

# 一个水稻穗特异表达锌指蛋白基因的克隆与结构分析

李余生 黄 骥 禹山林 侯夫云 张红生\*

(南京农业大学 作物遗传与种质创新国家重点实验室, 江苏 南京 210095; \* 通讯联系人, E-mail: hszhang@njau.edu.cn)

## Cloning and Structure Analysis of a Panicle Specific Zinc Finger Protein Gene from Rice (*Oryza sativa*)

LI Yu sheng, HUANG Ji, YU Shan lin, HOU Fu yun, ZHANG Hong sheng\*

(State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China;

\* Corresponding author, E-mail: hszhang@njau.edu.cn)

**Abstract:** Through bioinformatics and the RT-PCR assay, a novel zinc finger protein gene, belonging to TF A type, was cloned from rice and designated as *OsZPT3 1*, which is specifically expressed in rice panicles, but not detected in culms, leaves, anthers and roots at the adult stage. *OsZPT3 1* is 1104 bp in length and its complete opening read frame (ORF) is 1092 bp encoding a protein of 363 amino acid residues. The *OsZPT3 1* contains three C<sub>2</sub>H<sub>2</sub> zinc finger motifs and only the third zinc finger has a conserved QALGGH sequence, belonging to B type zinc finger. There was a Leu-rich region with consensus of LXLXL in the C terminal of *OsZPT3 1*, which might function as a transcription repressor. Promoter analysis revealed three MADS box transcription factor recognition sites existing in *OsZPT3 1* promoter. Taken together, it can be speculated that *OsZPT3 1* may play a key regulatory role in the development of panicle by repressing gene expression under the control of MADS box transcription factors.

**Key words:** rice; zinc finger protein; gene cloning; expression analysis

**摘要:** 利用生物信息学和 RT-PCR 方法从水稻中克隆鉴定了一个新的具有 TF A 型锌指结构的锌指蛋白基因, 其开放阅读框为 1092 bp, 编码 363 个氨基酸残基。组织表达谱结果表明, 该基因只在水稻幼穗组织中表达, 在成株期的根、叶、茎以及花药中不表达, 将其命名为 *OsZPT3 1*。序列分析表明 *OsZPT3 1* 具有 3 个 C<sub>2</sub>H<sub>2</sub> 型锌指结构, 在氨基酸的 C 端没有典型的转录抑制区域 DLN box, 但 LXLXL 的结构仍然表明 *OsZPT3 1* 可能是一个转录抑制因子。对 *OsZPT3 1* 启动子区域进行预测, 结果发现 3 个 MADS box 转录因子识别位点, 推测 *OsZPT3 1* 可能在 MADS box 转录因子的调节下, 通过抑制下游基因的表达在水稻穗部器官的生长发育过程中发挥重要的调控作用。

**关键词:** 水稻; 锌指蛋白; 基因克隆; 表达分析

中图分类号: Q943.2; S511.032

文献标识码: A

文章编号: 1001-7216(2006)04-0343-05

锌指蛋白 (zinc finger protein) 是一类具有“手指状”结构域的转录因子, 根据其氨基酸序列中半胱氨酸 (C) 和组氨酸 (H) 残基的数目和位置可将锌指转录因子分为 C<sub>2</sub>H<sub>2</sub>、C<sub>2</sub>HC、C<sub>2</sub>C<sub>2</sub>、C<sub>2</sub>HCC<sub>2</sub>C<sub>2</sub>、C<sub>2</sub>C<sub>2</sub>C<sub>2</sub>C<sub>2</sub> 等亚类, 其中大部分锌指蛋白属于 C<sub>2</sub>H<sub>2</sub> 型锌指蛋白<sup>[1]</sup>。根据锌指蛋白的结构特征或功能差异, 又可以分为 WRKY 蛋白、GATA 蛋白、PHD 蛋白、DOF 蛋白、TF A 蛋白等<sup>[2]</sup>。TF A 家族是植物中描述较为清楚的 C<sub>2</sub>H<sub>2</sub> 锌指蛋白, 也是除了 WRKY 蛋白之外包含成员最多的 C<sub>2</sub>H<sub>2</sub> 锌指蛋白<sup>[1]</sup>。TF A 型锌指蛋白最早在非洲爪蟾中发现, 其锌指区为 CX<sub>2-4</sub>CX<sub>3</sub>FX<sub>5</sub>LX<sub>2</sub>HX<sub>3-5</sub>H 的结构, 其中的两个 Cys 和两个 His 与锌原子形成配位键, 进而形成一个包含发夹和一个螺旋的紧密指状结构<sup>[2]</sup>。植物 TF A 锌指蛋白中所含有的锌指结构的数目一般从 1 个到 4 个不等, 在锌指结构的螺

