

# 水稻胞质核糖体蛋白基因 *OsRPS7* 的克隆与序列分析

顾志敏, 王建飞, 黄 骥, 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)

**摘要:**以已公布的黑麦胞质核糖体蛋白基因 *ScRPS7* 的 cDNA 序列为信息探针, 在中国华大水稻基因组数据库中搜索与之高度同源的基因组重叠群。采用计算机拼接和 RT-PCR 方法克隆了水稻胞质核糖体蛋白基因的全长 cDNA 序列, 命名为 *OsRPS7*。该 cDNA 序列全长 919bp, 编码 192 个氨基酸; 其与黑麦、拟南芥和芸薹的 S7 核糖体蛋白的氨基酸一致率分别为 88%、72% 和 72%。对 *OsRPS7* 的基因组结构和基因的功能进行了分析和预测。

**关键词:**水稻; 核糖体蛋白 S7; 基因克隆; 基因结构

中图分类号: Q785

文献标识码: A

文章编号: 0253-9772(2004)02-0181-05

## Molecular Cloning and Sequence Analyzing of Cytoplasmic Ribosomal Protein Gene *OsRPS7* from Rice (*Oryza sativa* L.)

GU Zhi-Min, WANG Jian-Fei, HUANG Ji, ZHANG Hong-Sheng

(National Key Laboratory of Crop Genetics and Germplasm Enhancement,  
Nanjing Agricultural University, Nanjing 210095, China)

**Abstract:** Using the cDNA of rye cytoplasmic ribosomal protein *ScRPS7* as a query probe, a highly homologous rice genomic contig was obtained from Huada rice genome database. The full-length cDNA sequence of rice cytoplasmic ribosomal protein S7 was assembled by informatics based on the contig. Furthermore, with the two primers designed according to this assembled cDNA, the full-length cDNA of rice ribosomal protein was cloned by RT-PCR and named as *OsRPS7*. The cDNA was 919bp in length and contained a complete open reading frame of 576bp, encoding a protein of 192 amino acid residues. The deduced amino acids of *OsRPS7* showed 88%, 72% and 72% identity with those from *Secale cereale*, *Arabidopsis thaliana* and *Brassica oleracea*, respectively. The genome structure of *OsRPS7* was analyzed, and its function was predicted in this paper.

**Key words:** rice; ribosomal protein S7; gene cloning; genomic structure

核糖体是生物细胞内重要的细胞器, 在细胞内通过解码 mRNAs, 然后与 tRNA 一起合成相应的蛋白质。真核生物的核糖体为 80S, 由 40S 和 60S 两个亚基构成, 包含 3~4 种 RNA 分子和 80 多种不同的蛋白质<sup>[1]</sup>。许多核糖体蛋白既具有结构上的功能, 同时又是翻译过程中必不可少的因子。随着多种物种的核糖体蛋白基因的相继克隆成功, 通过

同源分析查找核糖体蛋白保守氨基酸序列并推测其功能已成为核糖体蛋白研究的重要手段之一<sup>[2]</sup>。目前有关核糖体和核糖体蛋白的研究主要集中在昆虫、哺乳动物、真菌和细菌, 而植物方面的报道较少。

本研究采用生物信息学的方法, 以黑麦胞质核糖体蛋白基因 *ScRPS7*<sup>[3]</sup> 的 cDNA 序列为信息探针, 通过同源序列搜索、计算机拼接和 RT-PCR 的

收稿日期: 2003-03-13; 修回日期: 2003-07-09

基金项目: 教育部高等学校博士点基金资助项目(20010307003)[This work was supported in part by a grant from Ph. D program foundation of Education of Ministry of China (20010307003)]

作者简介: 顾志敏(1979-), 男, 硕士研究生, 专业: 作物遗传育种。E-mail: zhmg\_u\_student@sina.com

通讯作者: 张红生(1962-), 男, 教授, 博士生导师, 专业: 作物遗传育种。Tel: 025-84396075, E-mail: hszhang@njau.edu.cn

