

豆科 AGAMOUS 同源基因结构及功能分析

李庆忠^{1,2}, 陈江华¹

(1.中国科学院西双版纳热带植物园, 云南 昆明 650223; 2.中国科学院大学, 北京 100049)

摘要: AGAMOUS(AG)基因是控制高等植物花发育的重要基因, 已在 20 多种植物基因组中发现同源基因。作为 MADS-box 家族的一员, AG 基因结构具有高度的保守性。AG 及其同源基因在植物生长发育中的功能已经十分清晰。本文研究 AG 同源基因在豆科几个代表物种中的分布, 对其基因结构和蛋白序列进行分析比对。结果表明, AG 同源基因在不同的豆科物种中具有高度的序列同源性及结构保守性。进一步通过蒺藜苜蓿 *Medicago truncatula* 的 AG 同源基因表达模式分析发现, 其表达是与功能相互验证的。

关键词: AGAMOUS; 豆科; 蒺藜苜蓿

Doi: 10.3969/j.issn.1009-7791.2016.04.003

中图分类号: Q943.2

文献标识码: A

文章编号: 1009-7791(2016)04-0315-06

Structure and Function Analysis of AGAMOUS Homologous Genes in Fabaceae

LI Qing-zhong^{1,2}, CHEN Jiang-hua¹

(1.Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming 650223, Yunnan China; 2.University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: AGAMOUS (AG) is an important gene which controls the flower development in higher plants. Since it was found in 1990s, dozens of homologous genes were identified in more than 20 kinds of plants. As a member of MADS-box family, it has a highly conservative structure and function. Now its functions in the growth and development of plants are very clear. In this study, structures and functions of AG and its homologous genes in some model plants in Fabaceae were analysed. The results proved that structures and functions of AG and its homologous genes were higher conserved. Deeply research about gene expression in *Medicago truncatula* showed that expression of AG related with its functions. This study can not only use for deeply research of AG, but also provide an important experimental data for flower strains breeding.

Key words: AGAMOUS; Fabaceae; *Medicago truncatula*

自 20 世纪 90 年代以来, 基于模式植物拟南芥 *Arabidopsis thaliana* 及金鱼草 *Antirrhinum majus* 的众多花发育的同源异型突变体被发现。在表型分析的基础上, 学界认为存在着 A、B、C 三类基因分别控制四轮花器官的发育^[1-3]。1991 年 Coen 和 Meyerowitz^[4]共同提出花发育 ABC 模型, AGAMOUS (AG)就是其中的 C 类基因, 负责控制雄蕊及心皮的发育^[5-6]。Angenent 等^[7]提出的 ABCD 模型及 Pelaz 等^[8]提出的 ABCDE 模型, 使得对 AG 基因的认识更加清楚。目前, 已在 20 多种植物中(包含裸子植物和被子植物)发现 AG 基因的同源基因。

AG 基因属于 MADS-box 家族, 从 5'端到 3'端依次可分为 N 末端、M 区、I 区、K 区及 C 末端^[9]。其中 M 区高度保守, I 区和 K 区相对保守。C 末端则包含有相对保守的 AG I 区和 AG II 区^[10], 是 AG 基因的标志性序列。AG 及其同源基因有 4 方面作用: (1)与 A 功能基因相互拮抗, 抑制 A 功能基因在

收稿日期: 2016-09-17

基金项目: 国家自然科学基金(31471171)

作者简介: 李庆忠, 硕士研究生, 从事植物功能基因研究。E-mail: liqingzhong3@126.com

注: 陈江华为通信作者。E-mail: jhchen@xtbg.ac.cn

内两轮花器官中的表达^[11], 如拟南芥的 *APETALA2* (*AP2*)和 *APETALA1*(*AP1*)基因^[12-13]; (2)通过诱导 *KNU* 基因的表达间接抑制 *WUSCHEL*(*WUS*)基因表达或者通过募集 PcG 蛋白到 *WUSCHEL*(*WUS*)基因的启动子区域直接抑制其功能, 从而终止花分生组织活性^[14-16]; (3)指定雄蕊及心皮属性^[17]; (4)调控胚珠的发育^[18]。豆科模式植物蒺藜苜蓿 *Medicago truncatula* 基因组中已知 *MtAGa*、*MtAGb* 两个 C 功能基因的同源基因^[19-20], 其时空表达模式与 C 功能基因极为相似, 但单突变表型都不明显, 双突变也只具有部分 C 功能基因缺失的表型。

