

小麦 *TaMBD4* 基因的克隆、结构及表达的初步分析

孟凡荣¹, 刘昊英¹, 司志飞¹, 尹钧², 李永春^{2,*}

河南农业大学¹生命科学学院, ²国家小麦工程技术研究中心, 郑州 450002

摘要: 以拟南芥 MBD 基因的 EST 为基础, 采用电子克隆并结合 RT-PCR 方法分离克隆了包含开放阅读框的小麦甲基结合域蛋白基因 *TaMBD4*。序列分析显示, *TaMBD4* 蛋白有典型的甲基结合域。组织表达特性分析表明, *TaMBD4* 在干种子和胚乳中的表达量高于其它组织。*TaMBD4* 的 cDNA 和基因组 DNA 比较分析显示, 此基因包括 1 个内含子, 进一步分析表明这个内含子为 2 个 GGCAGT 序列的串联重复, 推测该内含子可能与 *TaMBD4* 基因的转录后调控相关。

关键词: DNA 甲基化; 小麦; *TaMBD4*; 克隆; 基因结构和表达

Preliminary Study on Cloning, Structure and Expression of *TaMBD4* Gene in Wheat

MENG Fan-Rong¹, LIU Hao-Ying¹, SI Zhi-Fei¹, YIN Jun², LI Yong-Chun^{2,*}

¹College of Life Science, ²National Engineering Research Centre for Wheat, Henan Agricultural University, Zhengzhou 450002, China

Abstract: In order to provide an insight into molecular characteristics and putative epigenetic functions of MBD in wheat, a gene with complete open reading frame (ORF) designated as *TaMBD4* was identified in wheat by *in silico* cloning approach and its fragment was cloned by RT-PCR. Amino acid sequence analysis showed that *TaMBD4* contained a conserved methyl-CpG binding domain, which might play a putative role in epigenetic mechanism. Spatial expression analysis showed that the expression level of *TaMBD4* was higher in dry seeds and endosperm than that in other tissues tested. The comparative analysis between cDNA and genomic DNA of the gene showed that a short intron was included. Interestingly, we found that the intron was composed of 2 tandem simple sequence repeats, which might be involved in the posttranscriptional regulation of the *TaMBD4*.

Key words: DNA methylation; wheat; *TaMBD4*; cloning; genomic structure and expression

DNA 甲基化是生物体内普遍存在的一种基因修饰现象。植物基因组中 DNA 甲基化一般发生在 mCpG、mCpNpG 和 mCpNpN (N=A、C 或 T) (Gruenbaum 等 1981; Matzke 等 1989; Tariq 和 Paszkowski 2004), 在植物组织特异性或发育调控性基因的表达中起重要作用, 高等植物细胞在不同的分化与发育阶段具有不同的甲基化模式, DNA 甲基化的动态变化是维持植物个体正常生长发育的分子调控机制之一(Cao 等 2000; Finnegan 等 2000; Takeda 和 Paszkowski 2006)。已有研究表明, DNA 甲基化与基因表达调控密切相关。启动子及基因编码区域的 DNA 甲基化或脱甲基化可以抑制或激活基因的转录, DNA 甲基化调控基因转录的分子机制有 2 种可能: 一是 DNA 甲基化直接干扰转录因子与启动子的结合(Eden 和 Cedar 1994); 二是甲基化的 DNA 上特异结合了甲基结合蛋白(methylation binding domain protein,

MBD), 进而招募转录抑制复合体(包括共抑制因子和组蛋白脱乙酰基酶等), 诱发组蛋白的脱乙酰化, 导致染色体结构的变化和转录的抑制(Springer 和 Kaeppler 2005; Ballestar 和 Esteller 2005)。多数的实验都支持第 2 种机制。甲基结合域蛋白 MBD 是与甲基化 DNA 特异结合的反式作用因子, 在基因表达调控和维持生物体正常生长发育中起作用。在脊椎动物中, 共分离到 5 个编码 MBD 蛋白的基因, 其中 MBD1、MBD2、MBD3 和 MeCP2 直接参与 DNA 甲基化的转录抑制; 而 MBD4 蛋白有 MBD 保守域和糖基化酶两个保守域, 主要在修复系统中起作用, 与转录失活无关(Nan 等 1998; Ng 等 1999, 2000; Jørgensen 和