旋区一般都有一个 QALGGH 的高度保守序列, 为植物锌指蛋白特有<sup>[3]</sup>。自从在矮牵牛中分离了第一个植物 TF A 锌指蛋白基因 *ZPT2 1* 以来, 已经在矮牵牛<sup>[3-4]</sup>、拟南芥<sup>[5]</sup>、小麦<sup>[6]</sup>、棉花<sup>[7]</sup> 等植物中相继分离了近 40 个植物 TF A 锌指蛋白基因, 其中大部分具有 QALGGH 保守结构, 少数具有类似 QALGGH 的结构。目前报道的植物 TF A 型锌指蛋白主要参与植物的生长发育以及环境胁迫下基因表达的调控<sup>[8]</sup>。

本研究采用生物信息学方法, 以本实验室从水

收稿日期: 2005-09-05; 修改稿收到日期: 2005-11-07。

基金项目: 国家自然科学基金资助项目 (30470921); 江苏省自然科学基金资助项目 (BK2005090); 长江学者和创新团队发展计划资助项目。

第一作者简介: 李余生 (1965-), 男, 助理研究员, 博士研究生。

稻中分离的 TF A 锌指蛋白 OsZFP<sup>[9]</sup> 为信息探针,通过搜索水稻基因组数据库、计算机拼接以及 RT-PCR 的方法从水稻中分离了一个新的 TF A 型锌指蛋白基因 *OsZPT3 1*。*OsZPT3 1* 具有 3 个锌指结构,组织表达分析表明 *OsZPT3 1* 基因特异性地在水稻穗中表达,表明它可能参与水稻穗部器官发育的调节。

## 1 材料与方法

### 1.1 植物材料

太湖流域粳稻地方品种韭菜青由南京农业大学水稻研究所提供。成株期的水稻叶片、根、茎以及 5~8 cm 长的幼穗和花药经液氮速冻后于 -80 °C 下保存,用于总 RNA 提取。

### 1.2 总 RNA 的提取和 cDNA 第一链的合成

RNA 的抽提采用 Trizol 试剂 (Invitrogen), 2 μg 的总 RNA (经 DNase 处理) 用于 cDNA 第一链的合成 (Promega), 作为 RT-PCR 的模板。

### 1.3 cDNA 的序列拼接

以本实验室从水稻中分离的具有 2 个锌指结构的 TF A 型锌指蛋白 OsZFP<sup>[9]</sup> 的氨基酸序列 (GenBank 注册号: AY077725) 为信息探针,运行 BLAST 程序搜索位于 GenBank 的水稻基因组数据库,寻找包含水稻 TF A 型锌指蛋白的水稻基因组序列。采用黄骥等<sup>[10]</sup> 推荐的方法进行序列拼接和开放阅读框 (ORF, open reading frame) 预测。根据拼接好的 ORF 搜索水稻 EST 数据库,拼接获得完整的 cDNA 序列。

