

苦瓜中与 *AGAMOUS* 相似基因启动子的克隆和表达载体的构建

彭书明^{1,2}, 雷泞菲^{1,2}, 陈放¹, 唐琳^{1,*}

¹四川大学生命科学学院, 成都 610064; ²成都理工大学材料与化学化工学院, 成都 610059

摘要: *McAG2* 基因是苦瓜中分离得到的与 *AGAMOUS* 相似基因, 在花器官和果实中特异表达, 参与花的第三轮和第四轮器官的形成。文章采用染色体 DNA 步移技术克隆得到长为 1 417 bp 的 *McAG2* 基因 5' 上游片段。序列分析显示此片段含有典型的 TATA-box、CAAT-box、丰富的激素应答调控元件。为了研究这些调控元件, 还构建了 *McAG2* 基因 5' 侧翼缺失和内含子缺失表达载体。

关键词: 苦瓜; *AGAMOUS* 相似基因; 启动子; 载体构建

Cloning of an *AGAMOUS*-like Gene Promoter from *Momordica charantia* L. and Construction of Expression Vector

PENG Shu-Ming^{1,2}, LEI Ning-Fei^{1,2}, CHEN Fang¹, TANG Lin^{1,*}

¹College of Life Sciences, Sichuan University, Chengdu 610064, China; ²College of Materials and Chemistry and Chemical Engineering, Chengdu University of Technology, Chengdu 610059, China

Abstract: *McAG2* gene is an *AGAMOUS*-like gene, which was cloned from *Momordica charantia*. It especially expressed in floral organ and young fruit and had a crucial role in the third and fourth whorl floral organ development. The 1 417 bp 5' upstream region of the *McAG2* gene was isolated by the DNA walking technology. Sequence analysis reveals that it contains classical TATA-box, CAAT-box, which are conserved in eukaryotic gene promoters, and other regulation elements regarding to hormones. For further study on those regulator elements, the vectors were constructed with the deletion in the 5' region and intron of *McAG2*.

Key words: *Momordica charantia*; *AGAMOUS*-like gene; promoter; vector construction

MADS-box 基因是一类转录因子, 控制植物发育过程。根据拟南芥和金鱼草研究中的结果, Coen 和 Meyerowitz (1991) 以及 Weigel 和 Meyerowitz (1994) 提出了花发育的 ABC 模型, 后来又在矮牵牛花和拟南芥中发现了 D 功能基因 (Colombo 等 1995) 和 E 功能基因 (Becker 和 Theißen 2003), 花发育的 ABC 模型进一步发展为 ABCE 模型 (Colombo 等 1995)。 *AGAMOUS* 基因就是最早从模式植物拟南芥中克隆得到的 C 功能基因, 它在花发育过程中的性别器官形成中起作用 (Bowman 等 1989), 参与花的第三轮和第四轮器官形成。近年来, 大量与 *AGAMOUS* 相似基因被克隆, 它们的功能研究已较为深入, 但其转录调控的研究则相对较少。据报道, 与 *AGAMOUS* 相似的基因在基因组 DNA 中含有不同数目的内含子, 最长达 8 个。这些内含子长短不一, 内含子 1 一般都在 5' 端非翻译区 (Yanofsky 等 1990; Bradley 等 1993; Savidge 等 1995; Rutledge 等 1998; Davies 等 1999; Brunner 等 2000); 内含子 2 最长, 达 3~5 kb。 *AGAMOUS* 基因内含子 2 含有

LEAFY 蛋白结合相关的增强序列, 其转录需要 LEAFY 蛋白结合在内含子 2 上 (Busch 等 1999), Hong 等 (2003) 的研究也证实内含子 2 具有转录调控元件。而对于 5' 非翻译区内的内含子和其他内含子, 以及基因上游序列的调控的研究很少 (de Folter 和 Angenent 2006)。因此, 本文在苦瓜中得到与 *AGAMOUS* 相似基因的全长 cDNA 序列 (命名为 *McAG2*, GenBank 登录号为 DQ299943) 的基础上 (Peng 等 2008), 采用 DNA 步移技术克隆得到 *McAG2* 基因上游调控序列, 为深入研究苦瓜 *McAG2* 基因的表达建立了基础。

