

进化基因组学在昆虫天然免疫研究中的应用前景^{*}

朱顺义^{**} 应牡英

(中国科学院动物研究所 农业虫害鼠害综合治理国家重点实验室 北京 100080)

Evolutionary genomics-based approaches for studies of insect innate immunity. ZHU Shun-Yi^{*}, YING Mu-Ying
(State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100080, China)

Abstract Evolutionary genomics-based approaches are increasingly changing traditional research pattern in Entomology field. Studies on insect innate immunity do not rely only on experimental methods. New research directions will emerge from three model organisms (*Drosophila melanogaster*, *Anopheles gambiae* and *Apis mellifera*) whose full genome sequences have been determined. In this commentary, we discuss how to establish the full set of antimicrobial peptide genes of a specific insect species on the basis of the concept of evolutionary divergence and convergence; and how to employ evolutionary information to highlight putative functional surface of the Spatzle protein, a key component of Toll pathway controlling insect innate immunity response.

Key words Evolutionary genomics, innate immunity, model insects

摘要 整合基因组学和进化论而发展起来的进化基因组学正在逐渐改变传统昆虫学的研究模式。对昆虫天然免疫的研究已不再仅仅依靠实验学方法。3种全基因组序列被破译的模式昆虫(黑腹果蝇、冈比亚按蚊和意大利蜜蜂)将为这些研究引入新的方向。该文将以模式昆虫为代表,简要介绍如何利用进化上的趋同和趋异概念建立一特定昆虫物种的抗微生物肽基因蓝图;以及如何利用基因组数据和进化分析方法鉴定控制昆虫 Toll 信号通路关键组分 Spatzle 配体的进化优势位点。

关键词 进化基因组学, 天然免疫, 模式昆虫

在 21 世纪的后基因组时代, 破解新发现基因的功能以及演绎控制生命过程中的蛋白质和各种信号通路活动的分子机制对于阐明生命现象的本质及其靶向蛋白和通路的药物设计具有重要的理论价值和指导意义。在这诱人的研究领域, 进化信息将扮演关键的作用。而基于进化观点的实验性研究也将成为 21 世纪生命科学研究的新的亮点。3 种完全变态的模式昆虫(黑腹果蝇 *Drosophila melanogaster*、冈比亚按蚊 *Anopheles gambiae* 和意大利蜜蜂 *Apis mellifera*) 全基因组序列的确定首次为人类从进化的角度比较研究昆虫发育、分化、免疫和生态适应等重要生命科学的基本问题提供了良机^[1-3]。就昆虫天然免疫而言, 进化基因组学的引入将极大地帮助我们系统了解不同昆虫物种之间免疫适

应的分子机制以及多个免疫效应分子之间可能存在的网络联系。这类研究毫无疑问也将对人类的健康产生重要的影响。例如, 一特定昆虫物种抗微生物肽基因蓝图的建立能够直接提供发展抗感染性疾病药物的模板; 而昆虫免疫分子机理的阐明有望在分子和遗传方面为人类天然免疫的机理提供全新的视野。最后, 开发昆虫天然免疫的知识也将为人类设计新型微生物杀虫剂提供理论指导。本文将简要介绍进化基因组学在昆虫天然免疫研究方面的应用前景, 主要集中在免疫效应分子及控制其表达的 Toll

