

植物MYB2转录因子

陈焕新, 关秋玲, 李秋莉*

辽宁师范大学生命科学学院, 辽宁大连 116029

MYB2 Transcription Factors of Plant

CHEN Huan-Xin, GUAN Qiu-Ling, LI Qiu-Li*

School of Life Sciences, Liaoning Normal University, Dalian, Liaoning 116029, China

摘要: MYB转录因子是植物转录因子中最大的家族之一, 以含有保守的MYB结构域为共同特征。MYB2转录因子是MYB家族的一个亚家族, 本文介绍植物MYB2蛋白的结构、与顺式作用元件结合的位点、进化以及生物学功能的研究进展。

关键词: 植物MYB家族; MYB2转录因子; 结构; 结合位点; 进化; 功能

植物在长期进化过程中, 为适应和抵抗各种生物和非生物的胁迫, 自身形成了多种机制和复杂的信号网络, 从而能迅速地感知胁迫, 主动调控以忍耐胁迫反应。基因表达的转录调控在其响应胁迫的过程中起作用的转录因子(transcription factor, TF)是植物中最重要的一类调节基因。

转录因子也称为反式作用因子, 是指能够与真核基因的顺式作用元件发生特异性相互作用并对转录有激活或抑制作用的DNA结合蛋白(吴乃虎2001)。转录因子是具有序列特异位点的结合活性或含有同已知DNA结合结构域相同特征的蛋白质, 因而能保证目的基因以特定的强度、在特定的时间与空间表达。它通过与靶位点上顺式作用元件结合, 调节靶基因的转录速度。植物许多基因的表达都受到特定的转录因子与特定的顺式作用元件相互作用的调控。自1987年Paz-Ares首次报道玉米转录因子的克隆以来, 研究者们相继从高等植物中分离出了调控干旱、高盐、低温、激素、病原反应及生长发育等相关基因表达的转录因子已达数百种(刘强等2000)。典型的转录因子一般具有4个功能区: DNA结合区、转录调控区、核定位信号区和寡聚化位点。转录因子通过这些功能区域与顺式元件相互作用或者与其他转录因子的功能区域相互作用来调控基因的表达。

1982年, Klempnauer等在禽成髓细胞瘤病毒(avian myeloblastosis virus)中鉴定出一个能直接导致急性成髓细胞白血病(acute myeloblastic leukemia)的癌基因, 称为*v-myb*, 不久发现在正常动物细胞中也存在相应的原癌基因*c-myb*, 随后研究的结果表明, *v-MYB*、*c-MYB*蛋白都定位在细胞核中, 与

核基质和染色质紧密相连, 而且都具有DNA结合活性和转录调节功能。玉米的*cl*基因所编码的蛋白是第一个从植物中发现的MYB转录因子(Paz-Ares等1987), 后来的研究又相继在拟南芥和玉米发现大量的MYB转录因子, 它们在转录调节中起着多方面的作用。MYB转录因子是植物转录因子家族中最大的家族之一, 有着广泛的生理功能, 几乎参与植物发育和代谢的各个方面。植物MYB2转录因子是MYB大家族中的一个小的亚族, 不同植物的MYB2基因有着不同的生物学功能, 但它们都是在转录水平上调控植物各个阶段的生长发育。本文介绍植物MYB2蛋白的结构、与顺式作用元件结合的位点、进化以及生物学功能的研究进展。

1 植物MYB2转录因子的结构特点

MYB家族有高度保守的结合域——MYB结构域, 每个结构域含有50~53个氨基酸残基, 形成螺旋-转角-螺旋(helix-turn-helix)的构型(图1); 每个结构域中有3个保守的色氨酸残基, 间隔为18~19个氨基酸, 是疏水核心的重要成分, 对于维持螺旋-转角-螺旋的构型起着非常重要的作用(Jin和Martin 1999)。根据所含MYB结构域的数量, 植物中的MYB类转录因子可分为单一MYB结构域蛋白(single-MYB)、2个串联MYB结构域蛋白(R2R3-MYB)和3个串联MYB结构域蛋白(R1R2R3-MYB)3个亚类。其中数量庞大、功能多样的是R2R3-MYB蛋白, 仅在拟南芥中就有120多种, 估计在其

