

# 油菜 HY5 基因的电子克隆及生物信息学分析

郭继平<sup>1,2</sup>

(1.衡水学院 生命科学系,河北 衡水 053000;2.哈尔滨工业大学 食品科学与工程学院,黑龙江 哈尔滨 150090)

**摘要:**利用已知的芜菁 HY5 编码基因序列为探针,采用电子克隆的方法获得了油菜 HY5 基因 cDNA 的序列。采用生物信息学工具对其编码蛋白序列进行了研究,并将其编码蛋白同芜菁和拟南芥的 HY5 蛋白进行了同源性比较,构建了其系统发育树。结果表明:油菜 HY5 基因编码 164 个氨基酸,其分子量为 18 037.8 Da,是亲水蛋白,同芜菁 HY5 具有高度的同源性,高于和拟南芥的同源性;该试验初步获得了油菜 HY5 基因的编码序列,为进一步研究其功能打下基础。

**关键词:**HY5;油菜;光受体;电子克隆

**中图分类号:**S 634.3 **文献标识码:**A **文章编号:**1001-0009(2013)23-0116-03

光是影响植物生长发育的重要外界因子之一。植物对光的感受及光信号传导以及光形态建成一直是相关领域的热点研究内容之一<sup>[1-2]</sup>。植物通过众多调控因子来进行对光信号的感受和相应生理变化的调节。其中从拟南芥的突变体中克隆得到的 HY5 (Long Hypocotyl 5)是一个对植物光形态建成起重要调节作用的调控因子<sup>[3]</sup>,它参与了植物对红光<sup>[4]</sup>、蓝光和紫外光的感受和信号传导。

目前,少数一些植物如拟南芥、水稻等的 HY5 基因已经公布,但是总体还很有限,蔬菜植物芸薹属中芜菁 (*Brassica rapa*) HY5 的编码基因已经被克隆,但是油菜 (*Brassica napus*) HY5 编码基因序列还未被克隆。该研究采用已知的芜菁 HY5 编码基因 cDNA 为探针,利用已经公布的油菜表达序列标签 (EST) 数据库对油菜

HY5 基因进行了电子克隆,初步获得了其 cDNA 编码序列,并采用生物信息学的方法对其蛋白序列以及蛋白性质进行了研究。同时研究了其与芜菁和拟南芥 HY5 蛋白同源性,以及同其它物种 HY5 之间的进化关系,以期更好的研究油菜这种常见叶菜类蔬菜对光的感受,以更好地服务于生产实践提供理论支持。

## 1 材料与方法

### 1.1 试验材料

电子克隆的探针序列采用芜菁 (*Brassica rapa* subsp. *rapa* cultivar Tsuda) HY5 基因的 cDNA 序列 (GenBank: EU386772.1)。

### 1.2 试验方法

1.2.1 油菜 HY5 基因的电子克隆技术路线 种子序列选取芜菁的 HY5 基因的 cDNA 序列 (GenBank: EU386772.1), BLAST 搜索油菜: *Brassica napus* (taxid: 3708) EST 数据库,在线 (<http://pbil.univ-lyon1.fr/cap3.php>) 拼接得到的片段,得到的 contig 再分别对油菜 EST 数据库进行 BLAST 检索至不能进行延伸。将得到 contig 进行开放读码框的查找 (<http://www.ncbi>).

**作者简介:**郭继平(1979-),女,博士,副教授,研究方向为基因工程。E-mail: guojiping888@163.com.

**基金项目:**河北省高等学校自然科学研究计划资助项目 (Z2012045);河北省科学技术研究与发展计划资助项目 (12212702)。

**收稿日期:**2013-06-19

**Abstract:** Taking four varieties of peppers leaves of 'Niujiiao pepper', 'Qiemen pepper', 'Nanweijinfeng 2' and '8819 line pepper' as test materials, with water culture, the effect of 1 mmol/L and 0.01 mmol/L phosphorus (P) on peroxidase (POD) activity and their isoenzymic bands were studied. The results showed that POD activity in leaves of 'Niujiiao pepper', 'Qiemen pepper' and '8819 line pepper' increased firstly, decreased secondly and then increased again with time, while the POD activity of 'Nanweijinfeng 2' increased during the whole test time, and the changing trend of POD isozyme zymogram was accordance with POD activity. The enzymatic activity under low phosphorus was stronger than that of normal phosphorus treatment and the isozyme bands and expression quantity of POD were more than that of normal phosphorus treatment at the same time. The result suggested that POD was related with plant against low phosphorous stress.