### 1.4 RT-PCR 扩增

根据拼接好的完整 cDNA 序列设计特异性引物,以来自水稻不同组织的 cDNA 为模板进行 PCR 扩增。本研究中用于克隆 *OsZPT3 1* 的引物为 RZF26-5: 5' CATCCAGATATGGCACCAAG 3' 和 RZF26-3: 5' GGTTCACTGAAGATTGCTCG 3'。PCR 程序为: 94 °C 下预变性 5 min, 94 °C 下变性 30 s, 58 °C 下复性 50 s, 72 °C 下延伸 1.5 min, 33 个循环; 72 °C 下延伸 10 min。PCR 产物经过 1% 琼脂糖凝胶电泳检测后回收,连接到 pGEM-T 载体 (Promega), DNA 测序委托上海博亚公司进行。

### 1.5 序列分析

cDNA 序列的 ORF 查找以及翻译采用 BioXM 软件 (ver 2.5) 进行;氨基酸的多序列比较采用 Clustalx 软件进行;GeneDOC 软件 (ver 2.5) 用于显示比较的结果;功能域预测采取 SMART 在线程序

(http://smart.embl-heidelberg.de/) 进行;启动子区域转录因子结合位点预测采用 MatInspector 程序 (http://www.genomatix.de) 进行。

## 2 结果与分析

### 2.1 *OsZPT3 1* 基因的克隆与序列分析

以水稻 TF A 锌指蛋白 OsZFP<sup>[9]</sup> 的氨基酸序列为信息探针,运行 BLAST 程序搜索位于 GenBank 的水稻基因组数据库,发现在水稻不同染色体上至少分布有 75 个编码 TF A 型锌指蛋白的基因。通过对这些基因进行开放阅读框预测和 EST 拼接,初步得到了这些基因的完整 cDNA。根据拼接好的序列设计 PCR 引物,选择来自水稻不同组织的 RNA 进行 RT-PCR 扩增和功能分析。本文报道一个位于第 12 染色体上的编号为 26 的 TF A 型锌指蛋白基因的克隆和功能分析。该基因仅在水稻幼穗中表达,在成株期的根、茎、叶片以及花药中都没有检测到表达 (图 1)。测序结果表明该基因长 1104 bp,开放阅读框为 1092 bp,编码 363 个氨基酸残基,分子量 40 kD,被命名为 *OsZPT3 1*。

经 SMART 程序预测,*OsZPT3 1* 推导的氨基酸序列包括 3 个 C<sub>2</sub>H<sub>2</sub> 型锌指基序,锌指区符合 CX<sub>2-4</sub>CX<sub>3</sub>FX<sub>5</sub>LX<sub>2</sub>HX<sub>3-5</sub>H 的结构,其中第 1 和第 3 个锌指基序中具有 QALGGH (或类似) 的保守结构,但第 2 个锌指基序中没有 QALGGH 结构 (图 2)。在第 2 和第 3 个锌指基序之间,发现一个富含碱性氨基酸的可能的核定位信号 (NLS) KRKRSFRSKR,表明 *OsZPT3 1* 编码产物可能定位于细胞核内,符合转录因子的特征。在 *OsZPT3 1* 的 C 端没有发现大多数植物 TF A 型锌指蛋白具有的富含亮氨酸的 DLN box<sup>[5]</sup>,而 DLN box 已经被证实具有转录抑制的功能<sup>[11-12]</sup>。在 *OsZPT3 1* 的整个氨基酸序列中存在 3 个富含亮氨酸的疏水区 L-

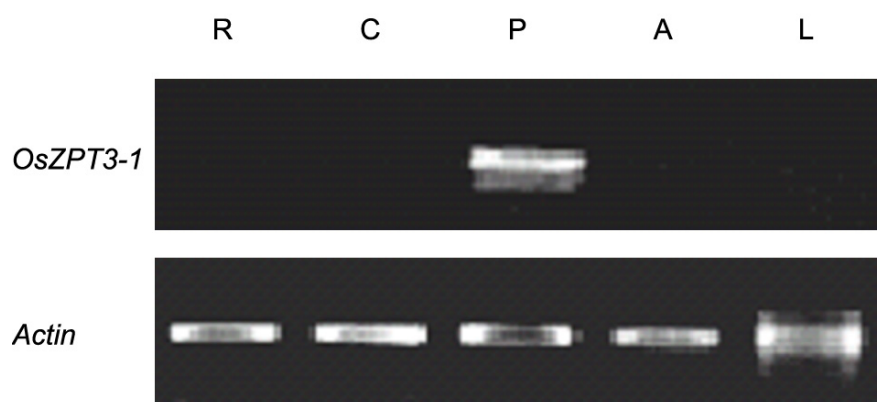


图 1 水稻 *OsZPT3 1* 的组织表达特征

Fig. 1. Expression of *OsZPT3 1* in various rice tissues.