方法从水稻幼苗组织中分离了水稻核糖体蛋白 S7 基因的完整 cDNA 序列,并对其基因组结构和推测的氨基酸序列进行了分析。

## 1 材料和方法

### 1.1 植物材料

粳稻地方品种韭菜青由南京农业大学水稻研究所提供。种子经 10% 的次氯酸钠消毒,用蒸馏水冲洗后,在 30℃ 下浸种催芽,点播到置有纱布的瓷盘中,恒温光照培养箱(25℃)培养两周,取幼苗于液氮中速冻后,保存在-70℃冰箱备用。

### 1.2 RNA 的提取和 cDNA 第一链的合成

总 RNA 的提取按照 RNA 提取试剂盒(RNAex system IV for Plant,上海华舜公司)说明书。取 2 $\mu$ g 总 RNA 用于 cDNA 第一链的合成(Reverse transcript system, Promega),用 RNase 消化 cDNA 产物。

### 1.3 cDNA 序列的拼接

以已公布的黑麦胞质核糖体蛋白基因 *ScRPS7* (GenBank 注册号: AF118149) 的 cDNA 序列为信息探针,运行 blast 程序搜索华大水稻基因组数据库(<http://btn.genomics.org/rice>),寻找与之高度同源的水稻重叠群(contig)。采用黄骥等<sup>[4]</sup>推荐的方法对重叠群进行外显子预测和可译框(open reading frame, ORF)的拼接。根据拼接好的 ORF 搜索水稻 dbEST 数据库,下载与之部分匹配的 EST,进一步拼接获得全长 cDNA。

### 1.4 RT-PCR 扩增

根据拼接好的全长 cDNA 序列设计特异性引物 r1: 5'-cttcgctgcacggtgctag-3', r2: 5'-cgtcgactaac-tatgacatg-3', 以合成的 cDNA 第一链为模板,以 94℃ 预变性 5min; 94℃ 变性 40s, 60℃ 退火 50s, 72℃ 延伸 1min, 34 个循环; 72℃ 延伸 7min 为反应程序进行 PCR 扩增。扩增片段经 1% 琼脂糖凝胶电泳检测回收后连接到载体 pGEM-T (Promega), DNA 测序由上海申能博彩生物科技有限公司进行。

### 1.5 序列分析

cDNA 序列的 ORF 查找和翻译采用 DNAssist 2.0 软件,氨基酸序列的同源性比较采用 GeneDoc 软件,进化树的绘制和编辑采用 Clustalx 1.81 和 TreeView 1.6 软件,启动子区转录因子结合位点预测在 MatInspector(<http://www.genomatix.de>)程

序下进行。

## 2 结果与讨论

### 2.1 *OsRPS7* 的 cDNA 克隆与序列分析

以黑麦 *ScRPS7* 的 cDNA 序列为信息探针,在中国华大水稻基因组数据库中发现水稻基因组重叠群 1325 与之高度同源;根据外显子预测和计算机拼接的完整 cDNA 设计 PCR 引物,通过 RT-PCR 的方法克隆了水稻核糖体蛋白基因的 cDNA,该 cDNA 全长 919bp (图 1);包含 5' 端 102bp 和 3' 端 238bp 的非翻译区;在 5' 端第一个起始密码子 ATG 上游存在一个与之同读码框的终止密码子 TAG (图 1 加框处),因此克隆的 cDNA 中的可译框编码的是水稻全长胞质核糖体蛋白 *OsRPS7*。将该 cDNA 命名为 *OsRPS7* (GenBank 注册号: AF529175),其与黑麦 *ScRPS7* 的 cDNA 序列的核苷酸相似率为 78%。

*OsRPS7* 的可译框编码 192 个氨基酸,推测的氨基酸序列中含有 36 个可与 rRNA 相互作用的赖氨酸和精氨酸等保守氨基酸残基<sup>[2]</sup>。在 N 端 58~74 位还存在一个核定位序列(Nuclear Targeting Sequence, NTS) (图 1 下划线处),具有介导蛋白质核定位的功能<sup>[5]</sup>。

