

研究报告

Research Report

大豆中 *GmRAD* 与 *GmDIV* 基因的 *in silico* 鉴定及表达分析

魏海超 郝荣华 李方正 杨桂文* 冯献忠*

山东省高校系统生物学重点实验室, 山东师范大学生命科学学院, 济南, 250014

* 通讯作者, fengxianzhong@yahoo.com; yanggw@sdu.edu.cn

摘要 RAD 与 DIV 属于 MYB 转录因子家族成员, 在金鱼草研究中发现它们对花对称性的形成起着重要作用。本研究利用全基因组信息鉴定了大豆中的 *GmRAD* 与 *GmDIV* 基因, 分析了其蛋白的理化性质、二级结构, 基因结构与其在染色体上的分布, 并对其系统发生和表达方式进行了研究。研究表明: 大豆中存在 5 个 *GmRAD* 基因和 5 个 *GmDIV* 基因, 它们分别属于 MYB 家族的 R1 和 R2R3 类, 分布在大豆的 9 条染色体上; 进化分析显示 *GmRAD* 和 *GmDIV* 基因具有不同的进化历程, *GmDIV* 比 *GmRAD* 基因更趋于单元发生; 表达分析证明 *GmDIV* 和 *GmRAD* 成员具有不同的表达模式, 其中 *GmRAD3* 和 *GmRAD5* 以及 *GmDIV1*、*GmDIV3*、*GmDIV4* 和 *GmDIV5* 在花组织中有较高的表达。

关键词 大豆, *GmRAD* 基因, *GmDIV* 基因, 生物信息学分析, 花发育

In silico Identification and Expression Analysis of the *GmRAD* and *GmDIV* Genes in Soybean

Wei Haichao Hao Ronghua Li Fangzheng Yang Guiwen* Feng Xianzhong*

Key Laboratory of Systems Biology in Universities of Shandong, College of Life Science, Shandong Normal University, Jinan, 250014

* Corresponding authors, fengxianzhong@yahoo.com; yanggw@sdu.edu.cn

DOI: 10.3969/mpb.010.000155

Abstract Two MYB transcription factors, RADIALIS (RAD) and DIVARICATA (DIV), play an important role in the control of flower symmetry in *Antirrhinum*. In this study, *GmRAD* and *GmDIV* genes in soybean were identified through a genome-wide bioinformatics analysis. The biochemical properties and secondary structures of their proteins, as well as their gene structures and their distributions in chromosomes were all analyzed. Their phylogeny and expression pattern were also investigated. This research shows that in soybean there are five *GmRAD* genes and five *GmDIV* genes, which belong to R1 class and R2R3 class of MYB family respectively, and they are distributed in nine chromosomes of soybean genome. Phylogenetic analysis indicates *GmRAD* genes and *GmDIV* genes evolved through different paths during speciation, and *GmDIV* genes are more likely to originate from monophyly than *GmRAD* genes. Expression analysis shows that members of these genes have different expression patterns, and *GmRAD3*, *GmRAD5*, *GmDIV1*, *GmDIV3*, *GmDIV4* are highly expressed in soybean flowers.

Keywords Soybean (*Glycine max*), *GmRAD* gene, *GmDIV* gene, Bioinformatics analysis, Flower development

花是被子植物进行有性繁殖的主要器官, 千姿百态的花型是植物吸引动物传媒和被特异授粉者识别的重要形态学特征。被子植物始祖的花为辐射对称性花, 伴随着特化传粉者的出现才有两侧对称性花的出现。两侧对称花型被认为是由原始辐射对称花型

经过多次反复独立进化而来(Endress, 1999; Stebbins, 1974)。研究植物花型发育的分子调控机理, 对于认识高等植物发育进化的历程和发展人工改造植物花器官形状的新方法具有重要的意义。

在过去的十多年中, 对模式植物金鱼草的研究

发现了由 *CYCLOIDEA* (*CYC*)、*DICHOTOMA* (*DICH*)、*RADIALIS* (*RAD*)和 *DIVARICATA* (*DIV*)共同控制的两侧对称花型发育的分子调控网络。野生型金鱼草的花由 2 枚背部花瓣、2 枚侧部花瓣和 1 枚腹部花瓣组成；其中背部花瓣和侧部花瓣的形状是不对称的，而腹部花瓣为对映体对称。两个植物特有的 TCP 家族转录因子 *CYC* 和 *DICH* 共同控制了金鱼草的背部花瓣属性，当这两个基因发生突变时，所有的花瓣均呈现腹部花瓣的特征，花型由两侧对称变成辐射对称(Luo et al., 1996; 1999)。植物 MYB 转录因子 *RAD* 同样参与了背部花瓣属性的控制，*rad* 突变体与 *cyc:dich* 突变体具有类似的表型，在花原基发育的早期 *RAD* 在背部区域表达，但是略晚于 *CYC*、*CYC* 和 *DICH* 基因可以激活 *RAD* 的表达(Corley et al., 2005)。另一个 MYB 转录因子 *DIV* 蛋白决定了腹部花瓣的特征，*DIV* 基因早期在所有花瓣原基中表达，晚期集中在腹部花瓣原基中表达，在 *cyc:dich:div* 三突变体中，所有花瓣均表现为侧部花瓣的特征(Almeida et al., 1997; Corley et al., 2005; Galego and Almeida, 2002)。*DIV* 是一个由 307 个氨基酸构成的 MYB 转录因子，含有 R2R3 两个 MYB 保守域，*RAD* 是一个只含有一个 MYB 保守域的 93 个氨基酸的 MYB 转录因子。*RAD* 和 *DIV* 蛋白具有高度相似的保守结构，可能通过目标序列的竞争结合，或者两者直接结合使 *DIV* 所决定的背部属性在背部及侧部受到抑制(Corley et al., 2005)。在野生型金鱼草花发育过程中，两个 TCP 转录因子 *CYC* 和 *DICH* 在花原基的背部表达，激活 *RAD* 基因的表达；而 *RAD* 抑制 *DIV* 在背部和侧部花原基中作用，从而使花瓣的背部属性在近轴的背部花瓣和侧部花瓣得到不同程度的体现，形成具有三种不同形态的两侧对称花(Corley et al., 2005; Galego and Almeida, 2002; Cui et al., 2010)。金鱼草中两侧对称花型发育的分子调控网络为认知植物花器官形态发育提供了基础，近年的研究证实与金鱼草同属唇形目近缘的苦苣苔科和车前科中，由 TCP 转录因子和 MYB 转录因子共同调控花型发育的基因调控网络与种类次生辐射对称花的形成相关，并且证实此基因调控网络的蜕变导致虫媒传粉向风媒传粉症候的演化(Zhou et al., 2008; Preston et al., 2011)。