^{*} 国家自然科学基金(30570381; 9068009)。

^{**} Email: zhusy@ioz.ac.cn

收稿日期: 2006-12-20, 修回日期: 2006-12-27。

信号通路配体 2 个方面。

1 进化基因组学建立模式昆虫天然免疫效应分子的基因蓝图

抗微生物肽作为动物天然免疫主要的效应分子,能够快速高效清除病原微生物的感染。自从 20 世纪 80 年代瑞典科学家 Boman 等人首次从蝴蝶体内分离到抗微生物肽以来,昆虫已成为这类分子最丰富的来源^[4]。由于其腐生的生存环境,一种昆虫需要利用多种不同的抗微生物肽应付复杂病原物的潜在感染。以果蝇为例,传统的实验方法已经鉴定了 20 个抗微生物肽,包括抗真菌的 Drosomycin 和 Metchnikowin,抗革兰氏阴性细菌的 Cecropin 和 Drosocin,以及抗革兰氏阳性细菌的 Defensin 等^[5]。最近,运用蛋白质组学的方法对果蝇血淋巴的组分进行分析,发现了大量新的低分子量的果蝇免疫诱导分子^[6]。由于不同进化谱系的昆虫占据着不同的生态位并因此承受着不同微生物的选择压力,不同种类的昆虫已经发展了自身特有的抗菌免疫组分。法国 Entomed 公司从 100 种昆虫中分离和鉴定了 175 个抗微生物肽,从未发现有相同的肽分子^[7]。因此,地球上 200 万种昆虫构成了一个巨大的抗微生物肽资源库。充分利用这些资源将为新型抗生素的开发提供极其丰富的模板。

虽然生物化学和分子生物学等传统的实验性方法在寻找昆虫新的抗微生物肽研究中发挥了重要的作用^[8]。新近发展的实验基因组学和蛋白质组学的方法也获得了很大的成功^[6,9~11]。但是由于不同的刺激源和刺激方式往往诱导不同的抗微生物肽的表达,这种独特的表达方式使得单独依靠实验方法在 mRNA 和蛋白质水平上鉴定一特定昆虫全套抗微生物肽方面存在明显的缺点。以果蝇为例,昆虫病原真菌球孢白僵菌自然感染选择性激活了抗真菌肽 Drosomycin 的表达;而革兰氏阴性细菌胡萝卜软腐欧文氏菌优先诱导抗细菌肽 Diptericin 的表达;相反,注射球孢白僵菌的孢子则诱导了 Drosomycin 和 Diptericin 的同时表达^[12]。疟原虫

则诱导了溶菌酶的表达^[13]。因此,基于全基因组序列分析为核心的进化基因组学策略具有明显的优势。在这方面,Christophides 等人进行了有益的尝试。通过运用序列比对方法,他们分析了冈比亚按蚊的全基因组序列,鉴定了 242 个免疫相关基因。这一结果为深入研究冈比亚按蚊免疫系统对疟原虫的影响奠定了基础。但是由于这一研究主要利用了序列相似性信息导致仅有 8 个典型的抗微生物肽被发现^[14]。

从进化的角度看,一个物种在长期与病原微生物斗争中发展的特定抗微生物肽基因库是趋异和趋同进化共同作用的结果。具有显著序列相似性的同一折叠类型的多肽是趋异进化的产物。相反,进化趋同则是指多肽分子内特定的局部区域序列和/或结构的相似性。这些基序通常定义了不同折叠类型的多肽具有相似的功能特性。因此,以进化趋同和趋异概念为指导,通过分析相关模式昆虫的全基因组序列,我们就能够建立这些物种全套的抗微生物肽基因蓝图。例如,通过提取无脊椎动物抗微生物肽的共有序列特征,Yount 和 Yeaman 鉴定了不同折叠类型的抗微生物肽共有一特定的 γ -基序,这一基序介导了这些肽的生物活性,是功能趋同的结果^[15]。同时,在同一折叠类型中(如 CS $\alpha\beta$ 型抗微生物防御肽),所有在 γ -基序区域的点突变导致的功能多样性则是趋异进化的结果^[16]。运用这一策略,我们目前已经建立了 3 种模式昆虫的全套抗微生物防御肽基因(图 1)。进化分析这些防御肽分子揭示了抗真菌肽 Drosomycin 在冈比亚按蚊和意大利蜜蜂进化谱系发生了基因丢失,但在黑腹果蝇基因组完成了基因复制和阳性达尔文选择。抗真菌肽 Drosomycin 的丢失合理地解释了冈比亚按蚊对真菌高度敏感的原因^[17]。而它的复制和适应性进化则使得果蝇获得有效的抵抗真菌感染的能力。目前,运用相似的策略合并实验生物学方法我们正在建立上述 3 种模式昆虫的抗微生物肽的全基因组蓝图。这些研究不仅为新型抗生素的研制和开发提供模板和资源,同时也为从整体和网络的角度深入研究抗微生物肽介导

