

水母雪莲 Myb 转录调控因子 SmP 基因的克隆及序列分析

金治平¹ 赵德修^{1*} 乔传令² 瞿文全¹ 陈亚琼¹ 付春祥¹

¹(中国科学院植物研究所 北京 100093)

²(中国科学院动物研究所 北京 100080)

摘 要 采用 TD-PCR(Touch down PCR)法从水母雪莲(*Saussurea medusa* Maxim)红色系愈伤组织 cDNA 文库中筛选并克隆了雪莲 Myb 转录调控因子 SmP(*S. medusa* Maxim Myb-related P gene)基因。序列分析表明该基因全长 969bp, 包括一个 771bp 的完整开放阅读框架(ORF), 编码一个 256 氨基酸残基的蛋白质。氨基酸序列的同源性分析表明在 N-端具有两个典型的 R2R3-Myb DNA 结合结构域。C-端富含亲水的丝氨酸(18.38%), 且以寡聚体的形式存在, 具有转录调控因子激活结构域常见的特征。

关键词 水母雪莲, cDNA 文库, Myb 转录调控因子, SmP 基因

中图分类号 Q78 文献标识码 A 文章编号 1000-3061(2003)03-0368-04

Myb 转录调控因子广泛存在于高等植物中^[1], 其主要特征是含有 1~3 个由 51 或 52 个氨基酸残基组成的不完全重复的 Myb-DNA 结合结构域(R1、R2、R3)。每个结构域由 3 个 α -螺旋组成, 其中第二与第三个螺旋形成螺旋-转角-螺旋(bHTH)结构, 该结构与 λ 抑制因子 DNA 识别结合结构域相似^[2,3]。根据所含 Myb-DNA 结合结构域的数目可分为 R1、R2R3、R1R2R3 三个家族, 其中 R2R3 家族对植物次生代谢具有十分重要的调控作用。玉米 P 基因是植物中最早被克隆并且功能研究得最清楚的 R2R3-Myb 转录调控因子之一^[4,5]。玉米 P 基因通过激活 CHS(Chalcone Synthase) DFR(Dihydroflavonol 4-reductase) 两个类黄酮生物合成关键酶基因的表达, 有效地调控玉米 3-脱氧类黄酮的生物合成。将玉米 P 基因导入类黄酮合成能力缺失的墨西哥甜玉米悬浮细胞突变系中异位表达, 使本不具有类黄酮合成能力的突变系获得了类黄酮合成能力^[6]。3-脱氧类黄酮高车前素(Hispidulin)和金合欢素(Jaceosidin)是我国珍稀名贵传统中药材水母雪莲(*Saussurea medusa* Maxim)的主要有效成分, 具有抗腹水型肝癌等多种功能^[7]。本文在建立了水母雪莲红色系愈伤组织 cDNA 文库及获得了雪莲 R2R3-Myb 转录调控因子 SmP 基因特异探针的基础上, 采用 PCR 法筛选文库, 以期获得与玉米 P 基因功能类似的 Myb 转录调控因子 SmP 基因, 为雪莲类黄酮生物合成进行分子调控, 提高高车前素和金合欢素产量以达到雪莲细胞工业化大规模培养奠定基础。

1 材料与方法

1.1 材料与试剂

雪莲红色系愈伤组织 cDNA 文库由本实验室构建, 由 20 个独立的亚文库(A~T)组成。大肠杆菌 XL1-Blue 与 BM25.8 均由 CLONTECH 公司提供。PCR 扩增及其他常规试剂均为国产试剂。

1.2 引物设计

SmP 基因特异引物根据本实验室获得的该基因特异探针序列设计:

pSmP5' 5'-GTT GCA GGT TAC GAT GGG TG-3'

pSmP3' 5'-CAT TGT CAC TTC TTC CAG GCA-3'

检测阳性克隆插入片段大小的 PCR 引物根据文库载体 λ TripIEx2MCS 位点已知序列设计:

P₁ 5'-AAG CGC GCC ATT GTG TTG-3'

P₂ 5'-AAG TGA GCT CGA ATT GCG G-3'

1.3 方法

1.3.1 噬菌体培养、培养物的回收、文库滴度测定及将噬菌体转化为质粒等基本操作均参考文献[8]进行。

1.3.2 第一轮筛选, 确定初级阳性池及目的基因在初级阳性池中的丰度: 每个亚文库取 1 μ L 为模板, 以 pSmP5'、pSmP3' 为引物, 以 25 μ L 体系, 先 95 $^{\circ}$ C 加热 10min 裂解噬菌体, 加入 PCR 扩增混合物按以下程序进行 PCR: 94 $^{\circ}$ C 预变性 4min; 94 $^{\circ}$ C 变性 1min, 65~50 $^{\circ}$ C 退火 1min, 每进行一个循环退火温