R - 根; C - 茎; P - 幼穗 (5~8 cm); A - 花药; L - 叶片.

R, Root; C, Culm; P, Young panicle; A, Anther; L, Leaf.

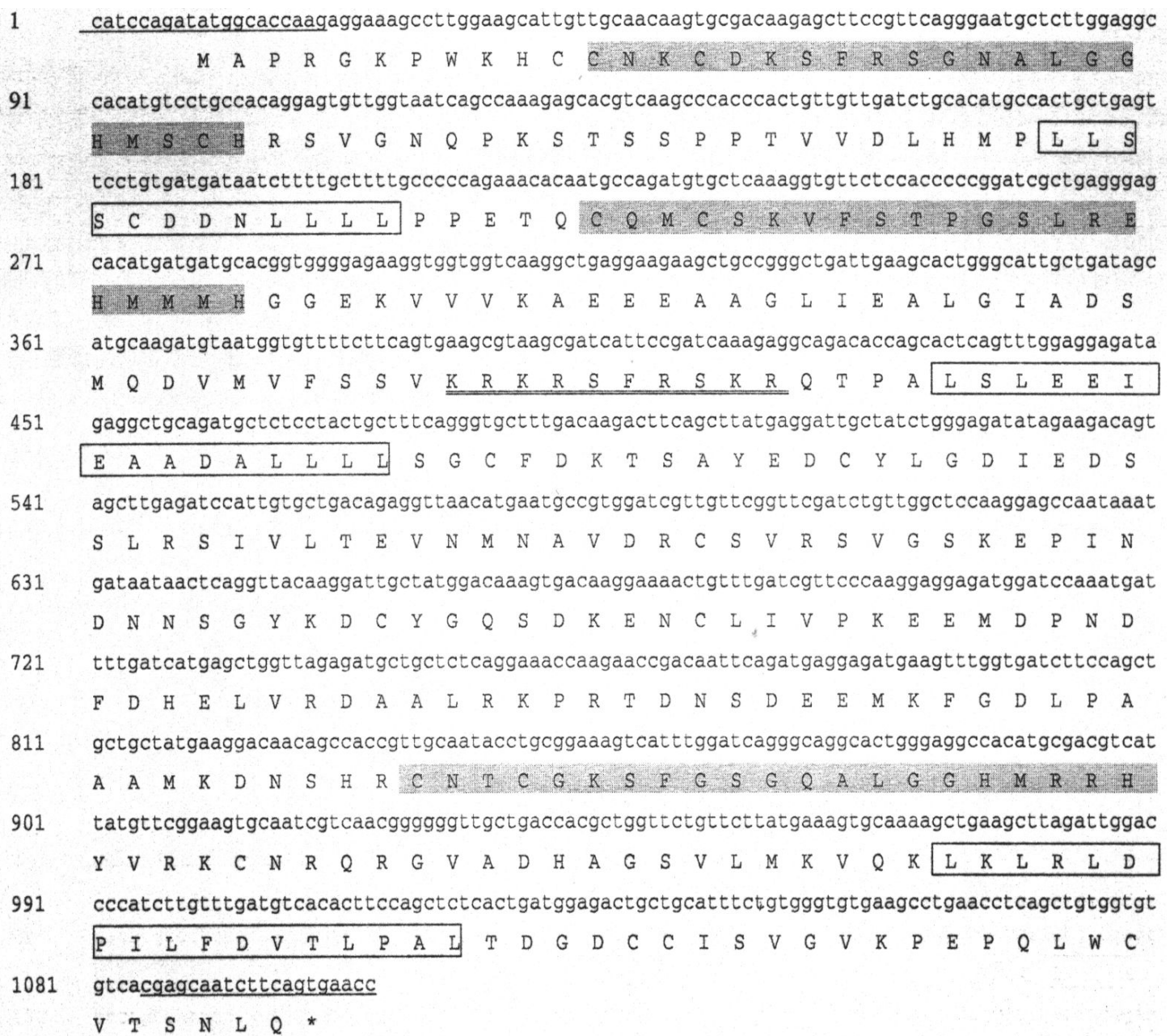


图 2 水稻 *OsZPT3 1* 的核苷酸及推导的氨基酸序列

Fig. 2. Sequence of nucleotide and deduced amino acids for *OsZPT3 1*.

下划线处表示 RT-PCR 引物序列，阴影处表示锌指结构，双下划线处表示可能的核定位信号，加框处为富含亮氨酸残基的 L-box；\* 表示终止密码子。

The RT-PCR primers are underlined, zinc finger structures shaded in gray, putative NLS double underlined, Leu-rich hydrophobic regions (L-box) boxed and the stop codon marked by asterisk.

box 这些疏水结构可能与蛋白质间的相互作用有关<sup>[5]</sup>。

## 2.2 *OsZPT3 1* 与其他植物 TF A 型三锌指蛋白的序列比较

将 *OsZPT3 1* 与其他植物 TF A 型三锌指蛋白进行了氨基酸序列比较，发现 *OsZPT3 1* 与它们的氨基酸一致率为 16% ~ 32% (表 1)。虽然整个氨基酸序列的相似率较低，但这些 TF A 型锌指蛋白的锌指结构非常保守，大部分都具有 QALGGH 的保守结构(或类似结构)(图 3)。根据锌指结构中 QALGGH 及其侧翼序列的差异，将锌指结构分为 3 个类型：A 型、B 型和 M 型<sup>[13]</sup>，每种类型的锌指结构都符合 CX<sub>2</sub>CX<sub>3</sub>FX<sub>5</sub>LX<sub>2</sub>HX<sub>3</sub>H 的序列特征。