

东亚飞蝗气味结合蛋白 I 的生物信息学分析

陆 峥, 张妍妍, 洪 芳, 蒋国芳

(南京师范大学生命科学学院, 江苏省生物资源技术重点实验室, 江苏 南京 210046)

[摘要] 利用生物信息学方法分析了东亚飞蝗 (*Locusta migratoria manilensis*) OBP1 的核苷酸与氨基酸序列 (GenBank 登录号 FJ215322.1/ACB0696.1), 并对其组成成分、疏水/亲水区域、信号肽、跨膜结构域、蛋白质二级结构, 以及分子进化关系等进行了预测与推断. 结果显示, 该蛋白由 148 个氨基酸组成, 预测相对分子质量为 16 122.6, 理论等电点 (pI) 为 4.99, 分子式为 $C_{696}H_{1128}N_{184}O_{223}S_{15}$; 含有蛋白激酶 C 磷酸化位点和酪氨酸激酶 II 磷酸化位点等; 具有标准的昆虫信息素/气味结合蛋白结构域.

[关键词] 东亚飞蝗, 气味结合蛋白, 生物信息学, 蛋白结构和功能

[中图分类号] Q811.4 [文献标识码] A [文章编号] 1001-4616(2010)02-0063-08

Bioinformatic Analysis on the *Locusta migratoria manilensis* Odorant-Binding Proteins I

Lu Zheng Zhang Yanyan, Hong Fang, Jiang Guofang

(Jiangsu Key Laboratory for Bioresource Technology, School of Life Science, Nanjing Normal University, Nanjing 210046, China)

Abstract The nucleotide and amino acid sequences of odorant-binding protein I (GenBank accession number FJ215322.1/ACB0696.1) from *Locusta migratoria manilensis* were analyzed and predicted by the tools of bioinformatics in the following aspects: the composition, equipotential hydrophobicity or hydrophilicity, the signal peptide, transmembrane topological structure, protein secondary structure, the composition of homologous protein and phylogenetic relationship. The results showed that the LmigOBP1 gene encoded a 148 amino acid polypeptides with a standard insect pheromone/odorant binding protein domain, the estimated relative molecular weight of 16 122.6, the theoretical pI of 4.99 and the structural formula of $C_{696}H_{1128}N_{184}O_{223}S_{15}$, which included Protein kinase C phosphorylation sites and Casein kinase II phosphorylation sites.

Key words *Locusta migratoria manilensis*, odorant binding protein (OBP), bioinformatics, protein structure and function

昆虫可以准确而灵敏地感受周围环境中的信号物质以及种群内部的交换信息, 其触角系统不仅具有嗅觉功能, 还具有触觉、味觉功能, 在昆虫觅食, 寻找配偶, 交配, 产卵等行为中发挥重要作用. 这主要依赖于其触角内的各种蛋白质介导, 主要包括气味结合蛋白、化学感受蛋白、气味受体、气味降解酶以及感觉神经元膜蛋白等.

国内外学者对昆虫气味结合蛋白 (odorant binding proteins, OBPs) 已进行了较长时间的研究, 目前认为它是一类低相对分子质量 (15 000 左右)、等电点偏酸性 (pH 4.0~5.0)、球状的水溶性蛋白, 多肽链全长约 120~160 个氨基酸, N 末端有一段 20 个氨基酸左右的信号肽, 序列中有 6 个保守的半胱氨酸^[1]. OBPs 主要集中在存在于昆虫嗅觉感受器内部的淋巴液中, 作用是结合并运送脂溶性的气味分子通过水性的淋巴液到达嗅觉神经元树突末梢上的膜结合受体, 引起昆虫的嗅觉^[2-3]. 根据识别的气味分子不同, 气味结合蛋白又可以分为性信息素结合蛋白 (pheromone binding protein, PBP) 和普通气味结合蛋白 (general odorant binding protein, GOBP)^[4-6].

东亚飞蝗 (*Locusta migratoria manilensis*) 是我国飞蝗的 3 个亚种之一, 在我国大致分布在北纬 42° 以南的

收稿日期: 2009-09-08

基金项目: 国家自然科学基金 (30670257).