度下降 0.5℃,72℃延伸 1min,30 循环,94℃变性 1min,50℃退火 1min,72℃延伸 1min,15 循环,72℃延伸 5min。1.6% 琼脂糖电泳检测,确定阳性池。选取信号最强的阳性池,作 10 倍梯度稀释到 10⁶ 倍,每个梯度稀释物取 1μL 为模板,参照上述程序进行 PCR 检测,确定在下一级筛选时本级阳性池的最大稀释倍数 N (能够检测到阳性信号的极限稀释倍数)并估计 SmP 基因在本级阳性池中的丰度: A = N × 10³ / T × 100% (A:目的基因丰度,N:最大稀释倍数,T:阳性池滴度 pfu/mL)。

1.3.3 第二轮筛选 产生次级阳性池并确定次级阳性池中目的基因的丰度:将选中的初级阳性池按其最大稀释倍数 N 稀释,取 1μL 稀释物铺制 10 个 9cm LB 平板,培养并回收洗脱物。参照 1.3.2,检测 10 个平板回收物,确定次级阳性池及其丰度 A₂、最大稀释倍数 N₂。如果丰度低于 1% 则继续进行第三轮或更多轮的筛选,直到目的基因丰度大于 1%。

1.3.4 单噬菌斑检测 当目的基因的丰度达到 1% ~ 2% 或更高时,以每个平板约 100 个单噬菌斑的密度铺制 3 ~ 5 个平板,从平板上挑取生长并分离良好的单噬菌斑,置于 100μL SM 缓冲液中,涡旋震荡,稍离心,于 4℃ 冰箱过夜。取 3μL 洗脱物为模板,参照 1.3.2 程序扩增检测,获得阳性克隆。

1.3.5 阳性克隆插入片段检测 取 3μL 阳性克隆洗脱物为模板,以 P_{Al}、P₂ 为引物,做 PCR 检测插入片段大小。扩增条件:95℃加热 10min,裂解噬菌体,加入 PCR 反应混合物,并按以下程序扩增:94℃预变性 4min,94℃变性 1min,56℃退火 1min,72℃延伸 4min,25 循环,72℃延伸 10min。

1.3.6 噬菌体转化为质粒 参照文献 [8] 进行。

1.3.7 数据分析 在 GenBank 中分析数据。

2 结 果

2.1 三级筛选阳性池的获得及各级阳性池目的基因丰度的确定

初级筛选从 20 个亚文库中获得 D 亚文库一个强阳性池,H、J、M 亚文库信号极弱,其余的没有信号。D 亚文库稀释到 10⁴ 倍即检测不到阳性信号,最大稀释倍数 N₁ = 10³。由于 D 亚文库滴度约为 10¹¹,因此 SmP 基因的丰度 A₁ = 10⁵,属低丰度表达基因。在来自 D 亚文库的 10 个次级洗脱物中,D5 有强阳性信号,当将 D5 稀释到 10⁴ 倍时就检测不到明显的阳性信号,次级阳性池 D5 滴度约为 10¹⁰,因此,阳性池 D5 中 SmP 基因 A₂ = 10⁻⁴,N₂ = 10³。第三轮筛选获得强阳性池 D5d,滴度接近 10⁹,稀释 10⁴ 倍还可检测到阳性信号,D5d 中 SmP 基因的丰度 A₃ = 10⁻²,N₃ = 10⁴。

2.2 单噬菌斑筛选

当经过三轮筛选后,阳性池滴度越来越低,但目的基因的丰度却越来越高,达到 10⁻² 左右。阳性池 D5d 1μL 10⁴ 倍稀释物感染铺制的平板上,只有 100 个左右单噬菌斑,从中挑取 88 个单噬菌斑,检测到了 2 个阳性克隆 P16 与 P57 (图 1)。

2.3 阳性克隆的插入片段大小检测

P16 与 P57 两个克隆的插入片段均在 1kb 左右(图 2),由于用的是特异引物,因此它们极可能是同一个克隆。

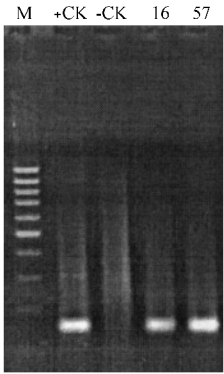


图 1 单噬菌斑检测

Fig.1 Detection of the single plaques

M. 100bp DNA ladder ;

+ CK. Positive control ; - CK. Negative control ;

