

辣椒推定 WRKY 基因 cDNA 的生物信息学分析

王育娜^{1,4}, 贺俐², 曹蕾³, 赖燕¹, 牟少亮¹, 何水林¹

(¹福建农林大学生命科学学院, 福州 350002; ²井冈山大学生命科学学院, 吉安 343009;

³漳州农业学校, 漳州 363000; ⁴洛阳师范学院生命科学系, 洛阳 471022)

摘要:【研究目的】该研究在利用同源克隆法分离到一个推定的辣椒 WRKY 转录因子的基础上, 采用多种生物信息学方法对该序列进行特征分析和功能预测, 以获得该基因及其编码蛋白更多的功能提示。【主要方法】应用 Blast、ORF Finder、Wolf Psort、DNAMAN 等软件或数据库, 进行序列的相似性比较, 开放读码框预测、保守域和编码蛋白质的功能等分析。【研究结果】所得到的序列是 WRKY 基因家族中的一个全长 cDNA 序列, 与其它植物具有较高的同源性, 该基因编码的蛋白具有 WRKY 保守结构域和锌指结构特征, 亚细胞定位预测结果显示该蛋白定位于细胞核, 具有 WRKY 转录因子的三维结构特征。【结论】所获得序列为辣椒 WRKY 基因家族的新成员, 具有 WRKY 转录因子家族的基本特征, 进化上有高度保守性, 可能在调控植物的生长发育以及干旱、高温、高盐等逆境应答过程中起重要调节作用。因此, 这一辣椒 WRKY 新基因有进一步研究的价值。

关键词:辣椒; WRKY; 转录因子; 生物信息学

中图分类号: S1 **文献标识码:** A

Bioinformatic Analysis of a Pepper Putative WRKY Transcriptional Factor

Wang Yuna^{1,4}, He Li², Cao Lei³, Lai Yan¹, Mou Shaoliang¹, He Shuilin¹

(¹College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002;

²College of Life Science, JingGangShan University, Jiangxi, 343009;

³Zhangzhou Agriculture School, Zhangzhou, Fujian 363000;

⁴Department of Life Science, Luoyang Normal University, Luoyang 471022)

Abstract: 【OBJECTIVE】A pepper cDNA library was screened to look for the WRKY transcriptional factor homologous molecules. Then a novel cDNA was found during the screen and cloned. Bioinformatics analysis was performed in order to study the possible function of this cDNA sequence. 【METHOD】Various bioinformatics methods including blastn, ORF Finder, Wolf Psort and DNAMAN were used to analyze and predict the feature and possible functions of the novel cDNA sequence. 【RESULTS】The result shows this cDNA sequence was the full-length of a novel gene and there was a conserved WRKY domain predicted by NCBI, the homologous analysis shows the putative WRKY sequence was highly identity with other plants and predicted to locate in nuclei. It may function in plant development and play a role during resistance to various environmental. Therefore, the functions of the novel gene are valuable for further investigations.

Key words: pepper, WRKY, transcriptional factor, bioinformatics

基金项目: 国家自然科学基金项目“PEAS 启动子中 G、W 盒结合蛋白 cDNA 的克隆与功能鉴定”(30571128)。

第一作者简介: 王育娜, 女, 1980 年出生, 河南洛阳人, 博士, 主要从事植物分子生物学研究。通信地址: 471022 洛阳师范学院生命科学系, Tel: 0379-65639522, E-mail: wangna799@sohu.com。

通讯作者: 何水林, 男, 1965 年出生, 湖南邵阳人, 教授, 博导, 副院长, 现主要从事植物分子生物学与基因工程研究。通信地址: 350002 福建农林大学生命科学学院, Tel: 0591-83789362, E-mail: hsl324@yahoo.cn。

收稿日期: 2008-09-24, **修回日期:** 2008-10-06。

WRKY 转录因子最初是由 Ishiguro 和 Nakamura 1994 年从甜马铃薯 (*Ipomoea batatas*) 中分离到的, 命名为 SPFI^[1]。随后, 在几种植物中发现了相似的蛋白, 如棉花中的 GaWRKY1^[2] 和欧芹的 PcWRKY1, 2, 3^[3]。这类蛋白最主要的结构特点是其 DNA 结合域至少都含有一个 WRKY 结构域, 该结构域包含大约 60 个氨基酸残基, N 端的七肽 WRKYGQK 十分保守, 但并不是一成不变, C 端有一个 C₂H₂ 或 C₂HC 的锌指结构^[4]。WRKY 蛋白在高等植物中构成一个基因超家族 (gene superfamily)。全基因组已经测序完成的拟南芥中有 74 个 WRKY 成员^[5], 水稻中有 100 余个成员^[6,7]。研究表明 WRKY 转录因子的基因工程在提高植物应对各种胁迫方面具有重要的作用^[8]。而关于辣椒 WRKY 转录因子在植物逆境中功能的报道相对较少。辣椒作为一种重要的蔬菜与经济作物, 不断受到外界环境的影响导致品质的降低^[9], 因此研究清楚辣椒体内 WRKY 转录因子的功能和作用模式, 达到综合改良植物抗性的目的就势在必行。