### 2.2 *OsRPS7* 与其他物种核糖体蛋白 S7 的序列比较

将 *OsRPS7* 与不同植物来源的核糖体蛋白 S7 进行了同源性比较(图 2),发现它与同属禾本科的黑麦的同源性最高,氨基酸的相似率和一致率分别为 95% 和 88%;而与十字花科拟南芥和芸薹<sup>[6]</sup>的核糖体蛋白 S7 的氨基酸相似率分别为 90% 和 88%,氨基酸的一致率均为 72%。*OsRPS7* 与其他物种的 S7 蛋白也有较高的同源性,如与人类<sup>[7]</sup> (P23821)、烟草天蛾<sup>[8]</sup> (P48155) 和酵母<sup>[9]</sup> (S52729) 的核糖体蛋白 S7 的氨基酸相似率分别为 73%、68% 和 64%,说明核糖体蛋白 S7 是一种进化上较为保守的蛋白质。

以 *OsRPS7* 的氨基酸序列搜索 GenBank 的 nr 数据库,运行 Clustalx 程序绘制进化树(图 3),结果表明:水稻、黑麦、拟南芥和芸薹同是植物,亲缘关系较近,归为同一类;烟草天蛾、黏虫、黑腹果蝇和冈比亚按蚊同是昆虫,归为一类;而人和小鼠同是哺乳动物,归在一类。

```

ggagcgcactgc 12
gaagcgaaccclleccgcgccclccgclcllecgclgcacgglgclagclcgccgccgclccgclccgcgcgcclcaaggglalcaaacg 102
atgtatcacagcaaggaggaagatccagaaggacaagggctggagccaactgagttgaggacaactgttgctcaggcaatffttgacatt 192
M Y T A R R K I Q K D K G L E P T E F E D T V A Q A F F D L
gagaalgggaalcaggagclgaagaglgactlgaaggacclllacalcaatgggacagllcagalggallaccllgccaacaggaagget 282
E N G N Q E L K S D L K D L Y I N G A V Q M D L P G N R K A
gllatalalcatglccatacaggclcgccaaggaacatalaagaagalccalglgaggcllglagggaaactlgagaagaaallcaglggg 372
V I I H V P Y R L R K A Y K K I H V R L V R E L E K K F S G
aaggatgtggttclgggtgcaactagaagaataglgaggeccccaaagaaaggetcagctgttgcctccctcgtaccgcacactact 462
K D V V L V A T R R I V R P P K K G S A V V R P R T R T L T
gclgllcatgalggcalclggaggatgllglgtalccagcagagallglgggaagcglcagataccacllggalggtagaaaaalc 552
A V H D G I L E D V V Y P A E I V G K R V R Y H L D G R K I
algaagalcllclggaccnaagagcgaacaacacccagtlacaagclcgacaclllcagctclclacaggaggctclglgggaag 642
M K I F L D P K E R N N T E Y K L D T F S S V Y R R L C G K
gacgtcgttgcactaccceatgaccgagactgcgtaaacatgtcatagttagtcgacggctatctffggccgcaaatgctgactc 732
D V V F D Y P M T E T A *
ttcctaaaaatagactataactlgtggtttcctatatlgcgaaaaglltgacagtcctgtgatctgcatactgattgctagaactggttta 892
aattgtcaggtgtttccaaactgcagttgtagaactgttgottgttcaatggttagtatttcaacgtgcaaaaaaaaaaaaaaaaa 912
aaaaaaa 919

```

图 1 水稻 *OsRPS7* 的 cDNA 的核苷酸及推测的氨基酸序列

Fig. 1 The sequence of nucleotide acids and deduced amino acids for *OsRPS7*

下划线示核定位序列; \* 示终止密码子

NTS is underlined; stop codon is indicated by asterisk (\*)