材料与方法

苦瓜 (*Momordica charantia* L.) 品种 '长白' 种

收稿 2008-04-22 修定 2008-07-14

资助 教育部博士点基金 (2002061088)。

* 通讯作者 (E-mail: tanglin@scu.edu.cn; Tel: 028-85417281)。

植于四川大学生命科学学院药用植物园。大肠杆菌(*Escherichia coli*) Top10 由实验室保存。测序和亚克隆 pMD18-T 载体购自宝生物工程(大连)有限公司。植物原生质体瞬时表达载体 pBI221 由四川大学杨毅先生惠赠。

少量胶回收试剂盒购自上海华舜生物工程有限公司, H. Q. & Q. 质粒小量提取试剂盒购自安徽优晶生物工程有限公司。TaKaRa Ex TaqTM, T4 DNA 连接酶, 各种 DNA 限制性内切酶 *DraI*、*EcoRV*、*PvuI*、*StuI* 等均购自 TaKaRa 公司; Tryptone 和 Yeast Extract 为 OXOID 产品; 琼脂糖、琼脂粉、Tris、SDS 为 Amersham 产品; 氨卞青霉素购自 TaKaRa 公司; IPTG、X-gal 为 Fluka 公司产品; 其他化学药品为进口或国产分析纯试剂。

苦瓜总 DNA 的大量提取采用 CTAB 方法(略有修改)(Doyle 和 Doyle 1987)。DNA 经 RNase (DNase free) 除去其中的 RNA 后储存于 -20 °C 中备用。

基因组 DNA 步移文库的构建参照 BD GenomeWalkerTM Universal Kit 中的方法进行。基因组 DNA 分别采用 *DraI*、*EcoRV*、*PvuI* 和 *StuI* 平末端酶酶切, 酶切产物与接头连接, 构建了 4 个基因组步移文库。

苦瓜 *McAG2* 基因启动子克隆分 2 轮。第一轮 PCR 反应: 在 4 个 0.5 mL 的 EP 管(C1~C4)中分别以构建好的 DNA 文库(DL1、DL2、DL3 和 DL4)为模板, 以基因特异引物 SP1 (5' GGCTGGAGAAAAC-GATGAGAGCAACCT 3')和接头引物 AP1 (5' GTAATACGACTCACTATAGGGC 3')在 PTC-100TM (MJ) 热循环仪上进行 PCR 反应。反应程序为: 95 °C 预变性 5 min; 随后紧接着 30 个循环为 94 °C 30 s, 62 °C 30 s, 72 °C 3 min; 然后, 72 °C 保温 7 min。反应完成后, 将上述 4 管 PCR 产物稀释 10 倍(D1~D4)用于第二轮 PCR 扩增。第二轮 PCR 反应: 以基因特异引物 SP2 (5' TTGAGCAGCCCATTGCGTCTCT-TACAG 3')和接头引物 AP2 (5' ACTATAGGGCA-CGCGTGGT 3')进行 PCR 扩增, 反应程序同上, 仅改变循环数为 35 次。所得 PCR 产物经 1% 琼脂糖凝胶电泳检测, 产物胶回收、连接和转化, 验证后送北京三博生物技术公司测序。

为进一步研究 *McAG2* 基因启动子可能的转录调控元件和 *McAG2* 基因的转录起始位点至翻译起始位点之间的内含子的调控作用, 分别对 *McAG2* 启动子序列进行 5' 侧翼缺失和 5' 上游 ATG 转录翻译起始位点和转录起始位点之间的缺失。构建缺失表达载体所需引物序列见表 1, 相应的引物位置

表 1 启动子缺失表达载体引物

Table 1 Primers of expression vectors with deleted promoter

上游引物(下划线示 <i>PstI</i> 酶切位点)	下游引物(下划线示 <i>BamHI</i> 酶切位点)
PFa: 5' TTCTGCAGTTTAAATGTCGAGTGAC 3'	PRa: 5' GCAGGATCCTTTCTTCTTGATTCTGGA 3'
PFb: 5' G TA CTGCAGGGCTGAGGATGCTTCTAT 3'	PRb: 5' GCAGGATCCGATATGTAGAGAAAGATGG 3'
PFc: 5' GGGCTGCAGGTATTTGAACATATACAT 3'	PRc: 5' GAGGATCCGGAAGGTAGCTAAAAGGTG 3'
PFd: 5' GGCCTGCAGATGGATTCTTACATCTTTA 3'	
PF e : 5' GATCTGCAGTCCATTTCCCTCTTCTCCT 3'	