目前, 我们从辣椒 cDNA 文库中用同源克隆法获得一个基因, 经分析它含有 WRKY 保守结构域。为了解该基因所编码的蛋白的结构和功能, 我们对其进行了一系列生物信息学分析。

1 材料与方法

1.1 基因基本信息分析

用 NCBI 的 ORF Finder 软件 ([http://www.ncbi.](http://www.ncbi.nlm.nih.gov/gorf/gorf.html)

[nlm.nih.gov/gorf/gorf.html](http://www.ncbi.nlm.nih.gov/gorf/gorf.html)), 读码框翻译后分析最有可能的开放读码框。

1.2 核苷酸和氨基酸序列的相似性比较

应用 NCBI/ Blastn 程序, 即 (<http://www.ncbi.nlm.nih.gov>) 中的 Blastn 进行序列对库检索, 分析此序列与其它植物序列的同源性。NCBI/ blastx 程序和 NCBI/ tblastx 程序分析其氨基酸序列与其它植物序列的同源性。

1.3 功能分析

蛋白质的疏水性轮廓分析和二级结构预测采用 DNAMAN 软件 (<http://lynonn.com/download/index/html/>) 保守结构域分析用 NCBI 中的 CDD (Conserved Domain Database) 进行序列对库检索, 寻找蛋白质序列中存在的保守区域, 预测其功能。使用 Wolf Psort 在线软件 (<http://wolffpsort.org/>) 进行亚细胞定位预测。使用 ESyPred3D Web Server 1.0 预测候选基因蛋白结构。

2 结果与分析

2.1 基本信息分析

用 NCBI 的 ORF Finder 软件进行开放读码框的搜寻, 选出一条全长为 501 个核苷酸的开放读码框序列, 起始密码子为 ATG, 中止密码子为 TGA, 推测编码一个 166 个氨基酸, 分子量 19.034 kDa, 等电点 (PI) 5.93 的蛋白质。在 129-160 氨基酸之间有典型 C₂H₂ 锌指结构的序列 (图 1)。



图 1 CaWRKY 基因的全长 cDNA 序列及推定的氨基酸序列

2.2 核苷酸和氨基酸序列的相似性比较

在 NCBI 网站上对该基因的核苷酸序列和氨基酸序列进行 blast 同源序列搜索,发现在氨基酸水平上与其它 WRKY 蛋白, 如与烟草推定的 WRKY10 (CAI38917)和 WRKY11(CAI38918)有大约 78%的同源性,与葡萄属植物 *Vitis vinifera* 假定蛋白(CAN73664)

和 (CAO43843) 有 51%的同源性,与拟南芥 WRKY51 (NP-568995.2)有大约 55%的同源性。

2.3 保守结构域分析

通过用 NCBI 中的 CDD 程序分析发现此氨基酸序列在 98-163 的位置含有 WRKY 保守结构域 (图 2)。



图 2 辣椒 WRKY 蛋白保守域分析

2.4 蛋白质二级结构预测(H=helix,E=strand,-=no prediction):

-----H-----HEE--H--H-----
 EEE-----HHHHHH-EEEEH---HHH-
 HH-----E-----EEEEH-----

2.5 亚细胞定位分析

应用 Wolf Psort 软件对推定的 WRKY 基因进行亚细胞定位分析 (<http://wolfsort.org/>)。发现基因定位

于细胞核, 预测数据如下:k used for kNN is: 14, queryProtein details nucl: 12.5, cyto_nucl: 7.0。

2.6 疏水性分析

氨基酸侧链的一个重要特性是其(亲)疏水性,分析蛋白质的亲、疏水性模式可揭示某些蛋白质的结构和折叠信息,对 WRKY 蛋白进行疏水性分析(图 3)发现亲水氨基酸较多。