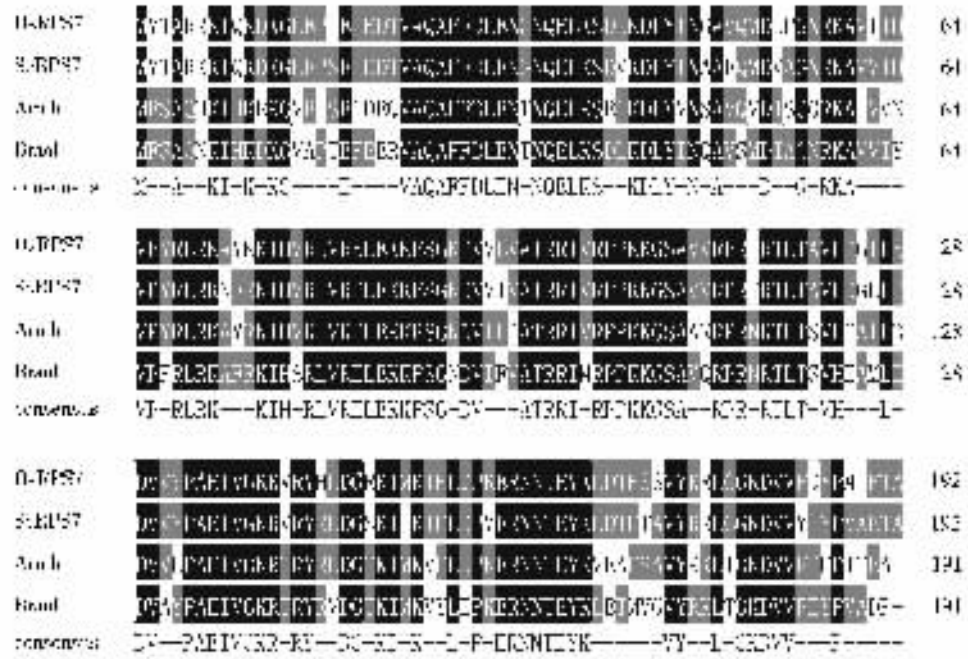


图 2 水稻 *OsRPS7* 氨基酸序列与黑麦 (AAD26256)、拟南芥 (AAM63913)、芸薹 (Q9XH45) 的相似性比较

Fig. 2 Alignment of amino acid sequences of ribosomal protein S7 from *Oryza sativa* (*OsRPS7*, AAN04468), *Secale cereale* (*ScRPS7*, AAD26256), *Arabidopsis thaliana* (*Arath*, AAM63913), *Brassica oleracea* (*Braol*, Q9XH45).

黑框表示序列一致性, 灰框表示序列相似性

Black boxes indicate sequence identity, and light gray boxes sequence similarity in the alignment.

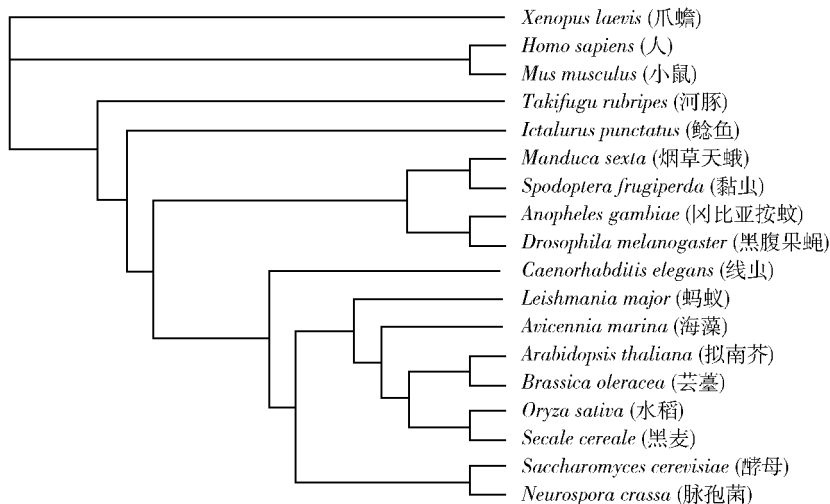


图 3 核糖体蛋白 S7 的进化树

进化树以水稻和其他物种已克隆的核糖体蛋白的氨基酸序列联配而构建。

Fig. 3 Phylogenetic tree of ribosomal protein S7

The phylogenetic tree was constructed from the amino acid sequence alignment of rice ribosomal protein S7 and other cloned ribosomal protein S7.