收稿 2008-07-18 修定 2008-09-09

资助 国家自然科学基金(30370806)。

* 通讯作者(E-mail: liqiuli@dl.cn; Tel: 0411-84258681)。



图1 小鼠c-MYB蛋白的R3结构域:螺旋-转角-螺旋构型(Furukawa等1996)

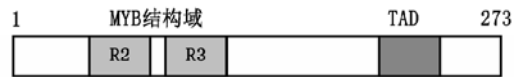


图2 AtMYB2蛋白的结构简图

根据文献(Yoo等2005)绘制。AtMYB2基因的cDNA全长1138 bp,其中83~904碱基是它的CDS区,共编码273个氨基酸。图中R2和R3表示AtMYB2蛋白有2个MYB结构域,TAD表示它的转录激活区域。

他高等植物中也会有相应的甚至更大数量的R2R3-MYB转录因子。

MYB2转录因子含有2个串联MYB结构域(图2),属于R2R3-MYB类转录因子。比较12种植物的MYB2蛋白的氨基酸序列结果表明,MYB2的两个串联MYB结构域中氨基酸序列比较保守,R2中含有3个保守的色氨酸残基,分别间隔19个氨基酸,R3中只含有后2个保守的色氨酸残基,间隔18个氨基酸,而第一个色氨酸残基被异亮氨酸或苯丙氨酸所取代(图3)。

这些保守的色氨酸残基构成了MYB2蛋白的疏水核心,MYB2蛋白借助此种结构插入靶DNA分子大沟与目的基因结合。

研究表明,每一重复子的C端螺旋是MYB蛋白结合DNA的识别螺旋,其中R3的识别螺旋特异性地与其识别序列的核心结合,而R2的识别螺旋与核苷酸的结合专一性较差(Ogata等1995)。此外,大多数转录因子在C端与DNA结合域之间都具有一个富含酸性氨基酸的转录激活功能域。Urao等(1996)通过在拟南芥叶片原生质体中进行瞬时表达实验,分析AtMYB2基因被干旱诱导后转录激活活性的结果表明,AtMYB2转录因子通过序列