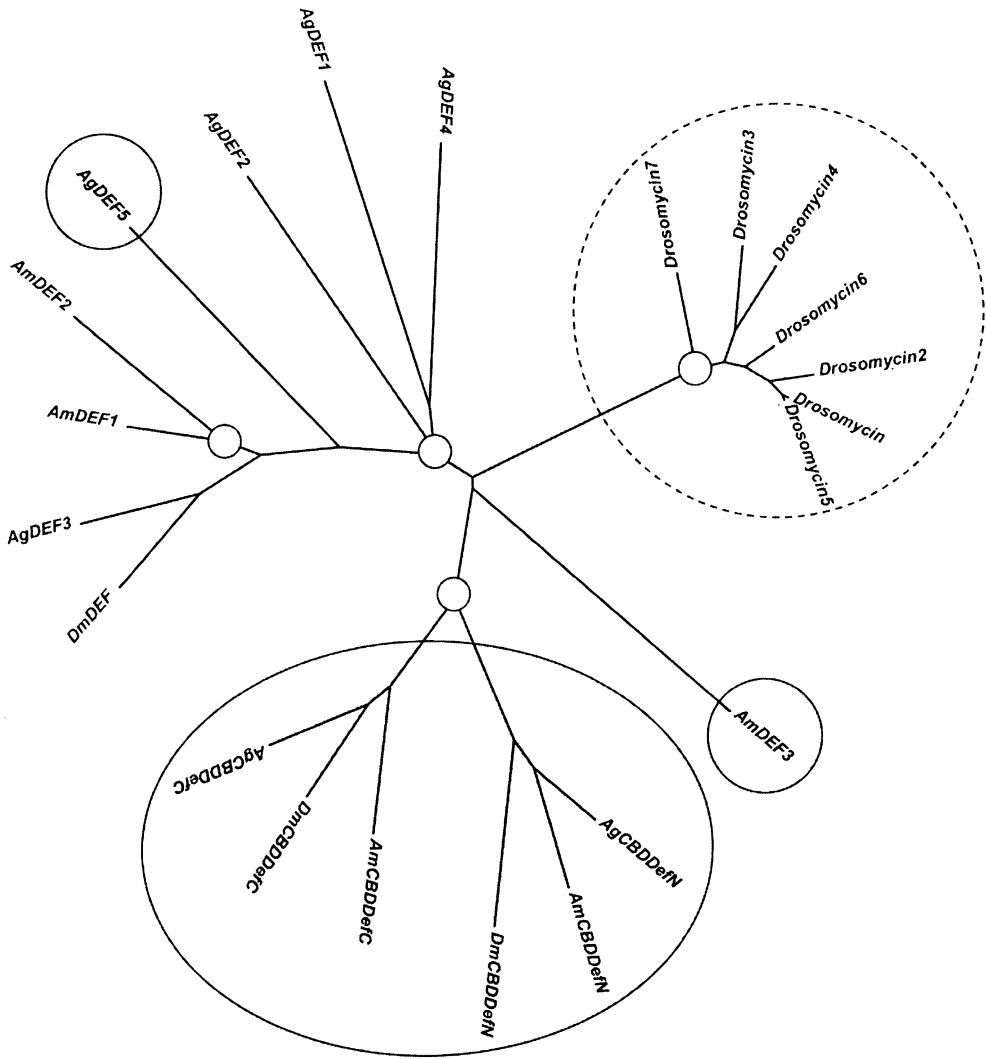


图 1 黑腹果蝇、冈比亚按蚊和意大利蜜蜂全套防御肽的进化树

运用进化基因组学方法搜索 3 种模式昆虫 Cs β 型抗微生物防御肽, 发现了 8 个全新的成员(用实线圈表示)。进化分析鉴定了 3 个模式昆虫抗微生物防御肽之间的进化关系(圆圈代表基因复制事件。其产生的后代形成了并系同源物)。虚线圈代表 Drosomycin 在黑腹果蝇基因组内的复制。Dm; *D. melanogaster* Ag; *A. gambiae* Am; *A. mellifera*

