

文章编号: 1000-5013(2009)01-0053-03

杀虫晶体蛋白对小菜蛾杀虫活性的识别预测

蔡福营, 樊 云, 张光亚, 林 毅

(华侨大学 化工学院, 福建 泉州 362021)

摘要: 考察对小菜蛾具有杀虫活性和无杀虫活性的晶体蛋白二肽组成的差异. 结果发现, 对小菜蛾具有杀虫活性的晶体蛋白中, SG, LS, SN, NI, GS, SV, SL, TV, EF, NQ 二肽分数较高; 而对小菜蛾无杀虫活性的晶体蛋白中, AN, LQ, DA, AQ, TF, SY, EI, LE, PT, RA 二肽分数较高. 总体分数上存在差异最大的是 S* 型二肽. 在此基础上, 发展一种识别杀虫晶体蛋白(ICP) 对小菜蛾杀虫活性的统计学方法, 通过对两组共 27 条杀虫晶体蛋白进行有无杀虫活性的初步识别, 该方法识别平均正确率分别是 71.4% 和 100% .

关键词: 苏云金芽孢杆菌; 杀虫晶体蛋白; 二肽组成; 杀小菜蛾活性; 识别预测

中图分类号: Q 934. 124; Q 516. 04 **文献标识码:** A

苏云金芽孢杆菌(*Bacillus thuringiensis*, Bt) 是一种革兰氏阳性细菌, 它在形成芽胞时伴随着产生晶体蛋白. 有些晶体蛋白被证实具有特异的杀虫活性, 通常被称为杀虫晶体蛋白(Insecticidal Crystal Protein, ICP) ^[1]. 不同的杀虫晶体蛋白对鳞翅目、鞘翅目、双翅目、膜翅目等农业害虫具有特异的毒杀作用, 而对人、畜和害虫天敌是无害的^[2], 可作为绿色生物农药推广. 小菜蛾(*Plutella xylostella*) 属于鳞翅目, 是十字花科的重要害虫, 由于其发生世代多且重叠, 为害隐蔽, 且容易产生抗药性, 因此对小菜蛾的杀虫活性测试和生物农药的筛选变得越来越重要^[3]. 许多氨基酸组成相似的蛋白质在二肽组成上存在较大差异, 而且同一个氨基酸在发挥功能的过程中要受到与它在序列和空间上相近氨基酸的影响. 二肽能反映它们在上下列的相互影响, 二肽组成比氨基酸组成蕴含了更多的生物学信息^[4]. 本文以杀虫晶体蛋白的二肽组成为研究对象, 构建杀虫晶体蛋白对小菜蛾杀虫活性的预测模型.

1 材料与方法

1.1 数据来源

对相关的文献和在线数据库(www.lifesci.sussex.ac.uk) 进行查阅, 收集了 21 条对小菜蛾具有明显杀虫活性的晶体蛋白, 分别是 Cry1Aa14, Cry1Ab13, Cry1Ab15, Cry1Ab16, Cry1Ab17, Cry1Ac1, Cry1Ac14, Cry1Af1, Cry1Ba3, Cry1Bd, Cry1Ca9, Cry1Cb2, Cry1Fb5, Cry1Gbl, Cry1Ie1, Cry1Ia8, Cry1Id1, Cry2Aa9, Cry2Ab3, Cry2Ab4 和 Cry2Ab10. 另外, 收集了 6 条没有任何杀虫活性的晶体蛋白, 分别是 Cry1Ca7, Cry1Db2, Cry1Gb2, Cry1Ia9, Cry9Ea2 和 Cry9Ea3.