16. Single plaque16 ; 57. Single plaque57

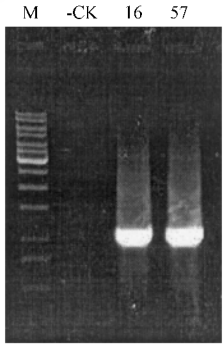


图 2 阳性克隆 P16、P57 插入片段检测

Fig.2 Detection of the inserts of the positive clones

M. 1kb DNA ladder ;

- CK. Negative control ;

16. Single plaque16 ; 57. Single plaque57

2.4 序列分析

P16 为一全长 969bp 的 cDNA 序列,有一个长 771bp 的完整 ORF (Open reading frame),起始于第 53 个碱基,终止于第 823 个碱基,编码一个 256 氨基酸的蛋白质,分子量 28.623kD,等电点 pI = 8.91,命名为 S. medusa Maxim Myb-related P gene,简称 SmP 基因。该序列已经在 GenBank 登录,登录号为 AF542042。

SmP 蛋白质 N-端具有两个典型的 Myb-DNA 结合结构域 (R1 12 ~ 63aa) 与 (R2 64 ~ 115aa) (图 3B)。R2 结构域自第 17aa 开始,每间隔 19 个 aa 有一个保守的疏水氨基酸残基色氨酸 (W, 17aa, 37aa, 57aa),共计 3 个 W; R3 结构域中第一个为苯丙氨酸 (F, 70aa) 所取代,自 70aa 起每间隔 18aa 有一个 W。

残基(89aa,108aa)(图3B)。蛋白质C-端(121aa~256aa)富含亲水氨基酸,尤其是丝氨酸(S,18.38%),并且以寡聚体的形式存在,这是转录调控因子激活结构域常见的特征^[9]。

在GenBank中做同源性分析表明,SmP基因氨基酸序列与其他Myb转录调控因子的同源区域都集中在N-端R2R3-DNA结合结构域,在C-端同源性极低。但有趣的是,与SmP基因该结构域同源性最高的是水稻OsMYB(73%)烟草LBM1(70%)马铃薯ThM18(70%)拟南芥AtMYB15(69%)矮牵牛Ph2MYB(69%)日本狐葡萄VIMYBB1-2(68%),而不是玉米的P基因ZmP(59%)(图3A)。尽管获得SmP基因特异探针的部分兼并引物是根据ZmP基因与其他R2R3-MYB相关基因的R2、R3保守区域设计的^[10]。LBM1和VIMYBB1-2的功能已得到确认,LBM1在植物受到损伤等胁迫时,可以激活苯丙氨酸裂解酶的表达而启动类黄酮生物合成途径,从而提高植物的防御能力^[11],而VIMYBB1-2则通过调节UFGT(UDP-glucose:flavonoid 3-O-glucosyltransferase)的表达而调控花色苷的生物合成^[12]。SmP与LBM1和VIMYBB1-2在C-端自120aa以后,还有一个共同点:虽然它们的同源性并不高,但都富含丝氨酸残基,分别为18.38%、14.29%、18.80%,并且都以寡聚体的形式存在,这似乎暗示SmP基因可能与它们有相似的调节类黄酮生物合成的功能。

3 讨 论

采用PCR进行文库筛选是一种简便、经济、高效的文库筛选方法^[13-15]。但由于文库提供的模板(噬菌体、菌液等)一般都未经过DNA的提取与纯化,其中往往含有文库缓冲液及细菌培养基中的某些成分,如额外的Mg²⁺等,使得扩增条件优化比较困难。本文对前人的方法略做改进,采用具有很宽适用范围的TD-PCR体系^[16],退火温度设定在相差15℃的范围内(跨引物估计的T_m值)从高到低递减,有效地提高了PCR的特异性与效率,避免了繁琐的条件优化工作,成功地从雪莲cDNA文库中获得了低丰度表达的Myb转录调控因子SmP基因。本文对选中的阳性池都做了目的基因的丰度检测,为下一级筛选时对阳性池的稀释倍数提供了依据,有效地解决了因稀释倍数不足或过高而使目的基因不能有效富集的问题。但PCR的高灵敏性也使得防止污染十分重要。