通讯联系人: 蒋国芳, 教授, 博士生导师, 研究方向: 进化遗传学和分子生态学. E-mail: cnjgl208@163.com

冲积平原地带,北起河北、山东、陕西以南,西至四川、甘肃南部以南,均有蝗害报导.迄今,对东亚飞蝗的气味结合蛋白及其基因的研究在国内外少有报道^[7-8].本文用生物信息学方法研究东亚飞蝗气味结合蛋白基因编码蛋白的理化性质、空间结构及系统进化等,为进一步从实验水平研究其生理功能奠定基础,为研制对东亚飞蝗两性均有吸引力的高效引诱剂提供理论依据,为东亚飞蝗的防治提供新的思路与途径.

1 材料与方法

1.1 获取核苷酸与蛋白质序列

在 GenBank 数据库中检索得到东亚飞蝗的 1 种气味结合蛋白 I 型 (Lm igOBP1) 的 mRNA 序列 (FJ215322.1) 与氨基酸序列 (ACI30696.1).

1.2 重要数据库和生物学软件

主要数据库为美国国立生物技术信息中心 (National Center for Biotechnology Information, NCBI) GenBank 数据库 (<http://www.ncbi.nlm.nih.gov/GenBank>). 蛋白质基本性质分析工具 ExPASy 数据库中的 ProtParam Tool (<http://www.expasy.org/tools/protparam.html>), 疏水性分析软件 ProtScale 程序 (<http://www.expasy.org/cgi-bin/protscale.pl>), 蛋白质跨膜结构预测软件 TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM>), 蛋白质信号肽预测工具 SignalP 3.0 Server (<http://www.cbs.dtu.dk/services/SignalP>), 蛋白质卷曲螺旋预测工具 COILS Server (http://www.ch.embnet.org/software/COILS_form.html), 蛋白质亚细胞定位预测工具 PSORT II Prediction (<http://psort.in.s.u-tokyo.ac.jp/form2.html>), 蛋白质 Domain 分析工具 SMART (<http://smart.embl-heidelberg.de>), 蛋白质二级结构预测工具 PredictProtein (<http://www.predictprotein.org>); 序列匹配分析软件 ClustalX-1.81, 系统进化分析软件 Mega 4.1.

1.3 用于本文分析的各种昆虫气味结合蛋白的序列相关信息

用于本文分析的各种昆虫气味结合蛋白 OBP1 的核苷酸与氨基酸序列 GenBank 登录号: 东亚飞蝗 (GenBank 登录号 FJ215322.1/ACB0696.1), 美洲大蠊 (*Periplaneta americana*) (FJ215311.1/ACB0685), 褐飞虱 (*Nilaparvata lugens*) (FJ215306.1/ACI30680), 暗黑鳃金龟 (*Holotrichia parallela*) (AB026555.1/BAC07272), 疟蚊 (*Anopheles gambiae*) (AF393487.1/AAL60412), 棉铃虫 (*Helicoverpa armigera*) (AJ278991.1/CAC08211), 蜜蜂 (*Apis mellifera*) (NM_001040226.1/NP_001035316), 双斑蟋 (*Gryllus bimaculatus*) (FJ215321.1/ACB0695), 豆黄鳃金龟 (*Hoplatypus picea*) (AB025571.1/BAC07270), 异发丽金龟 (*Phyllopertha diversa*) (AB026552.1/BAA88061), 蓖麻蚕 (*Samia cynthia ricini*) (AB040977.1/BAF91328), 斜纹夜蛾 (*Spodoptera litura*) (EF159978.1/ABM54823), 烟夜蛾 (*Helicoverpa assulta*) (AY864774.1/AAW65076), 小菜蛾 (*Plutella xylostella*) (EU163980.1/ABW05104), 亚洲玉米螟 (*Ostrinia furnacalis*) (DQ673101.1/ABG66419), 草地螟 (*Loxostege sticticali*) (EU413989.1/ACB47481).