豆科是仅次于菊科和兰科的第三大科植物, 具有食用、药用、建材及观赏价值, 与人类生活息息相关。目前还没有关于豆科植物 AG 基因的研究。本研究对豆科一些代表物种的 AG 基因进行筛选与序列分析, 并对蒺藜苜蓿中的 4 个 AG 基因进行深入研究, 为理解 AG 基因的进化及功能提供参考。

1 材料与方法

1.1 豆科各代表物种同源基因序列的获得

以拟南芥 *Arabidopsis thaliana* 的 AG 基因氨基酸序列为基础, 分别在拟南芥、鹰嘴豆 *Cicer arietinum*、大豆 *Glycine max*、百脉根 *Lotus corniculatus*、蒺藜苜蓿 *Medicago truncatula*、菜豆 *Phaseolus vulgaris*、木豆 *Cajanus cajan* 的网站上(<https://www.arabidopsis.org>、<http://legumeinfo.org/>、<http://www.kazusa.or.jp/lotus>、<http://bioinfo3.noble.org/doblast>)通过氨基酸序列 blast 获得 AG 同源基因序列。

1.2 蛋白比对、筛选与共线性分析及系统进化树构建

去除匹配度 80%以下的基因, 将剩余基因用 Vector NTI 软件进行氨基酸比对, 通过 AG 基因特有的保守结构域进行筛选及标注。以氨基酸序列为基础, 使用 Clustal X1.81 及 MEGA5.0 软件构建系统进化树。共线性分析在 PLAZA 官网(<http://bioinformatics.psb.ugent.be/plaza/>)在线完成, 通过 Photoshop 软件对所得结果进行处理。

1.3 蒺藜苜蓿 AG 同源基因表达模式分析

以蒺藜苜蓿 AG 同源基因的 CDS 序列为参照序列, 在 NOBLE 的苜蓿基因表达谱网站(http://mtgea.noble.org/v3/blast_search_form.php)进行 blast, 筛选后保存表达图谱。

2 结果与分析

2.1 AG 同源基因的获得及氨基酸同源性比对分析

以拟南芥 AG 基因的氨基酸序列为参照序列, 分别在拟南芥、鹰嘴豆、大豆、百脉根、蒺藜苜蓿、菜豆、木豆的基因组中筛选得到 103、52、36、8、48、79、76 个基因(表 1)。进一步排除氨基酸匹配度 80%以下的基因, 通过 AG 基因保守结构域筛选后, 最终分别筛选到 3、2、7、2、4、2、2 个基因, 共 22 个(表 2)。筛选所得 22 个基因都属于 phytozome_10_2.59223110 和 phytozome_10_2.59319177 两个家族。

表 1 AG 及其同源基因结果统计		
Table 1	Statistical results of AG and its homologous genes	
物种名称	Blast 结果	筛选出基因数
拟南芥 <i>Arabidopsis thaliana</i>	103	3
鹰嘴豆 <i>Cicer arietinum</i>	52	2
大豆 <i>Glycine max</i>	36	7
百脉根 <i>Lotus corniculatus</i>	8	2
蒺藜苜蓿 <i>Medicago truncatula</i>	48	4
菜豆 <i>Phaseolus vulgaris</i>	79	2
木豆 <i>Cajanus cajan</i>	76	2