大豆具有典型的两侧对称花型，但是其花瓣的对称性排列方式与金鱼草截然相反。大豆的花瓣是由 1 枚对称的背部旗瓣、2 枚不对称的侧部翼瓣和 2 枚不对称的腹部龙骨瓣所构成。两侧对称花型在

豆科和玄参科中花瓣排列方式也体现了它们独立起源的事实(Donoghue et al., 1998; Stebbins, 1974)。尽管 TCP 转录因子在控制豆科百脉根两侧对称花的形成过程中起着决定作用(Feng et al., 2006; Wang et al., 2008)，但是对于其下游的 MYB 转录因子是否也被特异的激活导致其两侧对称花的发育仍然不清楚。本研究利用生物信息学的手段鉴定了大豆中 *RAD* 与 *DIV* 的同源基因，对其蛋白结构、基因结构和理化性质的分析，利用 RNA-seq 的数据对它们在 9 个不同时期不同器官的表达情况进行了研究，为进一步研究大豆中 *GmRAD* 与 *GmDIV* 基因的功能提供了基础。

1 结果与分析

1.1 大豆 *GmRAD* 与 *GmDIV* 同源基因的鉴定

利用金鱼草中 *RAD* 与 *DIV* 的氨基酸序列，在大豆蛋白数据库 phytozome v7.0 中进行 Blastp 筛选，将所有可能的候选基因构建系统发育树，最终确定了大豆中与金鱼草中 *RAD* 同一分支的基因有 5 个：*GmRAD1* (*Glyma02g18210.1*)、*GmRAD2* (*Glyma04g16390.1*)、*GmRAD3* (*Glyma03g28050.1*)、*GmRAD4* (*Glyma06g46590.1*)和 *GmRAD5* (*Glyma19g30810.1*)；大豆中与金鱼草中 *DIV* 同一分支的基因有 6 个：*GmDIV1* (*Glyma02g39000.1*)、*GmDIV2* (*Glyma20g16630.1*)、*GmDIV3* (*Glyma14g37050.1*)、*GmDIV4* (*Glyma11g22960.1*)、*GmDIV5* (*Glyma18g07250.1*)和 *Glyma15g36670.1*。研究中发现，*Glyma15g36670.1* 与 *GmDIV2* (*Glyma20g16630.1*)一致区域的相似度为 67%，但是此基因仅含有一个 MYB 保守区，只有 108 个氨基酸，推测其可能是 *GmDIV* 基因功能演化的一个产物，在下面研究中未被包括。

利用 Clustal X (Version 2.1)分别对 *GmRAD* 和 *GmDIV* 蛋白序列进行比较分析结果如图 1 所示。这些基因所编码的蛋白分别与金鱼草中的 *RAD*、*DIV* 有较高的相似度，与 *RAD* 的氨基酸序列相比，*GmRAD1* 的相似度为 66%，*GmRAD2* 为 69%，*GmRAD3* 为 72%，*GmRAD4* 为 67%，*GmRAD5* 为 58%；与 *DIV* 的氨基酸序列相比较，*GmDIV1* 的相似度为 59%，*GmDIV2* 为 48%，*GmDIV3* 为 60%，*GmDIV4* 为 58%，*GmDIV5* 为 60%。*GmRAD* 的氨基酸序列比 *GmDIV* 的氨基酸序列分别与 *RAD* 和 *DIV* 的氨基酸序列总体上表现出更高的相似性。

利用 SMART (<http://s.mart.embl-heidelberg.de/>)