2.3 水稻 OsRPS7 的基因组结构分析

用生物信息学的方法对水稻 OsRPS7 的基因组结构进行了分析,并与拟南芥的核糖体蛋白基因 (AY086866)进行了比较(表 1),发现两个基因在基

因组结构上有很高的相似性,均包含 6 个外显子和 5 个内含子,外显子的大小相近,可译框都始于第二个外显子,说明不同植物核糖体蛋白 S7 基因的结构是高度保守的。

表 1 水稻与拟南芥 S7 蛋白基因结构比较

Table 1 Comparison of the S7 gene structure between rice and Arabidopsis

物种 Species	基因结构 Gene structure										
	E <sup>1)</sup> -1	I <sup>1)</sup> -1	E-2	I-2	E-3	I-3	E-4	I-4	E-5	I-5	E-6
水稻 Rice	未知 Unknown	707 <sup>2)</sup>	87	80	74	100	124	870	183	556	123
拟南芥 Arabidopsis	未知 Unknown	246 <sup>2)</sup>	118	76	73	113	125	234	184	83	118

1) E 和 I 分别表示外显子和内含子; 2) 表中数字表示外显子或内含子中的核苷酸数目。

1) E and I indicate exon and intron respectively; 2) The arabians indicate the number of nucleotide acids in intron or exon.

2.4 水稻 OsRPS7 启动子区转录因子结合位点预测

应用 MatInspector 程序预测了 OsRPS7 起始密码子上游 2kb 的基因组序列中的调控元件,发现该区域存在 1 个 W-box 元件,1 个 I-box 元件,2 个 G-box 元件和 2 个 ABA 应答元件(表 2)。

2.5 讨论

核糖体广泛存在于高等植物的细胞质、叶绿体和线粒体中。核糖体的生成可以分为 3 步,首先核

糖体蛋白在细胞质中合成并运送到核内;然后在核仁内与 rRNA 组装成核糖体;最后核糖体运送到胞质中加工成熟。由于真核生物的核糖体是由大约 80 种核糖体蛋白组成,所以核糖体的组装需要这 80 余种核糖体蛋白的均衡产生和一种共同协调的合成机制<sup>[10]</sup>。对于原核生物核糖体蛋白 S7 的研究结果表明,S7 是主要的 16S rRNA 结合蛋白之一,也是核糖体 A 位、P 位交叉连接 tRNA 分子的主要蛋白质<sup>[11]</sup>,位于 30S 核糖体小亚基的头部,面对着解码

中心区。在昆虫和哺乳动物等真核生物的核糖体中, S7 位于 40S 小亚基的表面, 所处区域被认为可能与蛋白质合成启动因子 eIF-3 交联<sup>[2]</sup>。本文首次

从单子叶模式植物水稻中克隆了核糖体蛋白 S7 基因, 将有助于进一步研究植物核糖体的结构和核糖体蛋白 S7 的生物学功能。

表 2 *OsRPS7* 启动子区可能的转录因子结合位点

Table 2 The possible binding sites for transcriptional factor in the promoter of *OsRPS7*

序列 Sequence	位置 Position	可能元件 Putative element	可能性 Probability
ggactTTGA <sup>1)</sup> ctatga	1247~1261 <sup>2)</sup>	W-box	0.92
aacaaGAAAagattctg	723~739	I-box	0.81
cgagagtGACGTggcat	264~280	G-box	0.92
gagacgagagtGACGTggcatctatt	258~284	ABA 应答元件	0.71
ggaagtGACGTggcat	217~233	G-box	0.92
agaggaaagtGACGTggcatctgctc	211~237	ABA 应答元件	0.71

1) 核心序列用大写字母表示, 1) Basepairs in capital denote the core sequence.

2) 阿拉伯数字表示调控元件和起始密码子之间的核苷酸数, 2) The arabians indicate the numbers of nucleotide acids between elements and start codon.