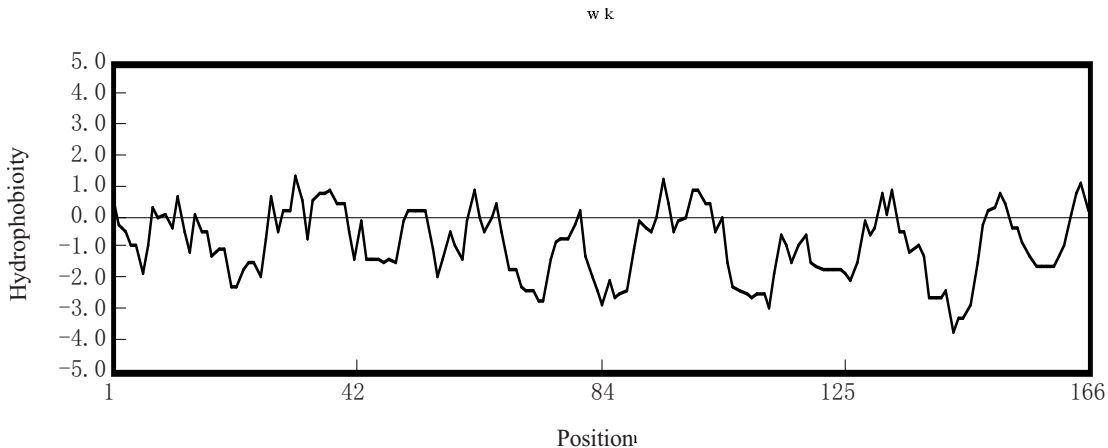


图 3 CaWRKY 推定蛋白的疏水轮廓

2.7 蛋白质三维结构预测

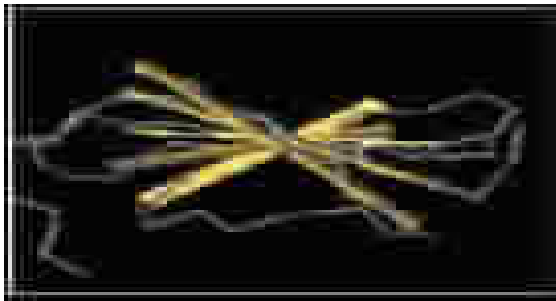


图 4 CaWRKY 蛋白三级结构的预测结果

3 讨论

生物信息学是 20 世纪末随着人类基因组计划的顺利进行而产生的,它是以计算机科学、数学、生物、物理

等多学科为理论基础形成的一门交叉性科学,在新基因的克隆,结构分析以及功能预测中发挥重要作用^[11]。

笔者从经紫外处理的辣椒果实 cDNA 文库中克隆出一个 WRKY CDNA 序列,在此基础上进行生物信息学分析,结果显示该基因编码含一个保守结构域的 WRKY 蛋白,一级结构分析表明,WRKY 与烟草及葡萄等物种的同源性较高,表明它是辣椒中一个新的 WRKY 成员,该蛋白核定位预测结果显示定位于细胞核信号,预示它是一个有功能的核调节蛋白。

综上所述,通过生物信息学分析我们初步揭示了 WRKY 及其编码产物的基本特性,提示其具有的重要生物学功能,为下一步深入探讨其功能和分子机制研究奠定了基础。

参考文献

- [1] Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol Gen Genet*, 1994, 244(6):563-571.
- [2] Rushton P J, Torres J T, Parniske M, et al. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parley PR1 genes. *EMBO J*, 1996, (15):5690-5700.
- [3] Xu Y H, Wang J W, Wang S, et al. Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)- δ -cadinene synthase-A. *Plant Physiol*, 2004, 135 (1): 507-515.
- [4] Eulgem T, Somssich I E. Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol*, 2007, 10(4):366-371.
- [5] Kalde M, Barth M, Somssich I E, Lippok B, et al. Members of the Arabidopsis WRKY group III transcription factors are part of different plant defense signaling pathways. *Molecular Plant Microbe Interactions*, 2003, 16(4):295-305.
- [6] Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary Biology*, 2005, 5(3):1-12.
- [7] Wu K L, Guo Z J, Wang H H, et al. The WRKY family of transcription factors in rice and Arabidopsis and their origins. *DNA RES*, 2005, 12(1):9-26.
- [8] Shimono M, Sugano S, Nakayama A, et al. Rice WRKY45 plays a crucial role in benzothiadiazole-inducible blast resistance. *Plant Cell*, 2007, 19(6):2064-2076.
- [9] 贾菊生. 新疆辣椒疫病及防治研究. *植物病理学报*, 1992, 22(3): 257-262.
- [10] Rehm BH. Bioinformatic tools for DNA/ protein sequence analysis, functional assignment of genes and protein classification [J]. *Appl Microbiol Biotechnol*, 2001, 57(5-6):579.
- [11] Leung AK, Andersen JS, Mann M, et al. Bioinformatic analysis of the nucleolus [J]. *Biochem J*, 2003, 376 (Pt 3): 553. *crobiol Biotechnol*, 2001, 57(5-6):5791.