



大豆 B 类花器官特征基因的生物信息学分析

王曦昕^{1,2}, 蒋炳军², 袁 珊², 武婷婷², 孙 石², 张亮生³, 韩天富^{1,2}

(1. 东北农业大学 农学院, 黑龙江 哈尔滨 150030; 2. 中国农业科学院 作物科学研究所, 北京 100081; 3. 浙江大学 农业与生物技术学院, 浙江 杭州 310058)

摘要: 大豆花器官较小且闭花授粉, 导致人工杂交效率低, 品种遗传基础狭窄。MADS-box 基因家族成员 *APETALA3* (*AP3*) 和 *PISTILLATA* (*PI*) 作为花发育 ABCDE 模型中的 B 类基因, 共同控制花瓣与雄蕊的形成, 在花器官形态建成中发挥重要作用。为明确大豆中 B 类基因的进化模式, 并探究其在不同大豆品种中的遗传变异, 本研究通过生物信息学手段对大豆 B 类基因进行了基因结构、系统进化、保守基序、共线性、单倍型分析及物理性质预测等研究。结果显示: 大豆基因组中 11 个 B 类基因, 分别属于 3 个亚组, 其中 *GmAP304* (*Glyma. 04G027200*)、*GmAP306* (*Glyma. 06G027200*)、*GmAP303* (*Glyma. 03G111500*)、*GmAP316* (*Glyma. 16G105600*) 和 *GmAP312* (*Glyma. 12G118100*) 属于 *paleoAP3* 亚组, *GmTM601* (*Glyma. 01G169600*) 和 *GmTM611* (*Glyma. 11G073700*) 属于 *TM6* 亚组, *GmPII3* (*Glyma. 13G034100*)、*GmPII4* (*Glyma. 14G155100*)、*GmPI04* (*Glyma. 04G245500*) 和 *GmPI06* (*Glyma. 06G117600*) 属于 *PI* 亚组, 且各亚组 B 类基因结构相对保守; 大豆 B 类基因编码的蛋白中共有 10 个 motif, 其中 motif 1、motif 4 和 motif 5 与 MADS 结构域重叠。表达量数据分析显示, *GmAP304*、*GmAP306*、*GmPI04*、*GmPI06*、*GmPII3*、*GmPII4*、*GmTM601* 和 *GmTM611* 在大豆花中有较高表达。结合单倍型分析发现, *GmAP304*、*GmAP312*、*GmPI06*、*GmPII3* 和 *GmPII4* 在栽培大豆中只包含 1 种单倍型, 比其它大豆 B 类基因更加保守, 可能在控制花器官形成中行使更为重要的功能。本研究结果有助于完善大豆花器官发育模型, 进一步阐明 B 类基因的功能, 为大豆花器官的改造提供分子靶点, 对提高大豆育种效率、拓宽品种遗传基础具有重要意义。

关键词: 大豆; *APETALA3*; *PISTILLATA*; MADS-box; 花器官特征基因 开放科学(资源服务)标识码(OSID):



Bioinformatics Analysis of B-class Floral Organ Identity Genes in Soybean

WANG Xixin^{1,2}, JIANG Bingjun², YUAN Shan², WU Tingting², SUN Shi², ZHANG Liangsheng³, HAN Tianfu^{1,2}

(1. College of Agriculture, Northeast Agricultural University, Harbin 150030, China; 2. Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China; 3. College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China)