1.2 二肽组成分析

包含 n 个氨基酸的蛋白序列 $\langle a_1, a_2, a_3, \dots, a_n \rangle$, 其二肽数量为 $n-1$ 个, 分别表示为 $\langle a_1, a_2 \rangle, \langle a_2, a_3 \rangle, \dots, \langle a_{n-1}, a_n \rangle$. 二肽分数 $x_i = n_i / N$. 其中, i 表示 400 种二肽的某一种, n 表示序列中某种二肽数量, N 表示该序列中二肽总数量. 对 20 个具明显杀小菜蛾活性的晶体蛋白和 6 个无杀小菜蛾活性的晶体蛋白, 首先计算出每条序列的二肽组成, 然后计算出其平均值, 分别为 x_{iS} 和 x_{iW} . 其中, 下标 S 和 W 分别表示具杀小菜蛾活性和无杀小菜蛾活性. 二肽分析由 CDP 软件完成^[4].

收稿日期: 2008-02-13

通信作者: 林 毅(1976-), 男, 副教授, 主要从事应用与环境微生物的研究. E-mail: lyhxm@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(40601046); 福建省自然科学基金资助项目(B0510011); 福建省高等学校新世纪优秀人才计划支持项目(2006)

1.3 小菜蛾杀虫活性的预测原理

对于一个未知杀虫活性的晶体蛋白, 首先计算出其 400 种二肽分数; 然后, 计算出它和具有杀小菜蛾活性晶体蛋白的 400 种二肽的绝对偏差之和 α_s , 以及它与无杀小菜蛾活性晶体蛋白 400 种二肽的绝对偏差之和 α_w . 若 $\alpha_s < \alpha_w$, 则视为有杀小菜蛾活性; 反之, 则视为无杀小菜蛾活性.

2 结果与分析

2.1 杀虫晶体蛋白二肽组成分析

CDP 分析结果显示, 影响小菜蛾有无杀虫活性的 ICP 二肽组成有较大差异, 如图 1 所示. 由图 1(a) 可知, 有 22 种二肽在有杀小菜蛾活性的晶体蛋白中, 其二肽分数大于无杀小菜蛾活性的晶体蛋白. 这

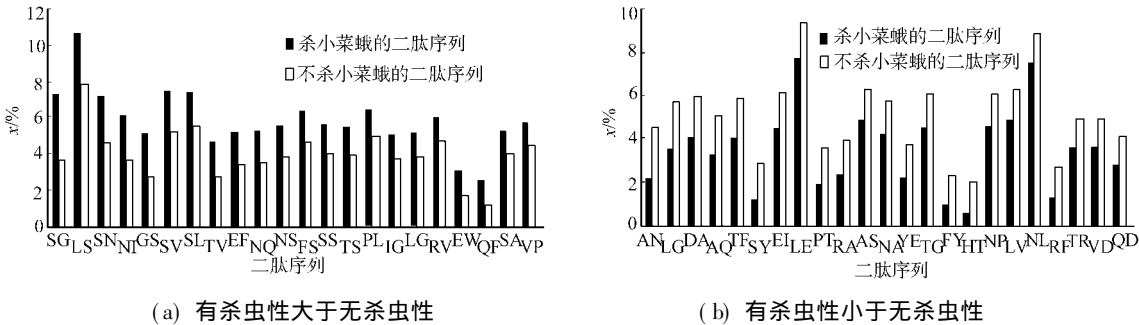


图 1 小菜蛾杀虫活性二肽组成分析

Fig.1 The analysis of the dipeptide composition

些二肽是 SG, LS, SN, NI, GS, SV, SL, TV, EF, NQ, NS, FS, SS, TS, PL, IG, LG, RV, EW, QF, SA 和 VP. 而从图 1(b) 可知, 有 23 种二肽正相反, 分别是 AN, LQ, DA, AQ, TF, SY, EL, LE, PT, RA, AS, NA, YE, TG, FY, HT, NP, LV, NL, RF, TR, VD 和 QD.

