏相似性, 在杆状病毒与鳞翅目昆虫之间同样没有相似性。但发现在家蚕的 *IAP* 基因的上游 151 bp、下游 281 bp 处有一对长度为 17 bp 的反向重复序列, 在 EpNPV 的 *IAP3* 基因上游 987 bp 与下游 266 bp 处也发现一对反向重复序列, 故推测鳞翅目昆虫的 *IAP* 基因可能是通过转座的方式水平转移到杆状病毒基因组。

序列相似性搜索表明 *Chi* 基因的同系物主要存在于杆状病毒、细菌、动物及植物中, 而鳞翅目昆虫的基因组中大多含有两个 *Chi* 基因。通过对 *Chi* 基因构建系统发育进化树(图 2), 发现杆状病毒、细菌及植物的 *Chi* 基因分别聚集成独立分支, 鳞翅目昆虫的两类 *Chi* 基因相互分离, 一支 *Chi1* 基因与其他动物相聚集, 另一支 *Chi-h* 基因与细菌和杆状病毒这两个分支相邻近, 显示鳞翅目昆虫的 *Chi-h* 基因属于外源基因, 可能来源于细菌或杆状病毒。用家蚕 *Chi-h* 蛋白质序列在数据库中搜索, 结果显示, 除昆虫外, 相似程度较高的前 3 个非昆虫类的物种分别为 *S. proteamaculans* 586、*S. liquefaciens* 和 *Enterobacter sp.* 都属于细菌。故推测鳞翅目昆虫的 *Chi-h* 基因来源于细菌。若用 BmNPV 的 *Chi* 蛋白质序列在数据库中搜索, 除杆状病毒外, 相似程度较高的前 3 个非杆状病毒的物种都属于鳞翅目昆虫, 可以认为杆状病毒中的 *Chi* 基因可能来源于宿主鳞翅目昆虫。以往的研究认为杆状病毒的 *Chi* 基因是由细菌通过基因水平转移而来的<sup>[23]</sup>, 但由于当时分

子数据的局限性, 还未发现鳞翅目昆虫中的 *Chi-h* 基因。杆状病毒和细菌都是鳞翅目昆虫的主要病原微生物, 通过本文分析发现杆状病毒的 *Chi* 基因并不是直接来源于细菌, 而是通过鳞翅目昆虫这一中间体。鳞翅目昆虫的 *Chi-h* 基因是由细菌转移而来的, 鳞翅目昆虫再转给了杆状病毒。通过分析不同物种的 *Chi* 基因的两侧序列, 发现 *Chi* 基因的上下游序列在杆状病毒之间有很高的相似性, 在杆状病毒与鳞翅目昆虫之间没有相似性, 鳞翅目昆虫与细菌之间也缺乏相似性。在分析的几个物种的 *Chi* 基因的上下游序列中也没有发现转座特征性序列, 故 *Chi* 基因不可能是通过转座或重组的方式水平转移的, 极有可能是在进化过程中随机插入而转移到不同的物种中。

尿苷二磷酸葡萄糖转移酶(*UGT*)基因的同族物在杆状病毒、细菌、酵母及动植物种均有发现。在比对结果中, 除杆状病毒外, 相似程度较高的前 3 个非杆状病毒物种分别为 *B. mori*、*Z. filipendulae* 和 *A. aegypti*, 都属于杆状病毒的宿主昆虫, 故推测杆状病毒中 *UGT* 基因来与宿主昆虫的 *UGT* 基因具有共同的祖先。根据对 *UGT* 基因的进化分析(图 3), 大多物种的 *UGT* 基因符合传统生物系统发育进化树, 只有杆状病毒的 *UGT* 基因所在分支与动物, 特别是鳞翅目昆虫的 *UGT* 基因所在分支聚集在一起, 同样可以判断出杆状病毒的 *UGT* 基因来源于宿主鳞翅目昆虫。对家蚕 *UGT* 基因的上下游序列(GenBank 登录号: NP\_001135960)分析表明, 上下游序列中有一段长度为 148 bp 的相似性序列, 同时这段序列中含有一对短的反向重复序列。已有研究指出<sup>[24]</sup>, 家蚕的部分 *UGT* 基因总是与反转座子连锁, 并可能处于反转座子之中, 且表现出高度同源性, 因而这些反转座子的存在也反映了鳞翅目昆虫的 *UGT* 基因可能通过转座子转移到杆状病毒的基因组中。

### 2.3 水平转移基因的特征分析

BmNPV 的全基因组序列(GenBank 登录号: NC\_001962), 全长共 128 413 bp, 包含有 143 个阅读框, 全部基因 GC 含量平均值为 40.7%, 说明 BmNPV 是一个低 GC 含量的物种。基于 BmNPV 143 个基因的核苷酸组成, 进行水平转移基因的核苷酸组成的统计分析。从结果可以看出, 全部基因的 GC 含量平均