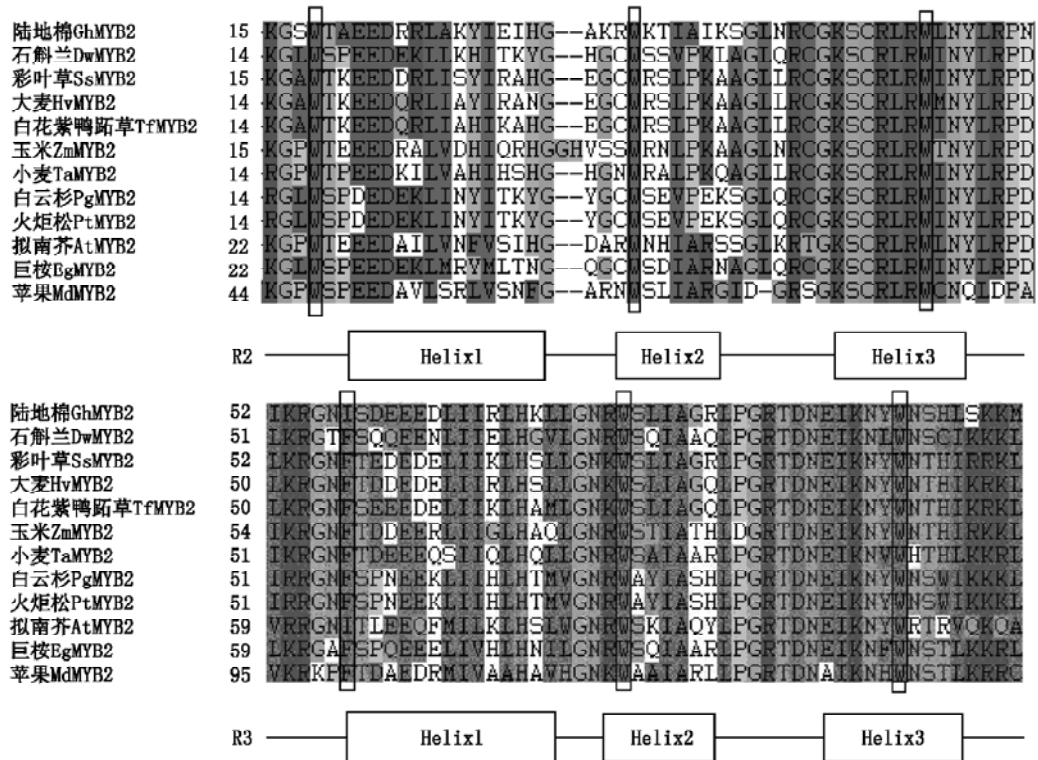


图3 12种植物MYB2蛋白的保守氨基酸序列比较

颜色一致的格表示氨基酸序列相同,黑色方框围起的氨基酸序列是保守的色氨酸残基,其中在R3结构域中,第一个色氨酸残基被异亮氨酸或苯丙氨酸所取代。12种植物MYB2蛋白的氨基酸序列均是来自于GenBank。

特异性方式结合到MYB的识别位点上, 转录激活报告基因(MBS × 5::GUS, MBS位点包含一段TAACTG序列)。对 AtMYB2 的 C 端区域基因删除后, 报告基因的转录激活降低, 这进一步证明 AtMYB2 的 C 端酸性区域含有其转录激活的足够氨基酸序列。他们的结果表明, AtMYB2 蛋白作为一个转录激活子, 它的 C 端酸性区域是其行使转录激活功能的部位。

2 MYB2 转录因子的识别位点

转录因子只有与顺式作用元件上的识别位点相结合, 才能行使它的生物学功能, 通过凝胶阻滞分析和定点突变等常规实验方法, 人们已经确定了一些 MYB2 转录因子的识别序列。

Iwasaki 等(1995)对拟南芥的一个脱水响应基因 *rd22* 的研究发现, *rd22* 基因启动子中一段 67 bp 的 DNA 序列上有一个 AtMYB2 转录因子的识别位点 TGGTTAG (其互补序列是 CTAACCA)。Hoeren 等(1998)通过对低氧诱导下, 拟南芥的 AtMYB2 转录因子与 *ADH1* (乙醇脱氢酶)基因启动子的相互作用的研究, 证明 AtMYB2 在 *ADH1* 基因启动子上的识别位点是 MBS-2 box 上的一个 GT 基序(5' TGGTTT 3'), 进一步的实验表明, GT 基序中的碱基序列 AAC 是 AtMYB2 转录因子识别 *ADH1* 基因启动子的核心序列。Wang 等(2004)在研究棉花 GaMYB2 蛋白在表皮毛生长中的作用时, 使用

PLACE 软件分析方法找到了 GaMYB2 的结合位点 (CAGTTG), 进而通过定点突变实验, 确定了 GaMYB2 的作用位点。

3 植物 MYB2 转录因子的进化

1996 年, Lipsick 提出了一个 MYB 基因的进化模型, 由于后来又在植物中发现了 *R1R2R3-MYB* 基因, Jin 和 Martin (1999)对此模型进行了适当的补充。这一模型认为大约在 10 亿年以前产生了 MYB 结构域, MYB 区的复制导致了含有多个不完全重复 MYB 区的蛋白(*R1R2R3-MYB* 蛋白)的产生, 接下来的整个基因的扩增(在动物身上扩增比较有限)产生了数目繁多的 MYB 蛋白, 由于 R1 的丢失, 从而产生了 *R2R3-MYB* 蛋白。

Chen 等(2006)通过拟南芥中的 MYB 基因家族与水稻的 MYB 基因家族进行系统发生学的比较, 认为 MYB 基因家族在从单子叶向双子叶植物进化过程中经历了一个快速的进化扩张过程。MYB 家族其他基因比 *R2R3-MYB* 基因更加原始, 或者说是进化更快。因此, MYB 超基因家族可以作为很好的完整的体系, 用来研究高等植物中复杂的基因家族的进化。MYB2 在不同物种中的进化如图 4 所示, 从图 4 可以看出, 2 种裸子植物聚在一起, 6 种单子叶植物聚在一起, 拟南芥、巨桉、苹果等双子叶植物与裸子植物、单子叶植物亲缘关系较远, 这与传统的分类方法完全相同, 但陆地棉则聚到了单子