具杀小菜蛾活性的晶体蛋白中, 二肽分数较高的丝氨酸 Ser(S), 在与杀虫能力呈正相关的 20 种二肽中出现了 11 次. 无杀小菜蛾活性的晶体蛋白中, 二肽分数较高的天氨酸 Asp(D) 和丙氨酸 Ala(A), 所形成的二肽其出现的频率也较高. 按 400 个二肽中第 1 个氨基酸组成将其分为 20 类, 分别表示为 A^* , C^* , ..., Y^* , “*” 表示 20 个氨基酸中的任一个. 这 20 类二肽在具杀小菜蛾活性和无活性的晶体蛋白中, 二肽分数存在较大差异. 在有杀小菜蛾活性的晶体蛋白中, S^* 型二肽分数最高, N^* 次之; 而在无活性晶体蛋白中, A^* 二肽最高, D^* 次之(图 2). 丝氨酸能够形成二硫键, 促进杀虫晶体蛋白结构的稳定. 它也是一种很强的亲水氨基酸, 可以增强杀虫晶体蛋白与害虫中肠膜上受体糖链结合. 这些因素可能是 S^* 型二肽在具有杀虫活性的晶体蛋白中二肽分数高的原因.

2.2 小菜蛾杀虫活性识别预测

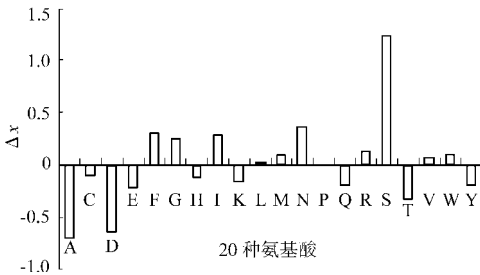
对 27 条晶体蛋白的序列进行计算, 结果如表 1 所示. 由表 1 可知, $\alpha_s > \alpha_w$ 的 12 个序列为 Cry1Ba3, 表 1 小菜蛾杀虫活性预测结果

Tab.1 Predictive results of the insecticidal activity against *Plutella xylostella*

ICP	α_s	α_w	ICP	α_s	α_w	ICP	α_s	α_w	ICP	α_s	α_w
Cry1Aa12	27.31	36.94	Cry1Ac14	25.35	37.62	Cry1Gb1	34.29	33.84	Cry1Ca7	32.97	31.65
Cry1Aa14	25.88	35.58	Cry1Af1	37.72	49.51	Cry1Ie1	52.53	50.62	Cry1Db2	34.49	32.88
Cry1Ab13	25.36	38.91	Cry1Ba3	34.35	33.56	Cry1Ia8	50.06	47.36	Cry1Gb2	34.50	34.09
Cry1Ab15	25.36	38.91	Cry1Bd1	35.49	36.56	Cry1Id1	51.53	50.58	Cry1Ia9	49.88	46.93
Cry1Ab16	25.69	39.21	Cry1Ca9	33.04	31.65	Cry2Aa9	64.23	72.88	Cry9Ea2	50.08	34.09
Cry1Ab17	24.71	38.59	Cry1Cb2	32.90	35.36	Cry2Ab3	62.63	70.26	Cry9Ea3	50.13	33.91
Cry1Ac	24.78	37.46	Cry1Fb5	33.24	37.19	Cry2Ab4	62.61	70.18			

图 2 20 类氨基酸在杀小菜蛾蛋白分数差异

Fig.2 Percent difference of the twenty amino acid residues of the insecticidal crystal protein



Cry1Ca9, Cry1Gb1, Cry1Ie1, Cry1Ia8, Cry1Id1, Cry1Ca7, Cry1Db2, Cry1Gb2, Cry1Ia9, Cry9Ea2, Cry9Ea3. 即这些晶体蛋白没有杀虫活性. 6 个序列 Cry1Ba3, Cry1Ca9, Cry1Gb1, Cry1Ie1, Cry1Ia8, Cry1Id1 是具有杀虫活性的序列, 占全部 21 条有杀虫活性的 28. 6%, 杀虫活性预测正确率为 71. 4%. 对于 6 条无杀小菜蛾活性的晶体蛋白序列, 全都能正确识别, 正确率为 100%, 预测结果良好.