经氨基酸序列比对分析, 发现这 22 个基因的氨基酸序列保守性很高, 特别是 M 区几乎完全保守, I 区和 K 区相对保守, 完全符合 MADS-box 基因家族的特征。此外, 这 22 个基因均含有 AG 同源基因特有的相对保守的 AGI 区和 AGII 区(图 1)。

2.2 基因结构分析、共线性分析及系统进化树构建

作为 MADS-box 家族的一员, AG 及其同源基因的结构保守性强, 这种保守性在豆科中同样存在。筛选所得的 19 个豆科 AG 同源基因的结构极其类似。首先, 除百脉根外的 17 个基因均含有 7 个外显子和 6 个内含子(图 2), 且每个基因 5'端和 3'端一般含有一个较大的外显子, 而内部的 5 个外显子则相对较小; 另一方面, 虽然 AG 同源基因内含子的大小差异很大, 但普遍都含有一个相对较大的内含子。

表2 AG及其同源基因

Table 2 Detailed information of AG and its homologous genes

植物名	编号	基因名称	染色体	范围	CDS	蛋白数	基因家族
拟南芥	AT4G18960	AGAMOUS(AG)	4	10383917-10388272	759	252	phytozome_10.2.59223110
<i>Arabidopsis thaliana</i>	AT3G58780	AGAMOUS-LIKE 1(AGL1), SHATTERPROOF 1(SHP1)	3	21739150-21741766	822	273	phytozome_10.2.59223110
	AT2G42830	AGAMOUS-LIKE 5(AGL5), SHATTERPROOF 2(SHP2)	2	17820602-17823806	747	248	phytozome_10.2.59223110
鹰嘴豆	Ca_00445		1	3643879-3651187	738	245	phytozome_10.2.59223110
<i>Cicer arietinum</i>	Ca_08639		6	9594967-9600418	741	246	phytozome_10.2.59319177
大豆	Glyma.05G163200	Glyma05G29595	5	35435742-35444074	732	243	phytozome_10.2.59319177
	Glyma.08G120600	Glyma08G12730	8	9301551-9308870	732	243	phytozome_10.2.59319177
	Glyma.13G223300	Glyma13G29510	13	33610531-33620180	732	243	phytozome_10.2.59223110
	Glyma.15G088600	Glyma15G09500	15	6826179-6835890	732	243	phytozome_10.2.59223110
	Glyma.02G287700	Glyma02G45730	2	46797959-46805018	741	246	phytozome_10.2.59223110
	Glyma.08G310100	Glyma08G42300	8	42948545-42957616	744	247	phytozome_10.2.59223110
	Glyma.18G105800	Glyma18G12590	18	11773697-11782733	741	246	phytozome_10.2.59223110
百脉根	Lj4g3v2628690		6	24265195-24272227	744	247	phytozome_10.2.59319177
	Lj3g3v1981270		4	36523180-36528338	630	209	phytozome_10.2.59223110
蒺藜苜蓿	Medtr2g017865	MtAGa	2	5545792-5553523	783	260	phytozome_10.2.59223110
	Medtr8g087860	MtAGb	8	36350163-36354700	735	244	phytozome_10.2.59223110
	Medtr3g452380		3	19142782-19134712	735	244	phytozome_10.2.59223110
	Medtr3g005530		3	250434-244289	672	223	phytozome_10.2.59223110
菜豆	Phvul.002G243200		2	40988219-40993833	729	242	phytozome_10.2.59319177
	Phvul.006G169600		6		738	245	phytozome_10.2.59223110
木豆	C.cajan_31945		Scaffold	316809-325078	723	240	phytozome_10.2.59223110
<i>Cajanus cajan</i>			127675				
	C.cajan_05842		2	14248178-14252726	738	245	phytozome_10.2.59223110

