

大豆 *GmMP* 基因的功能预测及表达分析

李冬梅,张可欣,李文滨

(东北农业大学 大豆生物学教育部重点实验室/农业部东北大豆生物学与遗传育种重点实验室,黑龙江 哈尔滨 150030)

摘要:通过在大豆基因组数据库中检索拟南芥 *AtMP* (ARF5) 在大豆中的同源基因,获得了 *GmMP* 基因序列。对 *GmMP* 基因编码的氨基酸序列及启动子序列进行生物信息学分析,结果表明:*GmMP* 基因 CDS 序列全长 2 802 bp, 编码 933 个氨基酸。*GmMP* 编码的蛋白为疏水性蛋白。结构域分析表明:*GmMP* 含有 B3 和 AUXIN RESPONSE FACTOR 结构域,同时该基因是 ARF 家族的成员。*GmMP* 预测的启动子区域含有与激素、胁迫、光应答、生物钟调控和转录因子结合相关的顺式作用元件。系统进化分析表明 MP 在豆科植物进化过程中比较保守。组织特异性表达分析结果显示 *GmMP* 在叶片中表达量最低,在茎尖中表达量最高,推测其可能参与生长素的代谢途径。

关键词:大豆; *GmMP* 基因; 生长素; 生物信息学

中图分类号:S565.1 **文献标识码:**A **DOI:**10.11861/j.issn.1000-9841.2017.01.0028

Functional Prediction and Expression Analysis of *GmMP* in Soybean

LI Dong-mei, ZHANG Ke-xin, LI Win-bin

(Key Laboratory of Soybean Biology in Chinese Ministry of Education / Key Laboratory of Northeast Soybean Biology and Genetic Breeding Agriculture Ministry / Northeast Agricultural University, Harbin 150030, China)

Abstract: *GmMP* gene sequence was obtained by searching the homologous gene of *AtMP* in the soybean genome database. Bioinformatics analysis of amino acid sequence encoded by *GmMP* and promoter sequences showed that the full-length CDS sequence was 2 802 bp, encoding 933 amino acids protein. Structure domains analysis indicated that *GmMP* contained B3 and AUXIN RESPONSE FACTOR domain, belong to B3 and AUXIN RESPONSE superfamily. Cis-elements associated with hormones, stresses, light responses, circadian rhythms and transcription factors binding were existed in the predicted promoter region. Phylogenetic analysis results suggested that MP was conserved in the leguminous plants. Results of tissue specific analysis revealed that expression level of *GmMP* was the lowest in leaves, the highest in SAM, it was speculated to participate in the auxin metabolic pathways.

Keywords: Soybean; *GmMP* gene; Auxin; Bioinformatics

生长素在植物生长发育的过程中起到重要的调控作用^[1]。生长素响应因子(auxin response factor, ARF)作为一类调控生长素响应基因表达的转录因子,参与信号转导及植物生长发育等多个过程^[2-3]。研究发现,ARF 在拟南芥种子发育和萌发过程中起到关键作用^[4]。拟南芥 ARF 基因家族可被分为 I、II、III、IV 和 V 组,共有 23 个 ARFs,其中大部分都含有一个保守的 N-末端的 DNA 结合结构域(DBD),一个可变的转录调控区(MR)和羧基端的二聚化结构域(CTD)^[5-6]。家族中 *AtMP* (ARF5) 作为转录激活子对生长素响应基因起转录激活作用^[7-8]。

近年来 ARF 转录因子的作用已被广泛报道,但大豆的 ARF 转录因子报道很少。本研究中,通过在大豆基因组数据库中检索得到与拟南芥 *AtMP* (ARF5) 高度同源的 *GmMP* 基因,通过对 *GmMP* 基因及其编码氨基酸的理化性质与结构域进行分析,

并根据 *GmMP* 启动子区域存在的顺式作用元件以及该基因在不同组织中的表达水平对其功能进行预测,旨在为 *GmMP* 基因功能的深入研究提供理论依据。

1 材料与方法

1.1 材料

选用大豆品种垦丰 16 作为试验材料;RNA 提取试剂 Trizol、cDNA 第一链合成试剂盒、SYBR Green 荧光定量试剂盒购自 Invitrogen 公司,引物合成由上海生工生物公司完成。