3 讨论

对晶体蛋白的杀虫活性分析一般采用传统的生物学测定法, 目前已建立了多种目标昆虫的生物测定模式^[5]. 生物学测试过程繁琐, 耗时、耗力、耗财, 考虑很多因素, 如害虫本身的生殖、生存潜能等生物学特性, 以及温度、湿度和光照等主要环境因子, 且对于一些害虫仍未建立一套行之有效的方法^[6]. 林毅等^[7]利用均匀设计方法, 构建了晶体蛋白氨基酸组成与其对鳞翅目、鞘翅目和双翅目的活性之间关系的支持向量机模型, 并得到了很好的预测结果. 但该模型无法具体应用于杀小菜蛾活性的预测.

通过考察晶体蛋白的二肽组成特征与其对小菜蛾的杀虫活性之间的内在联系, 提供了一种识别晶体蛋白对小菜蛾杀虫活性的新方法. 该方法在对杀小菜蛾晶体蛋白实施改造的过程中, 可作为一种实验前期预测方法, 以期减少人工筛选工作量, 提高研究效率和改造成功率. 目前, 通过实验方法对小菜蛾生物学测试的数据量有限, 较少的数据量对所建模型预测结果会有一定的影响. 本模型只是建立在简单的统计学基础上, 如果能找到一种更好的判别公式, 势必会大大提高模型的预测精度.

参考文献:

[1] HOFTE H, WHITELEY H. Insecticidal crystal proteins of *Bacillus thuringiensis*[J]. Microbiol Rev, 1989, 53 (36) : 242-255.

[2] CORINA M, BERON N. New strategy for identification of novel Cry type genes from *Bacillus thuringiensis* strains [J]. Applied and Environmental Microbiology, 2005, 9(13) : 761-765.

[3] TABASHINK B, FINSON N, JOHNSON M. Cross resistance to *Bacillus thuringiensis* toxin Cry1F in the diamond back moth[J]. Appl Environ Microbiol, 1994, 60(12) : 462-467.

[4] 张光亚, 方柏山. 基与二肽组成识别嗜热和常温蛋白的研究[J]. 生物工程学报, 2006, 22(2) : 293-298.

[5] 高梅影, 李荣森, 戴顺英. 苏云金芽孢杆菌抗蝗虫生物测定模式建立[J]. 微生物学报, 1999, 39(6) : 515-520.

[6] 倪珏萍, 曹 斌. 关于苏云金芽孢杆菌不同检测方法的探讨[J]. 世界农药, 2002, 24(6) : 25-26.

[7] 林 毅, 蔡福营, 张光亚. 苏云金杆菌杀虫晶体蛋白活性预测的支持向量机模型[J], 生物工程学报, 2007, 23 (1) : 127-132.

The Discrimination Prediction with Insecticidal Activity of Insecticidal Crystal Protein Against *Plutella xylostella*

CAI Fu-ying, FAN Yun, ZHANG Guang-ya, LIN Yi

(College of Chemical Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: In this work, the dipeptide composition of crystal proteins with and without insecticidal activity against *Plutella xylostella* was systematically analyzed. We found that crystal proteins with insecticidal activity against *Plutella xylostella* contained more dipeptides such as SG, LS, SN, NI, GS, SV, SL, TV, EF and NQ, whereas less dipeptides such as AN, LQ, DA, AQ, TF, SY, EI, LE, PT and RA, and generally, S* was the most different dipeptide. Based on this information, a statistical method for discriminating 27 crystal proteins with or without insecticidal activity against *Plutella xylostella* was developed and their accuracy was 71.4% and 100%, respectively.

Keywords: *Bacillus thuringiensis*; insecticidal crystal proteins; dipeptide composition; activity against *Plutella xylostella*; discrimination

(责任编辑: 黄仲一 英文审校: 陈国华)