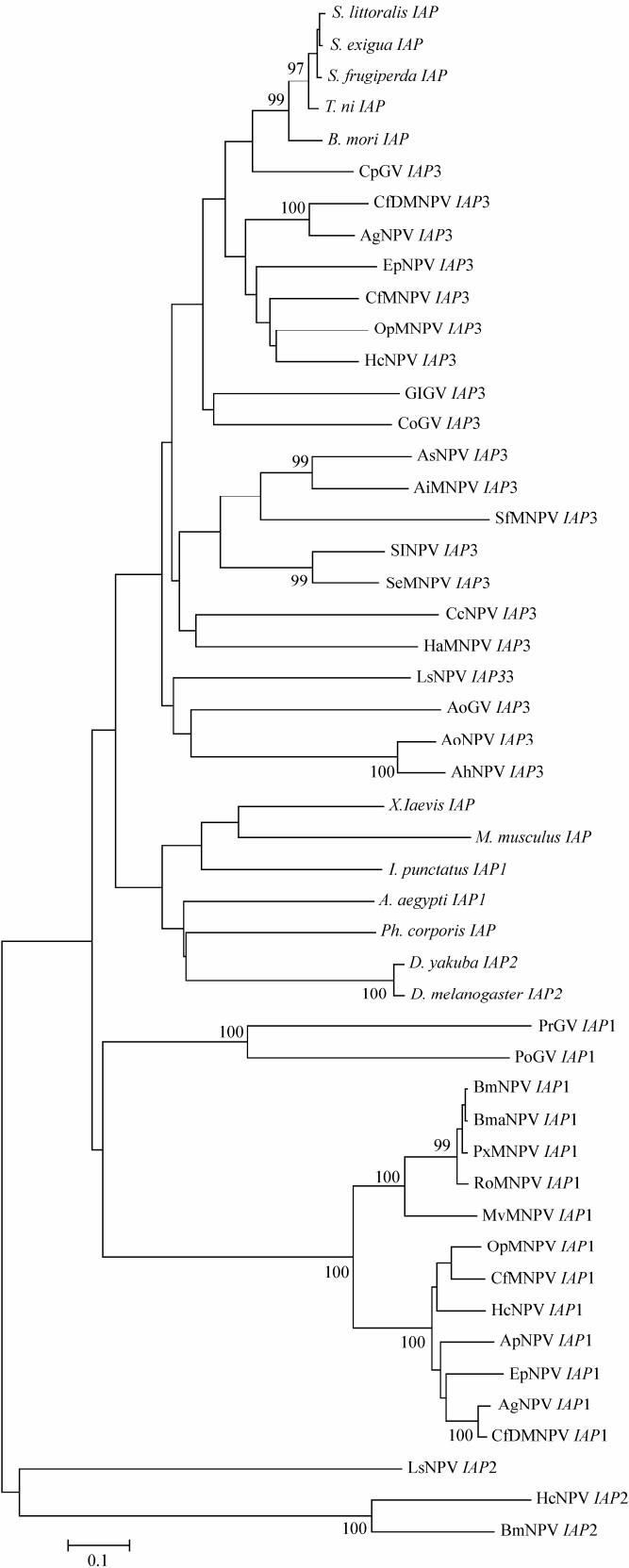


图 1 基于 IAP 氨基酸序列的进化树  
节点处的数字为 100 次引导值中该节点存在的置信度，只显示大于 95%。以上序列的数据来源及缩写符号见附表 1。

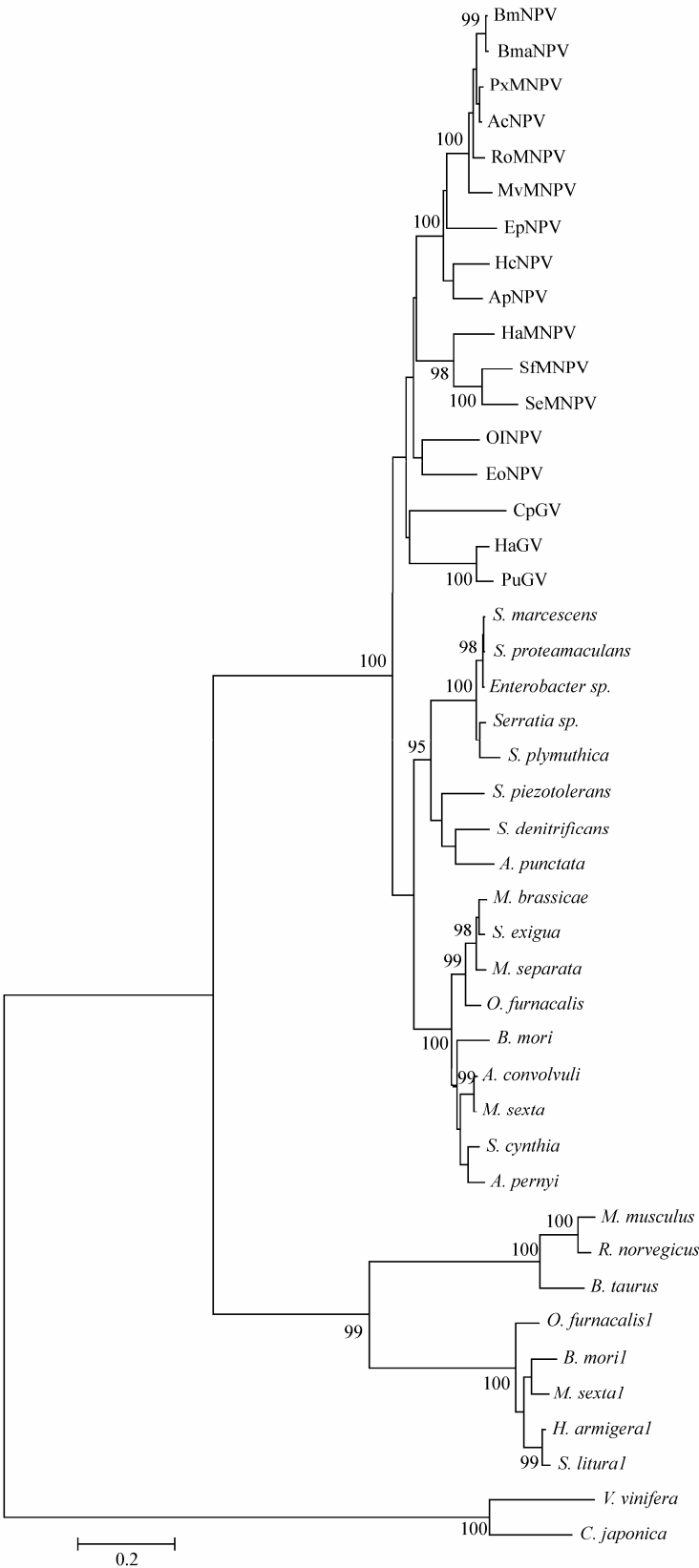


图 2 基于 *Chi* 氨基酸序列的进化树  
节点处的数字为 100 次引导值中该节点存在的置信度，只显示大于 95%。以上序列的数据来源及缩写符号见附表 2。



2)。分析发现, 全基因组中所有基因有 4 个显著的最优密码子。对 3 个水平转移基因密码子使用频率与全基因组密码子使用频率进行比较。一般认为比值若在 0.5~2.0, 表明两者对密码子偏爱性比较接近, 反之则差异较大。结果显示 *UGT* 基因对全基因组密码子使用频率比值大于 2 小于 0.5 的密码子有 16 种, *IAP* 基因对全基因组密码子使用偏爱性差异较大的密码子有 17 种, *Chi* 基因则有 14 种。表明这些水平转移基因与全基因组存在着密码子偏好性上的显著差异, 这也进一步支持了它们来源于外源物种的可能性。