1.2 方法

1.2.1 大豆 *GmMP* 基因的生物信息学分析 采用 ProtParam 在线网站 (<http://web.expasy.org/prot-param/>) 预测 *GmMP* 氨基酸的理化性质。采用 PredictProtein (<https://www.predictprotein.org/>) 数据库对 *GmMP* 蛋白的二级结构进行预测。NCBI 在线

预测 GmMP 蛋白质的保守结构域,SWISS MODEL (<https://www.swissmodel.expasy.org/>) 在线预测 GmMP 蛋白质的三级结构。PlantCare 在线网站 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 预测 *GmMP* 启动子区域顺式作用元件。SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>)预测 GmMP 蛋白可能存在的信号肽。利用 NCBI 数据库 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) 进行 *GmMP* 同源基因的检索,ClustalX2 软件进行氨基酸序列的比对,MEGA 6.2 构建系统进化树。

1.2.2 大豆 *GmMP* 基因的组织特异性分析 对不同生长时期的大豆组织进行取材,液氮研磨后提取 RNA,反转录合成 cDNA。根据基因 *GmMP* 的编码序列设计定量 PCR 引物, *GmMPDL*-F: TGGCGT-GAGGCGTGTTAG; *GmMPDL*-R: CCCTTGGATTG-TAGAAAATAGTG;使用大豆 *GmTUB3* 基因作为内参, *GmTUB3*-F: CAGCAGCCAATAGAGAATTCGG; *GmTUB3*-R:TCACTGGGTTGTGATTACCAAAC。

以 cDNA 第一链为模板进行荧光定量 PCR 分析。反应体系包括 1 μ L cDNA,10 μ L 2 \times SYBR

Mix,1 μ L *GmMP* 基因特异性引物,8 μ L ddH₂O。PCR 反应按以下条件进行:95℃,15 min 预变性;95℃,20 s 变性,60℃,20 s 退火,72℃,30 s 延伸;38 个循环。采用 Bio-Rad CFX Manager 软件分析数据,统计 *GmMP* 基因在不同组织中的转录本丰度。

2 结果与分析

2.1 *GmMP* 启动子序列分析

在 Phytozome 数据库^[9]中检索大豆 *GmMP* 基因起始密码子上游 1 500 bp 序列,利用 PlantCare 在线网站分析序列中所包含的顺式作用元件,结果表明: *GmMP* 预测的启动子区域含有 1 个与生长素响应相关的顺式作用元件;多个参与逆境胁迫诱导和光应答相关的顺式作用元件,包括厌氧诱导作用元件、热胁迫应答元件、干旱胁迫应答元件、防御与胁迫应答元件;参与光应答的顺式作用元件;参与生物钟调控的顺式作用元件以及 MYB 转录因子结合位点(表 1)。预测结果表明, *GmMP* 可能是大豆生长发育过程中的关键因子。

表 1 *GmMP* 基因启动子顺式调控元件的位置与功能预测
Table 1 Location and putative function of cis-elements in *GmMP* promoter

基序	正(+)反(-)链	位置	序列	预测功能
Motif	Sense (+) antisense (-) strand	Location/bp	Sequence	Putative function
ARE	-	- 839	TGTTTT	厌氧诱导作用元件
CCAAT-box	+	- 1134	CAACGG	MYBHv1 结合位点
ATCT-motif	-	- 1454	AATCTAATCT	光应答作用元件
Box4	-	- 417,459,603,740	ATTAAT	光应答作用元件
CCAAT-box	-	- 1356	CAACGG	MYB 结合位点
CGTCA	+	- 1159	CGTCA	茉莉酸甲酯作用元件
G-box	-	-1481	CACGTT	光应答作用元件
HSE	- +	- 511, - 553	AAAAAATTTC	热胁迫应答元件
MBS	-	- 150, 936	TAACTG	MYB 结合位点
TGA-element	-	- 246	AACGAC	生长素响应元件
Circadian	-	- 336, - 849	CAANNNNATC	生物钟调控元件