Abstract: Soybean has small-sized flowers and the property of self-pollination, leading to low efficiency of artificial hybridization and narrow genetic basis of varieties. *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), belonging to MADS-box gene family and class B genes in the floral ABCDE model, jointly control the formation of flower petals and stamens and play important roles in floral organ morphogenesis. In order to identify the evolutionary pattern of B-class genes and explore the genetic diversity among soybean varieties, we analyzed gene structure, phylogenetic tree, conservative motif, collinearity, haplotypes of B-class genes in soybean by bioinformatics methods. It was shown that there are 11 B-class genes in soybean genome. Of them, *GmAP304* (*Glyma. 04G027200*), *GmAP306* (*Glyma. 06G027200*), *GmAP303* (*Glyma. 03G111500*), *GmAP316* (*Glyma. 16G105600*) and *GmAP312* (*Glyma. 12G118100*) belong to the *paleoAP3* subgroup, *GmTM601* (*Glyma. 01G169600*) and *GmTM611* (*Glyma. 11G073700*) belong to the *TM6* subgroup, and *GmPII3* (*Glyma. 13G034100*), *GmPII4* (*Glyma. 14G155100*), *GmPI04* (*Glyma. 04G245500*) and *GmPI06* (*Glyma. 06G117600*) belong to the *PI* subgroup. The gene structures of B-class genes in each subgroup are relatively conservative. Ten motifs are found in the proteins encoded by B-class genes in soybean, among which motif 1, motif 4, and motif 5 overlap with the MADS domain. Transcriptome data indicated that *GmAP304*, *GmAP306*, *GmPI04*, *GmPI06*, *GmPII3*, *GmPII4*, *GmTM601* and *GmTM611* are highly expressed in flowers of soybean. Moreover, haplotype analysis showed that five B-class genes including *GmAP304*, *GmAP312*, *GmPI06*, *GmPII3* and *GmPII4* each contain only one haplotype in cultivated soybean, indicating that they are more conservative than other B-class genes, and may play more important roles in controlling floral organ formation. These results are helpful for improving the developmental model of flower organs in soybean and further elucidating the functions of B-class genes, which will facilitate the identification of molecular targets for the modification of flower organs, and will be of great significance for improving soybean breeding efficiency and broadening the genetic basis of new varieties.

Keywords: soybean; *APETALA3*; *PISTILLATA*; MADS-box; floral organ identity genes

收稿日期: 2023-09-28

基金项目: 国家重点研发计划(2023YFD1201300); 现代农业产业技术体系建设专项(CARS-04)。

第一作者: 王曦昕(1997—), 女, 硕士研究生, 主要从事大豆分子育种研究。E-mail: wangxixin717ww@163.com;

通讯作者: 蒋炳军(1977—), 男, 博士, 副研究员, 主要从事大豆分子育种研究。E-mail: jiangbingjun@caas.cn;

韩天富(1963—), 男, 博士, 研究员, 主要从事大豆遗传育种研究。E-mail: hantianfu@caas.cn。

大豆是典型的自花授粉作物,在自然条件下其蝶形花的龙骨瓣闭合,加之花器官较小,天然异交率很低^[1-5],人工去雄及授粉操作复杂费力,杂交成功率不高^[6],导致大豆品种的遗传基础狭窄^[7-8],杂种优势利用困难^[9]。

典型的被子植物的花器官可以看作由花萼、花瓣、雄蕊和心皮由外向内组成的四个同心圆,受5类花器官特征基因的共同调控。其中A类花器官特征基因控制花萼的形成;A类和B类基因共同控制花瓣的产生;B类和C类功能基因联合调控雄蕊的发育;C类功能基因单独控制心皮的形成;D类花器官特征基因与胚珠的生成有关^[10-12];E类基因为所有花器官形成所必需^[13]。在拟南芥等模式植物中,B类基因的突变将导致花瓣和雄蕊发育异常^[14-16],意味着有可能通过CRISPR/Cas9等基因编辑技术突变大豆B类基因,实现花器官改造和开放授粉,提高杂交育种效率。

B类花器官特征基因包括 *APETALA3* (*AP3*) 与 *PISTILLATA* (*PI*),属于MADS-box基因家族。*AP3/PI*的祖先基因,在被子植物多样化过程之前,经历了1次复制,形成paleoAP3谱系与PI谱系;paleoAP3谱系以核心双子叶植物为基础又复制了1次,产生了euAP3谱系与TM6系。尽管TM6系基因在许多真双子叶类群中存在,但在拟南芥中未发现代表性基因^[17]。在茄科植物(如非洲菊、矮牵牛、番茄和本氏烟草)中,*AP3*基因的复制使基因功能出现亚功能化,其中,euAP3系基因在花瓣与雄蕊的发育过程中起关键作用,TM6型基因在雄蕊发育中的功能与euAP3基因部分重叠^[18-19]。蒺藜苜蓿的2个*AP3*基因在花瓣与雄蕊中具有不重叠的互补功能,euAP3系基因 *MtNMH7* 在第二、三轮内层细胞中表达,TM6系基因 *MtTM6* 在花瓣与雄蕊的外层细胞中表达^[20]。

大豆是古四倍体,基因同源拷贝数多,四分之三的基因具有旁系同源基因,而这些同源基因中约有一半出现了亚功能化^[21-23]。已报道的大豆B类基因共有12个,包括 *GmMADS4*、*GmMADS5*、*GmMADS6*、*GmMADS7*、*GmMADS8*、*GmMADS9*、*GmMADS10*、*GmMADS110*、*GmMADS121*、*GmMADS133*、*GmMADS147* 和 *GmMADS175*^[24-25]。前人对大豆B类基因的报道多是集中在基因结构、MADS结构域中的保守基序与表达量,以及B类基因 *GmAP3* 对花期与根瘤的影响等方面^[26-27],而对B类功能基因在大豆花器官形成中的功能鲜有研究。本研究对大