表 2 BmNPV 全基因组中全部基因及 3 个 *HGT* 基因的密码子使用及偏性比较表

| 氨基酸  | 密码子 | 使用频率(100%) |            |            |            | <i>UGT</i> /Total | <i>IAP</i> /Total | <i>CHI</i> /Total |
|------|-----|------------|------------|------------|------------|-------------------|-------------------|-------------------|
|      |     | Total      | <i>UGT</i> | <i>IAP</i> | <i>CHI</i> |                   |                   |                   |
| Phe  | UUU | 3.94       | 4.54       | 5.8        | 3.25       | 1.15              | 1.47              | 0.83              |
|      | UUC | 1.05       | 0.39       | 0.68       | 1.08       | 0.37              | 0.65              | 1.03              |
| Leu  | UUA | 2.13       | 1.58       | 1.37       | 1.45       | 0.74              | 0.64              | 0.68              |
|      | UUG | 3.53       | 3.55       | 1.37       | 3.07       | 1.00              | 0.39              | 0.87              |
|      | CUU | 0.77       | 1.78       | 1.02       | 0.54       | 2.30              | 1.32              | 0.70              |
|      | CUC | 0.76       | 0.59       | 0.00       | 1.45       | 0.78              | 0.00              | 1.90              |
|      | CUA | 0.74       | 0.99       | 0.68       | 0.36       | 1.33              | 0.92              | 0.49              |
|      | CUG | 1.30       | 1.78       | 0.00       | 1.81       | 1.37              | 0.00              | 1.39              |
| Ile  | AUU | 3.17       | 2.76       | 3.07       | 2.17       | 0.87              | 0.97              | 0.68              |
|      | AUC | 1.44       | 1.97       | 1.71       | 0.72       | 1.37              | 1.19              | 0.50              |
|      | AUA | 1.96       | 1.58       | 1.02       | 1.99       | 0.81              | 0.52              | 1.02              |
| Met  | AUG | 2.74       | 3.16       | 2.05       | 1.45       | 1.15              | 0.75              | 0.53              |
| Val  | GUU | 1.68       | 1.18       | 2.05       | 2.35       | 0.70              | 1.22              | 1.40              |
|      | GUC | 1.15       | 1.58       | 1.37       | 2.71       | 1.37              | 1.18              | 2.35              |
|      | GUA | 1.25       | 2.76       | 2.73       | 1.63       | 2.21              | 2.18              | 1.30              |
|      | GUG | 2.52       | 2.76       | 2.05       | 2.53       | 1.10              | 0.81              | 1.00              |
| Ser  | UCU | 0.96       | 0.79       | 0.00       | 0.18       | 0.82              | 0.00              | 0.19              |
|      | UCC | 0.66       | 0.79       | 0.68       | 0.72       | 1.20              | 1.04              | 1.10              |
|      | UCA | 0.64       | 0.59       | 0.00       | 0.54       | 0.93              | 0.00              | 0.85              |
|      | UCG | 1.41       | 1.38       | 1.71       | 1.63       | 0.98              | 1.21              | 1.16              |
| Pro  | CCU | 0.69       | 0.99       | 1.02       | 0.18       | 1.43              | 1.48              | 0.26              |
|      | CCC | 1.05       | 1.97       | 0.68       | 1.63       | 1.87              | 0.65              | 1.54              |
|      | CCA | 0.68       | 0.20       | 0.00       | 0.54       | 0.29              | 0.00              | 0.80              |
|      | CCG | 1.25       | 0.99       | 1.37       | 1.45       | 0.79              | 1.09              | 1.16              |
| Thr  | ACU | 1.15       | 0.79       | 1.37       | 0.72       | 0.68              | 1.18              | 0.63              |
|      | ACC | 1.31       | 1.18       | 1.71       | 0.90       | 0.90              | 1.30              | 0.69              |
|      | ACA | 1.30       | 1.38       | 0.00       | 1.45       | 1.06              | 0.00              | 1.11              |
|      | ACG | 2.01       | 2.96       | 1.71       | 0.90       | 1.47              | 0.85              | 0.45              |
| Ala  | GCU | 1.00       | 0.59       | 0.68       | 1.99       | 0.59              | 0.68              | 1.98              |
|      | GCC | 1.61       | 3.55       | 2.05       | 2.53       | 2.21              | 1.27              | 1.58              |
|      | GCA | 0.94       | 0.99       | 1.71       | 0.54       | 1.05              | 1.81              | 0.58              |
|      | GCG | 1.60       | 3.94       | 1.37       | 2.53       | 2.47              | 0.85              | 1.58              |
| Tyr  | UAU | 2.11       | 1.97       | 2.73       | 1.27       | 0.94              | 1.29              | 0.60              |
|      | UAC | 2.64       | 2.76       | 1.37       | 3.8        | 1.04              | 0.52              | 1.44              |
| STOP | UAA | 0.32       | 0.00       | 0.34       | 0.18       | 0.00              | 1.07              | 0.57              |
|      | UAG | 0.04       | 0.20       | 0.00       | 0.00       | 4.53              | 0.00              | 0.00              |

(续表 2)

| 氨基酸 | 密码子 | 使用频率(100%) |      |      |      | UGT/Total | IAP/Total | CHI/Total |
|-----|-----|------------|------|------|------|-----------|-----------|-----------|
|     |     | Total      | UGT  | IAP  | CHI  |           |           |           |
| His | UGA | 0.07       | 0.00 | 0.00 | 0.00 | 0.00      | 0.00      | 0.00      |
|     | CAU | 0.98       | 1.58 | 1.02 | 0.54 | 1.62      | 1.05      | 0.56      |
|     | CAC | 1.30       | 1.58 | 2.05 | 0.72 | 1.21      | 1.57      | 0.56      |
| Gln | CAA | 2.68       | 1.78 | 2.05 | 2.53 | 0.66      | 0.76      | 0.94      |
|     | CAG | 1.10       | 1.38 | 1.02 | 0.54 | 1.26      | 0.93      | 0.50      |
| Asn | AAU | 3.58       | 3.16 | 2.05 | 1.99 | 0.88      | 0.57      | 0.56      |
|     | AAC | 4.49       | 4.34 | 5.8  | 3.98 | 0.97      | 1.29      | 0.89      |
| Lys | AAA | 5.52       | 5.13 | 3.07 | 5.42 | 0.93      | 0.56      | 0.98      |
|     | AAG | 1.60       | 1.58 | 1.02 | 0.90 | 0.98      | 0.64      | 0.56      |
| Asp | GAU | 2.58       | 1.58 | 2.39 | 3.8  | 0.61      | 0.93      | 1.47      |
|     | GAC | 3.37       | 3.55 | 7.51 | 4.52 | 1.05      | 2.23      | 1.34      |
| Glu | GAA | 3.75       | 2.56 | 5.46 | 3.25 | 0.68      | 1.45      | 0.87      |
|     | GAG | 1.48       | 0.59 | 2.39 | 1.99 | 0.40      | 1.61      | 1.34      |
| Cys | UGU | 1.13       | 0.39 | 4.44 | 0.18 | 0.35      | 3.92      | 0.16      |
|     | UGC | 1.30       | 0.39 | 2.39 | 0.54 | 0.30      | 1.84      | 0.42      |
| Trp | UGG | 0.80       | 0.79 | 1.37 | 3.07 | 0.98      | 1.70      | 3.83      |
| Arg | CGU | 0.85       | 0.59 | 1.71 | 0.54 | 0.69      | 2.00      | 0.64      |
|     | CGC | 1.37       | 1.38 | 1.71 | 0.90 | 1.01      | 1.24      | 0.66      |
|     | CGA | 0.76       | 0.59 | 0.34 | 0.36 | 0.78      | 0.45      | 0.48      |
|     | CGG | 0.46       | 0.20 | 0.34 | 0.00 | 0.43      | 0.74      | 0.00      |
| Ser | AGU | 0.94       | 0.79 | 0.00 | 0.90 | 0.84      | 0.00      | 0.96      |
|     | AGC | 1.58       | 1.18 | 0.34 | 1.27 | 0.75      | 0.22      | 0.80      |
| Arg | AGA | 1.00       | 0.39 | 1.02 | 0.72 | 0.39      | 1.02      | 0.72      |
|     | AGG | 0.40       | 0.00 | 0.00 | 0.54 | 0.00      | 0.00      | 1.35      |
| Gly | GGU | 0.73       | 0.39 | 0.68 | 1.08 | 0.54      | 0.93      | 1.48      |
|     | GGC | 1.67       | 1.58 | 1.02 | 5.42 | 0.94      | 0.61      | 3.25      |
|     | GGA | 0.67       | 1.38 | 1.02 | 1.08 | 2.07      | 1.53      | 1.62      |
|     | GGG | 0.29       | 0.20 | 0.34 | 0.90 | 0.68      | 1.18      | 3.13      |