的昆虫天然免疫的调控机理以及从进化角度揭示昆虫生态适应的免疫学机制提供新的线索。

2 进化基因组学演绎昆虫天然免疫信号通路关键分子的功能表面

真菌(如绿僵菌和白僵菌)是作为广谱的微生物杀虫剂, 广泛用于控制农业害虫对作物的危害。但是, 由于一些种类的害虫体内存在天然的抗真菌机制, 使得这些真菌杀虫剂抗虫效

果并不理想。现在知道, 果蝇抗真菌免疫反应是通过 Toll 信号通路控制的 Drosomycin 直接靶向真菌来实现的^[18]。理论上讲, 竞争性抑制果蝇 Toll 通路的配体 Spatzle 与 Toll 受体的结合就能够降低或阻断 Drosomycin 的表达, 提高果蝇对真菌的敏感性, 从而增强真菌杀虫剂的效能。为了设计和开发这类竞争性抑制剂, 配体 Spatzle 蛋白与 Toll 受体相互作用的功能表面的

信息是关键。在这方面,模式昆虫基因组序列^[1-3]能够提供足够的信息用于进化指导 Spatzle 蛋白的突变研究。比较基因组学研究发现控制昆虫天然免疫反应的 4 种信号通路途径,包括 Toll, Imd, Janus kinase (JAK)/STAT 和 JNK 在这 3 种模式昆虫中是高度保守的,各个效应分子存在明显的直系同源的进化关系,尤其是细胞内效应分子可以找到 1:1 对应关系^[19]。先前的遗传学和生物化学研究已经表明果蝇 Spatzle 通过直接与 Toll 受体的结合而激活细胞内的信号级联放大过程并最终导致靶基因 Drosomycin 的表达^[20]。

为了开展这类研究,我们首先利用先前已

知的 6 个黑腹果蝇和 1 个冈比亚按蚊 Spatzle 成熟肽 C106 的氨基酸序列排列为基础^[21],建立了这一家族的特征序列基序,并用这一基序分别搜索了黑腹果蝇,冈比亚按蚊和意大利蜜蜂的蛋白质组序列。合并 TBLAST 搜索,我们发现了另外 11 个新的 Spatzle 基因(4 个从冈比亚按蚊,7 个从意大利蜜蜂)。以此数据为基础构建 Spatzle 蛋白的分子进化树并基于该树的进化分组提取相应的进化信息,最后将这些进化照亮的位点定位到 Spatzle 的模型结构上。结构分析证实这些进化上优势位点在结构上成簇排列,表明可能的功能作用(图 2)。该位点由 10 个氨基酸残基组成,其中 9 个残基定位在二