大多数双锌指蛋白的第 1 个锌指为 A 型，具有 (F/Y/H)QALGGH (R/K) 的保守结构，第 2 个锌指为 B 型，具有 GQALGGH (M/K)R 的保守结构。所有

已发现的三锌指蛋白的第 1 个锌指都属于 M 型，具有 QALGGH 的类似结构，第 3 个锌指都为 B 型，而多数三锌指蛋白的第 2 个锌指为 A 型。M 型大多出现在多锌指蛋白中，具有 QALGGH 的类似结构，极少数多锌指蛋白中存在与 QALGGH 完全不同的锌指结构<sup>[13]</sup>。*OsZPT3-1* 的前两个锌指都属于 M

表 1 *OsZPT3 1* 与其他植物三锌指蛋白的相似性比较

Table 1 Comparison of homology between *OsZPT3 1* and other plant TF zinc finger proteins.

蛋白名称	来源	GenBank 注册号	一致率
Protein name	Origin	GenBank No.	Identity/%
ZPT3 1	矮牵牛 <i>Petunia</i>	AB000453	32
ZPT3 2	矮牵牛 <i>Petunia</i>	AB000454	21
ZPT3 3	矮牵牛 <i>Petunia</i>	AB035133	23
ZAT1	拟南芥 <i>Arabidopsis</i>	X98669	29
SlZPT3 1	麦瓶草 <i>Silene latifolia</i>	DQ017764	16
SlZPT3 2	麦瓶草 <i>Silene latifolia</i>	DQ017765	19



表 2 *OsZPT3 1* 启动子区域可能的转录因子结合位点

Table 2 . Putative *cis* elements in the promoter of *OsZPT3 1* .