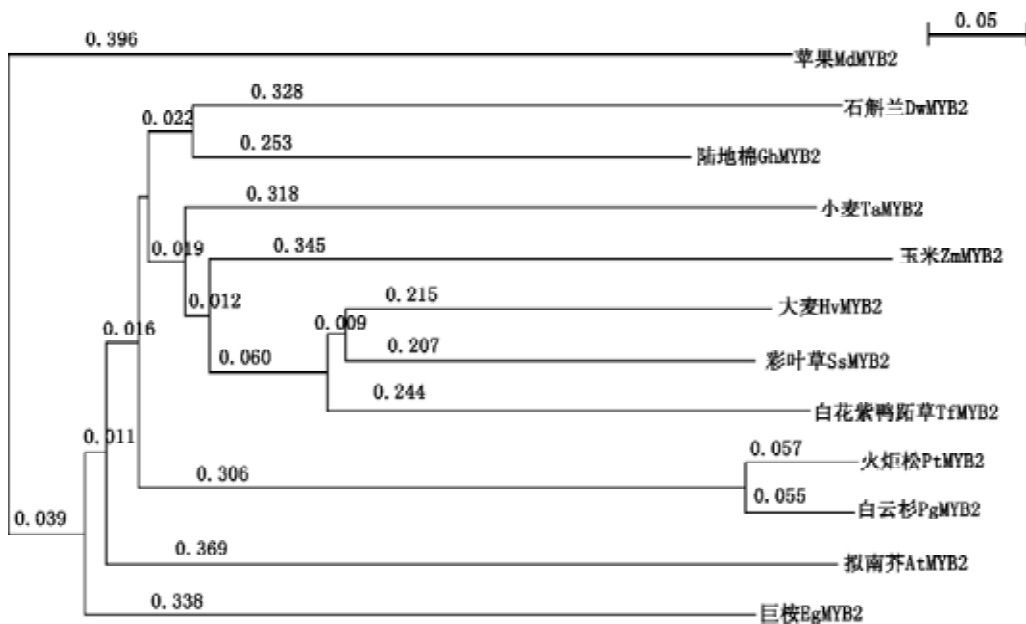


图 4 MYB2 在不同物种中的进化

叶植物群中,这与传统的分类方法相比出现了差异。

Kamiya等(2002)通过研究拟南芥20个生态型的 *AtMYB2* 基因的2.2 kb区域的DNA多样性,发现该区域核酸多样性系数为0.0027,低于拟南芥其他基因。*AtMYB2*的MYB结构域核酸多样性系数为0.0036,高于non-MYB区(0.0013)。通过HKA检测,比较这20个拟南芥 *AtMYB2* 基因多样性比率和分歧度,发现两者差异性比较大。低的多样性和高的分歧度主要出现在基因的non-MYB区,而non-MYB结构区,不同于一般的突变学说, *AtMYB2* 基因的这种多样性和分歧度主要是由强烈的纯化选汰(strong purifying selection)和自适应过程(adaptive process)两个方面的结果导致的。

4 植物MYB2转录因子的功能

植物MYB转录因子有广泛的生理功能,几乎参与植物发育和代谢的各个方面。单一MYB结构域蛋白,如拟南芥的LHY、CCAI、CPC和RTBP1,水稻的RTBP21,玉米的IBP21等,是一类重要的端粒结合蛋白,在维持染色体结构的完整性和调节基因转录上起重要作用(Yu等2000)。含有3个MYB结构域的蛋白,如烟草中的R1R2R3-MYB蛋白在细胞分裂的G₂/M期发挥重要的调节作用,调控 *cycle B* 基因和许多在G₂/M期表达的基因(Ito 2005)。R2R3-MYB成员是植物中数目最多的一类MYB蛋白,MYB2转录因子属于R2R3-MYB蛋白,它们参与应答激素刺激、环境胁迫,参与次生代谢的调节,还参与控制细胞的分化等。