豆B类花器官特征基因的系统发育、基因结构等进行了比较,并利用已公开的644份(包括栽培大豆548份,野生大豆96份)大豆全基因组重测序数据,对大豆B类花器官特征基因的单倍型进行分析,旨在为通过生物技术手段改造大豆花器官的结构,提高大豆异交率、拓宽遗传基础、推动杂种优势利用提供理论依据。

1 材料与方法

1.1 大豆B类基因的确定

根据NCBI网站(<https://www.ncbi.nlm.nih.gov/genbank/>)上拟南芥的*AP3* (NP_191002.1)和*PI* (NP_197524.1)基因的蛋白序列,在Phytozome数据库(<https://Phytozome-next.jgi.doe.gov/>)中进行序列BLAST,查找对应基因*AT3G54340*和*AT5G20240*的相关信息。通过同源蛋白搜索,将相似度在65%以上的大豆基因定义为与拟南芥B类花器官特征基因同源的大豆B类基因。

1.2 系统进化树构建

利用MEGA7软件包中的Clustal W程序,比对大豆及74个常见裸子植物及被子植物中B类功能基因的蛋白序列,之后构建Neighbor-Joining系统进化树,Bootstrap为1000。使用进化树在线美化工具iTOL(<https://itol.embl.de/>),对所构建的进化树进行美化。

1.3 基因表达热图构建

在Phytozome数据库中搜索大豆B类花器官特征基因在花器官、三出复叶、单叶、根、根瘤与种子中的表达量数据,使用软件TBtools绘制表达热图。

1.4 基因结构与保守基序分析

绘制11个大豆B类基因的外显子相对位置,使用MEME软件(<https://meme-suite.org/meme/tools/meme>)分析大豆B蛋白的motif,之后使用TBtools展示基因保守基序。

1.5 共线性分析

从EnsemblPlants(<https://plants.ensembl.org/index.html>)网站下载大豆基因组文件与注释文件,使用TBtools对大豆B类基因进行共线性分析并进行可视化。

1.6 单倍型分析

利用已公开的644份大豆材料(包括栽培大豆548份,野生大豆96份)的全基因组重测序数据(NCBI索引号:SRP020131,SRP062560,SRP045129和PRJNA589345)^[28-31],参照Yue等^[32]的方法,进

行 B 类花器官特征基因的单倍型进行分析。简言之,使用带有默认参数的 bwa v0.7.10 软件,将序列数据映射到 Williams 82 基因组 (v275)。使用 GenomeAnalysisTK(<https://gatk.broadinstitute.org/>) 的 UnifiedGenotyper (-stand_call_conf 30.0 -stand_emit_conf 10.0) 功能调用 SNPs/indels。进一步提取 B 类基因的多态性信息,选择位于 CDS 区且包含错义突变的 SNP/InDel 位点定义单倍型并进行单倍型分析^[33]。

2 结果与分析

2.1 大豆 B 类花器官特征基因的确定

以拟南芥 AP3 (AT3G54340) 和 PI (AT5G20240) 为参考基因,在 Phytozome 数据库中查找到它们在大豆中的同源基因,共 11 个(表 1)。序列比对结果表明, *Glyma. 04G027200* (命名为 *GmAP304*, 即大豆 04 号染色体上的 AP3 基因) 和 *Glyma. 06G027200* (*GmAP306*) 与拟南芥 AP3 (AT3G54340) 同源性最高,相似度分别为 78.4% 和 77.9%; *Glyma. 04G245500*

(*GmPI04*) 和 *Glyma. 06G117600* (*GmPI06*) 与拟南芥 PI (AT5G20240) 同源性最高,相似度分别为 83.7% 和 83.1%; *Glyma. 01G169600* 和 *Glyma. 11G073700* 与 AT3G54340 的相似度分别为 76.0% 与 71.9%, 在后续的进化分析中被归类于 *paleoAP3* 谱系演化后产生的 *TM6* 家系,故以 *TM6* 基因命名。