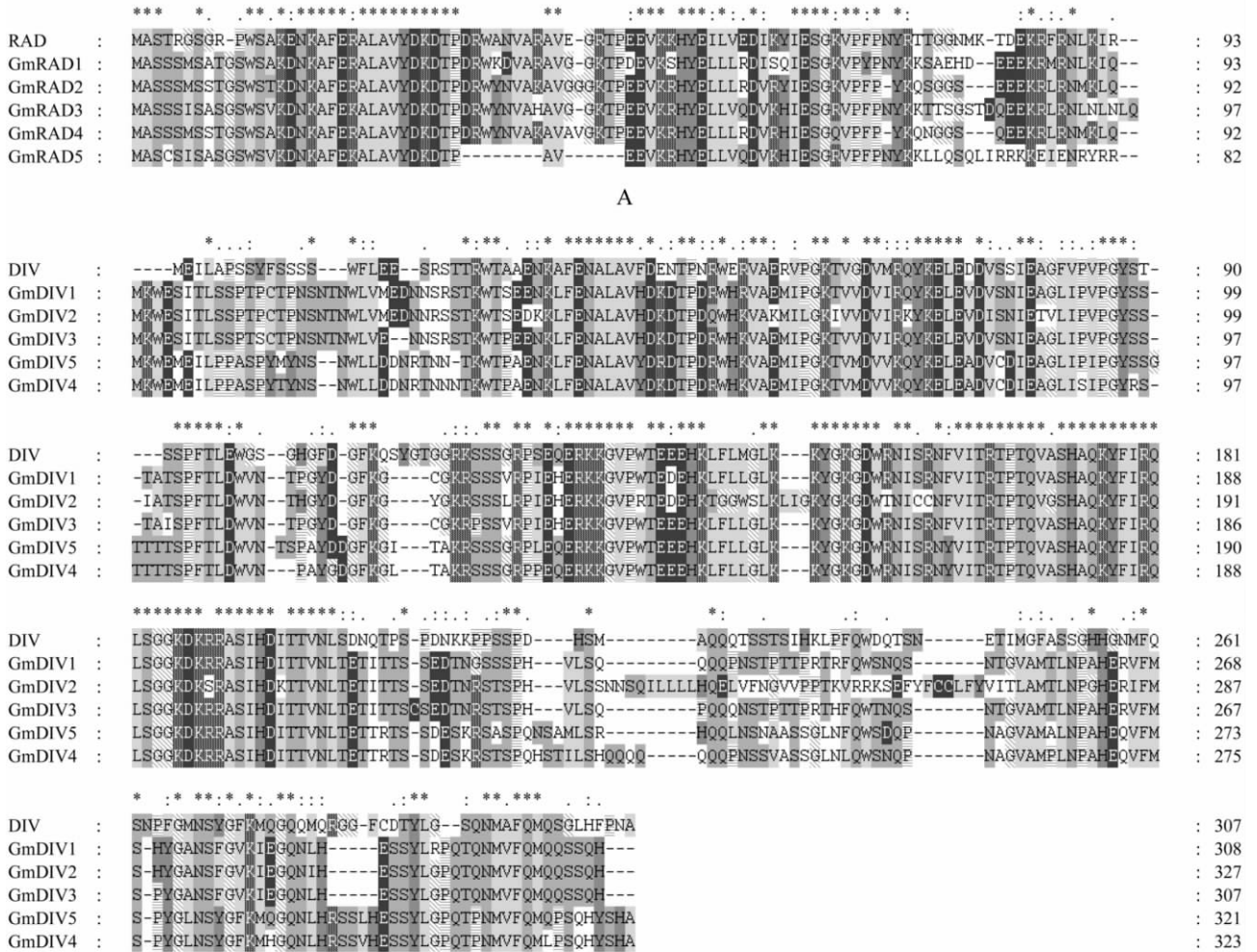


图 1 金鱼草与大豆中 RAD 和 DIV 氨基酸序列比对

Figure 1 Amino acid alignments of RAD and DIV protein sequences in *Antirrhinum majus* and *Glycine max*

对这些基因的保守区进行了分析,分析的结果显示在 GmRAD 中,GmRAD1~GmRAD4 均一个具有典型的 MYB 结构域,位置为 9~59 位氨基酸,而 GmRAD5 在此区间有 14 个氨基酸的缺失(图 1),已经无法检测到典型的 MYB 结构域,推测此缺失可能导致其功能发生较大的变化。GmDIV1~GmDIV5 都具有典型的 R2R3MYB 保守域的结构,其第一个保守域的位置在 29~81 位氨基酸之间,第二个保守域的位置在 136~188 位氨基酸之间,其中 GmDIV2 两个保守区之间的距离最大为 72 个氨基酸。

1.2 大豆 GmRAD 和 GmDIV 的理化性质和二级结构分析

为了进一步确定大豆中 GmRAD 与 GmDIV 蛋白质的理化性质,我们分析了它们编码的氨基酸组成及理化性质(表 1)。GmRAD1~GmRAD4 氨基酸的数目在 92~97 之间,分子量为 10 515.8~10 965.2 Da。

GmRAD5 由于中间 14 个氨基酸缺失,理化性质发生较大差异,其氨基酸数目为 82,分子量为 9 581 Da。GmDIV1~GmDIV5 氨基酸数目在 307~327 之间,分子量在 34 429.1~37 015.9 Da 之间。与金鱼草中 RAD 与 DIV 相比,GmRAD 与 GmDIV 蛋白质都在碱性范围内,表明蛋白分子中富含较多的碱性氨基酸,脂肪族氨基酸指数在 55~80 之间,疏水指数在 -1.05~-0.5 之间,蛋白的亲水性较高。这些蛋白的不稳定指数在 42~85 之间,均大于 40,为不稳定蛋白。

对 GmRAD 与 GmDIV 二级结构的比较发现(表 2),大豆中 GmRAD 与 GmDIV 蛋白与金鱼草结构高度相似,均由 α -螺旋、无规则卷曲和扩展链结构三种形式组成,且构成 α -螺旋的氨基酸数目都比较相近(GmRAD, 39~48; GmDIV, 62~89)。金鱼草中 RAD 含有 4 个 α -螺旋区,大豆的 RAD 同源蛋白具有相似的结构(GmRAD5 序列不完整,只含有 3 个

表 1 大豆与金鱼草中 RAD 与 DIV 理化分析

Table 1 The biochemical analysis of RAD and DIV in *Antirrhinum majus* and *Glycine max*