植物次生代谢产物是丰富的天然药库,但植物次生代谢途径十分复杂,大量的酶参与,这使得采用单个酶基因对其次生代谢进行分子调控变得很困难。类黄酮生物合成的大多数关键步骤都受R2R3-Myb调控因子的调节^[17],并且能在转录水平上对多个靶基因同时进行调控,如玉米P基因^[6],因此Smp基因的克隆对雪莲类黄酮生物合成的分子调控具有重要意义。但SmP基因是否对雪莲类黄酮次生代谢途径有调控作用、其靶基因是什么基因、作用机制如何,还有待进一步的研究证明。

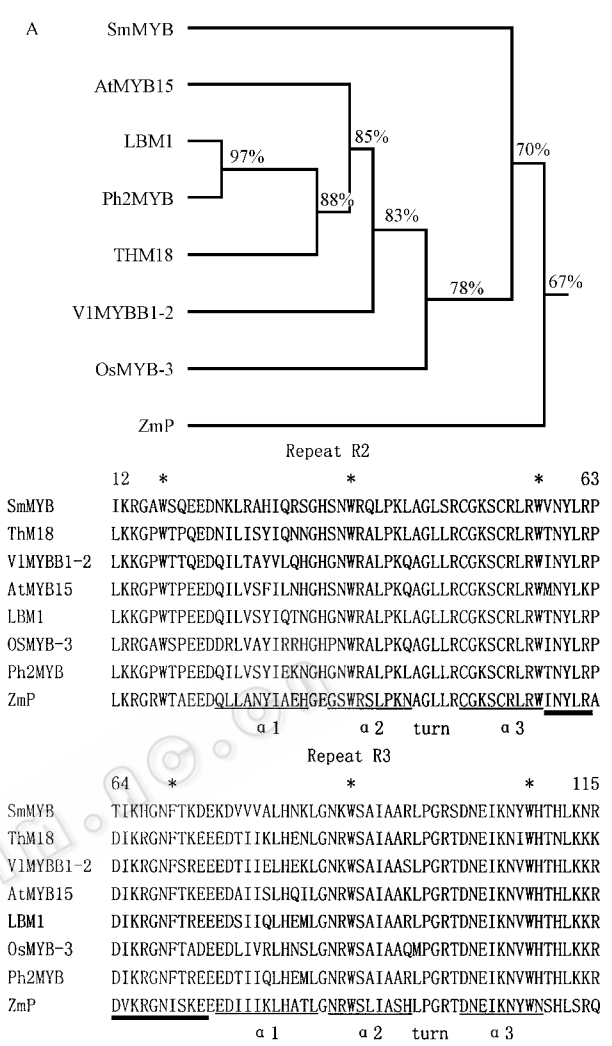


图3 SmP为具有R2R3 MYB DNA结合结构域蛋白

Fig.3 SmP shows features of an R2R3 MYB DNA binding domain protein

(A) Dendrogram of relationships among the R2R3 domains from MYB-related proteins. For construction of the tree, we used only the R2R3 MYB domain sequence (104 amino acid residues; see [B]) of each selected MYB-related protein. The matrix similarities was calculated with DNAMAN program. The number above the branches indicates the percentage of bootstrap after 1000 replicates. Sequences used are maize ZmP (CAA77939), tomato ThM18 (T07395), Kyoho grape VIMYBB1-2 (BAC07544), Arabidopsis AtMYB15 (NP-188966), tobacco LBM1 (BAA88221), rice OsMYB-3 (T03823), petunia Ph2MYB (S26604), and SmP (SmMYB: AAN17830).

(B) Sequence comparison of the MYB DNA binding domain of SmP with other MYB-related proteins from maize ZmP, tomato ThM18, Kyoho grape VIMYBB1-2, Arabidopsis AtMYB15, tobacco LBM1, rice OsMYB-3, petunia Ph2MYB, the accession numbers are the same as (A). Asterisks denote the conserved W residues. The three putative α -helices are underlined, and the closed box corresponds to the linker sequence between the R2 and R3 repeats. Amino acid residues in SmP are numbered from the translation start codon.