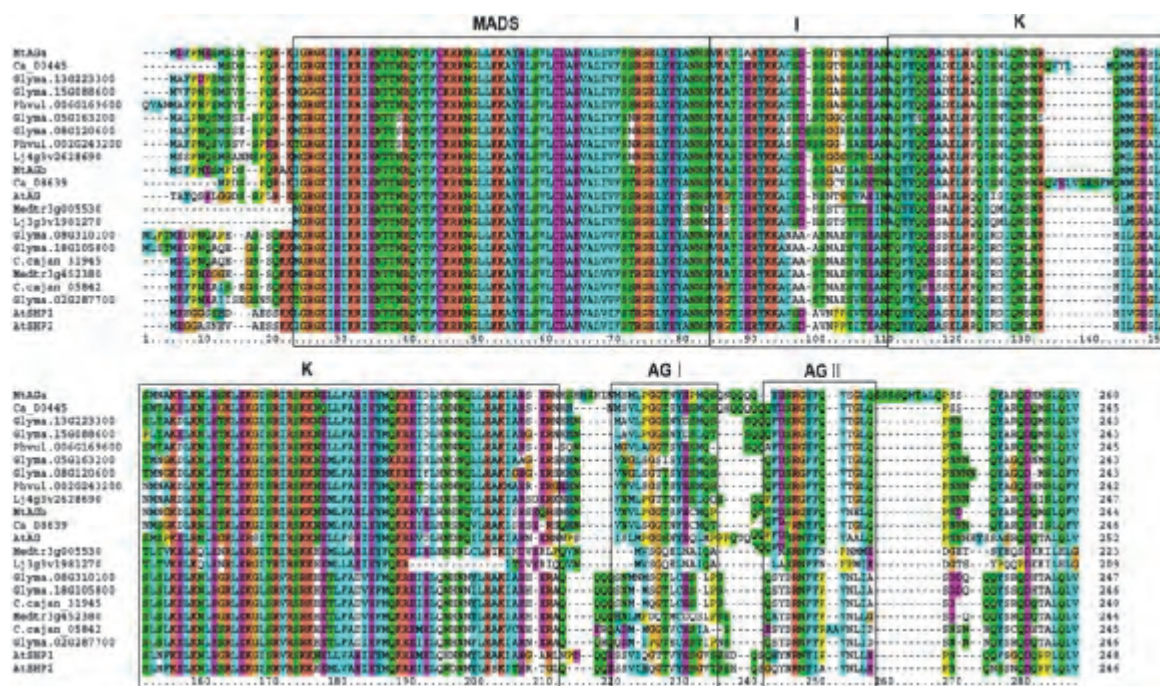


图1 AG同源基因氨基酸比对

Fig. 1 Amino acids alignments of AG and its homologous genes

通过 PLAZA 网站对拟南芥、大豆和蒺藜苜蓿的 AG 同源基因进行共线性分析, 发现 3 个物种的 14 个基因中有 11 个是共线的(图 3), 因此, 可以认为 AG 基因在进化过程中是十分保守的。同时, 系统进化树分析表明, AG 及其同源基因可分为两个亚家族, 分别是 AG 亚家族和 SHP 亚家族(图 4)。SHP 亚家族, 即 ABCDE 模型中的 E 类基因, 主要在内三轮花器官属性建立的过程中发挥作用, 该基因的正常表达是花瓣、雄蕊及心皮正常发育所必需的。此外, 进化树表明 AG 亚家族和 SHP 亚家族在进化上极可能来自于共同的原始基因。