基因 Gene	氨基酸数目 No. of amino acids	分子量(Da) Molecular weight (Da)	等电点 Theoretical pI	分子式 Formula	不稳定指数 Instability index	脂肪酸指数 Aliphatic index	亲水性平均系数 Grand average of hydrophobicity (GRAVY)
<i>RAD</i>	93	10 774.2	9.74	C ₄₇₅ H ₇₅₄ N ₁₄₂ O ₁₄₁ S ₂	42.89	56.67	-1.046
<i>GmRAD1</i>	93	10 588.8	7.94	C ₄₆₀ H ₇₃₁ N ₁₃₃ O ₁₄₈ S ₃	48.66	58.82	-1.040
<i>GmRAD2</i>	92	10 489.8	9.40	C ₄₆₀ H ₇₂₆ N ₁₃₂ O ₁₄₃ S ₃	83.32	55.11	-0.983
<i>GmRAD3</i>	97	10 965.2	8.89	C ₄₈₂ H ₇₅₉ N ₁₃₉ O ₁₅₂ S ₁	67.49	67.32	-0.884
<i>GmRAD4</i>	92	10 515.8	9.47	C ₄₆₀ H ₇₂₈ N ₁₃₆ O ₁₄₁ S ₃	81.50	60.43	-0.927
<i>GmRAD5</i>	82	9 581.0	9.63	C ₄₂₅ H ₆₈₆ N ₁₂₄ O ₁₂₄ S ₂	85.19	79.63	-0.783
<i>DIV</i>	307	34 429.1	8.82	C ₁₅₁₀ H ₂₂₈₉ N ₄₃₃ O ₄₇₁ S ₁₂	57.10	44.17	-0.839
<i>GmDIV1</i>	308	34 838.8	8.95	C ₁₅₂₃ H ₂₃₇₂ N ₄₄₄ O ₄₇₈ S ₉	50.92	60.39	-0.800
<i>GmDIV2</i>	327	37 015.9	8.93	C ₁₆₄₂ H ₂₅₆₆ N ₄₅₆ O ₄₉₆ S ₁₂	47.52	73.24	-0.539
<i>GmDIV3</i>	307	34 672.7	8.86	C ₁₅₂₃ H ₂₃₇₀ N ₄₃₈ O ₄₇₃ S ₉	51.38	61.86	-0.763
<i>GmDIV4</i>	323	36 597.9	9.06	C ₁₆₁₃ H ₂₄₉₁ N ₄₆₁ O ₄₉₆ S ₁₀	60.15	60.68	-0.835
<i>GmDIV5</i>	321	36 296.6	8.73	C ₁₅₉₇ H ₂₄₆₄ N ₄₅₄ O ₄₉₃ S ₁₂	55.29	60.50	-0.780

表 2 大豆与金鱼草中 RAD 与 DIV 二级结构分析

Table 2 The secondary structure analysis of RAD and DIV in *Antirrhinum majus* and *Glycine max*

基因 Gene	α-螺旋数目 Number of alpha helix domain	扩展链的数目 Number of extend strand domain	无规卷曲的数目 Number of random coil domain
<i>RAD</i>	4	1	5
<i>GmRAD1</i>	4	0	5
<i>GmRAD2</i>	4	2	5
<i>GmRAD3</i>	4	3	6
<i>GmRAD4</i>	4	1	5
<i>GmRAD5</i>	3	2	5
<i>DIV</i>	7	7	15
<i>GmDIV1</i>	7	9	14
<i>GmDIV2</i>	6	11	18
<i>GmDIV3</i>	6	10	14
<i>GmDIV4</i>	5	9	13
<i>GmDIV5</i>	6	8	16

α-螺旋) ;金鱼草 DIV 含有 7 个 α-螺旋 ,大豆中的 DIV 同源基因含有 5~7 个 α-螺旋。在 MYB 基因家族中 α-螺旋在 MYB 转录因子发挥识别功能中起着主要作用 ,因此含有的 α-螺旋的不同可能会预示着这些同源基因的功能具有多样性。

1.3 大豆 *GmRAD* 和 *GmDIV* 的基因结构与染色体上的分布

利用 GSDS 在线工具制作内含子外显子结构图(图 2) ,与金鱼草中 *RAD* 和 *DIV* 的基因结构相比 ,大豆中 *GmRAD* 与 *GmDIV* 基因的内含子长度具有较大的差异。在大豆中 *GmRAD* 基因的内含子长度明显小于金鱼草中的 *RAD* 基因内含子长度(829 bp)。其

中 *GmRAD1*~*GmRAD4* 的内含子长度分别为 351 bp、515 bp、639 bp 和 383 bp *GmRAD5* 含有两个内含子 ,长度依次是 38 bp 和 94 bp。金鱼草 *DIV* 基因的内含子长度为 260 bp ,与 *GmDIV* 的内含子长度也存在显著的差别。其中 *GmDIV1* 的内含子长度为 336 bp ,*GmDIV3*~*GmDIV5* 的内含子的长度分别为 372 bp、441 bp 和 420 bp *GmDIV2* 的结构比较特殊 ,含有 4 个内含子 ,长度依次为 :140 bp、176 bp、87 bp 和 320 bp。*GmRAD* 中内含子长度减少和 *GmDIV* 基因内含子长度增加的差异提示 ,两者可能经历了不同的进化选择过程。

利用大豆基因组数据库中基因的位置信息 ,制作了大豆中 *GmRAD* 与 *GmDIV* 在染色体上的分布

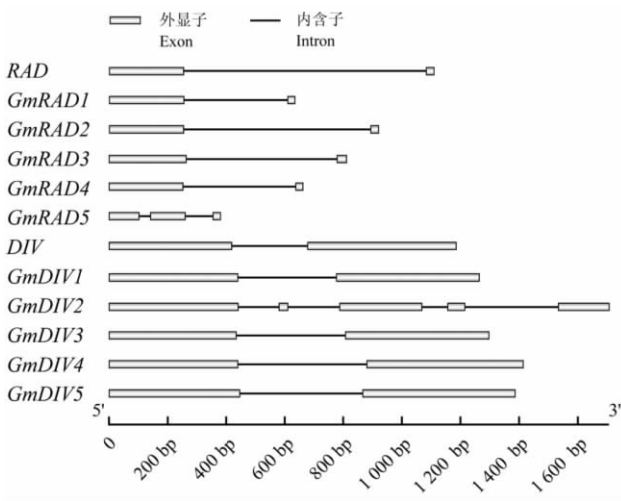
图2 金鱼草与大豆中 *RAD* 与 *DIV* 基因结构

Figure 2 The gene structures of *RAD* and *DIV* in *Antirrhinum majus* and *Glycine max*