2 结果与分析

2.1 东亚飞蝗 OBP1 的基本性质分析

2.1.1 氨基酸组成、相对分子质量、等电点分析

东亚飞蝗 OBP1 基因编码 148 个氨基酸, 相对分子质量为 16 122.6, 理论等电点为 4.99, 分子中的原子组成为 Carbon(C) 696, Hydrogen(H) 1128, Nitrogen(N) 184, Oxygen(O) 223, Sulfur(S) 15, 分子式为 C₆₉₆H₁₁₂₈N₁₈₄O₂₂₃S₁₅, 其中碱性氨基酸 (Arg + Lys) 18 个, 酸性氨基酸 (Asp + Glu) 26 个, 疏水性氨基酸 37 个, 极性氨基酸 64 个; 在体外哺乳动物网状细胞中的估计半衰期为 30 h, 在体外有机体内的半衰期大于 20 h, 在大肠杆菌内的半衰期大于 10 h, 蛋白的不稳定指数为 46.41, 属于不稳定蛋白 (不稳定系数大于 40 的为不稳定蛋白质); 疏水性指数为 66.69. 如图 1 所示的 7 种昆虫 (东亚飞蝗、美洲大蠊、褐飞虱、暗黑鳃金龟、疟蚊、棉铃虫、蜜蜂) 同类蛋白的氨基酸组成, 其中 Ala, Gly, Leu, Lys 在这几种昆虫中含量均较为丰富, 而 Arg, Asn 的含量较少, 几乎均不含有 Trp. 东亚飞蝗的 Ala, Gly 含量比其它昆虫丰富, 而 Asp, Cys, Tyr 含量比其它昆虫少. 表 1 是这 7 种昆虫气味结合蛋白特殊氨基酸组成和理化性质分析, 这 7 种昆虫分别属于直翅目、蜚蠊目、同翅目、鞘翅目、双翅目、鳞翅目、膜翅目. 表中的数据显示, 在整体层面上昆虫气味结合蛋白的特殊氨基酸组成具有趋同性, 但也存在明显差异. 其中东亚飞蝗气味结合蛋白的极性氨基酸含量

最低, 而疏水氨基酸含量最高.

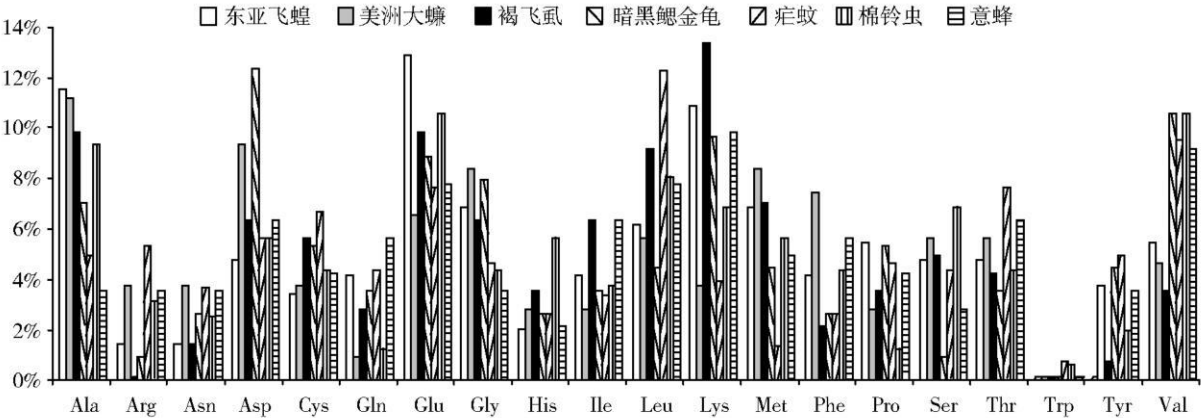


图 1 7 种昆虫(东亚飞蝗、美洲大蠊、褐飞虱、暗黑鳃金龟、疟蚊、棉铃虫、意蜂)气味结合蛋白的 20 种氨基酸组成

Fig.1 The 20 amino acids contents of odorant binding proteins among seven insects (*Locusta migratoria*, *Periplaneta americana*, *Nilaparvata lugens*, *Holotrichia parallela*, *Anopheles gambiae*, *Helicoverpa armigera*, *Apis mellifera*)

表 1 7 种昆虫 (东亚飞蝗、美洲大蠊、褐飞虱、暗黑鳃金龟、疟蚊、棉铃虫、意蜂) 气味结合蛋白的特殊氨基酸组成和理化性质分析
Table 1 A comparison of composition and physical characters of odorant binding proteins among seven insects (*Locusta migratoria*, *Periplaneta americana*, *Nilaparvata lugens*, *Holotrichia parallela*, *Anopheles gambiae*, *Helicoverpa armigera*, *Apis mellifera*)