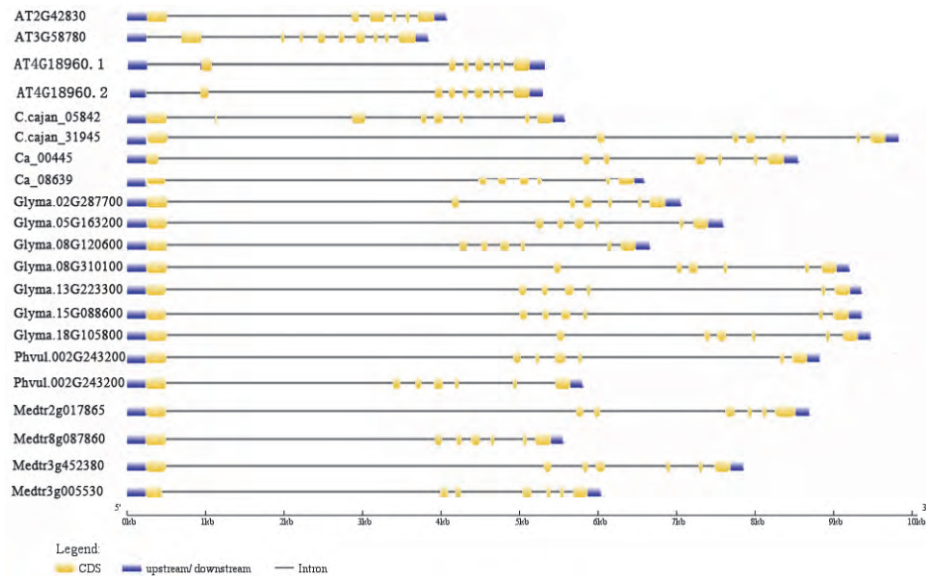


图 2 AG 同源基因结构图

Fig. 2 Structures of AG and its homologous genes



图 3 AG 同源基因共线性分析

Fig. 3 Synteny block of AG and its homologous genes

ath: 拟南芥 *Arabidopsis thaliana*; gma: 大豆 *Glycine max*; mtr: 蒺藜苜蓿 *Medicago truncatula*



图 4 AG 同源基因系统进化树

Fig. 4 Neighbor-joining tree of AG and its homologous genes

2.3 蒺藜苜蓿 AG 同源基因表达分析

蒺藜苜蓿中共发现 4 个 AG 基因的同源基因, 分别为 Medtr3g452380 和 Medtr3g005530, 以及 2014 年报道的 AGa、AGb 基因(由于序列的同源性, 导致 AGa 和 Medtr3g452380 查询到的是同一张表达谱, 因此查询所得表达数据相同)。查询表达谱发现 4 个基因在花中均有较高的表达, 其中 AGa、AGb 和 Medtr3g452380 三个基因在整个花发育过程中表达都很高, Medtr3g005530 则主要在 12 周的花中高表达, 这与 AG 基因的功能是对应的。另一方面, 4 个基因在生殖期茎尖的表达量都略高于营养期茎尖, 但是表达量都不高。此外, 蒺藜苜蓿的 4 个基因在发育早期的果荚、根中也都有较高的表达, 而在叶、茎、营养期的茎尖中 4 个基因表达量都很低(图 5)。

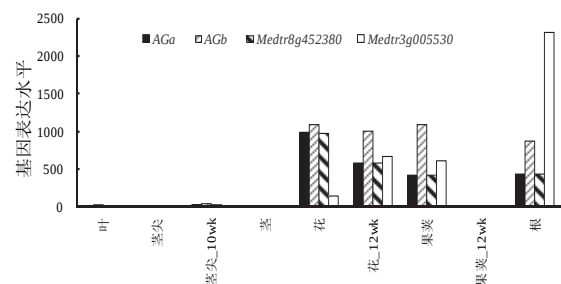


图 5 蒺藜苜蓿 AG 同源基因表达

Fig. 5 AG homologous genes expression in *Medicago truncatula*

3 讨论

筛选所得 22 个 *AG* 同源基因在 7 个代表物种中并非平均分配,大豆有 7 个,蒺藜苜蓿有 4 个,其余则只有 2~3 个同源基因。因此,可以推测 *AG* 同源基因的拷贝数一定程度上与物种基因组的大小及倍性有关。如由古四倍体演化而来的二倍体大豆基因组达到 1115 M^[21],因此含有最多的 *AG* 同源基因。而二倍体的拟南芥基因组仅 120 M,二倍体蒺藜苜蓿基因组不到 500 M,木豆基因组 833 M,鹰嘴豆 738 M,与大豆基因组的大小有着显著的差异,因此仅含有 2~4 个 *AG* 同源基因。