图(图3)。从图3中可以看出,大豆的5个 *GmRAD* 基因分别定位于2、3、4、6和19号五个染色体上,*GmDIV* 基因分别定位于2、11、14、18和20号五个染色体上。其中 *GmRAD1* 与 *GmDIV1* 定位在2号染色体上,其余基因分别分布在不同染色体的不同位置上,提示这些基因可能在大豆多倍化过程中可能经历了不同形式的选择。

1.4 大豆 *GmRAD* 与 *GmDIV* 的系统发育分析

利用金鱼草的 *RAD* 与 *DIV* 蛋白质序列,在拟南芥、蒺藜苜蓿和水稻中分别找到与金鱼草 *RAD* 和 *DIV* 同源性较高的基因分别3个、2个和3个。利用这些序列和 *GmRAD* 与 *GmDIV* 构建系统发育树

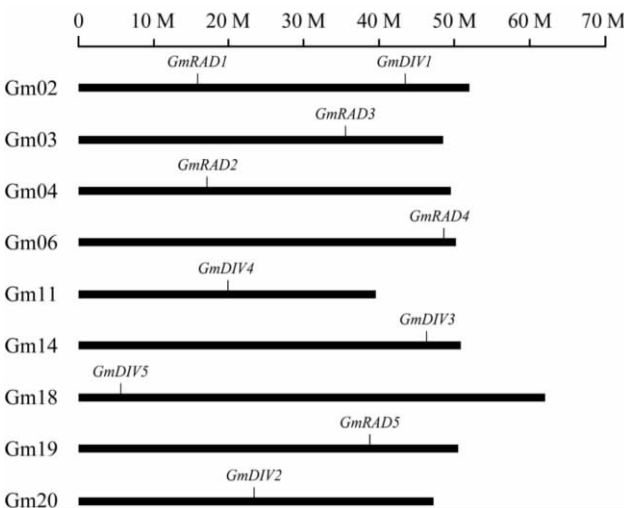
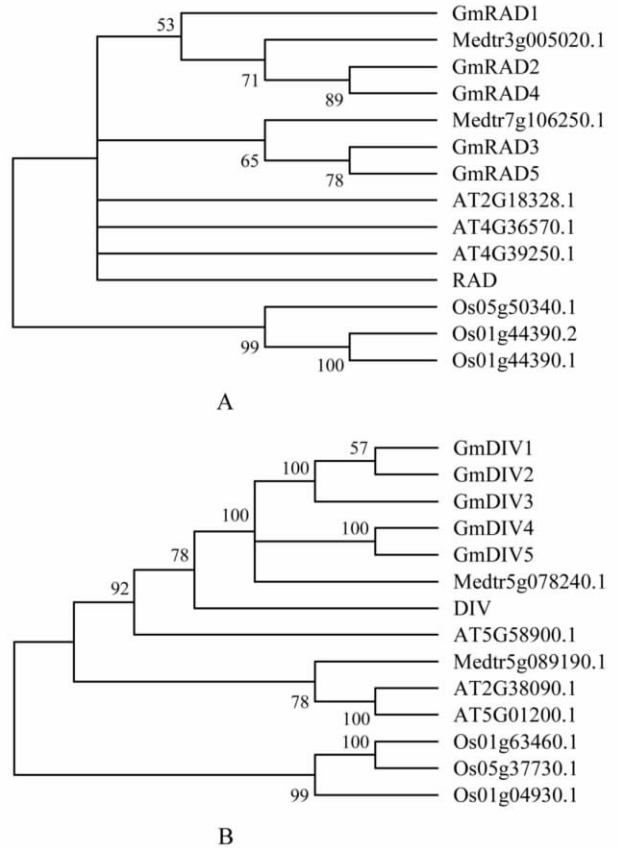
图3 大豆中 *GmRAD* 与 *GmDIV* 在染色体上分布

Figure 3 The distribution of *GmRAD* and *GmDIV* genes in *Glycine max* chromosomes

(图4)。从图4中可以看出,单子叶禾本科植物水稻的 *RAD* 和 *DIV* 均处于一个独立的进化枝上,说明双子叶中的 *RAD* 和 *DIV* 基因具有更高的亲缘关系;在双子叶枝中, *RAD* 和 *DIV* 的进化方式存在很大的差异, *DIV* 相较于 *RAD* 更趋于单元发生。在豆科模式

图4 利用 MEGA 5 构建了五个物种的 *RAD* 与 *DIV* 蛋白的 NJ 系统发育树

注:每一分枝显示 bootstrap (1 000 次)支持度大于 50%的序列;大豆, *GmRAD1*~*GmRAD5*, *GmDIV1*~*GmDIV5*; 金鱼草, *RAD*, *DIV*; 蒺藜苜蓿, *Medtr5g078240.1*, *Medtr5g089190.1*, *Medtr3g005020.1*, *Medtr7g106250.1*; 拟南芥, *AT4G36570.1*, *AT4G39250.1*, *AT2G18328.1*, *AT5G58900.1*, *AT5G01200.1*, *AT2G38090.1*; 水稻, *Os05g37730.1*, *Os01g04930.1*, *Os01g63460.1*, *Os01g44390.2*, *Os05g50340.1*, *Os01g44390.1*