见图 2。

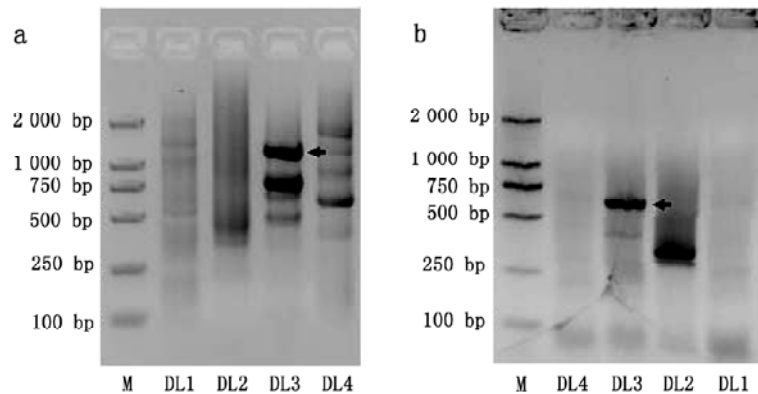
制备缺失片段时, 分别以下列引物对扩增缺失片段: P1 (PFa 和 PRa)、P2 (PFa 和 PRc)、P3 (PFb 和 PRc)、P4 (PFb 和 PRb)、P5 (PFb 和 PRa)、P6 (PFc 和 PRa)、P7 (PFd 和 PRa)、P8 (PF~~e~~ 和 PRa)。然后将上面获得缺失片段定向插入 pBI221 载体上, 取代原有的 CaMV 35S 启动子, 然后转化大肠杆菌, 经酶切和测序确定载体构建的

正确性。

结果与讨论

1 苦瓜 *McAG2* 基因启动子片段的克隆

用染色体 DNA 步移技术, 从步移文库 DL3 中克隆得到长 1300 bp 左右的片段(图 1-a), 该片段长 1305 bp, 与 *McAG2* 基因的 5' 端部分片段重合。根据测序结果, 在新得到片段 5' 端设计了 2 条特异

图1 *McAG2* 基因启动子的克隆Fig.1 Cloning of *McAG2* gene promoter

a: 第一轮2次扩增产物; b: 第二轮2次扩增产物。M: 分子量标记 DL2000; DL1: *PvuI* 酶切文库; DL2: *EcoRV* 酶切文库; DL3: *DraI* 酶切文库; DL4: *StuI* 酶切文库。

引物SP3 (5'TGTGAGTGGTGACTGTTAAGGTTTG 3')和SP4 (5'GACAGGGATGTTGGAATAAGTTGCA 3'), 从步移文库 DL3 又扩增得到约 600 bp 的片段, 除去重叠部分得到了约 290 bp (图 1-b) 的片段。经过 2 轮 DNA 步移技术, 总共克隆得到长约 1 417 bp 片段, 是苦瓜 *McAG2* 基因起始密码子 ATG 上游区域, 提交 GenBank 的登录号为 EU586159。

2 苦瓜 *McAG2* 基因 5' 上游调控序列的功能元件分析

将克隆得到的序列与 *McAG2* 基因的 cDNA 比对的结果显示, *McAG2* 基因上游序列 +147~+447 之间存有一长为 300 bp 的内含子, 其拼接点是 5' 端的 GT (+146~+147), 3' 端的 AG (+448~+449), 为典型的内含子拼接点。采用 PLACE 网站 (<http://www.dna.affrc.go.jp/PLACE>) 和 Neural Network Promoter Prediction 软件对克隆的苦瓜启动子进行了分析, 结果显示在 +1 处 C 为可能的转录起始密码子 (图 2)。预测转录起始位点上游 -30 bp 和 -488 bp 处有 2 个 TATA-box, 另外还含有 7 个 CAAT 盒 (-139、-395、-401、-576、-665、-677、-923)。除此之外, 还有一些其他调控元件: 如茁壮素应答元件 (306~314, KGTCCCAT)、促细胞分裂素蛋白结合元件 (-531~-525、-17~-11、85~91, TATTAG)、赤霉素应答顺式作用元件 (-348~-342, CCTTTT) 和赤霉素应答复合体中心元件 (-136~-129, TAACAAA) 等, 这些元件都和激素相关, 说明