拟南芥中的 *AtMYB2* 基因是第一个被发现的受ABA和脱水诱导表达的R2R3-MYB蛋白。*AtMYB2*蛋白与bHLH(basic-helix-loop-helix,碱性螺旋-环-螺旋)类蛋白RdBP1相互作用,协同调节 *Rd22* 基因(response to desiccation,脱水响应基因)的表达(Abe等1997)。还有研究认为, *AtMYB2* 是一个重要的参与缺氧反应调控的转录调控因子,它结合在一些特定的序列,如MYB结合位点(MBS-1、MBS-2)并诱导这些缺氧基因表达(Hoeren等1998)。

巨桉(*Eucalyptus grandis*)中的EgMYB2,也是R2R3-MYB转录因子家族亚族的一员,从桉树木质部形成阶段的基因文库中克隆出来。EgMYB2在巨桉的连锁图谱映射一个唯一的位点,与木质素含

量的数量性状位点共存。EgMYB2蛋白重组体能够与两个木质素合成基因(*CCR*和*CAD*)的启动子结合,这两个启动子包含有MYB2的结合位点,能在瞬时和稳定表达实验中调控*CCR*和*CAD*基因的转录。与野生烟草相比,转基因烟草过表达EgMYB2显示出表型变化,次生细胞壁明显变厚,木质素分布也有变化,木质素合成酶转录本的丰度不同程度上增加,并且苯丙素基因(phenylpropanoid genes)核心没有被影响。这些结果说明, EgMYB2协同调控木质素的合成和次生细胞壁的形成属于一个独立的途径(Goicoechea等2005)。

Wang等(2004)在研究棉花纤维基因控制时,发现棉花GaMYB2蛋白在表皮毛的生长中起调控作用。棉花纤维基因的启动子RDL1序列中含有可被MYB2识别的基序, GaMYB2蛋白与其结合,能够进一步驱动棉花纤维基因的表达。他们的实验还表明,棉花GaMYB2蛋白/纤维因子1,在酵母和植物中都能转录激活RDL1启动子。通过实时定量PCR和原位杂交分析表明, GaMYB2在棉花纤维早期生长过程中显著表达。转染进入拟南芥后, GL1::GaMYB2能够补偿GL1(拟南芥的一个调控表皮毛生长的MYB转录因子)基因突变体的表皮毛的形成。另外, 35S::GaMYB2也能诱导种子表皮毛的产生。进一步研究GL1基因和GaMYB2基因的第一个内含子时发现,它在表皮毛中有增强子的功能,在非表皮细胞中有阻遏子的功能。破坏GaMYB2基因和GL1基因的第二个内含子中的MYB基序,将会影响表皮毛的生成。这些结果表明,棉花和拟南芥使用相似的转录因子调控表皮毛生成,据此认为, GaMYB2可能是棉花纤维生长过程中关键的调节因子。

Leech等(1993)从苔藓丝状体组织的cDNA文库中克隆得到3个含有MYB蛋白结合区域的cDNA,这3个cDNA编码两类不同的MYB蛋白: Pp1和Pp2。这两个蛋白的C端酸性氨基酸构成 α -螺旋结构,说明他们和其他MYB基因一样,编码的蛋白有转录激活功能。分析发育中的小立碗藓(*Physcomitrella patens*)在转录水平时mRNA的含量发现, Pp1和Pp2两个基因转录的最高水平发生在野生型丝状体组织中,同时其有丝分裂指数也达到最大值。随着野生型组织的生长,两基因的表达量开始下降。两基因在突变的小立碗藓中表达水平

异常,还会导致其形态学上的变化,尤其是丝状体生长率极大地降低。最终的实验数据表明,在小立碗藓正常的配子形成的细胞生长过程中, *Pp1* 和 *Pp2* 两个基因的表达是必不可少的。

当然,MYB2转录因子在行使其生理功能时,不是单独起作用的。有实验表明,一种钙调蛋白 GmCaM4 能够与 AtMYB2 转录因子结合,促进拟南芥的抗盐能力。拟南芥中过表达的 GmCaM4 蛋白可以提高 AtMYB2 转录因子的转录效率,同时提高脯氨酸合成酶的含量,进而促进盐胁迫下植物体内脯氨酸的积累(Yoo等2005);在ABA和脱水诱导下,拟南芥的 AtMYB2 蛋白和 AtMYC 蛋白共同作用, *Rd22* 基因受调控,表达量最高,而在突变体中, AtMYB2 蛋白单独调控 *Rd22* 基因,其表达量并不十分明显(Abe等1997)。