2.2 GmMP 氨基酸的多序列比对及结构域分析

对 *GmMP* 编码氨基酸序列进行结构域分析,结果表明:大豆 *GmMP* 含有 B3 和 Auxin-response factor 结构域,是 B3 和 Auxin-response 的成员(图 1,2)。通过将 *GmMP* 与红豆、菜豆、木豆的氨基酸序列进行多序列比对,结果表明不同物种中的 MP 蛋白均含有 B3 和 Auxin-response 结构域,并且结构域中存在多个保守的氨基酸基序,这些特有的基序都是较为保守的。

2.3 GmMP 氨基酸的理化性质分析

ProtParam 网站分析结果显示 *GmMP* 编码的氨基酸理论等电点(pI)为 5.48,预测分子质量为 103.492 16 kD。其中氨基酸含量最高的为丝氨酸(ser)占 9.6%;色氨酸(Trp)和半胱氨酸(Cys)含量最少,均占 1.5%;负电荷总数为 100;正电荷总数为 77;在第 12~32、52~72、263~283、605~621 个氨基酸之间存在 4 个由内向外的跨膜螺旋,在第 11~27、49~71、263~289 个氨基酸之间存在 3 个

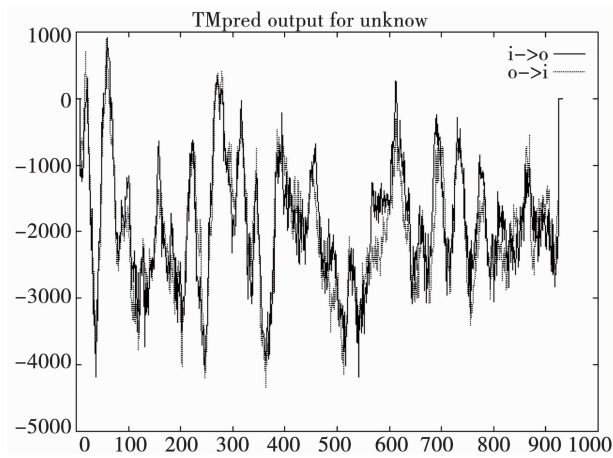


图3 GmMP 氨基酸的亲疏水性分析
Fig.3 Hydrophilic and hydrophobic analysis of GmMP protein

2.4 大豆 GmMP 蛋白的系统进化树

通过将 GmMP 蛋白序列与 NCBI 中相似度较高的其它物种 MP 蛋白进行同源性比对,并构建系统



图4 GmMP 编码蛋白的三维结构
Fig.4 3D structure of the protein encoded by GmMP

发生树(图 5)。结果表明:GmMP 与绿豆、菜豆、红豆等豆科植物 MP 蛋白位于同一分枝上,可以看出该类蛋白在豆科植物进化过程中比较保守。

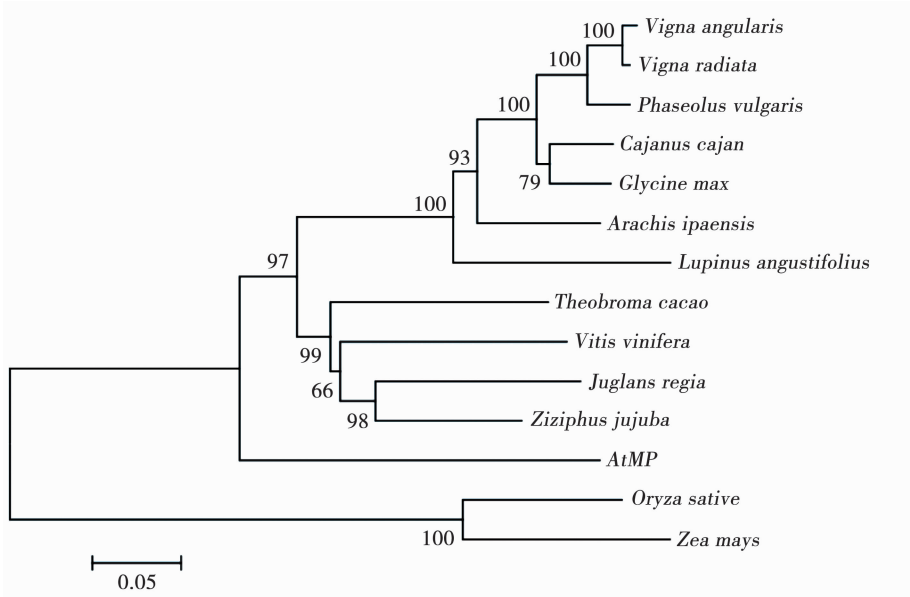


图5 GmMP 蛋白系统发生树分析
Fig.5 Phylogenetic tree analysis of GmMP protein

2.5 GmMP 基因的组织特异性分析

分别对大豆的种子、根、叶片、茎、茎尖、花和荚进行取材,利用荧光定量 PCR 的方法分析了 GmMP 基因在大豆不同组织中的表达水平,结果表明:GmMP 在茎尖中表达水平最高,其次是根、种子和荚,在花和叶片中检测到的表达量相对较低,其中叶片中表达量最低(图 6)。

3 讨 论

ARF 转录因子家族中的很多成员在植物生理过程中起重要作用,目前 ARF 转录因子家族在玉米 (Zea mays)^[10-11]、水稻 (Oryza sativa)^[12] 高粱^[13], 番

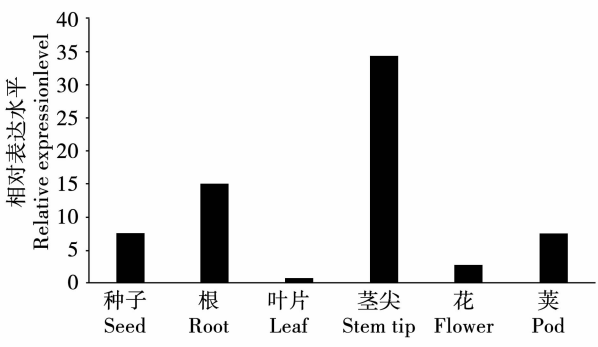