在豆科植物中,百脉根 (*Lotus japonicus*) 含有 4 个 B 类花器官特征基因,豌豆 (*Pisum sativum*) 有 2 个,紫花苜蓿 (*Medicago sativa*) 有 2 个,蒺藜苜蓿 (*Medicago truncatula*) 有 3 个,花生 (*Arachis hypogaea*) 有 8 个,菜豆 (*Phaseolus vulgaris*) 有 3 个,豇豆 (*Vigna unguiculata*) 有 3 个。与之相比,模式植物拟南芥只有 1 个 AP3 基因和 1 个 PI 基因,说明在不同物种中 B 类基因的进化存在差异,在花器官形态建成机制方面,豆科植物与模式植物拟南芥相比或更为复杂。造成豆科植物中 B 类基因数目差异的原因可能是,大豆、花生等豆科植物的染色体在进化过程中进行过加倍,其 B 类基因分化出更复杂的功能。

表 1 大豆中的 B 类花器官特征基因
Table 1 B-class floral organ identity genes in soybean

基因名称 Gene name	基因编号 Gene number	MADS-box 家族命名 Gene name in MADS-box family	基因类型 Type of B-class genes
<i>GmAP304</i>	<i>Glyma. 04G027200</i>	<i>GmMADS5</i>	<i>AP3</i>
<i>GmAP306</i>	<i>Glyma. 06G027200</i>	<i>GmMADS6</i>	<i>AP3</i>
<i>GmAP303</i>	<i>Glyma. 03G111500</i>	<i>GmMADS147</i>	<i>AP3</i>
<i>GmAP316</i>	<i>Glyma. 16G105600</i>	<i>GmMADS133</i>	<i>AP3</i>
<i>GmAP312</i>	<i>Glyma. 12G118100</i>	<i>GmMADS121</i>	<i>AP3</i>
<i>GmTM601</i>	<i>Glyma. 01G169600</i>	<i>GmMADS4</i>	<i>TM6</i>
<i>GmTM611</i>	<i>Glyma. 11G073700</i>	<i>GmMADS110</i>	<i>TM6</i>
<i>GmPI13</i>	<i>Glyma. 13G034100</i>	<i>GmMADS8</i>	<i>PI</i>
<i>GmPI14</i>	<i>Glyma. 14G155100</i>	<i>GmMADS9</i>	<i>PI</i>
<i>GmPI04</i>	<i>Glyma. 04G245500</i>	<i>GmMADS10</i>	<i>PI</i>
<i>GmPI06</i>	<i>Glyma. 06G117600</i>	<i>GmMADS7</i>	<i>PI</i>

2.2 大豆 B 类花器官特征基因的进化分析

B 类花器官特征基因的祖先基因经第一次复制,形成 *paleoAP3* 系与 *PI* 系, *paleoAP3* 系再次复制,形成 *euAP3* 系和 *TM6* 系^[17]。根据大豆 B 类基因与其他 74 个物种 B 类基因的全蛋白序列(附表 1)构建系统进化树如图 1 所示。在大豆中, *GmAP304*、*GmAP306*、*GmAP303*、*GmAP316* 和 *GmAP312* 基因均

属于 *euAP3* 家系, *GmTM601*、*GmTM611* 属于 *TM6* 家系, *GmPI04*、*GmPI06*、*GmPI13*、*GmPI14* 属于 *PI* 家系。 *euAP3* 系与 *TM6* 系亲缘关系较近。在之前对 B 类基因功能的报道中, *paleoAP3* 家系与 *TM6* 家系参与雄蕊发育, *euAP3* 家系参与花瓣与雄蕊的产生, 且 *TM6* 系与 *euAP3* 系基因可能会出现非功能化与亚功能化^[34-38]。