收稿 2007-08-29 修定 2007-12-24

资助 国家自然科学基金项目(30300195)。

* 通讯作者(E-mail: yongchunli71@yahoo.com.cn; Tel: 0371-63558215)。

Bird 2002)。近年来, 植物MBD的研究已取得初步进展。拟南芥、水稻、玉米基因组分别编码13、16和16个MBD蛋白(Zemach和Grafi 2003; Springer和Kaeppler 2005), 小麦MBD的研究非常少。为了探讨小麦中MBD基因的结构与功能, 本文以拟南芥编码MBD基因的EST为基础, 采用电子克隆并结合RT-PCR方法从小麦中分离克隆到一个包含完整开放阅读框(open reading frame, ORF)的MBD基因*TaMBD4*, 分析其编码蛋白的结构特性和研究*TaMBD4*的组织表达特性以及基因结构, 为进一步探讨小麦MBD基因的生物学功能积累了基础资料。

材料与方法

以普通小麦(*Triticum aestivum* L.)品种‘郑麦9023’干种子为材料, 进行小麦MBD基因的分离克隆。分别取干种子、萌动24 h的胚和胚乳、萌动48 h的胚根和胚芽、三叶期叶片及茎, 提取RNA用于研究*TaMBD4*基因的组织表达特性。

总RNA提取采用热酚法(倪中福等2000), 并参照Lohmann (1995)的方法纯化。

cDNA的合成采用20 μ L反应体系[含2 μ g总RNA、5 mmol·L⁻¹ Tris-HCl (pH 8.3)、75 mmol·L⁻¹ KCl、3 mmol·L⁻¹ MgCl₂、10 mmol·L⁻¹ DTT、50 mmol·L⁻¹ dNTP、50 pmol锚定引物、200 U M-MLV反转录酶], 37 °C温育2 h, 取2 μ L反应液用于PCR扩增。

以拟南芥MBD基因的EST作为基础序列, 用BLASTN对dbEST进行同源性搜索, 获得有片段重叠和同源性高的小麦ESTs; 严格聚类分析, 尽量避免含有旁系同源基因, 拼接后产生序列重叠群(contig), 以新获得的contig为目标序列重复进行上述分析步骤, 直至不能获得延伸为止; 最后, 通过BLASTX验证该拼接片段是否符合预测。

针对ORF区域设计特异PCR引物进行扩增, 引物序列为TaMBD4-L: 5' GCAAGATCCCAA-CGAAAG 3'; TaMBD4-R: 5' AGAGGCATCATA-AAGTCCATAC 3'。PCR反应采用40 μ L体系, 其中包括4 μ L反转录产物或者40~50 ng DNA、10 \times PCR缓冲液4 μ L、10 mmol·L⁻¹ dNTP 1.6 μ L、特异引物TaMBD4-L和TaMBD4-R各80 ng、Taq DNA聚合酶2 U。扩增程序为94 °C 5 min; 94 °C

1 min, 53 °C 1 min, 72 °C 1 min, 40个循环; 72 °C后延伸10 min。

PCR产物经1%琼脂糖凝胶电泳后回收克隆到T载体中进行测序分析。凝胶回收采用北京天为时代公司的DNA凝胶回收试剂盒, 克隆载体采用Promega公司的pGEM-T Easy载体, 测序在北京三博远志公司进行; cDNA序列在GenBank进行BLAST搜索, 进行同源性比较, cDNA序列的初步分析、氨基酸翻译采用DNAMAN软件, 序列的氨基酸比对分析采用Clustal X软件; 蛋白质序列以及结构分析通过ExPaSy (<http://www.expasy.org>)进行。

用扩增ORF时的引物进行RT-PCR分析, 为防止扩增达到平台期, 分别进行28、30和35个循环的扩增, 最后显示的是30个循环的检测结果; 选用小麦 β 肌动蛋白基因为对照, 引物序列为Actin-1: 5' CAGCAACTGGGATGATATGG 3'和Actin-2: 5' ATTTTCGCTTTCAGCAGTGGT 3'。

为了分析基因组中此基因的结构, 采用CTAB方法(Sagehai-Maroo等1984)提取小麦幼苗的总DNA。以TaMBD4-L和TaMBD4-R为引物, 基因组DNA为模板, PCR扩增了*TaMBD4*的基因组片段, 并经过回收、连接、转化, 选择阳性克隆送公司测序, *TaMBD4*的cDNA序列与基因组序列的比对用DNAMAN软件进行。