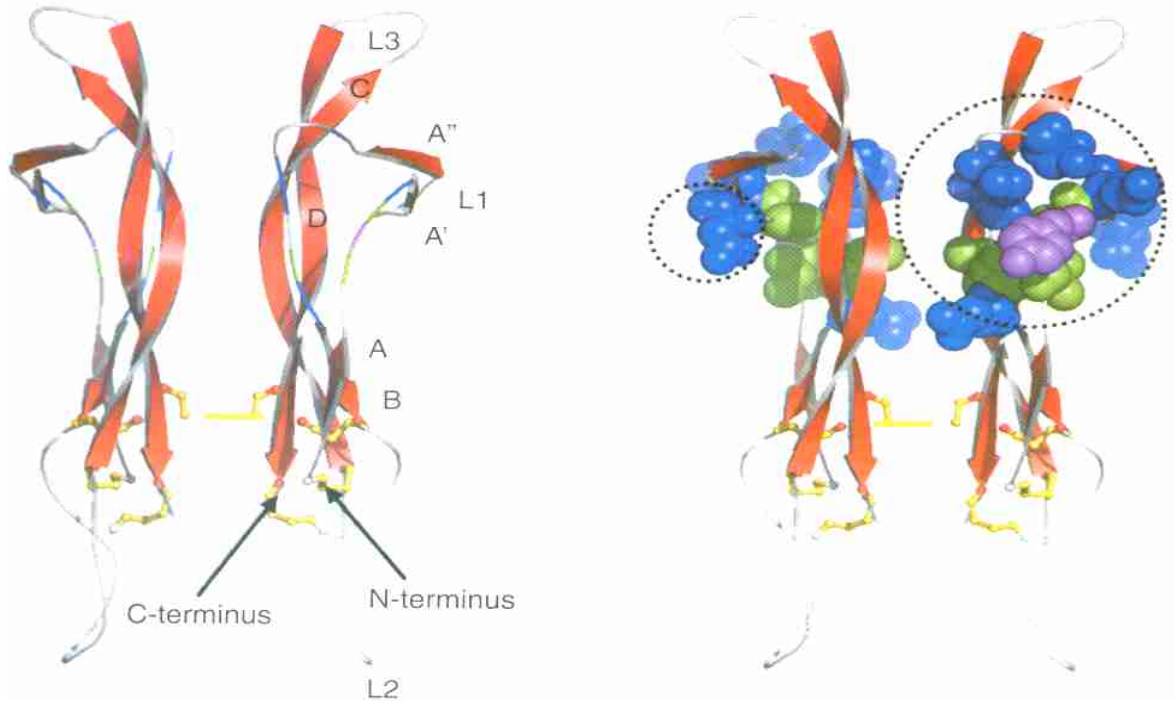


图 2 果蝇 Spatzle 的结构模型及进化印记分析鉴定的潜在功能表面

通过运用折叠识别和比较建模方法,我们获得了一个可靠的果蝇 Spatzle 的结构模型。其中 6 个保守的半胱氨酸残基形成的 3 对分子内二硫键构成了一个典型的胱氨酸结(Cystine Knot, CK)的结构基序。第 7 个半胱氨酸残基形成了 1 对分子间二硫键用于 2 个单体的连接。DSSP 用于定义 Spatzle 模型的二级结构元件。进化印记分析 3 个模式昆虫 Spatzle 鉴定了一个潜在的功能表面(见正文)。其步骤包括:(1)通用表达模式和 TBLASTN 用于基因组范围搜索 3 个模式昆虫 Spatzle 蛋白;(2)运用折叠识别和比较建模验证 Spatzle 新成员的 CK 结构;(3)多序列排列和进化树构建;(4)运用进化印记分析预测 Spatzle 家族进化上重要的结构和功能位点,并将这些位点定位到果蝇 Spatzle 的结构模型。(注:进化印记分析方法:该方法整合了序列、结构和功能信息。通过序列一致性的进化树将一个蛋白质家族近似地分成不同的功能亚组并鉴定印记位点。然后将印记位点残基定位到三维结构并基于其结构成簇来预测蛋白质家族潜在的功能位点^[22]。

级结构元件 β -链的连接区(A-A' 和 A''-B), 形成了一个典型的蛋白质结合表面(我们称为主功能区)。该表面中心包含了一个芳香族残基和 3 个疏水性残基, 周围是 5 个亲水性或电荷残基; 1 个定位于 L1 环区的酸性残基正好位于主功能区的反面。这可能有利于它和另一个单体 Spatzle 的主功能区一道通过电荷作用参与 Toll 受体的结合。为了评价这一进化优势位点的功能意义, 目前我们选用了定点突变策略人工修饰这些区域, 并用其靶基因 Drosomycin 的相对表达量作为突变的判定标准(Zhu *et al.*, unpublished data)。