**Key words:** peppers; low phosphorous stress; peroxidase (POD); isozyme

nlm.nih.gov/gorf/gorf.html),选取编码连续读码框最长且和种子序列编码蛋白长度接近的作为油菜 HY5 基因的电子克隆序列。

1.2.2 油菜 HY5 基因的生物信息学分析 利用多蛋白序列比对软件 Clustalw 2.1 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>)将油菜 HY5 蛋白和拟南芥、芜菁 HY5 进行相似性比对。利用 Protparam 程序对电子克隆得到的编码蛋白一级结构进行分析 (<http://www.expasy.org/tools/protparam.html>)。二级结构特点利用服务器 ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)) 进行分析。以 ProtScale 程序 (<http://www.expasy.org/cgi-bin/>

protscale.pl)进行疏水性分析。

## 2 结果与分析

### 2.1 油菜 EST 数据库的检索结果与序列拼接

以芜菁(*Brassica rapa*)的 HY5 基因搜索油菜 EST 数据库中得到 EST 片段进行拼接。得到 2 条 contigs,将 2 条 contigs 进行比对,发现 contig2 同种子序列芜菁 HY5 基因更为接近。因此,将此条序列(图 1)作为油菜 HY5 基因的电子克隆序列。

利用 NCBI 网站的开放读码框搜寻软件 ORF finder 对在线拼接成的油菜 HY5 基因的 cDNA 序列进行分析,得到一个长度为 495 bp 的开放读码框(图 1 中下划线的部分),编码 164 个氨基酸。

```
GATCCGTCAGATTCTCTATCAATCCGACGGCTGTAACACAGTAATCTATTCCTCCGACAAGACTC
GCAACTAGATTCTTCCCAGTTCTGTAAGTCCCGCTCTTTTCATCATCTCTATCTTCATCACCAGCTT
CTGTAATCGCATACTTTATCTATCCTCAAAAGGAGATTCAAAAAATGCAGGAGCAAACGACTAGC
TCTTTACCTGCAAGCTCTCTACCGTCAAGCAGCGAGAGATCCTCAAGCTCTGCTCCTCATTTGGAGAT
CAAAGAAGGAATTGAAAGCGATGAGGAGATACGGCGAGTTCCGGAGTTTGGAGGAGAAGCTACCG
GAAAAGAAATCTCTGGATCGGTGACCGGTCAGGACCAGACACAAGCAACGGTCGGAGGAGAGAGT
CAAAGGAAGAGAGGGAGGACTCCGGCTGAGAAAAGAGACCAAGCGGCTTAAGAGGTTGTGAGGAA
CAGAGTTTCAGCACAGCAAGCAAGAGAGAGGAAGAAAGCTTACTTGGGTGAGTTGGAAAATAGAGT
GAAAGACTTGGAGAATAGAACTCTGAACCTGAAGAGAGACTCTCTACTTTGCAGAACGAGAACCA
GATGCTTAGACAGATTCTGAAGAACACAACAGGAAACAAGAGAGAGGCGGTGGTTCTAACGCTGA
TGCGAGCCTCTGACCTCTTTGTTCTTATGTTATTTATGTTGATAAATTTACAGAGAATTGTACTAAT
AAATATTTACATTGCATGGTATGTGCTGTGACTTGAACCTTGTAGTTGTCTTTAGATTTTCTTAAT
TCGTTTTCTTTGTTGTCTGATAGATTGCTTAATTAATAA
```

图 1 电子克隆的油菜 HY5 基因 cDNA 序列

### 2.2 油菜、芜菁与拟南芥 HY5 基因编码蛋白的同源比较

将电子克隆得到的油菜 HY5 基因(*B. napus*)编码蛋白和芜菁(*B. rapa*)、拟南芥(*A. thaliana*)的 HY5 蛋白采用 Clustalw 2.1 进行同源比对。由图 2 可知,油菜