研究发现,*AG* 同源基因 CDS 长度及氨基酸数目保守性较高,氨基酸数目普遍集中在 250 个左右,可能是 *AG* 基因功能较为保守,需要众多保守的功能结构域。同时,22 个 *AG* 同源基因中常包含一个较长的内含子,该内含子的位置包含有 *AG* 基因的增强子,能与 *LEAFY* 等基因相互作用^[22];而且该内含子对 *AG* 基因的组织特异性表达有着极其重要的作用^[23]。

花器官发育 ABC 模型中,C 功能基因可以分为 *AG* 和 *PLE* 两个亚家族,如拟南芥的 *AGAMOUS* 和 *SHP1*、*SHP2*^[24-25]。通过进化树发现,22 个 *AG* 同源基因也明显分为两支,推测在漫长进化过程中的某一时刻 *AG* 基因由于环境或其他偶然因素发生了突变,从而产生新的拷贝,新的拷贝在随后的演化过程中逐渐使得 *AG* 基因功能分化,形成两个具有特异功能的亚家族共同承担原 C 功能基因的功能。

蒺藜苜蓿 4 个 *AG* 及其同源基因生殖期茎尖的表达量都略高于营养期茎尖,但是表达量较低。*AG* 及其同源基因在营养期是不表达的,因此营养期能观察到的表达极可能是受到其他同源性较高基因的干扰。生殖期茎尖表达量低则可能是由于取材的因素,*AG* 主要在植株进入生长期之后开始表达,并且 *AG* 基因的表达仅仅局限在发育中花原基的雄蕊原基和心皮原基中,而生殖期茎尖中包含有茎尖分生组织、叶原基等众多结构,导致 *AG* 及其同源基因表达量被稀释。另一方面,4 个基因除了在不同发育阶段的花中都高表达之外,在发育中的果荚和根中也都有表达。其中,*AGb* 基因在果荚中有着较高的表达量,*Medtr3g005530* 基因在根中有着最高的表达且其表达量随着果实发育存在明显的动态变化(数据未展示)。但是,目前还没有任何研究和证据表明 *AG* 及其同源基因在根及果荚发育过程中的作用。

AG 基因在生殖发育阶段通过控制雄蕊及心皮的发育直接参与植物的有性生殖。在花器官发育后期与 *WUSCHEL* 等基因合作共同调控花分生组织活性的维持与终止。通过对该基因功能的解析,有助于更深入地了解植物整个生长发育的过程;此外,*AG* 基因的进化过程也为认识植物进化提供依据。

参考文献:

- [1] Schwarz-Sommer Z, Huijser P, Nacken W, Saedler H, Sommer H. Genetic control of flower development by homeotic genes in *Antirrhinum majus*[J]. *Science*, 1990(250): 931—936.
- [2] Theissen G. Development of floral organ identity: stories from the MADS house[J]. *Current Opinion in Plant Biology*, 2001(4): 75—85.
- [3] Theissen G, Saedler H. Floral quartets[J]. *Nature*, 2001(409): 469—471.
- [4] Coen E S, Meyerowitz E M. The war of the wholes: genetic interactions controlling flower development[J]. *Nature*, 1991(353): 31—37.
- [5] Colombo L, Franken J, Koetje E, van Went J, Dons H J, Angenent G C, van Tunen A J. The petunia MADS-box gene FBP11 determines ovule identity[J]. *Plant Cell*, 1995(7): 1859—1868.
- [6] Theissen G, Becker A, Rosa A D, Kanno A, Kim J T, Münster T, Winter K U, Saedler H. A short history of MADS-box genes in plants[J]. *Plant Molecular Biology*, 2000(42): 115—149.
- [7] Angenent G C, Franken J, Busscher M, van Dijken A, van Went J L, Dons H J, van Tunen A J. A novel class of MADS-box genes is involved in ovule development in *Petunia*[J]. *Plant Cell*, 1995,7: 1569—1582.
- [8] Pelaz S, Ditta G S, Baumann E, Wisman E, Yanofsky M F. B and C floral organ identity functions require *SEPALLATA* MADS-box genes[J]. *Nature*, 2000(405): 200—203.

