

大白菜 microRNA 和靶基因的预测与鉴定

王金彦, 侯喜林*, 杨学东

(农业部华东地区园艺作物生物学与种质创制重点实验室/作物遗传与种质创新国家重点实验室, 南京 210095)

MicroRNAs (miRNAs) 是一类新鉴定的、非编码蛋白的单链小分子 RNA。大量研究表明 miRNA 在植物中通过在转录后对靶基因的降解或者抑制蛋白的翻译的方式来调控基因的表达。目前在 miRBase 数据库 (www.mirbase.org/) 中已经鉴定出数千个植物 miRNA 分子。大白菜 [*Brassica campestris* L. ssp. *pekinensis* (Lour.) Olsson] 是中国和东亚种植面积最广的蔬菜作物之一, 但是在 miRNA 数据库中仅报道有 23 个大白菜 miRNAs。

在本研究中通过比较基因组学方法, 利用 NCBI 公共数据库 (<http://www.ncbi.nlm.nih.gov/>) 中 EST (Expressed Sequence Tags) 和 GSS (Genomic Survey Sequences) 序列, 鉴定出大白菜 miRNA。

利用 NCBI 中的大白菜 EST 和 GSS 数据、miRBase 库中的 miRNA 序列预测大白菜中 miRNA; 利用非冗余蛋白数据库、KEGG (Kyoto Encyclopedia of Genes and Genomes) 数据库、GO (Gene Ontology) 数据库对 miRNA 的靶基因进行功能注释。利用 BLAST 比对程序、WATER 比对程序、CAP3 序列拼接程序、RNAfold 二级结构预测程序进行 miRNA 预测。miRNA 鉴定的标准: 不超过 3 个错配; 长度不小于 45 nt; 形成完美的发夹结构; miRNA 与 miRNA* 之间不超过 6 个错配; miRNA 与 miRNA* 之间不能出现颈环结构。miRNA 靶基因的鉴定标准: 不能超过 4 个错配; 不能连续出现 2 个错配; 1~9 nt 不能超过 1 个错配; 第 10 和 11 nt 不能有错配; 不能有 gap 存在。

利用自主研发的 PERL 程序以及整合的 miRNA 预测流程对大白菜中的 miRNA 进行预测, 共鉴定出 168 个大白菜 miRNA, 分别来自于 22 个 EST 和 119 个 GSS 序列。这 168 个 miRNA 隶属于 38 个 miRNA 家族, 其中 miR1533、miR156 和 miR2911 家族所占的比例最大。通过生物信息学分析, 发现 129 个 miRNA 作用于 1386 个大白菜靶基因。这 1386 个基因参与多个生物和代谢途径, 其中包括细胞增长、信号转导、胁迫响应和植物发育等。对靶基因的 GO 分析表明 688 个基因与分子功能相关, 532 个基因参与生物过程, 287 个基因与细胞组分相关。KEGG 分析表明, 大白菜中 miRNA 参与 214 个代谢途径, 包括植物—病原菌相互作用、脂肪酸代谢、氨基酸代谢、硝酸盐代谢、植物激素信号转导等。

结论: miRNA 在植物的生长发育、响应生物和非生物胁迫反应、信号转导、蛋白降解中发挥着重要作用。通过生物信息学分析, 共发现大白菜中存在有 168 个保守 miRNA, 经过分析发现这些 miRNA 作用于 1386 个白菜靶基因, 这些靶基因参与细胞增长、信号转导、胁迫响应和植物发育等过程。

关键词: 大白菜; miRNA; 比较分析; 靶基因; 代谢途径

中图分类号: S 634.1

文献标识码: A

文章编号: 0513-353X (2011) S-2534-01

收稿日期: 2011-08-18

基金项目: 江苏高校优势学科建设工程项目; 中央高校基本科研业务费专项资金项目 (KYZ200912)

* 通信作者 (E-mail: hxl@njau.edu.cn; Tel: 025-84395917)