动物	氨基酸 残基数	相对 分子质量	pI	氨基酸比例 /%			
				碱性氨基酸	酸性氨基酸	极性氨基酸	疏水氨基酸
东亚飞蝗	148	16 122. 6	4. 99	12. 16	17. 57	25. 0	43. 2
美洲大蠊	108	11 930. 4	4. 49	7. 41	15. 74	31. 5	42. 6
褐飞虱	143	15 660. 4	5. 74	13. 29	16. 08	25. 9	41. 3
暗黑鳃金龟	114	12 649. 3	4. 45	10. 53	21. 05	28. 1	37. 7
疟蚊	304	34 459. 5	5. 15	9. 21	13. 16	35. 9	39. 1
棉铃虫	162	18 155. 9	5. 3	9. 88	16. 05	25. 3	43. 2
意蜂	143	16 620. 4	6. 3	13. 29	13. 99	29. 4	41. 3

2. 1. 2 疏水性和跨膜区分析

利用 ExPASy 数据库中的 ProScale 程序分析东亚飞蝗的 OBPI, 可得如图 2 所示的疏水性图谱. 由图中可见其属于水溶性蛋白质, 疏水性最大值为 2. 889, 最小值为 - 2. 678 大约有 3 个疏水区, 分别为第 5~ 24 位、第 50~ 56 位和第 64~ 74 位. 蛋白质跨膜结构预测软件 TMHMM Server v. 2. 0 分析该蛋白可得图 3 的结果, 表明其 N 端附近的疏水区与蛋白跨膜区明显相关. 而后 2 个疏水区并非跨膜区.

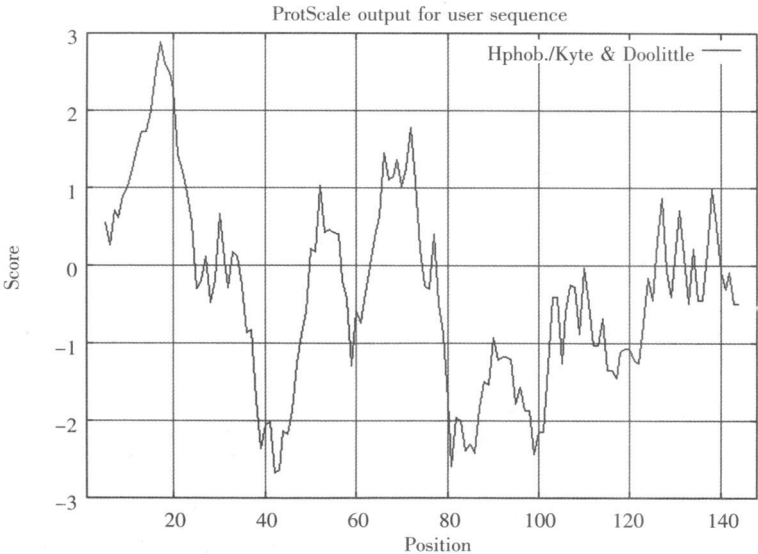


图 2 ProtScale 程序分析 LmigOBPI 的疏水性结果

Fig.2 The hydrophobicity / hydrophilicity analysis of LmigOBPI by using ProtScales