Figure 4 Phylogram based on NJ analysis of *RAD* and *DIV* proteins from five species using MEGA5

Note: Branches with bootstrap support (1 000 replicates) over 50% are indicated for each branch; The amino acid sequences from the five species are: *Glycine max*, *GmRAD1*~*GmRAD5*, *GmDIV1*~*GmDIV5*; *Antirrhinum majus*, *RAD*, *DIV*; *Medicago*, *Medtr5g078240.1*, *Medtr5g089190.1*, *Medtr3g005020.1*, *Medtr7g106250.1*; *Arabidopsis thaliana*, *AT4G36570.1*, *AT4G39250.1*, *AT2G18328.1*, *AT5G58900.1*, *AT5G01200.1*, *AT2G38090.1*; *Oryza sativa*, *Os05g37730.1*, *Os01g04930.1*, *Os01g63460.1*, *Os01g44390.2*, *Os05g50340.1*, *Os01g44390.1*

植物蒺藜苜蓿中 *RAD* 和 *DIV* 的进化模式呈现出明显的不同。在蒺藜苜蓿中有两个基因与金鱼草的 *RAD* 亲缘关系较近，而与 *DIV* 亲缘关系较近的只有一个基因。大豆中 *GmRAD* 基因同蒺藜苜蓿一样，可以区分为与 *RAD* 平行的两个进化枝，其中的 *GmRAD1*、*GmRAD2* 和 *GmRAD4* 枝经历了两次分化；它们与 *GmRAD3*、*GmRAD5* 可能来源于不同始祖基因，而大豆中 *GmDIV* 与单一的蒺藜苜蓿 *DIV* 蛋白聚类在同一进化枝上，并在随后经历了三次分化，说明 *GmDIV* 可能来源于相同的始祖基因。系统发育分析比较显示，大豆中 *GmRAD* 的进化途径与 *GmDIV* 的进化途径明显相同。

1.5 大豆 *GmRAD* 和 *GmDIV* 基因的表达分析

利用大豆转录组的数据 (Libault et al., 2010)，对 *GmRAD* 和 *GmDIV* 基因在 9 个不同表达时期及部位 (播种 84 h 后植株的根毛，播种 120 h 后植株的根毛，根瘤，茎顶端分生组织，花，绿色果荚，叶片，根和根尖) 进行了基因表达分析 (图 5)。分析结果显示，*GmRAD1*、*GmRAD3*~*GmRAD5* 主要在花瓣、叶片和绿色果荚中表达，*GmRAD2* 主要在叶片与根尖中表达；*GmDIV1*、*GmDIV3*~*GmDIV5* 主要在绿色果荚、叶片表达，在播种 84 h 后植株的根毛、根瘤中也有表达，而 *GmDIV2* 主要在播种 84 h 后植株的根毛和茎顶端分生组织中表达。以 RPM (Reads Per Million reads) 为

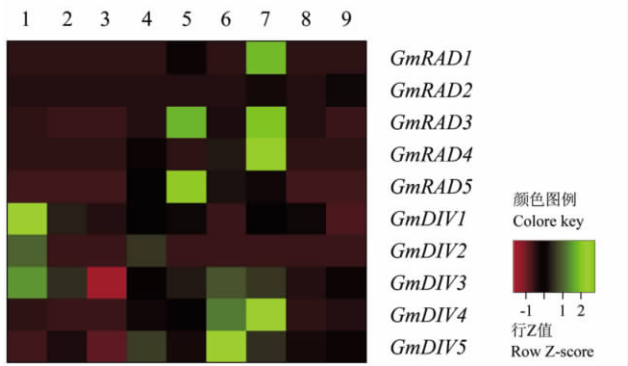


图 5 大豆中 *GmRAD* 与 *GmDIV* 基因表达谱分析
注: 1: 播种 84 h 后根毛; 2: 播种 120 h 后根毛; 3: 根瘤; 4: 茎顶端分生组织; 5: 花; 6: 绿色果荚; 7: 叶片; 8: 根; 9: 根尖; 方块内颜色显示基因表达水平, 红色最低, 黑色居中, 绿色最高
Figure 5 The expression profiles of *GmRAD* and *GmDIV* genes in soybean
Note: 1: Root hair 84HAS; 2: Root hair 120HAS; 3: Nodule; 4: SAM; 5: Flower; 6: Green pod; 7: Leaf; 8: Root; 9: Root tip; Color in square represents the level of gene expression, red is the lowest, black is the middle, green is the highest

标准, 在花组织中 *GmRAD1*~*GmRAD5* 的表达相对量分别为 0.97、0.00、41.52、0.00 和 13.52; *GmDIV1*~*GmDIV5* 的表达相对量分别为 21.25、0.00、64.70、29.9 和 31.8。在 *GmRAD* 基因中, *GmRAD3* 和 *GmRAD5* 更有可能参与了花发育的调控, 其余基因可能已经被招募到其它器官的发育进程中; 在 *GmDIV* 基因中, *GmDIV1*、*GmDIV3*、*GmDIV4* 和 *GmDIV5* 均有可能参与花发育的调控。