这段序列的启动活性很可能受激素调控。在 *McAG2* 基因启动子 (*McAG2P*) 中还有大量的重复序列, 包括间隔重复序列: AAAAAT (-857~-851、-850~-844、168~174、378~384、385~391)、TGATAAAG (-505~-497、-56~-48)、TTCAAAA (-640~-633、-629~-622)、TCCCTC (8~14、20~26)、AAAATG (-849~-843、-822~-816、372~378、379~385、386~394); 连续重复序列: GTATG (-801~-791)、GGAAA (-261~-251)、GATTT (429~439)、AAAATGA (372~393); 简单连续重复: TG (49~55)、AG (-539~-531)、TCC (110~119); 还有一个互补序列: TGTTTAAACA (-651~-641); 并存在 TC 富含区 (-349~-331、6~30) 和 AG 富含区 (如: -917~-889)。由此可见, *McAG2* 基因上游片段包含典型的 TATA 盒、CAAT 盒及其丰富的调控元件和多种重复序列。因此认为, *McAG2P* 对 *McAG2* 基因的表达有调控作用和启动子活性。

3 *McAG2* 基因启动子缺失表达载体构建和酶切鉴定

为了进一步研究这些预测的元件、重复序列等, 又做了以下实验。以 PFa 及 PRa 引物对, 以苦瓜幼叶中总 DNA 为模板扩增全长 *McAG2P* 片段。经亚克隆测序确定正确无误后, 以之为模板, 克隆其他缺失片段。再将 PCR 扩增得到的 8 条长度不等的缺失片段 (P1, 1 417 bp; P2, 937 bp; P3, 792 bp; P4, 1 120 bp; P5, 1 270 bp; P6, 1 100 bp; P7, 716 bp;