已经克隆的不同生物的核糖体蛋白 S7 基因所编码的氨基酸数目都在 190 个左右, 根据基因的基因组结构和推测的氨基酸序列比较, 说明核糖体蛋白基因是非常保守的。结合其启动子区域预测的调控元件, 表明核糖体蛋白 S7 基因的功能可能是多方面的。许多植物在受到低温胁迫时, 体内的一些代谢过程发生变化, 如优化蛋白质合成体系的翻译能力, 加快相关基因的合成, 从而适应环境的改变, 抵抗低温胁迫的伤害。黑麦 *ScRPS7* 基因在分生组织的表达量高于非分生组织; 幼苗受到低温胁迫时, *ScRPS7* 表达量急剧降低, 以后逐渐回升, 当低温胁迫 30 天后, 其表达量恢复到低温胁迫前的水平, 从而维持低温下黑麦幼苗的生长<sup>[3]</sup>。在水稻中是否也有类似的现象还不清楚, 须进一步研究该基因的表达特性和表达谱。

## 参 考 文 献 (References):

[1] Wool I G. The structure and function of Eukaryotic ribosomes. *Annu Rev Biochem*, 1979, 48: 719~754.

[2] Xu Wen-Yue, Huang Fu-Sheng, Zhang Jing-Ru, Duan Jian-Hua. cDNA cloning of ribosomal protein S7 from *Anopheles dirus*. *Immunological Journal*, 2001, 17(5): 400~401.  
徐文岳, 黄复生, 张敬如, 段建华. 大劣按蚊核糖体蛋白 S7cDNA 克隆. *免疫学杂志*, 2001, 17(5): 400~401.

[3] Thomas B, Michael U, Ju'rgen F. cDNA cloning of cytoplasmic ribosomal protein S7 of winter rye (*Secale cereale*) and its expression in low-temperature-treated leaves. *Biochim Biophys Acta*, 2000, 1492: 276~279.

[4] Huang Ji, Zhang Hong-Sheng, Cao Ya-Jun, Qian Xiao-Yin, Yang Jin-Shui. Strategy of in silico cloning of functional genes in rice (*Oryza sativa*). *Chinese J Rice Sci*, 2002, 16(4): 295~298.  
黄 骥, 张红生, 曹雅君, 钱晓茵, 杨金水. 水稻功能基因的电子克隆策略. *中国水稻科学*, 2002, 16(4): 295~298.

[5] Dingwall C, Laskey R A. Nuclear targeting sequences - A consensus. *Trends Biochem Sci*, 1991, 16: 478~481.

[6] Gossen K K, Katavic V, Taylor D C. Nucleotide sequence (Accession No. AF144752) of a cDNA from microspore derived embryos of *Brassica oleracea* is homologous to 40S ribosomal protein S7. (PGR99-096). *Plant Physiol*, 1999, 120: 934.

[7] Annilo T, Laan M, Stahl J, Metspalu A. The human ribosomal protein S7-encoding gene: isolation, structure and localization in 2p25. *Gene*, 1995, 165: 297~302.

[8] Jiang H., Wang Y, Kanost M R. Primary structure of ribosomal proteins S3 and S7 from *Manduca Sexta*. *Insect Mol Biol*, 1996, 5(1): 31~37.

[9] Takakura H, Tsunasawa S, Miyagi M, Wamer J R. NH<sub>2</sub>-terminal acetylation of ribosomal proteins of *Saccharomyces cerevisiae*. *J Biol Chem*, 1992, 267(8): 5442~5445.

[10] Mager W H. Control of ribosomal protein gene expression. *Biochem Biophys Acta*, 1988, 949: 1~15.

[11] Wimberly B T, White S W, Ramakrishnan V. The structure of ribosomal protein S7 at 1.9 Å resolution reveals a beta-hairpin motif that binds double-stranded nucleic acids. *Structure*, 1997, 5(9): 1187~1198.

[12] Menken A E, Schindler U, Cashmore A R. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZip proteins. *Trends Biochem Sci*, 1995, 20: 506~510.