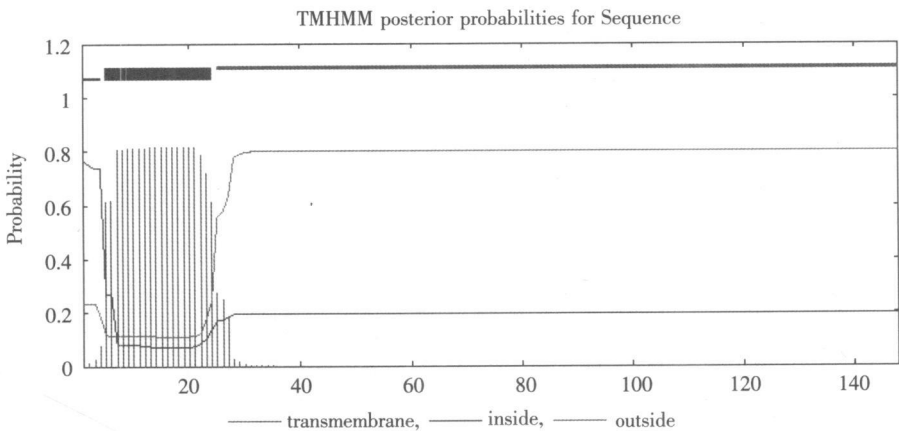


图 3 TMHMM Server,v.2.0 软件分析 LmigOBP1 的跨膜区结果

Fig.3 The transmembrane segment analysis of LmigOBP1 by using TMHMM Server,v.2.0

2. 1. 3 蛋白质信号肽分析

蛋白质信号肽预测工具 SignalP 3.0 Server分析东亚飞蝗的 LmigOBP1, 可得如图 4和图 5所示的信号肽分析图谱. 由分析结果可知, 该蛋白的信号肽位于第 1~ 23 位氨基酸, 最有可能的酶切位点在 23和 24 位氨基酸残基之间. 可见上述跨膜区分析结果所述的 N 端跨膜区应是该蛋白的信号肽区, 这与前人对尺蠖和烟夜蛾等气味结合蛋白的研究结果相一致^[9-11].

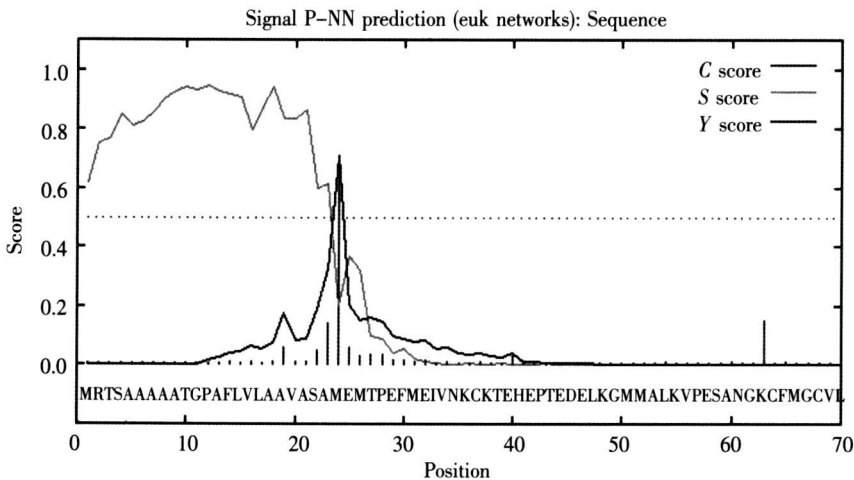


图 4 SignalP 3.0 Server 的神经式网络(NN)分析 LmigOBP1

Fig.4 The analysis of LmigOBP1 by using neural networks (NN) of SignalP 3.0 Server

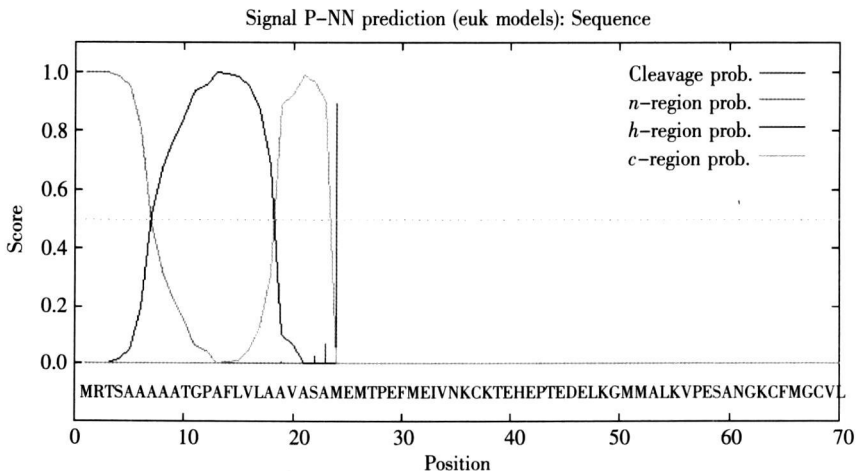


图 5 SignalP 3.0 Server 的隐马尔可夫模型(HMM)分析 LmigOBP1

Fig.5 The analysis of LmigOBP1 by using hidden Markov models (HMM) of SignalP 3.0 Server

运用蛋白质卷曲螺旋预测工具 COILS Server 分析 Lm IgOBP1, 从图 6 结果可见, Lm IgOBP1 在 3 种模式扫描下有一个共有的短的卷曲螺旋区 (95-105 个氨基酸), 且概率在 78% 以上.