注: Total 表示 BmNPV 基因组全部基因的密码子使用频率; UGT/Total 表示 UGT 基因对于 BmNPV 全部基因的密码子使用偏性; IAP/Total, CHI/Total 同上。

为了了解 BmNPV 的 3 个水平转移基因在进化过程中对编码区的选择压力估计,对部分物种的 3 个水平转移基因的 DNA 序列进行了选择压力分析(表 3~5)。分析表明,不同物种的 *Chi* 基因间的  $Ka/Ks$  相近,都是  $Ka<Ks$  的负向选择。不同物种的 *UGT* 基因间的  $Ka/Ks$  有差异,但都是  $Ka>Ks$  的正向选择。不同物种的 *IAP* 基因间的  $Ka/Ks$  不同,既有  $Ka>Ks$  的正向选择,也有  $Ka<Ks$  的负向选择,但在昆虫之间 *IAP* 基因都是  $Ka<Ks$  的负向选择。这些结果显示被自然选择所支持并保留下来的水平转移基因,由于其不同于传统的进化历程,在进化中会受到不同的选择压力的影响。

本文还分析了 3 个水平转移基因在不同的物种

中的基因排布,由于缺少昆虫的基因组数据,本文只分析了 3 个水平转移基因在不同杆状病毒基因组中的排布情况。对 *Chi* 基因,分析了 BmNPV、RoMNPV 和 PxMNPV 3 个核型多角体病毒,发现该基因在这 3 个病毒的基因组中有较高的保守性(图 4A),同样具有很高的保守性的基因还有杆状病毒的 *EGT* 基因、*IAP1* 基因和 *IAP2* 基因(图 4: B、C、D)。,基于 *EGT* 基因和 *Chi* 基因在分化后的不同杆状病毒基因组中的排布仍维持着很高的保守性,推测它们的水平转移发生在杆状病毒分化之前。对 *IAP3* 基因,我们分析了 EpNPV、HcNPV 和 NaNPV 3 个核型多角体病毒,发现该基因在这 3 个病毒基因组中的排布完全不同(图 4E)。该现象进一步确认

表 3 杆状病毒 *Chi* 基因与昆虫 *Chi-h* 基因及细菌的 *Chi* 基因的 Ka/Ks 值

| 基因名称              | 编号 | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      |
|-------------------|----|--------|--------|--------|--------|--------|--------|--------|--------|
| RoMNPV <i>Chi</i> | 1  |        |        |        |        |        |        |        |        |
| PxMNPV <i>Chi</i> | 2  | 0.1126 |        |        |        |        |        |        |        |
| BmNPV <i>Chi</i>  | 3  | 0.1836 | 0.3505 |        |        |        |        |        |        |
| Bm <i>Chi-h</i>   | 4  | 0.0852 | 0.1134 | 0.0871 |        |        |        |        |        |
| Ms <i>Chi-h</i>   | 5  | 0.0825 | 0.1161 | 0.0825 | 0.0383 |        |        |        |        |
| Of <i>Chi-h</i>   | 6  | 0.0852 | 0.0975 | 0.0835 | 0.0357 | 0.0397 |        |        |        |
| Spl <i>Chi</i>    | 7  | 0.0859 | 0.1780 | 0.1445 | 0.0660 | 0.0635 | 0.0665 |        |        |
| Spr <i>Chi</i>    | 8  | 0.0822 | 0.1285 | 0.1329 | 0.0615 | 0.0623 | 0.1107 | 0.0450 |        |
| Sm <i>Chi</i>     | 9  | 0.0755 | 0.1659 | 0.1749 | 0.0644 | 0.0625 | 0.0450 | 0.0528 | 0.0329 |

注：数据来源及缩写符号见附表 2。

表 4 杆状病毒 *EGT* 基因与昆虫 *UGT* 基因的 Ka/Ks 值

| 基因名称              | 编号 | 1      | 2      | 3      | 4      | 5      |
|-------------------|----|--------|--------|--------|--------|--------|
| AcNPV <i>egt</i>  | 1  |        |        |        |        |        |
| RoMNPV <i>egt</i> | 2  | 2.1330 |        |        |        |        |
| BmNPV <i>egt</i>  | 3  | 2.3098 | 1.8536 |        |        |        |
| Bm <i>UGT</i>     | 4  | 1.4023 | 1.4246 | 1.3934 |        |        |
| Zi <i>UGT</i>     | 5  | 1.1264 | 1.0969 | 1.1695 | 1.2562 |        |
| Aa <i>UGT</i>     | 6  | 1.3109 | 1.3377 | 1.2741 | 1.1321 | 1.0986 |

注：数据来源及缩写符号见附表 3。

表 5 杆状病毒 *IAP* 基因与鳞翅目昆虫 *IAP* 基因的 Ka/Ks 值

| 基因名称               | 编号 | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
|--------------------|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| HcNPV <i>IAP1</i>  | 1  |        |        |        |        |        |        |        |        |        |        |
| EpNPV <i>IAP1</i>  | 2  | 1.1340 |        |        |        |        |        |        |        |        |        |
| BmNPV <i>IAP1</i>  | 3  | 0.8380 | 0.7373 |        |        |        |        |        |        |        |        |
| BmNPV <i>IAP2</i>  | 4  | 1.2439 | 1.1667 | 0.8583 |        |        |        |        |        |        |        |
| MvMNPV <i>IAP2</i> | 5  | 1.1509 | 1.798  | 1.3645 | 1.0981 |        |        |        |        |        |        |
| HcNPV <i>IAP2</i>  | 6  | 1.3483 | 1.3961 | 0.6363 | 0.8722 | 1.0577 |        |        |        |        |        |
| HcNPV <i>IAP3</i>  | 7  | 0.5329 | 0.5343 | 0.9234 | 1.0092 | 0.6718 | 1.0832 |        |        |        |        |
| EpNPV <i>IAP3</i>  | 8  | 0.6254 | 1.4203 | 0.6846 | 0.9183 | 0.7279 | 0.0690 | 0.6225 |        |        |        |
| Sf <i>IAP</i>      | 9  | 0.4338 | 0.739  | 0.536  | 0.806  | 0.7119 | 1.1023 | 0.4547 | 0.5409 |        |        |
| Sl <i>IAP</i>      | 10 | 0.4936 | 0.7439 | 0.6246 | 0.8683 | 1.0964 | 1.2066 | 0.3919 | 0.4608 | 0.3882 |        |
| Bm <i>IAP</i>      | 11 | 0.4991 | 0.5797 | 0.9944 | 0.9785 | 1.1700 | 2.1225 | 0.3641 | 0.4720 | 0.2791 | 0.3499 |