5 结语

目前,对植物MYB2转录因子的研究多集中在其生理功能上,而关于MYB2转录因子与其上、下游基因之间的作用机制研究较少,如MYB2转录因子通过哪些作用位点与下游基因结合,调控哪些下游基因的表达;同时,MYB2基因的表达又受哪些基因或外界条件的调控,这方面的研究还有待进一步探索,以期阐明MYB2转录因子的作用机制及功能。

参考文献

- 刘强,张贵友,陈受宜(2000).植物转录因子的结构与调控作用.科学通报,45(14):1465~1474
- 吴乃虎(2001).基因工程原理(下册).第2版.北京:科学出版社,88~97
- Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997). Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. *Plant Cell*, 9: 1859~1868
- Chen Y, Yang X, He K, Liu M, Li J, Gao Z, Lin Z, Zhang Y, Wang X, Qiu X et al (2006). The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol*, 60: 107~124
- Furukawa K, Oda M, Nakamura H (1996). A small engineered protein lacks structural uniqueness by increasing the side-chain conformational entropy. *Proc Natl Acad Sci USA*, 93(24): 13583~13588
- Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N et al (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J*, 43(4): 553~567
- Hoeren FU, Dolferus R, Wu Y, Peacock WJ, Dennis ES (1998). Evidence for a role for AtMYB2 in the induction of the *Arabidopsis* alcohol dehydrogenase gene (*ADHI*) by low oxygen. *Genetics*, 149(2): 479~490
- Ito M (2005). Conservation and diversification of three-repeat Myb transcription factors in plants. *J Plant Res*, 118: 61~69
- Iwasaki T, Yamaguchi-Shinozaki K, Shinozaki K (1995). Identification of a *cis*-regulatory region of a gene in *Arabidopsis thaliana* whose induction by dehydration is mediated by abscisic acid and requires protein synthesis. *Mol Gen Genet*, 247: 391~398
- Jin H, Martin C (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol*, 41: 577~585
- Kamiya T, Kawabe A, Miyashita NT (2002). Nucleotide polymorphism at the *Atmyb2* locus of the wild plant *Arabidopsis thaliana*. *Genet Res*, 80(2): 89~98
- Klempnauer KH, Gonda TJ, Bishop JM (1982). Nucleotide sequence of the retroviral leukemia gene *v-myb* and its cellular progenitor *c-myb*: the architecture of a transduced oncogene. *Cell*, 31(2 Pt 1): 453~463
- Leech MJ, Kammerer W, Cove DJ, Martin C, Wang TL (1993). Expression of *myb*-related genes in the moss, *Physcomitrella patens*. *Plant J*, 3(1): 51~61
- Lipsick JS (1996). One billion years of Myb. *Oncogene*, 13: 223~235
- Ogata K, Morikawa S, Nakamura H, Hojo H, Yoshimura S, Zhang R, Aimoto S, Ametani Y, Hirata Z, Sarai A et al (1995). Comparison of the free and DNA-complexed forms of the DNA-binding domain of *c-Myb*. *Nat Struct Biol*, 2: 309~320
- Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H (1987). The regulatory *cl* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J*, 6(12): 3553~3558
- Urao T, Noji M, Yamaguchi-Shinozaki K, Shinozaki K (1996). A transcriptional activation domain of ATMYB2, a drought-inducible *Arabidopsis* Myb-related protein. *Plant J*, 10(6): 1145~1148
- Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY, Wang LJ, Chen XY (2004). Control of plant trichome development by a cotton fiber MYB gene. *Plant Cell*, 16(9): 2323~2334
- Yoo JH, Park CY, Kim JC, Heo WD, Cheong MS, Park HC, Kim MC, Moon BC, Choi MS, Kang YH et al (2005). Direct interaction of a divergent CaM isoform and the transcription factor, MYB2, enhances salt tolerance in *Arabidopsis*. *J Biol Chem*, 280(5): 3697~3706
- Yu EY, Kim SE, Kim JH, Ko JH, Cho MH, Chung IK (2000). Sequence specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. *J Biol Chem*, 275: 24208~24214