图6 GmMP 的组织特异性表达分析
Fig.6 Tissue-specific expression analysis of GmMP

茄 (*Solanum lycopersicum*)^[14]、大白菜 (*Brassica rapa*)^[15]中有过报道。本研究从大豆中首次发现与拟南芥 *AtMP* (*ARF5*) 高度同源的 *GmMP* 基因,生物信息学分析发现 *GmMP* 具有 B3 和 AFR 家族保守结构域和与其它已知的生长素家族蛋白类似的高级结构(-)^[16],因次预测 *GmMP* 基因与生长素响应因子相关。研究发现,MP(*ARF5*)在拟南芥中与生长素信号转导转录因子密切相关,与下游的靶基因有关,它不仅仅是生长素响应的调节器,在特定的发育决定中充当一个关键因子,还是细胞分裂素合成和信号的调节因子^[17]。本研究中,*GmMP* 预测的启动子区域含有 1 个与生长素响应相关的顺式作用元件,多个参与光应答相关的顺式作用元件,参与生物钟调控的顺式作用元件等,预测结果表明,*GmMP* 可能是大豆生长发育过程中的重要调节因子,这与前人报道一致。*GmMP* 在大豆 SAM 中表达水平最高,与拟南芥 *AtMP* 的表达模式类似^[18],因此推测 *GmMP* 可能具有 IAA 响应活性。此外,*GmMP* 预测启动子区域含有与脱落酸、逆境胁迫、光反应、胚乳表达和 MYB 转录因子结合相关等一系列顺式作用元件,预示 *GmMP* 可能在植物抵御逆境胁迫与大豆生长发育等重要的生物学过程中发挥重要作用。有关 *GmMP* 蛋白的具体功能及其作用的分子机制还有待于进一步研究。

参考文献

[1] Lee D J, Park J W, Lee H W, et al. Genome-wide analysis of the auxin-responsive transcriptome downstream of IAA1 and its expression analysis reveal the diversity and complexity of auxin-regulated gene expression [J]. Journal of Experimental Botany, 2009, 60 (13): 3935-3957.

[2] 蒋素梅,陶均,李玲. 早期生长素响应蛋白在生长素信号转导中的作用[J]. 植物生理学通讯, 2005, 41 (1): 125-130. (Jiang S M, Tao J, Li L. The role of early auxin response protein in auxin signal transduction[J]. Plant Physiology Communication, 2005, 41 (1): 125-130.)