注：数据来源及缩写符号见附表 1。

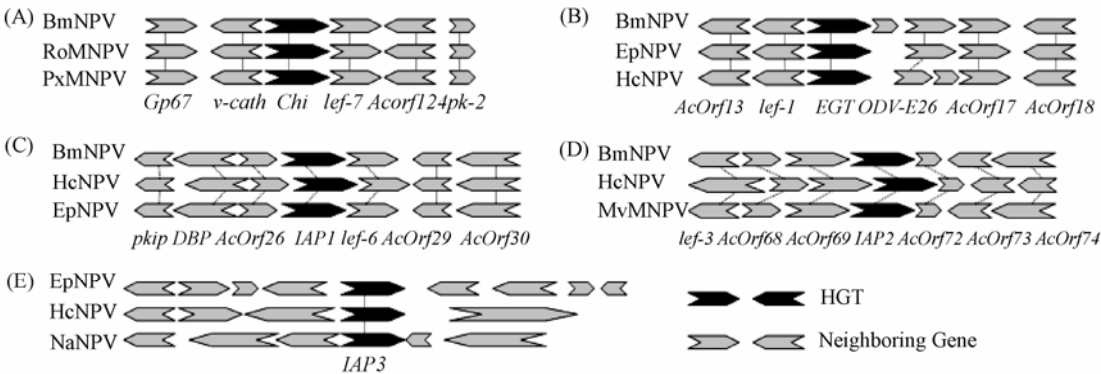


图 4 不同核型多角体病毒中水平转移基因(黑色)及旁侧基因(灰色)的组织结构  
(A): *Chi* 基因及旁侧基因在 3 个不同核型多角体病毒中的组织结构; (B): *EGT* 基因及旁侧基因在 3 个不同核型多角体病毒中的组织结构; (C): *IAP1* 基因及旁侧基因在 3 个不同核型多角体病毒中的组织结构; (D): *IAP2* 基因及旁侧基因在 3 个不同核型多角体病毒中的组织结构; (E): *IAP3* 基因及旁侧基因在 3 个不同核型多角体病毒中的组织结构。箭头方向表示基因的表达方向, 不同病毒中相同结构的基因用线相连。BmNPV: *Bombyx mori* NPV, RoMNPV: *Rachiplusia ou* MNPV, PxMNPV: *Plutella xylostella* multiple nucleopolyhedrovirus, EpNPV: *Epiphyas postvittana* NPV, HcNPV: *Hyphantria cunea* nucleopolyhedrovirus, MvMNPV: *Maruca vitrata* MNPV, NaNPV: *Neodiprion abietis* NPV.

*IAP3* 基因是外源基因, 转移方向是从鳞翅目昆虫到杆状病毒, 而且水平转移的过程发生在杆状病毒分化之后。

### 3 讨论

杆状病毒与宿主昆虫之间的基因水平转移现象早在生物实验中有所发现。用 *AcMNPV* 感染 *T.ni* 的 TN-368 细胞, 病毒在宿主细胞繁殖过程中形成很多多角体, 偶尔也会出现少多角体突变现象。实验观察这些杆状病毒的空斑形态突变发生过程中常伴有来源于宿主细胞的 DNA 插入病毒基因组的现象(即基因水平转移), 并且这些 DNA 插入大部分与转座子有关<sup>[25]</sup>。本文通过对家蚕和 *BmNPV* 基因组的分析, *BmNPV* 中有 6 个蛋白与家蚕的相关蛋白有很大的相似性。对 6 个基因进行系统发育分析, 结果显示有 3 个基因有水平转移的迹象, 并且这 3 个基因可能都是由宿主昆虫转移而来的。2001 年, 发现 *BmNPV* 的 *Chi* 基因与灵菌的 *Chi* 基因在氨基酸水平上有较高的同源性, 推测杆状病毒编码的 *Chi* 基因是由细菌通过基因的水平转移而来<sup>[23]</sup>。2003 年, 家蚕的 *Chi-h* 基因第一次被克隆和鉴定<sup>[26]</sup>, 与家蚕的另一个 *Chi1* 基因不同的是, *Chi-h* 基因与细菌和杆状病毒的 *Chi* 基因在氨基酸序列水平和蛋白结构上有很大的相似性, 通过系统发育分析, 推测家蚕的 *Chi-h* 基因起源于远古细菌或杆状病毒。上述研究中, 由于当时分子数据的缺乏, 没能给出 *Chi* 基因明确的进化关系。本文基于更全面完整的数据, 利用系统发育与相似性搜索相结合的方法分析了杆状病毒 *Chi* 的来源, 结果显示杆状病毒的 *Chi* 基因并不是直接从细菌中转移而来的, 而是由细菌先转移给昆虫, 昆虫再转移给杆状病毒。Hughes 等<sup>[15]</sup>构建了昆虫和昆虫病毒的 *IAP* 的进化关系, 并与用 DNA 聚合酶构建的进化关系相比较。结果指出 *CpGV* 的一个 *IAP* 基因是从昆虫中获取, 而 *NPV* 的 *IAP3* 基因可能起源于 *CpGV*, 并认为在 *IAP* 的进化历程中至少还有另外两起水平转移事件。本文分析了 *IAP3* 基因在不同的杆状病毒基因组中的排布, 发现不同的杆状病毒的 *IAP3* 基因的座位及其两侧的基因组织方式各不相同, 该现象不仅说明了杆状病毒的 *IAP3* 基因是外源的, 还显示了 *IAP3* 基因水平转移的时间是在杆状病毒分化之后, 且发生了多次的水平转移事件。

有趣的是, 这 3 个水平转移基因都对杆状病毒与昆虫宿主的相互关系产生重要影响。研究表明, 病毒的侵染使宿主的生长发育不再完全受自身系统的控制, 而是更有利于病毒的复制与增殖。杆状病毒 *Chi* 基因的表达产物主要分布于内质网, 能够水解昆虫体内的组成型几丁质, 促进宿主虫体的液化, 导致病毒的扩散, 易于病毒的水平传播。杆状病毒 *Chi* 基因缺失试验表明, *Chi* 基因的失活不影响病毒的复制与多角体的形成, 使 *Chi* 基因缺失病毒的半致死量和半致死事件都相应减少, 说明 *Chi* 基因的缺失能够增强病毒的敏感性, 提高其杀虫毒力<sup>[27, 28]</sup>。*BmNPV* 的蜕皮甾体尿苷二磷酸葡萄糖转移酶(*EGT*)基因编码的产物能够使家蚕分泌的蜕皮激素(MH)失活, 从而阻止幼虫的变态, 延长幼虫的取食时间<sup>[29]</sup>。杆状病毒的 *IAP* 基因是病毒侵染宿主所必需的基因, 否则宿主细胞会因为病毒的感染而发生凋亡, 病毒的 *IAP* 基因通过抑制 *caspase* 的加工过程和活性来控制宿主细胞凋亡<sup>[30]</sup>。