REFERENCES (参考文献)

- [1] Harad Kranz, Kai Scholz, Bernd Wissahaar. cMYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *The Plant Journal*, 2000, **21**(2), 231 – 235
- [2] Garbrielsen O S, A Sentence, P Fromageot. Specific DNA binding by c-Myb : evidence for a double helix-turn-helix-related motif. *Science*, 1991, **253** : 1140 – 1143.
- [3] Ogata K, MORIKAWA S, NAKAMURA H *et al.* Solution structure of a special DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell*, 1994, **79** : 639 – 648
- [4] Erich Grotewold, Bruce J Drummond, Ben Bowen, Thomas Peterson. The Myb-Homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell*, 1994, **76** : 543 – 553
- [5] Grotewold E, Athma P, Peterson T. Alternatively spliced products of the maize P gene encode proteins with homology to the DNA-binding domain of Myb-like transcription factors. *Proc Natl Acad Sci U S A*, 1991, **88**(11) : 4587 – 4591
- [6] Erich Grotewold, Mark Chamberlin, Maurice Snook *et al.* Engineering secondary metabolism in Maize cell by ectopic expression of transcription factors. *The Plant Cell*, 1998, **10** : 721 – 740
- [7] HAN S I (韩书亮). Study on the effect of four chemical constituents against cancer of *Saussurea Involucrata*. *Carinogenesis Teratogenesis and Mutagenesis* (癌变·畸变·突变), 1995, **16**(2) : 80 – 83
- [8] CLONTECH SMART™ cDNA Library Construction Kit User Manual, PT300-I(PR15738), 2001
- [9] Gregor Schmitz, Edith Tillmann, Filomena Carriero *et al.* The tomato Blind gene encodes a MYB transcription factor that controls the formation of lateral meristems. *PNAS*, 2002, **99**(2) : 1064 – 1069
- [10] Pablo D Rabinowicz, Edward L Braun, Andrea D Wolfe *et al.* Maize R2R3 Myb genes : Sequence analysis reveals amplification in the higher plants. *Genetics*, 1999, **153** : 427 – 444
- [11] Sugimoto K, Takeda S, Hirochika H. MYB-related transcription factor NtMYB2 induced by wounding and elicitors is a regulator of the tobacco retrotransposon Tto1 and defense-related genes. *Plant Cell*, 2000, **12**(12) : 2511 – 2528
- [12] Kobayashi S, Ishimaru M, Hiraoka K *et al.* Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. *Planta*, 2002, **215**(6) : 924 – 933
- [13] Isola N R, Ham H J, Cooper D L. Screening recombinant DNA libraries : a rapid and efficient method for isolating cDNA clones utilizing the PCR. *BioTechniques*, 1991, **11** : 580 – 582
- [14] Amaravadi L, King M W. A rapid and efficient, nonradioactive method for screening recombinant DNA libraries. *BioTechniques*, 1994, **16** : 98 – 103
- [15] Takumi T, Lodish H F. Rapid cDNA cloning by PCR screening. *BioTechniques*, 1994, **17** : 443 – 444
- [16] Hecker K H, Roux K H. High and low annealing temperatures increase both specificity and yield in TD and SD PCR. *BioTechniques*, 1996, **20** : 478 – 485
- [17] Memelink J, Menke F L H, Van Der Fits L *et al.* Transcriptional regulators to modify secondary metabolism Metabolic. *Engineering of Plant Secondary Metabolism*. Edited by R Verpoorte and A W Alfermann. Kluwer Academic Publishers, 2000

Cloning and Sequence Analysis of Myb Transcriptional Regulator SmP Gene of *Saussurea medusa* Maxim

JIN Zhi-Ping¹ ZHAO De-Xiu^{1*} QIAO Chuan-Ling² QU Wen-Quan¹ CHEN Ya-Qiong¹ FU Chun-Xiang¹

¹(Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China)

²(Institute of Zoology, Chinese Academy of Sciences, Beijing 100093, China)

Abstract A full-length cDNA encoding a MYB-related regulatory gene was isolated from a cDNA library prepared from mRNAs of the red line callus of *S. medusa* by TD-PCR. The cDNA, designated SmP, is 969 nucleotides long and has an open reading frame of 771 bp with a deduced amino acid sequence of 256 residues. The putative protein of SmP has two typical conserved R2R3-Myb DNA-binding domains in N-terminal and displays a rather high degree of similarity to OsMYB from rice and LBM1 from tobacco, showing 73% and 70% identity within the DNA-binding domains. However, the C-terminal domain of the SmP protein does not show obvious similarity to any other known protein sequence. It is rich in hydrophilic amino acids, especially in serine residues (18.38%), partly organized in homopolymeric stretches, a feature often found in activation domain of transcription factors.

Key words *S. medusa* Maxim, cDNA library, Myb transcription factor, SmP gene

Received : 12-16-2002

This work was supported by Grant from National Natural Science Foundation of China (No. 39970896).

* Corresponding author. Tel 86-10-62591431 ext 6201 ; Fax 86-10-82594821 ; E-mail : zhaodx@ns.ibcas.ac.cn