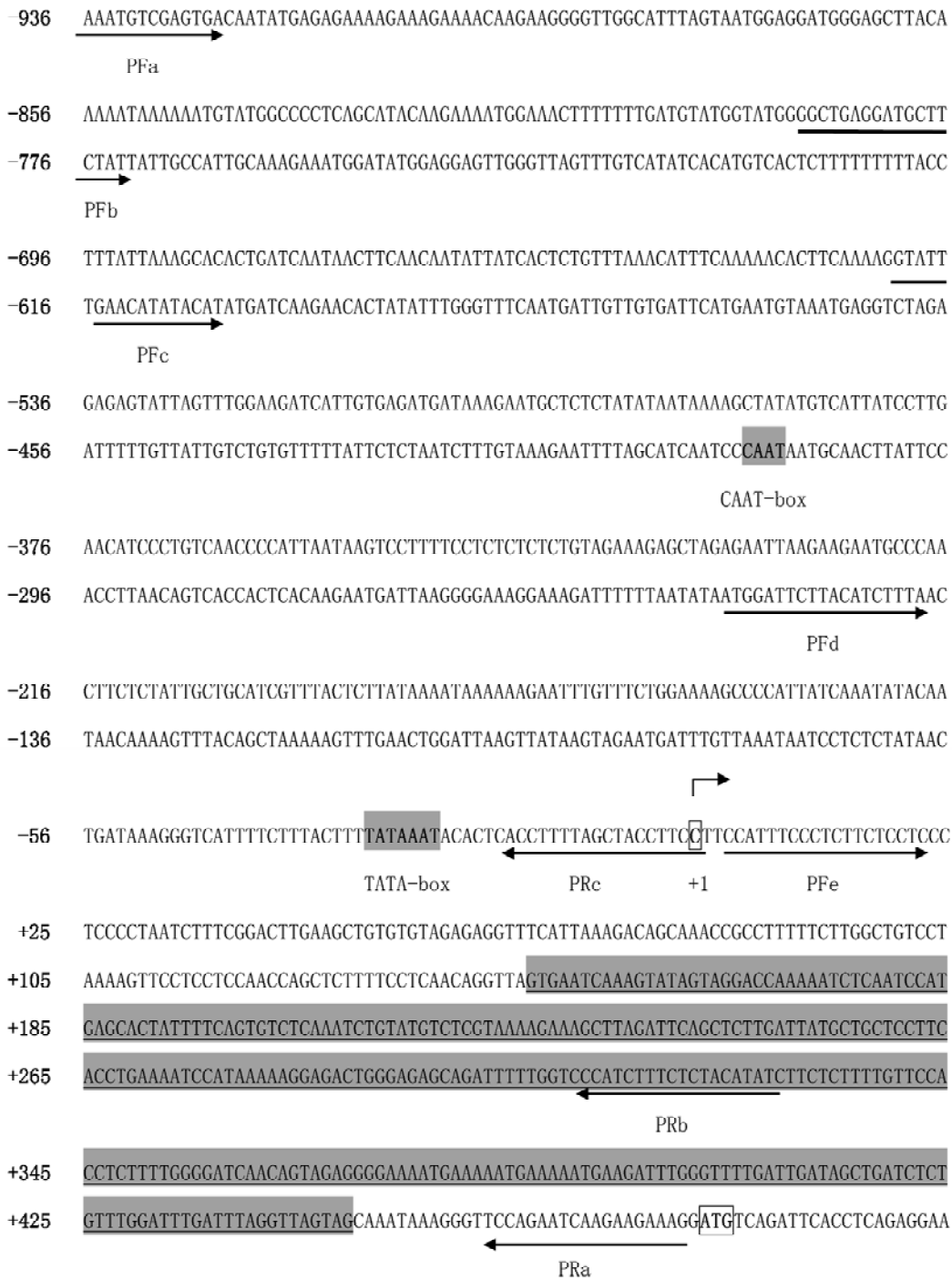


图2 *McAG2* 基因启动子序列

Fig.2 The promoter sequence of *McAG2* gene

加框粗黑体为转录起始和翻译起始位点; 阴影为 TATA-box 和 CAAT-box; 阴影加下划线为 5' UTR 中的内含子; 箭头为扩增引物。

P8, 476 bp)用 *Pst*I 和 *Bam*HI 双酶切定向插入质粒 pBI221 中, 构建了8个由 *McAG2*P1~P8 驱动报告基因 *GUS* 的植物表达载体(图3)。将构建好的瞬时表

达载体转化大肠杆菌 Top10 细胞后, 菌落PCR和质粒提取酶切验证(图4)得到构建好的 *McAG2*PBI1~8 瞬时表达载体。

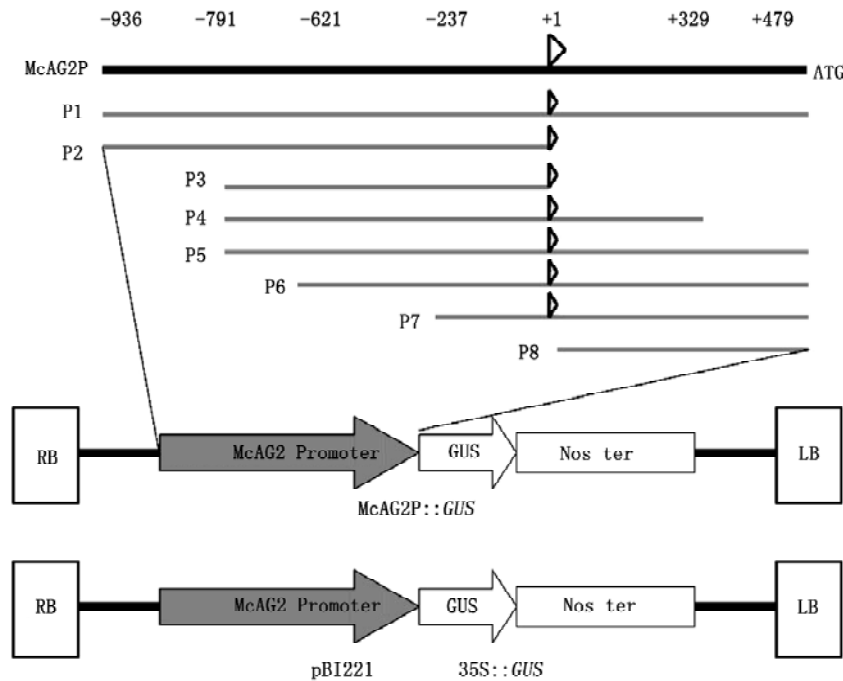


图3 McAG2Pn::GUS表达载体

Fig.3 Map of McAG2Pn::GUS expression vector

P1~P8: McAG2P 缺失片段; 三角形: 预测的转录起始位点; ATG: 翻译起始位点。构建所用载体为 pBI221, 由 35S 启动子启动 GUS 报告基因表达。