在原核生物中基因水平转移的主要方式有转化、接合和转导<sup>[31]</sup>, 但真核生物中水平转移的发生机制仍不明确。最近研究的真核生物中的水平转移事件主要发生在病毒或细菌与其真核宿主之间。系统发育分析显示病毒或细菌更容易从其真核宿主中捕获有功能的基因, 而不是从相反的方向发生水平转移<sup>[3]</sup>。研究表明大部分水平转移基因都没有功能, 寄生在真核生物中的 DNA 病毒也能够从宿主基因组中获取少数有功能的基因, 这些基因可能让病毒产生新的毒力、对细胞凋亡的抵御或对宿主免疫反应的应答起作用等。本研究中通过对家蚕和 *BmNPV* 基因组的分析, 发现涉及杆状病毒侵染宿主的过程的 3 个基因发生了水平转移事件, 并且杆状病毒的这些基因都是从宿主昆虫中水平转移而来的。分析表明杆状病毒从宿主昆虫中获取这些基因, 对控制宿主昆虫的生长发育有很大的影响, 使昆虫生理不受自己的调节系统控制, 从而提高了病毒对昆虫宿主的适应性, 并且有利于病毒的侵染和繁殖。

### 附录

附表 1~3 见文章电子版([www.Chinagene.cn](http://www.Chinagene.cn))。

### 参考文献(References):

- [1] 欧剑虹, 谢志雄, 陈向东, 倪丽娜, 沈萍. 水平基因转移.

- 遗传, 2003, 25(5): 623–627.
- [2] Than C, Sugino R, Innan H, Nakhleh L. Efficient inference of bacterial strain trees from genome-scale multilocus data. *Bioinformatics (Oxford, England)*, 2008, 24(13): i123–i131.
- [3] Whitaker JW, McConkey GA, Westhead DR. Prediction of horizontal gene transfers in eukaryotes: Approaches and challenges. *Biochem Soc Trans*, 2009, 37(Pt 4): 792–795.
- [4] 程延才, 夏庆友, 刘春, 赵萍, 查幸福, 徐汗福, 向仲怀. 家蚕 *chi*、*gluE* 和 *fruA* 基因与微生物相应基因的同源性及其基因水平转移初探. *遗传学报*, 2004, 31(10): 1082–1089.
- [5] Buckley CO, Stephens D, Herring PA, Jackson JH. % (G+C) variation and prediction by a model of bacterial gene transfer and codon adaptation. *Omic*s, 2002, 6(3): 259–272.
- [6] Dufraigne C, Fertil B, Lespinats S, Giron A, Deschavanne P. Detection and characterization of horizontal transfers in prokaryotes using genomic signature. *Nucleic Acids Res*, 2005, 33(1): e6.
- [7] Pál C, Papp B, Lercher MJ. Horizontal gene transfer depends on gene content of the host. *Bioinformatics (Oxford, England)*, 2005, 21(S2): ii222–ii223.
- [8] Tsirigos A, Rigoutsos I. A new computational method for the detection of horizontal gene transfer events. *Nucleic Acids Res*, 2005, 33(3): 922–933.
- [9] 吴建盛, 谢建明, 周童, 翁建洪, 孙啸. 基于支持向量机的细菌基因组水平转移基因预测. *生物化学与生物物理进展*, 2007, 34(7): 724–731.
- [10] 陈阳, 王守觉. 仿生模式识别在细菌基因组水平转移基因预测中的应用. *现代生物医学进展*, 2008, 8(8): 1518–1521.
- [11] Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. Versatile and open software for comparing large genomes. *Genome Biol*, 2004, 5(2): R12.
- [12] Li ZF, Wang L, Zhong Y. Detecting horizontal gene transfer with T-REX and RHOM programs. *Brief Bioinform*, 2005, 6(4): 394–401.
- [13] Hamady M, Betterton MD, Knight R. Using the nucleotide substitution rate matrix to detect horizontal gene transfer. *BMC Bioinformatics*, 2006, 7: 476.
- [14] 吴小锋. 家蚕核型多角体病毒的基因组结构及其表达模式. *病毒学报*, 2006, 22(4): 324–328.
- [15] Hughes AL, Friedman R. Genome-wide survey for genes horizontally transferred from cellular organisms to baculoviruses. *Mol Biol Evol*, 2003, 20(6): 979–987.
- [16] Katsuma S, Kawaoka S, Mita K, Shimada T. Genome-wide survey for baculoviral host homologs using the *Bombyx* genome sequence. *Insect Biochem Mol Biol*, 2008, 38(12): 1080–1086.
- [17] Wang J, Xia QY, He XM, Dai MT, Ruan J, Chen J, Yu G, Yuan HF, Hu YF, Li RQ, Feng T, Ye C, Lu C, Wang J, Li SG, Wong GKS, Yang HM, Wang J, Xiang ZH, Zhou ZY, Yu J. SilkDB: a knowledgebase for silkworm biology and genomics. *Nucleic Acids Res*, 2005, 33(S1): D399–D402.
- [18] Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings Bioinformatics*, 2008, 9(4): 299–306.
- [19] Guindon S, Delsuc F, Dufayard JF, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol (Clifton, NJ)*, 2009, 1(537): 113–137.
- [20] Edgar RC, Myers EW. PILER: identification and classification of genomic repeats. *Bioinformatics (Oxford, England)*, 2005, 21(S1): i152–i158.
- [21] 张瑞, 姚青, 彭建新, 洪华珠. 杆状病毒 IAP 基因的结构、功能及其进化. *微生物通报*, 2006, 33(1): 128–132.
- [22] 包人月, 吴金美, 吴小锋. 昆虫杆状病毒细胞凋亡抑制基因. *细胞生物学杂志*, 2006, 28(5): 676–680.
- [23] 贡成良, 薛仁宇, 曹广力, 石晓燕, 马志明. 家蚕核型多角体病毒几丁质酶基因. *蚕业科学*, 2001, 27(4): 283–287.
- [24] 曹广力, 贡成良, 薛仁宇, 朱越雄, 魏育红. 家蚕核型多角体病毒 *egt* 基因的分子进化分析. *昆虫学报*, 2008, 51(12): 1244–1254.
- [25] Fraser MJ, Smith GE, Summers MD. Acquisition of host cell DNA sequences by baculoviruses: Relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis Viruses. *J Virol*, 1983, 47(2): 287–300.
- [26] Daimon T, Hamada K, Mita K, Okano K, Suzuki MG, Kobayashi M, Shimada T. A *Bombyx mori* gene, *BmChi-h*, encodes a protein homologous to bacterial and baculovirus chitinases. *Insect Biochem Mol Biol*, 2003, 33(8): 749–759.
- [27] Hawtin RE, Arnold K, Ayres MD, de A Zanutto PM, Howard SC, Gooday GW, Chappell LH, Kitts PA, King LA, Possee RD. Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome. *Virol*, 1995, 212(2): 673–685.
- [28] Saville GP, Patmanidi AL, Possee RD, King LA. Deletion of the *Autographa californica* nucleopolyhedrovirus chitinase KDEL motif and *in vitro* and *in vivo* analysis of the modified virus. *J Gen Virol*, 2004, 85(4): 821–831.
- [29] Euans OP, O'Reilly DR. Purification and kinetic analysis of a baculovirus ecdysteroid UDP-glucosyltransferase. *Biochemistry*, 1998, 330(pt 3): 1265–1270.
- [30] Manji GA, Hozak RR, LoCount DJ, Friesen PD. Baculovirus inhibitor of apoptosis functions at or upstream of the apoptotic suppressor p35 to prevent programmed cell death. *J Virol*, 1997, 71(6): 4509–4516.
- [31] Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nat*, 2000, 405(6784): 299–304.