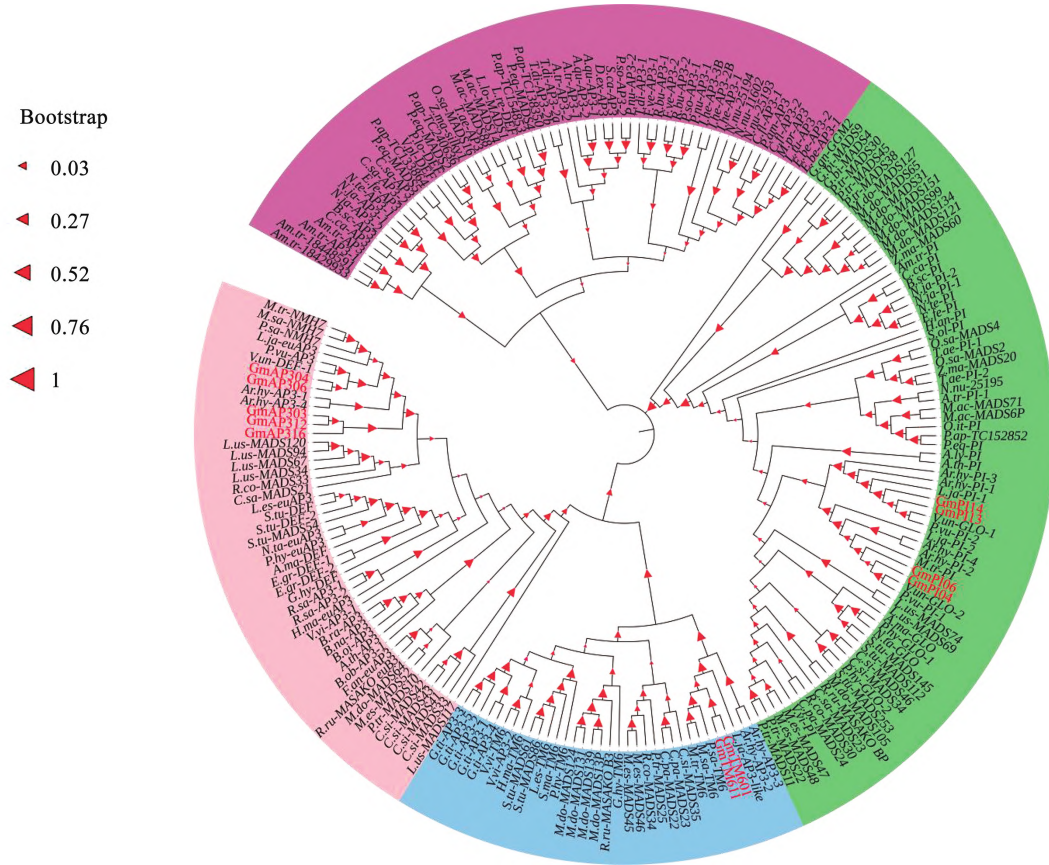


图1 74个常见物种B类花器官特征基因的系统发育树

Fig. 1 Phylogenetic tree of B-class floral organ identity genes in 74 plant species

2.3 大豆B类花器官特征基因的结构分析

大豆AP3基因 *GmAP304* 和 *GmAP306*, *TM6* 基因 *GmTM601* 和 *GmTM611*, *PI* 基因 *GmPI04*、*GmPI06*、*GmPII3* 和 *GmPII4* 均含有7个外显子,其中 *PI* 基因 *GmPI04* 包含3段3'UTR结构。*AP3* 基因 *GmAP303*、*GmAP316* 和 *GmAP312* 无3'UTR和5'UTR结构,均含

4个外显子,且外显子相对位置保守(图2)。*GmAP304* 与 *GmAP306* 的CDS序列相似度为96.49%,*GmAP303*、*GmAP316* 和 *GmAP312* 的CDS相似度高达99.42%,*GmTM601* 与 *GmTM611* 的CDS相似度为87.68%,*GmPI04* 与 *GmPI06*、*GmPII3* 与 *GmPII4* 的CDS相似度分别为96.15%和96.65%。