结果与讨论

1 小麦*TaMBD4*基因cDNA的克隆和序列分析

以拟南芥MBD基因EST作为基础序列, 用BLASTN对dbEST进行同源性搜索, 获得同源性高且有部分重叠的2个小麦EST (BJ241681和BJ236657), 进行拼接后得到一条包含完整ORF的序列, 命名为*TaMBD4*; 提取小麦品种‘郑麦9023’干种子的RNA, 经反转录得到的cDNA为模板, 以TaMBD4-L和TaMBD4-R为引物, 获得与目标大小一致的条带(图1)。经过回收、克隆和序列分析表明, *TaMBD4*序列长度为507 bp, 其中编码区为300 bp, 5'非翻译区为143 bp, 3'非翻译区为64 bp(图2)。推测编码蛋白的分子量为11.25 kDa, 等电点的理论值为9.34。

2 小麦*TaMBD4*蛋白的结构分析

BLASTX搜索结果显示, 小麦*TaMBD4*蛋白

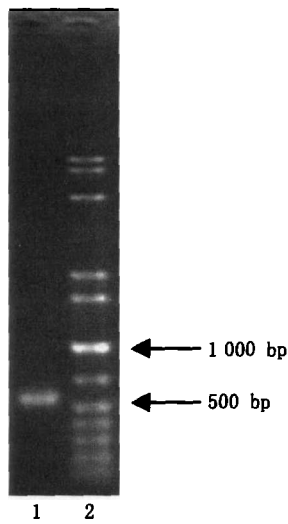


图1 目的基因的PCR产物

Fig.1 PCR product of targeted gene

1: 扩增产物; 2: DNA 分子量标准。

与来自不同物种的甲基结合蛋白氨基酸具有一定的相似性, 其中与拟南芥 *AtMBD13*、玉米 *MBD113* 和 *MBD114*、水稻 *MBD704* 和 *MBD705* 的同源关系相对较近(图3)。氨基酸序列分析表明, *TaMBD4* 编码蛋白质中具有甲基结合蛋白的保守功能域(第4~74个氨基酸), 组成有典型环的 α/β 夹层结构。按照 Springer 和 Kaepler (2005) 的分类方法, 应该属于第 VIII 亚类, 此类 MBD 蛋白的相对分子量都较小, 只包含单一的 MBD 保守域结构。

另外, 小麦 *TaMBD4* 蛋白还包含多个翻译后修饰位点: 2 个蛋白激酶 C 的磷酸化位点、1 个酪蛋白激酶磷酸化作用位点, 在小麦中属于新克隆的基因。

3 小麦 *TaMBD4* 基因的组织表达特性分析

为保证各材料总 RNA 反转录产物中的 cDNA

```

1   GCAAGATCCCAACGAAAGAGGAGTTCGAGACGATCCGCGAGAGCTTAGCCGAGGACCCC
60  TGGTTCTGTGGCAGGGACCCAGCGCCGGCCGTTCCCTGCAAGCAGCCGGAGGACATTCGG
120 TGCACAGCAGCTGCATCTGGGTCATGGACAAGCCAGGCATCCCGCGCCCGCCACCTGCC
      M D K P G I P R P P P A
180 ACGAAGCGGCTGGTGATCATGCGGCGTGACCTGTCCAAGATGGACACCTACTACTTGCTG
      T K R L V I M R R D L S K M D T Y Y L L
240 CCCAACGGGAAGCGCGCGCGTCCGGCAACGACGTGGAGAAGTTCCTCCAGGAGAACCCG
      P N G K R A R S G N D V E K F L Q E N P
300 GAGTATAGGGCGAACCTGCCGGCATCAAAGTTCTCCTTCGCCACGCCAAGATCGTTCCG
      E Y R A N L P A S K F S F A T P K I V P
360 GCGACTATTGGAGAGAGCTCACTATGGAGGGTTGCCAGGCCGAGGGGAAAAGTTTGAG
      A T I G E S S L W R V A Q A E G E K F E
420 GAGATGGACGTGTTCCGATTCTGATTACTATTACTATCCGAGCGGCAGCAACTAGGTACT
      E M D V F G F *
480 ATGGAAGTATGGACTTTATGATGCCTCT
  