3 前景与展望

在后基因组时代, 解析特定昆虫物种不同效应分子之间存在的网络联系以及控制这些联系的信号通路关键组分相互作用的分子机制构成了昆虫天然免疫研究的 2 个重要方向。基于进化基因组学确立的模式昆虫抗微生物肽的基因蓝图将为网络联系的研究提供基础的数据平台。虽然遗传学等传统的实验学方法在演绎控制天然免疫反应的信号通路方面取得了巨大的成功, 但是对于进一步阐明通路各组分的功能位点及其相互作用的分子机制则缺乏相应的能力。集中于进化信息分析照亮的潜在功能表面定点设计突变实验则具有明显的优势。本文仅仅是利用进化基因组学策略在基因组范围内开展模式昆虫天然免疫研究的 2 个例子。随着更多模式昆虫基因组序列的解析, 我们相信进化基因组学和实验生物学方法的整合将在后基因组时代的昆虫天然免疫研究中发挥越来越重要的作用。

参 考 文 献

- Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., *et al.* *Science*, 2000, **287**: 2 185~2 195.
- Holt R. A. Subramanian G. M., Halpern A., Sutton G. G., Charlab R., *et al.* *Science*, 2002, **298**: 129~149.
- The Honeybee Genome Sequencing Consortium. *Nature*, 2006, **443**: 931~949.
- Steiner H., Hultmark D., Engstrom A. B., Eernich H., Boman H. G. *Nature*, 1981, **292**(5 820): 292~246.
- Khush R. S., Lemaitre B. *Trends Genet.*, 2000, **16**(10): 442~449.
- Levy F., Rabel D., Charlet M., Bulet P., Hoffmann J. A. *Biochimie* 2004 **86**: 607~616.
- Andres E., Dimarcq J. L. *J. Intem. Med.*, 2004, **255**: 519~520.
- Shafer W. M. *Antibacterial Peptide Protocols (Methods in Molecular Biology)*. Humana Press. 1997. 32~35.
- Irving P., Ubecka J. M., Doucet D. Troxler L., Lagueux M. *Cell Microbiol.*, 2005 **7**(3): 335~350.
- Irving P., Troxler L., Heuer T. S., Belvin M., Kopczynski C. *Proc. Natl. Acad. Sci. USA*, 2001, **98**(26): 15 119~15 124.
- Vodovar N., Vinals M., Liehl P., Basset A., Degraud J. *Proc. Natl. Acad. Sci. USA*, 2005, **102**(32): 11 414~11 419.
- Lemaitre B., Reichhart J. M., Hoffmann J. A. *Proc. Natl. Acad. Sci. USA*, 1997 **94**(26): 14 614~14 619.
- Roxstrom-Lindquist K., Terenius O., Faye I. *EMBO Rpt.*, 2004 **5**(2): 207~212.
- Christophides G. K., Zdobnov E., Barillas-Mury C., Bimsey E., Blandin S., *et al.* *Science*, 2002, **298**: 159~165.
- Yount N. Y., Yeaman M. R. *Proc. Natl. Acad. Sci. USA*, 2004 **101**(19): 7 363~7 368.
- Zhu S., Gao B., Tytgat J. *Cell Mol. Life Sci.*, 2005 **62**(20): 2 257~2 269.
- Scholte E. J., Ng'habi K., Kihonda J., Takken W., Paaijmans K. *Science*, 2005 **308**: 1 641~1 642.
- Lemaitre B., Nicolas E., Michaut L., Reichhart J. M., Hoffmann J. A. *Cell*, 1996 **86**(6): 973~983.
- Evans J. D., Aronstein K., Chen Y. P., Hetru C., Imler J. L., *et al.* *Insect Mol. Biol.*, 2006, **15**(5): 645~656.
- Weber A. N., Tauszig-Delamasure S., Hoffmann J. A., Lelievre E., Gascan H., *Nat. Immunol.*, 2003, **4**(8): 794~800.
- Parker J. S., Mizuguchi K., Gay N. J. *Proteins*, 2001, **45**(1): 71~80.
- Zhu S., Tytgat J. *FASEB J.*, 2004 **18**(9): 940~947.