附表 1 IAPs 数据来源

| 分类        | GenBank 登录号  | 物种拉丁名  | 缩写符号        |
|-----------|--------------|--|-------------|
| 杆状病毒 IAP1 | NP_047432    | <i>Bombyx mori</i> NPV                         | BmNPV IAP1  |
|           | NP_203202    | <i>Epiphyas postvittana</i> NPV                | EpNPV IAP1  |
|           | YP_803434    | <i>Anticarsia gemmatalis</i> NPV               | AgNPV IAP1  |
|           | YP_473302    | <i>Hyphantria cunea</i> NPV                    | HcNPV IAP1  |
|           | YP_611079    | <i>Antheraea pernyi</i> NPV                    | ApNPV IAP1  |
|           | YP_002884258 | <i>Bombyx mandarina</i> NPV                    | BmaNPV IAP1 |
|           | YP_758493    | <i>Plutella xylostella</i> MNPV                | PxMNPV IAP1 |
|           | NP_703017    | <i>Rachiplusia ou</i> MNPV                     | RoMNPV IAP1 |
|           | YP_950748    | <i>Maruca vitrata</i> MNPV                     | MvMNPV IAP1 |
|           | NP_046197    | <i>Orgyia pseudotsugata</i> MNPV               | OpMNPV IAP1 |
|           | NP_848348    | <i>Choristoneura fumiferana</i> MNPV           | CfMNPV IAP1 |
|           | NP_663251    | <i>Phthorimaea operculella granulovirus</i>    | PoGV IAP1   |
|           | YP_003429403 | <i>Pieris rapae granulovirus</i>               | PrGV IAP1   |
|           | YP_473268    | <i>Hyphantria cunea</i> NPV                    | HcNPV IAP2  |
|           | NP_047474    | <i>Bombyx mori</i> NPV                         | BmNPV IAP2  |
| 杆状病毒 IAP2 | YP_758377    | <i>Leucania separata</i> NPV                   | LsNPV IAP2  |
|           | NP_203195    | <i>Epiphyas postvittana</i> NPV                | EpNPV IAP3  |
|           | YP_803428    | <i>Anticarsia gemmatalis</i> NPV               | AgNPV IAP3  |
| 杆状病毒 IAP3 | YP_473308    | <i>Hyphantria cunea</i> NPV                    | HcNPV IAP3  |
|           | NP_818735    | <i>Adoxophyes honmai</i> NPV                   | AhNPV IAP3  |
|           | YP_529789    | <i>Agrotis segetum</i> NPV                     | AsNPV IAP3  |
|           | YP_249643    | <i>Chrysodeixis chalcites</i> NPV              | CcNPV IAP3  |
|           | YP_002300603 | <i>Adoxophyes orana</i> NPV                    | AoNPV IAP3  |
|           | YP_002332813 | <i>Spodoptera litura</i> NPV                   | SINPV IAP3  |
|           | YP_758337    | <i>Leucania separata</i> NPV                   | LsNPV IAP3  |
|           | ACA02668     | <i>Spodoptera frugiperda</i> MNPV              | SfMNPV IAP3 |
|           | NP_046191    | <i>Orgyia pseudotsugata</i> MNPV               | OpMNPV IAP3 |
|           | NP_848342    | <i>Choristoneura fumiferana</i> MNPV           | CfMNPV IAP3 |
|           | ACI28829     | <i>Agrotis ipsilon</i> MNPV                    | AiMNPV IAP3 |
|           | NP_037870    | <i>Spodoptera exigua</i> MNPV                  | SeMNPV IAP3 |
|           | YP_002332667 | <i>Helicoverpa armigera</i> MNPV               | HaMNPV IAP3 |
|           | NP_872543    | <i>Adoxophyes orana granulovirus</i>           | AoGV IAP3   |
|           | NP_891863    | <i>Cryptophlebia leucotreta granulovirus</i>   | CIGV IAP3   |
|           | NP_148801    | <i>Cydia pomonella granulovirus</i>            | CpGV IAP3   |
|           | YP_654505    | <i>Choristoneura occidentalis granulovirus</i> | CoGV IAP3   |
| 昆虫        | NP_001037024 | <i>Bombyx mori</i>                             | BmIAP       |
|           | AAF35285     | <i>Spodoptera frugiperda</i>                   | SfIAP       |
|           | ABA62322     | <i>Spodoptera exigua</i>                       | SxIAP       |
|           | CAM96614     | <i>Spodoptera littoralis</i>                   | SlIAP       |
|           | AAF19819     | <i>Trichoplusia ni</i>                         | TnIAP       |
|           | XP_001657917 | <i>Aedes aegypti</i>                           | AaIAP       |
|           | AAO24632     | <i>Ictalurus punctatus</i>                     | IpIAP       |
| 其他真核类生物   | XP_002423703 | <i>Pediculus humanus corporis</i>              | PhoIAP      |
|           | NP_001108272 | <i>Xenopus laevis</i>                          | XlIAP       |
|           | NP_001156719 | <i>Mus musculus</i>                            | MmIAP       |

