研究报告

秀丽小杆线虫分泌蛋白组的计算机分析

吴红芝[1, 2], 李成云[2], 朱有勇[2], 毕玉芬[3]

云南农业大学1.园林园艺学院；2.教育部生物多样性与病虫害控制重点实验室；3.动物科学技术学院，昆明　650201

摘要

结合计算机技术和生物信息学的方法，采用组合的信号肽分析软件SignalP v3.0、TargetP v1.01、Big-PI Predictor和TMHMM v2.0，预测了秀丽小杆线虫（Caenorthaditis elegans ws123）的全基因组19855个ORF编码蛋白的信号肽，同时系统分析了信号肽的特征。结果表明，在19855个秀丽小杆线虫的蛋白中，有1990条为带有信号肽的分泌型蛋白，其中，1936条为典型的分泌型信号肽（即信号肽酶Ⅰ型信号肽），53条为信号肽酶Ⅱ型信号肽，1条为信号肽酶Ⅳ型信号肽；在Ⅰ型信号肽中，有41条为RR-motif亚组型信号肽。在1990条信号肽中，有742条没有典型的N-区，其余1248条包含典型的3个区。比较了秀丽小杆线虫与原核生物分泌蛋白信号肽中20种氨基酸残基在信号肽酶切位点的使用情况，表明：在Ⅰ型信号肽酶切位点，其信号肽中使用的氨基酸总体趋势与原核生物基本相似，但秀丽小杆线虫选用的氨基酸种类更多，变化更大；在Ⅱ型信号肽酶切位点，秀丽小杆线虫脂蛋白信号肽中使用的氨基酸的种类与原核生物有很大的不同。通过与真核单细胞生物比较，作为真核多细胞生物的秀丽小杆线虫，其分泌蛋白信号肽所占比例更高、种类更多，可知线虫信号肽的组成具有很高的多态性，表明该物种的分泌蛋白具有多种功能。此外，分析结果显示，脂蛋白信号肽在结构上比分泌型信号肽更为保守。在秀丽小杆线虫分泌蛋白中出现了少数氨基酸组成完全一致的信号肽，采用BLAST 2 SEQUENECES对具有相同信号肽的分泌蛋白进行了序列比对，结果表明具有相同信号肽的分泌蛋白同源性非常高，它们的存在是生物进化过程中基因倍加（duplication）及环境选择的结果，信号肽特征的详细描述必将对这些蛋白功能的研究提供重要的帮助。

关键词　秀丽小杆线虫；分泌蛋白组；信号肽；分泌途径

分类号

Computational Analysis of Signal Peptide- dependent Secreted Protein in *Caenorthaditis elegans* ws123

WU Hong-Zhi[1,2], LI Cheng-Yun[2], ZHU You-Yong[1], BI Yi-Fen[3]

1. *The College of Landscape and Horticulture* ; 2. *the Ministry of Education Key Laboratory for Agricultural Biodiversity for Pest Management*, 3. *The College of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201 China*

Abstract

<DIV>The internet-based softwares SignalP v3.0, TargetP v1.01, big-PI predictor and TMHMM v2.0 were combined to predict the signal peptides and the signal peptide-dependent secreted proteins from the 19855 ORFs in Caenorthaditis elegans ws123 genome. 1990 proteins were predicted to be secreted and to contain signal peptides among 19855 proteins, among which 1936 have SignalPase Ⅰ signal peptide(containing 41 with RR-motif signal peptide), 53 have SignalPase Ⅱ signal peptide and one has SignalPase Ⅳ signal peptide. The signal peptides of 742 secreted proteins include only H-domain and C-domain, but no typical N-domain; the signal peptides of other 1248 secreted proteins include all three domains. Although the amino acids constitution of the SignalPase Ⅰ signal peptides were similar in general between Caenorthaditis elegans and prokaryote, there were apparently small

differences, and the amino acid composition of Caenorthaditis elegans are more diverse and less conserved. But there are distinct differences on the amino acid composition of SignalPaseⅡ signal peptides. The signal peptides of Caenorthaditis elegans were more diverse than unicellular eukaryotic organism. The signal peptides of a few proteins were exactly the same. We used the BLAST 2 SEQUENECES aligning method to compare the homology among the secreted proteins with the same signal peptides. The alignment results indicated that the genes sharing the same signal peptide sequences were homologous to each other and were likely to have arisen from gene duplication.</DIV>

通讯作者 毕玉芬 biyifenynna@sina.com