```

图2 小麦 *TaMBD4* 基因核苷酸及推测的氨基酸序列

Fig.2 Nucleotide and deduced amino acid sequence of *TaMBD4* gene wheat

下划线为 MBD 保守域。

含量的一致性, 以小麦 β 肌动蛋白基因的 RT-PCR 扩增产物为对照。结果表明, 所检测的 7 个组织中的 *TaMBD4* 表达量有差异, 干种子和萌动 24 h 的胚乳中的表达量高, 萌动 48 h 的胚芽和萌动 24 h 的胚中的表达量次之, 三叶期叶片和茎、萌动 48 h 的胚根中的表达量极低(图4)。这与 Springer 和 Kaepler (2005) 关于玉米 MBD 基因时空表达特性的研究结果相似, 玉米 *MBD114*、*MBD119* 和

MBD120 在胚乳、种子、幼苗、根尖、幼叶和成熟叶中存在组织间特异表达现象, 其余玉米 MBD 基因在所检测各组织中均表达, 但不同组织间表达量有差异。Berg 等(2003)对拟南芥 MBD 的表达特性分析也表明, *AtMBD1*、*AtMBD2*、*AtMBD4*、*AtMBD5*、*AtMBD6* 和 *AtMBD11* 在花、花序、干果、叶片和茎秆等组织中均表达, 但不同基因的表达量在不同组织中也有明显差异。

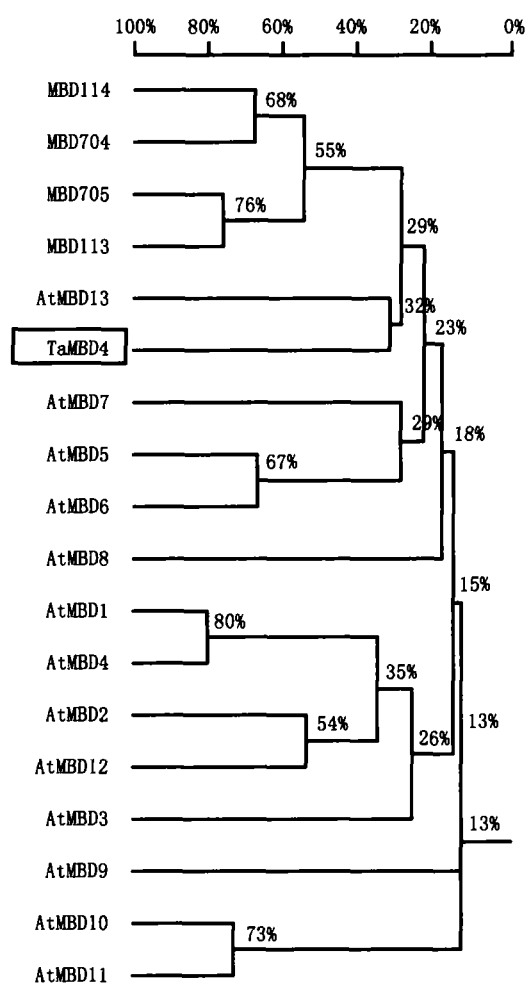


图3 小麦 TaMBD4 与来自拟南芥、玉米和水稻的 MBD 蛋白氨基酸序列同源性比较

Fig.3 Homophylic comparison of TaMBD4 in wheat with MBD proteins from Arabidopsis, corn, and rice

AtMBD1~13: 拟南芥; MBD113~MBD114: 玉米; MBD704 水稻; TaMBD4: 小麦。分支上数字代表相似性。

本文结果表明, *TaMBD4* 在干种子和胚乳的表达量高于其他组织, 这种表达模式可能与胚乳和干种子中基因的选择性沉默有关(Gehring 等 2004), 也可能与胚乳中基因的基因组印迹有关(Grossniklaus 等 2001), *TaMBD4* 的组织表达模式

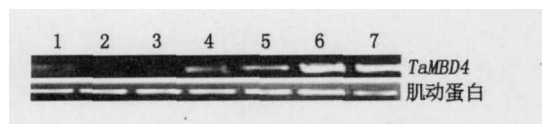


图4 小麦 *TaMBD4* 的组织表达特性分析
Fig.4 Expression patterns of *TaMBD4* in different wheat tissues

1: 叶片; 2: 茎; 3: 胚根; 4: 胚芽; 5: 胚; 6: 胚乳; 7: 干种子。

暗示其可能参与调控胚乳和干种子相关基因的沉默。至于 *TaMBD4* 与甲基化 DNA 的互作及其在小麦种子发育过程中的生物学功能的实验正在进行中。

4 小麦 *TaMBD4* 的基因组结构分析

用小麦基因组 DNA 为模板, 利用 PCR 方法扩增 *TaMBD4* 的基因组序列, 测得的 *TaMBD4* 的 DNA 序列与以前获得的 cDNA 序列进行比对的结果表明: 此基因在终止密码子后的 3' UTR 处有一个 12 bp 的内含子(图 5), 进一步分析表明, 此内含子是 2 个简单序列 GGCAGT 的串联重复, 推测此内含子可能与 *TaMBD4* 转录后调控有关。关于此基因内含子的调控功能及其在小麦生长发育过程中的生物学功能研究正在进行。