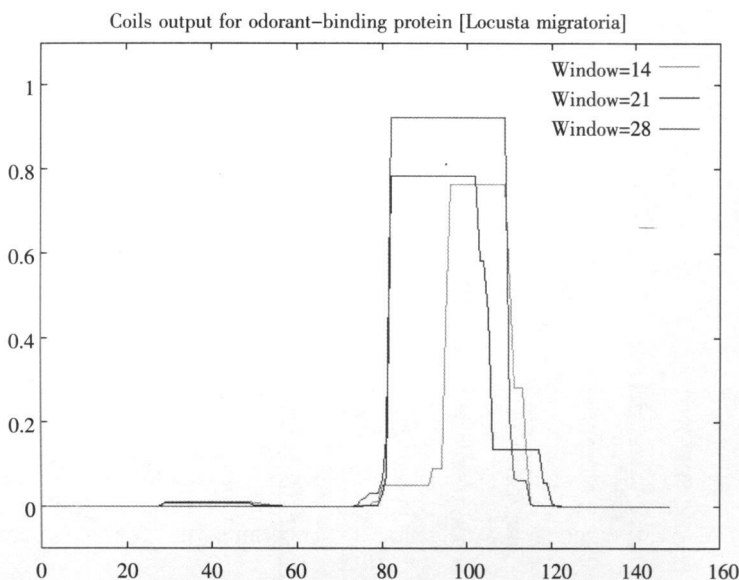


图6 COILS Server 软件分析 LmigOBP1 的卷曲螺旋预测结果
Fig.6 The coiled coil analysis of LmigOBP1 by using COILS Server

2.1.5 蛋白质亚细胞定位分析

利用网上在线的蛋白质亚细胞定位分析软件 PSORT II Prediction 对东亚飞蝗气味结合蛋白进行预测,结果显示其在各细胞器分布比例为:线粒体 39.1%,细胞质 21.7%,高尔基体 17.4%,内质网 17.4%,液泡 4.3%。

2.1.6 蛋白质 Domain分析

运用蛋白质 Domain 分析工具 SMART 分析 LmgOBP1, 结果显示该蛋白包括第 1~23 位的信号肽序列, 第 33~135 位的 PhBP 区域, 以及第 137~148 位的固有无序化区域, 其中的 PhBP 区域被证实是昆虫信息素/气味结合蛋白结构域。

2.1.7 蛋白质二级结构预测

蛋白质二级结构预测工具 PredictProtein 分析 Lm igOBP1 结果如下: 预测 α 螺旋 (H)、 β 折叠 (E) 和无规则卷曲 (空白部分) 的比例分别是 62.16: 1.35: 36.49, 可见 α 螺旋是其中最大量的结构元件 (如图 6 所示); 预测其在第 95 位氨基酸存在蛋白激酶 C 磷酸化位点, 在第 22 39 44 95 位氨基酸存在酪氨酸激酶 II 磷酸化位点.



图7 PredictProtein 对于 LmigOBP1 的二级结构预测结果

Fig.7 The secondary structure analysis of LmigOBP1 by using PredictProtein

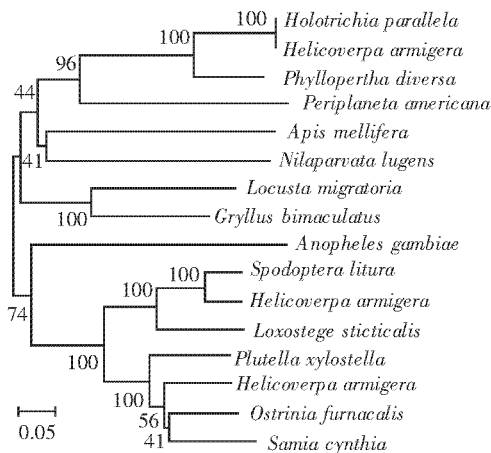


图 9 东亚飞蝗、双斑蟋、美洲大蠊、褐飞虱、暗黑鳃金龟、豆黄鳃金龟、异发丽金龟、疟蚊、蓖麻蚕、棉铃虫、斜纹夜蛾、烟夜蛾、小菜蛾、亚洲玉米螟、草地螟、意蜂 16 个物种的系统发生树 (NJ tree)