顺式元件 <i>cis</i> element	位置 <sup>1)</sup> Position <sup>1)</sup>	序列 <sup>2)</sup> Sequence <sup>2)</sup>	可能性 Possibility/%
RNA 聚合酶 结合位点 RNA polymerase binding site(TATA box)	- 231 ~ - 217	ttaaTATAaatatat	88
热激元件 Heat shock element	- 344 ~ - 330	caatgtatccAGAAc	81
MADS 识别位点 MADS recognition site	- 433 ~ - 413	ctgtaagcA AAAtggatactg	81
MADS 识别位点 MADS recognition site	- 615 ~ - 595	gccatCCAGatattggctgaag	83
MADS 识别位点 MADS recognition site	- 796 ~ - 776	agcatCCA Ggtaaagaacgcc	83
热激元件 Heat shock element	- 995 ~ - 881	ggtaaaattcAGAAc	81
热激元件 Heat shock element	- 1069 ~ - 1055	taatgcatcaAGAAc	81
TF A 锌指蛋白识别位点 TF A zinc finger protein recognition site	- 1153 ~ - 1131	atcactttctaCAGTggccaac	75

1) 数字表示顺式元件相对于起始密码子的位置 ;

2) 核心序列以大写字母表示。

1) The Arabic numerals indicate the numbers of the nucleotides between *cis* elements and the translation start codon ;

2) Bases in capital denote the core sequence of *cis* element .

型 ,其中第 2 个锌指完全没有 QALGGH 的保守结构 ,第 3 个锌指则与其他植物来源的三锌指蛋白相似 ,都属于 B 型锌指结构 (图 3)。

### 2.3 *OsZPT3 1* 启动子区域的转录因子结合位点分析

应用 MatInspector 程序预测了 *OsZPT3 1* 起始密码子上游 1500 bp 的基因组序列中的顺式作用元件 ,发现该区域存在 1 个 RNA 聚合酶 结合位

点 TATA box ,3 个热激元件 ,3 个 MADS box 转录因子识别位点以及 1 个植物 TF A 型锌指蛋白识别位点 (表 2)。

### 3 讨论

TF A 型锌指蛋白广泛存在于高等植物中 ,迄今为止 ,在矮牵牛、拟南芥等植物中至少已经分离了 40 个植物 TF A 型锌指蛋白基因 [8]。已经报道的植物 TF A 型锌指蛋白大多数都具有单个或两个锌指结构 ,而且这些锌指结构中都具有保守的 QALGGH 序列。已有研究表明 A、L、G、G、H 中的任何一个氨基酸发生突变都会使锌指结合目标 DNA 的能力丧失 ,而 Q 突变后锌指的结合能力会显著下降 [13]。OsZPT3 1 与大多数的三锌指或四锌指蛋白类似 ,除了 A 型或 B 型锌指结构具有 QALGGH 之外 ,一般都存在着 M 型的锌指结构 ,但所有植物 TF A 型锌指结构中的 L (Leu) 在任何一个 M 型锌指中都保持不变 (图 2)。OsZPT3 1 的两个 M 型锌指结构由于 QALGGH 序列与 A、B 型锌指有所差异 ,因此很可能其识别位点或结合 DNA 的能力与 A 型或 B 型锌指也存在差异。

大多数植物 TF A 型锌指蛋白的 C 端都有一个富含 Leu 的 DLN box [3] ,在拟南芥锌指蛋白 STZ 和矮牵牛锌指蛋白 ZPT2 3 中 ,通过片段缺失分析证实了 DLN box 担负锌指蛋白转录抑制的功能 ,而缺失 DLN box 的 ZPT2 3 具有转录激活活性 [11-12]。最近 ,Hiratsu 等 [14] 对于拟南芥单锌指蛋白 SUPERMAN 的 DLN box 作了细致的研究 ,发现 SUPERMAN 的 DLN box 序列 DLELRL 对于 SUPERMAN 的转录抑制活性是充分而必要的 ,通