[3] 白华举,陈军营,刘林,等. TIR1 终于被确证为生长素受体[J]. 植物生理学通讯, 2006, 42 (4): 731-735. (Bai H J, Chen J Y, Liu L, et al. TIR1 was finally confirmed as an auxin receptor [J]. Plant Physiology Communication, 2006, 42 (4): 731-735.)

[4] Okushima Y, Overvoorde P J, Arima K, et al. Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: Unique and Overlapping Functions of ARF7 and ARF19 [J]. Plant Cell, 2005, 17: 444-463.

[5] Perez-Rodriguez P, Riano-Pachon D M, Correa L G, et al. PlnT-FDB: Updated content and new features of the plant transcription factor database [J]. Nucleic Acids Research, 2010, 38:822-827.

[6] Zhang H, Jin J, Tang L, et al. PlantTFDB 2.0: Update and improvement of the comprehensive plant transcription factor database

[J]. Nucleic Acids Research, 2011, 39:1114-1127.

[7] Wang S, Tiwari S B, Hagen G, et al. AUXIN RE-SPONSE FACTOR7 restores the expression of auxin-responsive genes in mutant *Arabidopsis* leaf mesophyll protoplasts [J]. Plant Cell, 2005, 17 (7): 1979-1993.

[8] Wilmoth J C, Wang S, Tiwari S B, et al. NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation [J]. Plant Journal, 2005, 43 (1): 118-130.

[4] Sulieman S, Tran L S. Asparagine: An amide of particular distinction in the regulation of symbiotic nitrogen fixation of legumes [J]. Critical Reviews in Biotechnology, 2012, 33(3): 695-710.

[5] Manavalan L P, Guttikonda S K, Tran L S, et al. Physiological and molecular approaches to improve drought resistance in soybean [J]. Plant Cell Physiology. 2009, 50:1260-1276.

[6] Tran L S, Mochida K. Functional genomics of soybean for improvement of productivity in adverse conditions [J]. Functional & Integrative Genomics, 2010, 10(4): 447-462.

[7] Tran L S, Nakashima K, Shinozaki K, et al. Plant gene networks in osmotic stress response: From genes to regulatory networks [J]. Methods in Enzymology, 2007, 428:109-128.

[8] Tran L S, Nishiyama R, Yamaguchi-Shinozaki K, et al. Potential utilization of NAC transcription factors to enhance abiotic stress tolerance in plants by biotechnological approach [J]. GM Crops, 2010, 1(1): 32-39.

[9] Goodstein L D. The interstate delivery of psychological service: Opportunities and obstacles [J]. Psychological Services, 2012, 9(3): 231-239.

[10] Shen C, Wang S, Bai Y, et al. Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa* L.) [J]. Experimental Botany, 2010, 61: 3971-3981.

[11] Xing H, Pudake R N, Guo G, et al. Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize [J]. BMC Genomics, 2011, 12(1): 178-189.

[12] Wu J, Wang F, Cheng L, et al. Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum* [J]. Plant Cell Report, 2011, 30:2059-2073.

[13] Wang S K, Bai Y H, Shen C J, et al. Auxin-related gene families in abiotic stress response in sorghum bicolor [J]. Functional & Integrative Genomics, 2010, 10(4): 533-546.

[14] Mun J H, Yu H J, Shin J Y, et al. Auxin response factor gene family in *Brassica rapa*: Genomic organization, divergence, expression, and evolution [J]. Molecular Genetics & Genomics, 2012, 287:765-784.

[15] Schmutz J, Cannon S B, Schlueter J, et al. Genome sequence of the palaeopolyploid soybean [J]. Nature, 2010, 463:178-183.

[16] Yang S, Vanderbeld B, Wan J, et al. Narrowing down the targets: Towards successful genetic engineering of drought-tolerant crops [J]. Molecular Plant, 2010, 3:469-490.

[17] Liao Y, Zou H F, Wang H W, et al. Soybean *GmMYB76*, *GmMYB92*, and *GmMYB177* genes confer stress tolerance in transgenic *Arabidopsis* plants [J]. Cell Research, 2008, 18:1047-1060.

[18] Finet C, Fourquin C, Vinauger M, et al. Parallel structural evolution of auxin response factors in the angiosperms [J]. Plant Journal, 2010, 63:952-964.