Fig.9 Neighbor-Joining (NJ) tree of 16 species (*Locusta migratoria*, *Gryllus bimaculatus*, *Periplaneta americana*, *Nilaparvata lugens*, *Holotrichia parallela*, *Heptophylla picea*, *Phyllopertha diversa*, *Anopheles gambiae*, *Samia cynthia*, *Helicoverpa armigera*, *Spodoptera litura*, *Helicoverpa assulta*, *Plutella xylostella*, *Ostrinia furnacalis*, *Loxostege sticticalis* and *Apis mellifera*) inferred from amino acid sequence of odorant-binding protein

报道相对较少, 对其与寄主植物之间的化学通讯研究比较缺乏.

东亚飞蝗是中国著名的农业害虫. 飞蝗蝗蝻与成虫都具有趋向阳光和群聚拥挤的行为习性. 飞蝗蝗蝻的聚集群行为、群行跳跃活动、零星迁飞和成群迁飞等行为以及聚集信息素的产生和作用均在变型的决定、发生、形成、发展、转变和稳定的过程中起着重要作用^[25]. 在东亚飞蝗的整个生活史中, 气味结合蛋白在其觅食、觅偶、交配、产卵等行为中均发挥重要作用. 本文利用日益完善的生物信息学技术对已测得的 Lm α OBP1 的核苷酸与蛋白质序列进行分析, 以期有助于研究蝗虫蛋白质结构与功能的关系、发现类似功能基因、揭示基因起源与分子进化规律等. 从对于该蛋白的多项分析结果可见, 东亚飞蝗与其它昆虫同类蛋白的相对分子质量、理论等电点、酸性与碱性氨基酸比例、极性与疏水氨基酸比例基本一致, 但其中东亚飞蝗气味结合蛋白的极性氨基酸含量最低, 而疏水氨基酸含量最高; 其蛋白质结构也与诸多昆虫气味结合蛋白相关报道^[9-11]比较吻合, 其中的 PhBP 区域被证实是昆虫信息素/气味结合蛋白结构域.

进一步研究该蛋白需要大量的纯化蛋白, 而昆虫触角较小, 直接分离纯化大量气味结合蛋白用于研究比较困难. 随着生物化学和分子生物学技术的快速发展, 我们可以利用基因克隆和表达的方法获得大量的昆虫气味结合蛋白. 后续展开的对昆虫气味结合蛋白的更多方面研究, 可以揭示昆虫与环境化学信息联系的本质、昆虫对环境化学因子刺激的感受机制等, 有助于探究有害昆虫治理和资源昆虫利用的途径.

[参考文献]

- [1] 李卫华, 涂洪涛, 苗雪霞, 等. 昆虫嗅觉相关蛋白的研究进展 [J]. 昆虫知识, 2006, 43(6): 757-762
- [2] 缪森, 张龙, 王丽红. 昆虫气味结合蛋白研究进展 [J]. 云南大学学报: 自然科学版, 2003, 25(增刊): 161-165.
- [3] 钱荷英, 徐安英. 昆虫信息素结合蛋白研究进展 [J]. 安徽农学通报, 2007, 13(23): 128-130
- [4] 王桂荣, 郭予元, 吴孔明. 昆虫触角气味结合蛋白的研究进展 [J]. 昆虫学报, 2002, 45(1): 131-137.
- [5] 王桂荣, 吴孔明, 郭予元. 昆虫感受气味物质的分子机制研究进展 [J]. 农业生物技术学报, 2004, 12(6): 120-126.
- [6] 修伟明, 董双林, 王荫长. 昆虫信息素结合蛋白及其分子运输机制和生理功能研究进展 [J]. 昆虫学报, 2005, 48(5): 778-784.
- [7] Ban L, Scaloni Ambrosio C D, Zhang L, et al. Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria* [J]. CMLS Cell Mol Life Sci 2003, 60: 390-400
- [8] Jin X, Brandazza A, Navarrini A, et al. Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts [J]. CMLS Cell Mol Life Sci 2005, 62: 1156-1166
- [9] Watanabe Hayaki Tabunoki Hiroko M Ima Nami et al. Analysis of odorant-binding proteins in antennae of a geometrid species *Ascotis selenaria cretacea*, which produces lepidopteran Type II sex pheromone components [J]. Invert Neurosci 2007