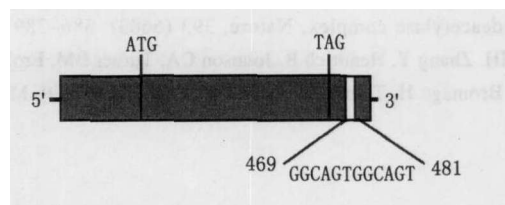


图5 小麦 *TaMBD4* 的基因组结构图

Fig.5 Structure graph of *TaMBD4* in wheat

黑色部分代表外显子区域, 空白部分代表内含子区域, 数字代表内含子的位置。

参考文献

- 倪中福, 孙其信, 吴利民(2000). 普通小麦不同优势杂交种及其亲本之间基因表达差异比较研究. 中国农业大学学报, 5 (1): 1~8
- Ballestar E, Esteller M (2005). Methyl-CpG-binding proteins in cancer: blaming the DNA methylation messenger. *Biochem Cell Biol*, 83 (3): 374~384
- Berg A, Meza TJ, Mahić M, Thorstensen T, Kristiansen K, Aalen RB (2003). Ten members of the *Arabidopsis* gene family encoding methyl-CpG-binding domain proteins are transcriptionally active and at least one, *AtMBD11*, is crucial for normal development. *Nucleic Acids Res*, 31 (18): 5291~5304
- Cao XF, Springer NM, Muszynskii MG, Phillips RL, Kaeppler S, Jacobsen SE (2000). Conserved plant genes with similarity to mammalian *de novo* DNA methyltransferases. *Proc Natl Acad Sci USA*, 97 (9): 4979~4984
- Eden S, Cedar H (1994). Role of DNA methylation in the regulation of transcription. *Curr Opin Gene Dev*, 4: 255~259
- Finnegan EJ, Peacock WJ, Dennis ES (2000). DNA methylation, a key regulator of plant development and other processes. *Curr Opin Gene Dev*, 10 (2): 217~223
- Gehring M, Choi Y, Fischer R (2004). Imprinting and seed development. *Plant Cell*, 16: S203~S213

- Grossniklaus U, Spillane C, Page DR, Köhler C (2001). Genomic imprinting and seed development: endosperm formation with and without sex. *Curr Opin Plant Biol*, 4: 21~27
- Gruenbaum Y, Naveh-Many T, Cedar H, Razin A (1981). Sequence specificity of methylation in higher plant DNA. *Nature*, 292: 860~862
- Jørgensen HF, Bird A (2002). MeCP2 and other methyl-CpG binding proteins. *Ment Retard Dev Disabil Res Rev*, 8 (2): 87~93
- Lohmann J, Schickle H, Bosch TC (1995). REN display, a rapid and efficient method for nonradioactive differential display and mRNA isolation. *Biotechniques*, 18 (2): 200~202
- Matzke MA, Primig M, Trnovsky J, Matzke AJM (1989). Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. *EMBO J*, 8 (3): 643~649
- Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, 393 (6683): 386~389
- Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Na Genet*, 23: 58~61
- Ng HH, Jeppesen P, Bird A (2000). Active repression of methylated genes by the chromosomal protein MBD1. *Mol Cell Biol*, 20: 1394~1406
- Saghai-Marouf MA, Soliman KM, Jørgensen RA, Allard RW (1984). Ribosomal DNA space length polymorphisms in barley: mendelian inheritance, chromosomal locations, and population dynamics. *Proc Natl Acad Sci USA*, 81: 8014~8018
- Springer NM, Kaeppler SM (2005). Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins. *Plant Physiol*, 138 (1): 92~104
- Takeda S, Paszkowski J (2006). DNA methylation and epigenetic inheritance during plant gametogenesis. *Chromosoma*, 115 (1): 27~35
- Tariq M, Paszkowski J (2004). DNA and histone methylation in plants. *Trends Genet*, 20 (6): 244~251
- Zemach A, Grafi G (2003). Characterization of *Arabidopsis thaliana* methyl-CpG-binding domain (MBD) proteins. *Plant J*, 34 (5): 565~572