2 讨论

植物中 MYB 转录因子是植物中最大的一类转录因子家族, 广泛参与了植物的发育和代谢调控, 所包含的 MYB 结构域从一个到四个不等, 其结构域数目和构成与其功能密切相关 (Feller et al., 2011)。MYB 保守区约含 51 或 52 个氨基酸残基, 可以形成三段螺旋结构, 且在第二段和第三段螺旋之间折转成一定的角度相连接, 形成了螺旋-转角-螺旋的结构, 构成了一个疏水核心; 第一个螺旋目前还没有明确的结构意义, 而第二个螺旋同 DNA 的磷酸骨架相接触, 稳固了与 DNA 的结合, 第三个螺旋为识别螺旋, 通过氨基酸侧链以及 DNA 碱基之间的氢键同 DNA 相识别 (Ogata et al., 1996; Dubos et al., 2010)。MYB 转录因子基因 *RAD* 与 *DIV* 作为 *TCP* 基因调控的下游基因, 它们在控制花对称性的作用机制已经在模式植物金鱼草中得到了阐明 (Corley et al., 2005), 本研究分析和鉴定了大豆中的 *GmRAD* 和 *GmDIV* 基因, 证实大豆中 *GmRAD* 和 *GmDIV* 基因至少存在 5 个拷贝, *GmRAD1*~*GmRAD4* 基因均具有一个典型的 MYB 结构域, *GmDIV1*~*GmDIV5* 均具有典型的 R2R3MYB 保守域的结构; *GmRAD5* 在 MYB 保守域存在的 14 个氨基酸的缺失, 同时发现存在着一个与 *GmDIV2* 同源性较高但失去一个 MYB 保守域的 Glyma15g36670.1 蛋白, 大豆中 *GmRAD* 和 *GmDIV* 基因的多样性为深入研究 MYB 基因的功能演化提供了一个很好的体系。

在大豆中 *GmRAD* 和 *GmDIV* 基因均存在多个拷贝, 这与大豆基因组的复制有关, 已有的研究结果表明, 大豆基因组至少发生了两次复制, 引起了整个基因组的高度重复 (约 75% 的基因以多拷贝形式出现) (Schmutz et al., 2010)。对 *GmRAD* 系统发育分析比较显示: 大豆中 *GmRAD1*、*GmRAD2*、*GmRAD4* 和 *GmRAD3*、*GmRAD5* 来源于不同的始祖基因, 其始祖基因的分化可能早于豆科蝶形花亚科的出现, 而大豆 *GmDIV1*~*GmDIV5* 基因则起源于相同的始祖基

因,说明这两者之间经历了不同的进化过程。在蝶形花亚科模式植物百脉根的研究中证实, *CYC*-like 基因同样决定和控制着背部花瓣的特征,在独立起源的不同类型两侧对称花中,控制花的两侧对称性的不同机制可能具有共同的分子起源(Feng et al., 2006)。随后在蝶形花亚科的豌豆中(Wang et al., 2008)证明 *PsCYC1*、*PsCYC2* 和 *PsCYC3* 同样参与了豌豆两侧对称花的形成,并与花瓣内部不对称性的形成有关。在豆科中 *CYC*-like 基因可能来源与两次基因加倍,第一次基因加倍发生在豆科蝶形花亚科分化之前,导致 LEGCYC GROUP 和 GROUP 两类 *CYC* 基因的产生;第二次基因加倍发生在在蝶形花亚科分化的早期,LEGCYC GROUP 内部又发生基因加倍事件,从而产生了 LEGCYC GROUP A 与 B 两类基因,因此在豆科中至少存在 3 个 *CYC*-like 基因(Citerne et al., 2003; Fukuda et al., 2003; Ree et al., 2004)。因此,豆科中控制花型 *TCP* 基因进化途径与 *GmRAD* 和 *GmDIV* 基因进化途径不尽一致,模式植物金鱼草中调控两侧对称花发育的分子调控网络的 MYB 转录因子在大豆中可能发生了功能冗余和亚功能化。大豆中控制两侧花对称性发育基因调控网络的因子在进化中呈现出功能多样性,其花型的调控可能存在更复杂的机制。

3 材料与方法

3.1 材料

本实验研究所用的金鱼草 RAD 与 DIV 蛋白质序列均来自 NCBI 网站(<http://www.ncbi.nlm.nih.gov>) 其登录号分别为 2CJJ_A 和 AAL78741 大豆、蒺藜苜蓿和水稻全基因组的数据、蛋白数据等均来自于 <http://www.phytozome.net/soybean> 网站,拟南芥全基因组的数据来自 TAIR 网站(<http://www.arabidopsis.org/>)。

3.2 大豆 *GmRAD* 与 *GmDIV* 基因鉴定及基因结构分析

利用金鱼草中 RAD 与 DIV 蛋白序列在大豆基因组数据库(<http://www.phytozome.com/>),蛋白数据库中进行 Blastp 查询,为了能得到全部的 *GmRAD* 与 *GmDIV* 相关基因, E 值设为 1 利用 SMART (<http://smart.embl-heidelberg.de/>) 在线工具对蛋白保守域进行了判定。利用 GSDS 在线工具(<http://gsds.cbi.pku.edu.cn/>) 制作了 *GmDIV* 与 *GmRAD* 基因的外显子-内含子结构图。根据 phytozome v7.0 中标注的基因信息,手工标注大豆 *GmRAD* 与 *GmDIV* 在染色

体中的位置。

3.3 大豆 *GmRAD* 与 *GmDIV* 理化性质分析

利用 <http://www.expasy.org> 和 <http://pbil.ibcp.fr/htm/index.php> 网站,对 *GmRAD* 与 *GmDIV* 同源蛋白的分子量、等电点、亲水性、稳定性和二级结构等方面进行预测,得到 *GmRAD* 与 *GmDIV* 的基本性质。

3.4 多序列比对及系统发育树的构建

将检索到的 *GmRAD* 与 *GmDIV* 相关序列保存成 FASTA 格式,利用 Clustal X (Version 2.1)对金鱼草、大豆、蒺藜苜蓿、拟南芥、水稻的 RAD 与 DIV 相关蛋白进行比对分析,利用 MEGA 软件(version 5.05) (Tamura et al., 2011)使用 Neighbor-Joining 的方法对蛋白质序列构建系统发育树,bootstrap 1 000 次。