7: 109-118

- [10] WangGuirong Wu Kongning, Guo Yuyuan. Partial Cloning and characterization of the cDNA of general odorant binding protein 1 gene in the antenna of *Helicoverpa armigera* [J]. Entomologia Sinica, 2001, 8(4): 289-297.
- [11] Lu Daguang Li Xiangui Li Xiaoxia et al. Identification and molecular cloning of putative odorant-binding proteins and Chem sensory protein from the bethylid wasp, *Sclerodema guani* Xiao et Wu [J]. Journal of Chemical Ecology, 2007, 33 1 359-1 375.
- [12] 巩中军, 原国辉, 郭线茹, 等. 烟实夜蛾触角普通气味结合蛋白 II cDNA 的克隆、序列分析及在大肠杆菌中的表达 [J]. 昆虫学报, 2005 48(1): 18-23.
- [13] 卜翠萍, 奚耕思, 梁爱萍, 等. 西方蜜蜂毒蕈碱型乙酰胆碱受体基因的生物信息学分析 [J]. 动物分类学报, 2009 34 (2): 301-308.
- [14] 金丰良, 董小林, 许小霞, 等. 斜纹夜蛾普通气味结合蛋白 II cDNA 的克隆及在原核细胞中的表达 [J]. 中国科学 C 辑: 生命科学, 2008 38(11): 1 076-1 183.
- [15] 王桂荣, 郭予元, 吴孔明. 棉铃虫普通气味结合蛋白 II 基因的表达及鉴定 [J]. 昆虫学报, 2002 45(3): 285-289.
- [16] 吴少英, 王桂荣, 吴孔明, 等. 烟青虫气味结合蛋白基因的克隆与序列分析 [J]. 中国农业科学, 2005 38(9): 1 817-1 824.
- [17] 钟国华, 李苗孟, 胡美英, 等. 斜纹夜蛾普通气味结合蛋白基因 ShGOBP1 的克隆及序列分析 [J]. 华南农业大学学报, 2008 29(2): 38-43.
- [18] 钟国华, 李苗孟, 胡美英, 等. 斜纹夜蛾触角气味结合蛋白基因 ShGOBP2 的克隆与序列分析 [J]. 华中农业大学学报, 2008 27(5): 578-584.
- [19] 钟涛, 尹姣, 刘怀, 等. 草地螟触角普通气味结合蛋白基因的克隆及序列分析 [J]. 植物保护, 2008 34(4): 31-36.
- [20] Ishida Y, Comela J, Leal W S. Identification and cloning of a female antenna-specific odorant-binding protein in the mosquito *Culex quinquefasciatus* [J]. Journal of Chemical Ecology, 2002 28(4): 867-871.
- [21] 朱彬彬, 姜勇, 牛长缨, 等. 家蝇气味结合蛋白基因 cDNA 片段的克隆与序列分析 [J]. 昆虫学报, 2005 48(5): 804-809.
- [22] Li Hongliang Zhang Yali Gao Qikang et al. Molecular Identification of cDNA, Immunolocalization, and Expression of a putative odorant-binding protein from an asian honey bee, *Apis cerana cerana* [J]. Journal of Chemical Ecology, 2008 34 1 593-1 601.
- [23] 李红亮, 聂文敏, 高其康, 等. 中华蜜蜂气味结合蛋白 ASP2 cDNA 的克隆及原核表达 [J]. 中国农业科学, 2008 41 (3): 933-938.
- [24] Meillour P N, Francois M C, Emmanuelle J. Identification and molecular cloning of putative odorant-binding proteins from the merican palm weevil *Rhynchophorus palm arum* L [J]. Journal of Chemical Ecology, 2004 30(6): 1 213-1 223.
- [25] 陈永林. 中国主要蝗虫及蝗灾的生态学治理 [M]. 北京: 科学出版社, 2007 1-361.

[责任编辑: 孙德泉]