- [9] Fan H Y, Hu Y, Tudor M, Ma H. Specific interactions between the K domains of AG and AGLs members of the MADS domain family of DNA binding proteins[J]. The Plant Journal, 1997,12(5): 999 —1010.
- [10] Kramer E M, Jaramillo M A, Di Stilio V S. Patterns of gene duplication and functional evolution during the diversification of the agamous subfamily of MADS-box genes in angiosperms[J]. Genetics, 2004(166): 1011—1023.
- [11] Bowman J L, Smyth D R, Meyerowitz E M. Genetic interactions among floral homeotic genes of *Arabidopsis*[J]. Development, 1991,112: 1—20.
- [12] Jack T, Sieburth L, Meyerowitz E. Targeted misexpression of AGAMOUS in whorl 2 of *Arabidopsis* flowers[J]. Plant Journal, 1997,11: 825—839.
- [13] Sablowski R. Flowering and determinacy in *Arabidopsis*[J]. Journal of Experimental Botany, 2007,58: 899—907.
- [14] Liu, X, Kim Y J, Müller R, Yumul R E, Liu C, Pan Y, Cao X, Goodrich J, Chen X. AGAMOUS terminates floral stem cell maintenance in *Arabidopsis* by directly repressing WUSCHEL through recruitment of Polycomb Group proteins[J]. Plant Cell, 2011,23(10): 3654—3670.
- [15] Sun B, Xu Y, Ng K H, Ito T. A timing mechanism for stem cell maintenance and differentiation in the *Arabidopsis* floral meristem[J]. Genes & Development, 2009. doi: 10.1101/gad.1800409.
- [16] Jack T. Molecular and genetic mechanisms of floral control[J]. Plant Cell, 2004,16(S1): S1—17.
- [17] Bowman J L, Smyth D R, Meyerowitz E M. Genes directing flower development in *Arabidopsis*[J]. Plant Cell, 1989(1): 37—52.
- [18] Brambilla V, Battaglia R, Colombo M, Masiero S, Bencivenga S, Kater M M, Colombo L. Genetic and molecular interactions between *BELL1* and MADS-box factors support ovule development in *Arabidopsis*[J]. Plant Cell, 2007,19: 2544—2556.
- [19] Joanna S, Edelin R. Two *euAGAMOUS* genes control C-function in *Medicago truncatula*[J]. PlosOne, 2014(9): e103770.
- [20] Hecht V, Foucher F, Ferrándiz C, Macknight R, Navarro C, Morin J, Vardy M E, Ellis N, Beltrán J P, Rameau C, Weller J L. Conservation of *Arabidopsis* flowering genes in model legumes[J]. Plant Physiology, 2005,137: 1420—1434.
- [21] 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, Jackson S A. Genome sequence of the palaeopolyploid soybean[J]. Nature, 2010,463(7278): 178—183.
- [22] Busch M A, Bomblies K, Weigel D. Activation of a floral homeotic gene in *Arabidopsis*[J]. Science, 1999,285(5427): 585—587.
- [23] Yang Y, Singer S D, Liu Z. Two similar but distinct second intron fragments from tobacco AGAMOUS, homologs confer identical floral organ-specific expression sufficient for generating complete sterility in plants[J]. Planta, 2010,231(5): 1159—1169.
- [24] Pinyopich A, Ditta G S, Savidge B, Wisman E, Yanofsky M F. Assessing the redundancy of MADS-box genes during carpel and ovule development[J]. Nature, 2003,424(6944): 85—88.
- [25] Liljegren S J, Ditta G S, Eshed Y, Savidge B, Bowman J L, Yanofsky M F. SHATTERPROOF MADS-box genes control seed dispersal in *Arabidopsis*[J]. Nature, 2000,404(13): 766—770.