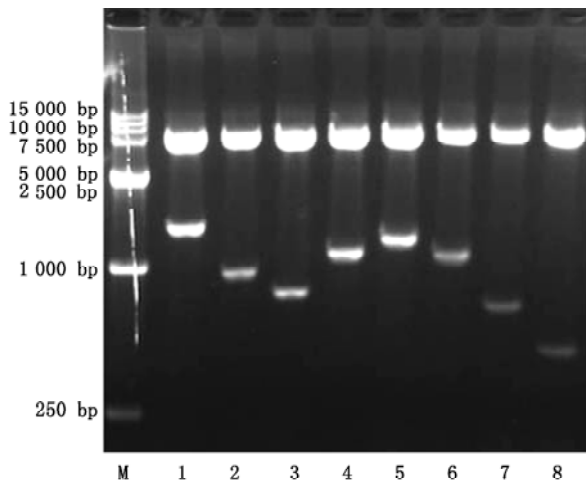


图4 McAG2 基因启动子缺失表达载体酶切鉴定

Fig.4 Identification of McAG2 gene with promoter deletion vectors using restriction digestion method

M: DNA 分子量标记 DL15000; 1~8: McAG2P1~P8 缺失表达载体用 PstI 和 BamHI 双酶切鉴定, 泳道中大片段为载体 DNA, 小片段为插入的缺失启动子 DNA。

参考文献

Becker A, Theißen G (2003). The major clades of MADS-box

genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol*, 29: 464~489
 Bowman JL, Smyth DR, Meyerowitz EM (1989). Genes directing flower development in *Arabidopsis*. *Plant Cell*, 1: 37~52
 Bradley D, Carpenter R, Sommer H, Hartley N, Coen E (1993). Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the *plena* locus of *antirrhinum*. *Cell*, 72: 85~95
 Brunner AM, Rottmann WH, Sheppard LA, Krutovskii K, DiFazio SP, Leonardi S, Strauss SH (2000). Structure and expression of duplicate AGAMOUS orthologues in poplar. *Plant Mol Biol*, 44: 619~634
 Busch MA, Bomblies K, Weigel D (1999). Activation of a floral homeotic gene in *Arabidopsis*. *Science*, 285: 585~587
 Coen ES, Meyerowitz EM (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, 353: 31~37
 Colombo L, Franken J, Koetje E, van Went J, Dons HJM, Angenent GC, van Tunen AJ (1995). The petunia MADS box gene *FBP11* determines ovule identity. *Plant Cell*, 7: 1859~1868
 Davies B, Motte P, Keck E, Saedler H, Sommer H, Schwarz-Sommer Z (1999). *PLENA* and *FARINELLI*: redundancy and regulatory interactions between two *Antirrhinum* MADS-box factors controlling flower development. *EMBO J*, 18: 4023~4034
 de Folter S, Angenent GC (2006). *Trans* meets *cis* in MADS science.

- Trends Plant Sci, 11 (5): 224~231
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*, 19: 11~15
- Hong RL, Hamaguchi L, Busch MA, Weigel D (2003). Regulatory elements of the floral homeotic gene *AGAMOUS* identified by phylogenetic footprinting and shadowing. *Plant Cell*, 15: 1296~1309
- Peng SM, Luo T, Zhou JY, Niu B, Lei NF, Tang L, Chen F (2008). Cloning and quantification of expression levels of two MADS-box genes from *Momordica charantia*. *Biologia Plant*, 52 (2): 222~230
- Rutledge R, Regan S, Nicolas O, Fobert P, Cote C, Bosnich W, Kauffeldt C, Sunohara G, Seguin A, Stewart D (1998). Characterization of an *AGAMOUS* homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis*. *Plant J*, 15 (5): 625~634
- Savidge B, Rounsley SD, Yanofsky MF (1995). Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. *Plant Cell*, 7: 721~733
- Weigel D, Meyerowitz EM (1994). The ABCs of floral homeotic genes. *Cell*, 78: 203~209
- Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldman KA, Meyerowitz EM (1990). The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature*, 346: 35~39