附表 2 *Chis* 数据来源

| 分类               | GenBank 登录号  | 物种拉丁名                                     | 缩写符号              |
|------------------|--------------|---|-------------------|
| 核型多角体病毒          | NP_047523    | <i>Bombyx mori</i> NPV                    | BmNPV <i>Chi</i>  |
|                  | YP_002884348 | <i>Bombyx mandarina</i> NPV               | BmaNPV <i>Chi</i> |
|                  | NP_054156    | <i>Autographa californica</i> NPV         | AvNPV <i>Chi</i>  |
|                  | YP_473218    | <i>Hyphantria cunea</i> NPV               | HcNPV <i>Chi</i>  |
|                  | NP_203279    | <i>Epiphyas postvittana</i> NPV           | EpNPV <i>Chi</i>  |
|                  | YP_611002    | <i>Antheraea pernyi</i> NPV               | ApNPV <i>Chi</i>  |
|                  | YP_001650934 | <i>Orgyia leucostigma</i> NPV             | OINPV <i>Chi</i>  |
|                  | YP_874243    | <i>Ecotropis obliqua</i> NPV              | EoNPV <i>Chi</i>  |
|                  | YP_758590    | <i>Plutella xylostella</i> MNPV           | PxMNPV <i>Chi</i> |
|                  | NP_703113    | <i>Rachiplusia ou</i> MNPV                | RoMNPV <i>Chi</i> |
|                  | YP_950825    | <i>Maruca vitrata</i> MNPV                | MvMNPV <i>Chi</i> |
|                  | YP_002332554 | <i>Helicoverpa armigera</i> MNPV          | HaMNPV <i>Chi</i> |
|                  | YP_001036314 | <i>Spodoptera frugiperda</i> MNPV         | SfMNPV <i>Chi</i> |
|                  | NP_037779    | <i>Spodoptera exigua</i> MNPV             | SeMNPV <i>Chi</i> |
| 颗粒病毒             | NP_148794    | <i>Cydia pomonella granulovirus</i>       | CpGV <i>Chi</i>   |
|                  | YP_001649087 | <i>Helicoverpa armigera granulovirus</i>  | HaGV <i>Chi</i>   |
|                  | YP_003422447 | <i>Pseudaletia unipuncta granulovirus</i> | PuGV <i>Chi</i>   |
| 细菌               | ACE78180     | <i>Serratia marcescens</i>                | Sma <i>Chi</i>    |
|                  | AAB97779     | <i>Enterobacter sp.</i>                   | Esp <i>Chi</i>    |
|                  | ABB86291     | <i>Serratia proteamaculans</i>            | Spr <i>Chi</i>    |
|                  | AAK72610.2   | <i>Serratia sp.</i>                       | Ssp <i>Chi</i>    |
|                  | CAD32933     | <i>Serratia plymuthica</i>                | Sp <i>Chi</i>     |
|                  | CAD58828     | <i>Aeromonas punctata</i>                 | Apu <i>Chi</i>    |
|                  | BAE16588     | <i>Agrius convolvuli</i>                  | Ac <i>Chi</i> -h  |
| 昆虫 <i>Chi</i> -h | ABA06506     | <i>Mamestra brassicae</i>                 | Mb <i>Chi</i> -h  |
|                  | BAE16586     | <i>Samia cynthia</i>                      | Sc <i>Chi</i> -h  |
|                  | BAE16585     | <i>Antheraea pernyi</i>                   | Ape <i>Chi</i> -h |
|                  | ABA06504     | <i>Spodoptera exigua</i>                  | Se <i>Chi</i> -h  |
|                  | NP_001108407 | <i>Bombyx mori</i>                        | Bm <i>Chi</i> -h  |
|                  | ABB88891     | <i>Manduca sexta</i>                      | Ms <i>Chi</i> -h  |
|                  | BAE16587     | <i>Ostrinia furnacalis</i>                | Of <i>Chi</i> -h  |
|                  | AAG37105     | <i>Bombyx mori</i> 1                      | Bm <i>Chi</i> 1   |
| 昆虫 <i>Chi</i> 1  | AAB53952     | <i>Manduca sexta</i> 1                    | Ms <i>Chi</i> 1   |
|                  | AAQ94193     | <i>Helicoverpa armigera</i> 1             | Ha <i>Chi</i> 1   |
|                  | ABI81757     | <i>Ostrinia furnacalis</i> 1              | Of <i>Chi</i> 1   |
|                  | AAQ91787     | <i>Spodoptera litural</i> 1               | Sl <i>Chi</i> 1   |
|                  | NP_075675.2  | <i>Mus musculus</i>                       | Mm <i>Chi</i>     |
| 其他真核类生物          | NP_997469    | <i>Rattus norvegicus</i>                  | Rn <i>Chi</i>     |
|                  | AAI02932     | <i>Bos taurus</i>                         | Bt <i>Chi</i>     |

附表 3 UGTs 数据来源

| 分类      | GenBank 登录号  | 物种拉丁名                                 | 缩写符号   |
|---------|--------------|---------------------------------------|--------|
| 杆状病毒    | NP_047420    | <i>Bombyx mori</i> NPV                | BmNPV  |
|         | YP_002884246 | <i>Bombyx mandarina</i> NPV           | BmaNPV |
|         | NP_054044    | <i>Autographa californica</i> NPV     | AcNPV  |
|         | YP_758481    | <i>Plutella xylostella</i> MNPV       | PxMNPV |
|         | BAG75097     | <i>Autographa nigrisigna</i> NPV      | AnNPV  |
|         | YP_249745    | <i>Chrysodeixis chalcites</i> NPV     | CcNPV  |
|         | NP_818769    | <i>Adoxophyes honmai</i> NPV          | AhNPV  |
|         | NP_258389    | <i>Spodoptera litura</i> NPV          | SINPV  |
|         | NP_872572    | <i>Adoxophyes orana granulovirus</i>  | AoGV   |
|         | BAB69953     | <i>Adoxophyes honmai granulovirus</i> | AhGV   |
| 昆虫      | NP_001135960 | <i>Bombyx mori</i>                    | BmUGT  |
|         | ACZ97418     | <i>Zygaena filipendulae</i>           | ZfUGT  |
|         | XP_001653185 | <i>Aedes aegypti</i>                  | AaUGT  |
|         | XP_970307    | <i>Tribolium castaneum</i>            | TcUGT  |
|         | XP_001601990 | <i>Nasonia vitripennis</i>            | NvUGT  |
|         | XP_001987323 | <i>Drosophila grimshawi</i>           | DgUGT  |
|         | XP_394494    | <i>Apis mellifera</i>                 | AmUGT  |
|         | XP_001948117 | <i>Acyrtosiphon pisum</i>             | ApUGT  |
|         | XP_002053428 | <i>Drosophila virilis</i>             | DvUGT  |
|         | XP_001954445 | <i>Drosophila ananassae</i>           | DaUGT  |
|         | XP_002031680 | <i>Drosophila sechellia</i>           | DseUGT |
|         | XP_002103915 | <i>Drosophila simulans</i>            | DsiUGT |
|         | NP_001097744 | <i>Drosophila melanogaster</i>        | DmUGT  |
|         | XP_002063948 | <i>Drosophila willistoni</i>          | DwUGT  |
|         | XP_002097170 | <i>Drosophila yakuba</i>              | DyUGT  |
|         | XP_002019745 | <i>Drosophila persimilis</i>          | DpUGT  |
|         | XP_001980604 | <i>Drosophila erecta</i>              | DeUGT  |
| 其他真核类生物 | NP_964007    | <i>Mus musculus</i>                   | MmUGT  |
|         | NP_001170917 | <i>Danio rerio</i>                    | DrUGT  |
|         | AAC50565     | <i>Homo sapiens</i>                   | HsUGT  |