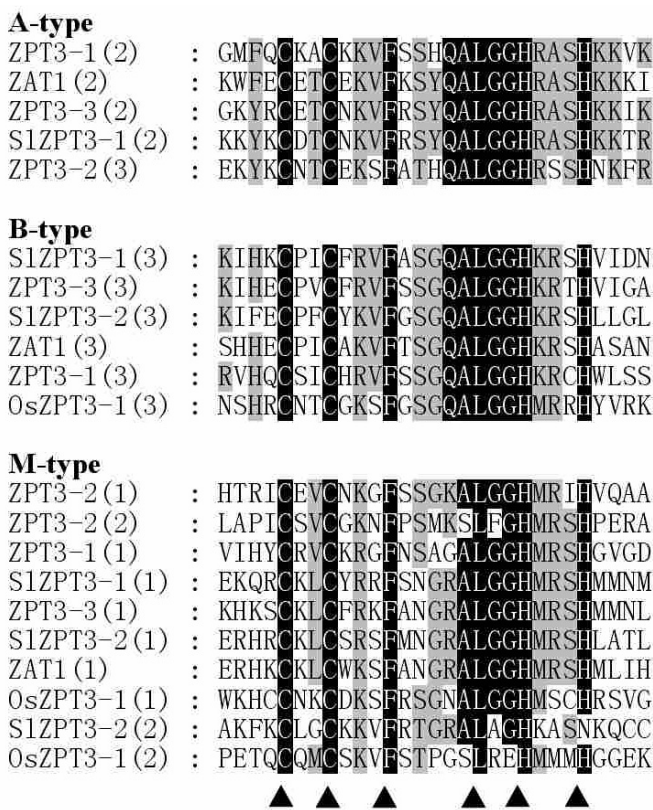


图 3 *OsZPT3 1* 与其他植物 TF A 型三锌指蛋白锌指结构的比较

Fig. 3 Comparison of *OsZPT3 1* with other plant TF A zinc finger proteins in zinc finger regions .

黑三角表示锌指结构中最保守的 6 个氨基酸残基 ,括号中的数字表示锌指序号。

The six highly conserved amino acid characteristic for TF A type zinc fingers are indicated by black triangles . Values in the parentheses mean the order of zinc finger in the proteins .

过点突变研究发现 DLN box 中的 Leu 对于它的抑制活性至关重要,任何一个 Leu 的替换都会导致抑制活性的下降,因此一个具有活性的 DLN box 应该具有 XLXLXL 的序列特征(第 1、5 位的氨基酸残基为 D/N/Q/S 和 R/T 时具有较好的抑制活性)。虽然 ZPT3 1 的 C 端没有典型的 DLN box,但其 C 端的 K L K L R L(L X L X L)结构表明其仍然可能是转录抑制因子<sup>[14]</sup>。

现有研究表明大多数植物的 TF A 型锌指蛋白参与植物的生长发育以及环境胁迫应答反应。与植物器官分化有关的绝大多数锌指蛋白都具有单个锌指结构,参与各种环境胁迫应答反应的 TF A 型锌指蛋白具有双锌指结构;而大多数三锌指或四锌指蛋白基因的表达也具有组织特异性,尤其是生殖器官的组织特异性<sup>[3,8]</sup>。本实验结果表明 *OsZPT3 1* 仅在水稻幼穗中检测到表达,而在成株期的叶片、根、茎以及花药中都没有检测到 *OsZPT3 1* 的表达,推测 *OsZPT3 1* 与水稻的穗部器官发育有关。在 *OsZPT3 1* 的启动子区域,我们发现了 3 个 MADS box 转录因子结合位点(表 2)。MADS box 基因在植物生殖器官发育过程中占据了核心的地位<sup>[15]</sup>,这可能暗示 *OsZPT3 1* 参与的水稻穗部器官的发育受到 MADS box 转录因子的调控。我们在启动子区域还发现了一个植物 TF A 型锌指蛋白的识别位点 A(G/C)T 序列<sup>[12]</sup>,表明 *OsZPT3 1* 可能受到另外的 TF A 型锌指蛋白的调节,或受到自身的转录调节。此外,启动子区域还发现了 3 个热激元件<sup>[16]</sup>,暗示了 *OsZPT3 1* 可能受热激转录因子的调节,参与植物的热胁迫应答反应。