HY5 基因同芜菁的同源性更高,仅是在第 5 位、59 位和 115 位氨基酸有差异。从油菜 HY5 蛋白 154 位起有 4 个甘氨酸(G),而芜菁有 7 个,拟南芥这个区域有 5 个连续的甘氨酸。

<i>B. napus</i>	MQEQTSSLPASSLPSSSERSSSSAPHLEIKEGIESDEEIRRVPFEGGEATGKEISG--- 57
<i>B. rapa</i>	MQEQTSSLPASSLPSSSERSSSSAPHLEIKEGIESDEEIRRVPFEGGEATGKEISG--- 57
<i>A. thaliana</i>	MQEQTSSLAASSLPSSSERSSSSAPHLEIKEGIESDEEIRRVPFEGGEAVGKETSGRES 60
	*****
<i>B. napus</i>	-SVTGQDQTQATVGGESQKRGRTPAEKETKRLKRLLRNRVSAQQARERKKAYLGELENR 116
<i>B. rapa</i>	-SGTGQDQTQATVGGESQKRGRTPAEKETKRLKRLLRNRVSAQQARERKKAYLGELETR 116
<i>A. thaliana</i>	GSATGQERTQATVG-ESQKRGRTPAEKENKRLKRLLRNRVSAQQARERKKAYLSELENR 119
	* * * * *
<i>B. napus</i>	VKDLENRNSELEERLSTLQENQMLRQILKNTTGNKRGGGG---SNADASL 164
<i>B. rapa</i>	VKDLENRNSELEERLSTLQENHMLRQILKNTTGNKRGGGGGGGSSNADASL 167
<i>A. thaliana</i>	VKDLENKNSLEERLSTLQENQMLRHILKNTTGNKRGGGGG--SNADASL 168
	*****

图 2 油菜与拟南芥、芜菁 HY5 基因编码蛋白的同源性比较

### 2.3 油菜 HY5 一级结构分析

对油菜 HY5 编码蛋白的一级结构进行分析,分析结果显示其由 18 种氨基酸组成。甘氨酸(G)含量最高,达到 12.8%,组氨酸、苯丙氨酸、酪氨酸均只有 1 个,含量为 0.6%,其分子量是 18 037.8 Da,分子式为  $C_{746}H_{1258}N_{248}O_{267}S_2$ ,理论上的等电点(pI)是 9.29。其蛋白分子由 2 521 个原子构成。在 280 nm 处消光系数为 1 490。计算其热稳定指数(II)结果为 61.67,这表明

油菜 HY5 蛋白是不稳定蛋白。其总体的亲水性为一 1.198,而平均脂溶指数是 61.34。

### 2.4 油菜 HY5 蛋白二级结构分析

对油菜 HY5 蛋白二级结构进行预测,结果表明,油菜 HY5 蛋白含有  $\alpha$ -螺旋为 42.07%,无规卷曲为 52.10%。分析其疏水性,其亲水系数分布与 -3~0.5 之间,可见油菜 HY5 蛋白是亲水蛋白。

### 2.5 油菜 HY5 蛋白系统发育树分析

从蛋白数据库下载野草莓(*Fragaria vesca* subsp. *vesca*)、蓖麻(*Ricinus communis*)、黄瓜(*Cucumis sativus*)、二穗短柄草(*Brachypodium distachyon*)、番茄(*Solanum lycopersicum*)、玉米(*Zea mays*)、水稻(*Oryza sativa* Japonica Group)的 HY5 蛋白序列,和油菜(*Brassica napus*)、芜菁(*Brassica rapa* subsp. *rapa*)、拟南芥(*Arabidopsis thaliana*)的 HY5 序列保存在一个 FASTA 文件中,进行 Neighbor-Joining 系统发育树分析,软件版本为 MEGA 5.05 (图 3)。HY5 蛋白系统进化树分析结果表明,单子叶植物(玉米、二穗短柄草和水稻)HY5 蛋白聚为一支,双子叶植物聚为一支,十字花科的油菜、芜菁和拟南芥聚为一支。和油菜 HY5 关系最近的是芜菁,其次是拟南芥。总体来看 HY5 的系统发育树同各个物种间的亲缘关系一致。