3.5 大豆 *GmRAD* 与 *GmDIV* 基因表达分析

根据已发表的对大豆转录组的测序结果(Libault et al., 2010),在其公布的大豆 51 529 个基因的表达数据中筛选大豆 *GmRAD* 与 *GmDIV* 基因的相关数据,检测大豆 *GmRAD* 与 *GmDIV* 基因在不同组织中的表达情况。将得到的数据使用 R 软件(Version 2.14.0)环境下的 heatmap.2 程序进行处理,制作成热图(heatmap)。

作者贡献

冯献忠、杨桂文为本研究的设计者及负责人,魏海超负责样品采集,试验设计与分析,及论文初稿的写作,郝荣华和李方正参与数据分析及论文修改。全体作者都阅读并同意最终的文本。

致谢

本研究由国家自然科学基金重大研究计划培育项目(91131008)和山东省杰出青年基金项目(JQ200909)共同资助。

参考文献

- Almeida J., Rocheta M., and Galego L., 1997, Genetic control of flower shape in *Antirrhinum majus*, *Development*, 124(7): 1387-1392
- Citerne H.L., Luo D., Pennington R.T., Coen E., and Cronk Q.C., 2003, A phylogenomic investigation of *CYCLOIDEA*-like *TCP* genes in the *Leguminosae*, *Plant Physiol.*, 131(3): 1042-1053
- Corley S.B., Carpenter R., Copsey L., and Coen E., 2005, Floral

- asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*, Proc. Natl. Acad. Sci., USA, 102(14): 5068-5073
- Cui M.L., Copsey L., Green A.A., Bangham J.A., and Coen E., 2010, Quantitative control of organ shape by combinatorial gene activity, Plos Biol., 8(11): e1000538
- Donoghue M.J., Ree R.H., and Baum D.A., 1998, Phylogeny and the evolution of flower symmetry in the *Asteridae*, Trends in Plant Science, 3(8): 311-317
- Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C., and Lepiniec L., 2010, MYB transcription factors in *Arabidopsis*, Trends in Plant Science, 15(10): 573-581
- Endress P.K., 1999, Symmetry in flowers: diversity and evolution, Int. J. Plant Sci., 160(S6): S3-S23
- Feller A., Machemer K., Braun E.L., and Grotewold E., 2011, Evolutionary and comparative analysis of MYB and bHLH plant transcription factors, Plant J., 66(1): 94-116
- Feng X., Zhao Z., Tian Z., Xu S., Luo Y., Cai Z., Wang Y., Yang J., Wang Z., Weng L., Chen J., Zheng L., Guo X., Luo J., Sato S., Tabata S., Ma W., Cao X., Hu X., Sun C., and Luo D., 2006, Control of petal shape and floral zygomorphy in *Lotus japonicus*, Proc. Natl. Acad. Sci., USA, 103(13): 4970-4975
- Fukuda T., Yokoyama J., and Maki M., 2003, Molecular evolution of *cycloidea*-like genes in *Fabaceae*, J. Mol. Evol., 57(5): 588-597
- Galego L., and Almeida J., 2002, Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers, Genes Dev., 16(7): 880-891
- Libault M., Farmer A., Joshi T., Takahashi K., Langley R.J., Franklin L.D., He J., Xu D., May G., and Stacey G., 2010, An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants, Plant J., 63(1): 86-99
- Luo D., Carpenter R., Copsey L., Vincent C., Clark J., and Coen E., 1999, Control of organ asymmetry in flowers of *Antirrhinum*, Cell, 99(4): 367-376
- Luo D., Carpenter R., Vincent C., Clark J., and Coen E., 1996, Origin of floral asymmetry in *Antirrhinum*, Nature, 383(6603): 794-799
- Ogata K., Kanei-Ishii C., Sasaki M., Hatanaka H., Nagadoi A., Enari M., Nakamura H., Nishimura Y., Ishii S., and Sarai A., 1996, The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation, Nat. Struct. Biol., 3(2): 178-187
- Preston J.C., Martinez C.C., and Hileman L.C., 2011, Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome, Proc. Natl. Acad. Sci., USA, 108(6): 2343-2348
- Ree R.H., Citerne H.L., Lavin M., and Cronk Q.C., 2004, Heterogeneous selection on *LEGCYC* paralogs in relation to flower morphology and the phylogeny of *Lupinus (Leguminosae)*, Mol. Biol. Evol., 21(2): 321-331
- Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.L., Song Q., Thelen J.J., Cheng J., Xu D., Hellsten U., May G.D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M.K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian Z., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X.C., Shinozaki K., Nguyen H.T., Wing R.A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G., Shoemaker R.C., and Jackson S.A., 2010, Genome sequence of the palaeopolyploid soybean, Nature, 463(7278): 178-183
- Stebbins G.L., 1974, Flowering plants: evolution above the species level, Belknap Press of Harvard University Press, University of Michigan, pp.399
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S., 2011, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol., 28(10): 2731-2739
- Wang Z., Luo Y., Li X., Wang L., Xu S., Yang J., Weng L., Sato S., Tabata S., Ambrose M., Rameau C., Feng X., Hu X., and Luo D., 2008, Genetic control of floral zygomorphy in pea (*Pisum sativum* L.), Proc. Natl. Acad. Sci., USA, 105(30): 10414-10419
- Zhou X.R., Wang Y.Z., Smith J.F., and Chen R., 2008, Altered expression patterns of TCP and MYB genes relating to the floral developmental transition from initial zygomorphy to actinomorphy in *Bournea (Gesneriaceae)*, New Phytol., 178(3): 532-543