综上所述,我们从水稻中分离了一个具有 3 个锌指结构的 TF A 型锌指蛋白基因 *OsZPT3 1*,发现它为穗特异表达基因。在启动子区域发现了 3 个 MADS box 转录因子识别位点,推测 *OsZPT3 1* 在穗中的表达受 MADS box 蛋白的调控。*OsZPT3 1* 可能在 MADS box 转录因子的调节下,通过抑制下游基因的表达在水稻穗部器官的生长发育过程中发挥重要的调控作用。

#### 参考文献：

[1] Benjamin L. Gene . New York :Oxford University Press ,

2000 : 654 656 .

- [2] Takatsuji H . Zinc finger transcription factors in plants . *Cell Mol Life Sci* ,1998 ,54(6) : 582 596 .
- [3] Takatsuji H . Zinc finger proteins : the classical zinc finger emerges in contemporary plant science . *Plant Mol Biol* ,1999 ,39(6) : 1073 1078 .
- [4] Takatsuji H , Mori M , Benfey P N , et al . Characterization of a zinc finger DNA binding protein expressed specifically in petunia petals and seedlings . *EMBO J* ,1992 ,11 : 241 249 .
- [5] Sakamoto H , Araki T , Meshi T , et al . Expression of a subset of the *Arabidopsis* Cys2/His2 type zinc finger protein gene family under water stress . *Gene* ,2000 ,248 : 23 32 .
- [6] Sakamoto A , Minami M , Huh G H , et al . The putative zinc finger protein WZF1 interacts with a *cis*acting element of wheat histone genes . *Eur J Biochem* ,1993 ,217 : 1049 1056 .
- [7] 王 东,杨金水.棉花类耐盐锌指蛋白基因的克隆与结构分析.复旦学报:自然科学版,2002,41(4):42-46.
- [8] 黄 骥,王建飞,张红生.植物 C<sub>2</sub>H<sub>2</sub> 锌指蛋白的结构与功能.遗传,2004,26(3):414-418.
- [9] 黄 骥,张红生,曹雅君,等.一个新的水稻 C<sub>2</sub>H<sub>2</sub> 型锌指蛋白 cDNA 的克隆与序列分析.南京农业大学学报,2002,25(2):110-112.
- [10] 黄 骥,张红生,曹雅君,等.水稻功能基因的电子克隆策略.中国水稻科学,2002,16(4):295-298.
- [11] Sugano S , Kaminaka H , Rybka Z , et al . Stress responsive zinc finger gene *ZPT2 3* plays a role in drought tolerance in petunia . *Plant J* ,2003 ,36(6) : 830 841 .
- [12] Sakamoto H , Maruyama K , Sakuma Y , et al . *Arabidopsis* Cys2/His2 type zinc finger proteins function as transcription repressors under drought , cold , and high salinity stress conditions . *Plant Physiol* ,2004 ,136(1) : 2734 2746 .
- [13] Kubo K , Sakamoto A , Kobayashi A , et al . Cys2/His2 zinc finger protein family of petunia : evolution and general mechanism of target sequence recognition . *Nucl Acids Res* ,1998 ,26 : 608 615 .
- [14] Hiratsu K , Mitsuda N , Matsui K , Ohme Takagi M . Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis* . *Biochem Biophys Res Comm* ,2004 ,321(1) : 172 178 .
- [15] Becker A , Theissen G . The major classes of MADS box genes and their role in the development and evolution of flowering plants . *Mol Phylog Evol* ,2003 ,29(3) : 464 489 .
- [16] Prandl R , Schoffl F . Heat shock elements are involved in heat shock promoter activation during tobacco seed maturation . *Plant Mol Biol* ,1996 ,31(1) : 157 162 .