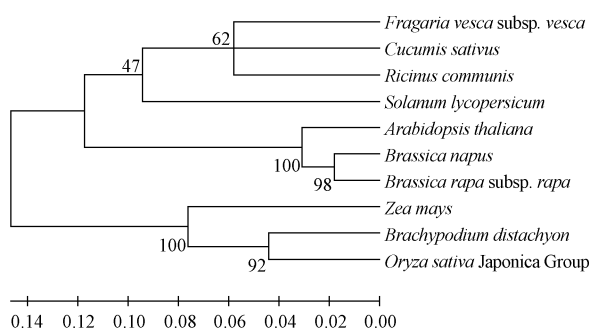


图 3 HY5 蛋白系统发育树分析

### 3 讨论

HY5 基因被证明是幼苗光形态建成的正调控因子。其突变体表型表明 HY5 突变体在外界环境因子包括光信号的诱导下,失去了正常表型应该有的下胚轴和根的反应<sup>[5]</sup>。对该研究克隆得到的 HY5 蛋白的分析表

明, HY5 基因编码定位在细胞核的一个 bZIP 蛋白。这与其它研究结果一致。许多 bZIP 蛋白已从多种植物中分离得到,多数情况下 bZIP 蛋白结合的 DNA 含有 ACGT 核心序列。虽然 ACGT 序列被称为顺式作用元件在植物中的各种刺激响应基因启动子的一个核心区。HY5 蛋白具有除 bZIP 域没有其它明显的域。潜在的转录激活域的缺乏意味着 HY5 蛋白要同其它转录因子协同作用<sup>[6]</sup>。

在后基因组时代,数据分析,而不是数据的收集将是目前生物学家最大的挑战。通过已有的测序数据分析基因组数据的生物学意义,进行功能分析,结构鉴定,或表达模式研究是非常有意义的。传统克隆基因的方法主要是 mRNA 差异显示技术和 RACE 技术,这 2 种技术对研究人员操作技术高,花费高,费时费力。而根据不同物种基因存在保守性,采用已经公布的表达序列标签进行基因的电子克隆可以快速简便低成本的获得基因序列。该研究初步获得了油菜中光调控因子 HY5 的基因编码序列,为进一步研究其功能,为油菜的栽培育种工作提供了参考。

### 参考文献

- [1] Wu D, Hu Q, Yan Z, et al. Structural basis of ultraviolet-B perception by UVR8[J]. Nature, 2012, 484(7393): 214-219.
- [2] Rizzini L, Favory J J, Cloix C, et al. Perception of UV-B by the Arabidopsis UVR8 protein[J]. Science, 2011, 332: 103-106.
- [3] Oyama T, Shimura Y, Okada K. The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus induced development of root and hypocotyl[J]. Genes and Development, 1997(11): 2983-2995.
- [4] 惠婕, 杜静, 黄丛林, 等. 植物光敏色素研究进展[J]. 北方园艺, 2010(7): 203-205.
- [5] Ang L H, Deng X W. Regulatory hierarchy of photomorphogenic loci: allele-specific and light-dependent interaction between the HY5 and COP1 loci[J]. Plant Cell, 1994(6): 613-628.
- [6] Foster R, Izawa T, Chua N. Plant bZIP proteins gather at ACGT elements[J]. FASEB Journal, 1994(8): 192-200.

## In silico Cloning and Bioinformatics Analysis on HY5 Gene of Brassica napus

GUO Ji-ping<sup>1,2</sup>

(1. Department of Life Sciences, Hengshui College, Hengshui, Hebei 053000; 2. College of Food Science and Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150090)

**Abstract:** With HY5 gene in *Brassica rapa* as probe, the cDNA of HY5 gene of *Brassica napus* was *in silico* cloned. Using bioinformatic tool, amino acid sequence coded by HY5 of *Brassica napus* was predicted and homology compared with *Arabidopsis thaliana* and *Brassica rapa*. Its phylogenetic tree was constructed. The results showed that HY5 of *Brassica napus* coded a protein of 164 amino acids. The calculated molecular weight of it was 18 037.8 Da. It was a hydrophilic protein. The homology of HY5 protein between *Brassica napus* and *Brassica rapa* was higher than that between *Brassica napus* and *Arabidopsis thaliana*. The HY5 gene coding sequence of *Brassica napus* was cloned preliminary, and laid the foundation for further study its function.

**Key words:** HY5; *Brassica napus*; photoreceptor; *in silico* cloning