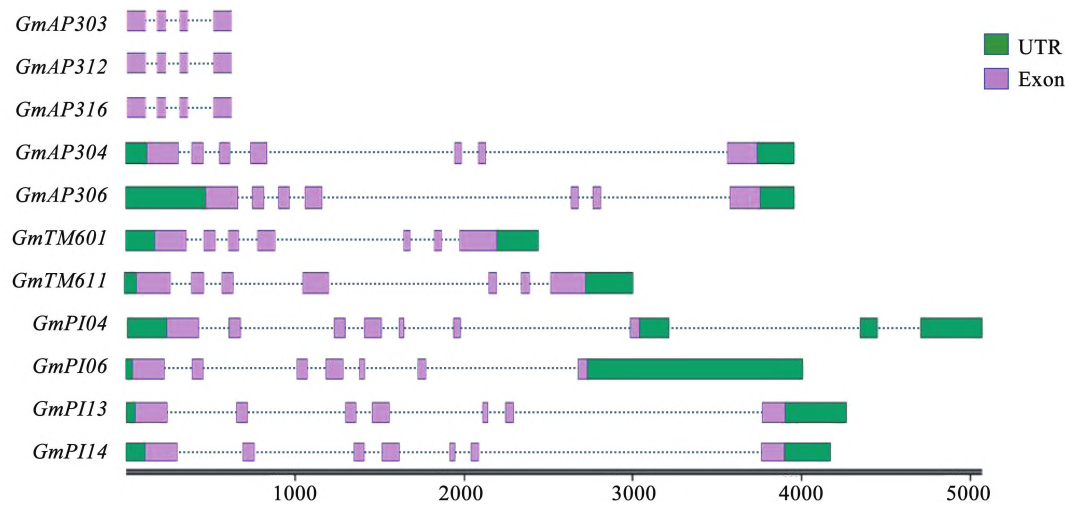


图2 大豆B类基因的结构

Fig. 2 Structure of B-class genes in soybean

使用 MEME 网站预测 11 个大豆 B 类基因编码蛋白的保守基序,发现大豆 B 类基因编码蛋白中共有 10 个 motif。其中 motif 1、motif 4 和 motif 5 与 MADS 结构域重叠,motif 2、motif 3 和 motif 7 与 K-box 结构域重叠。AP3 基因 *GmAP303*、*GmAP316* 和

GmAP312 的 MADS 结构域部分只包含 motif 1 和 motif 4,其余 B 类基因编码蛋白的 MADS 结构域部分与 motif 1、motif 4 和 motif 5 重叠(图 3)。由此推测 *GmAP303*、*GmAP316* 和 *GmAP312* 可能演化出新功能或失去功能。

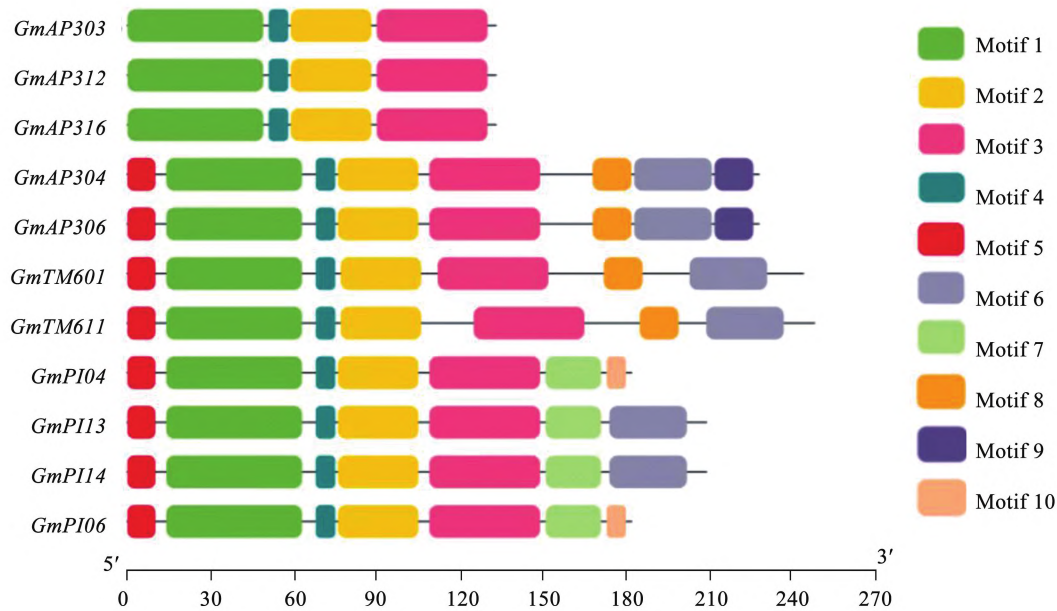


图 3 大豆 B 蛋白 motif 分析

Fig. 3 Analysis on motif of B-class protein in soybean

2.4 大豆 B 类花器官特征基因的表达分析

在 Phytosome 数据库中查找大豆品种,Williams 82 的 B 类基因在植株不同部位的表达数据,绘制热图结果如图 4 所示,在 11 个大豆 B 类花器官特征基因中,AP3 同源基因 *GmAP303*、*GmAP316* 和 *GmAP312* 在任何器官中均不表达。这可能是由于基因复制所导致的非功能化,即复制后的这 3 个 *euAP3* 家系基因被沉默。AP3 同源基因 *GmAP304* 和 *GmAP306*,PI 同源基因 *GmPI13* 和 *GmPI14* 在花器官中有较高表达量。TM6 同源基因 *GmTM601* 和 *GmTM611*,PI 同源基因 *GmPI04* 和 *GmPI06* 在未开放的花器官中有较高的表达量。对大豆 B 类基因在花器官不同部位的表达差异暂无研究。值得注意的是,*euAP3* 与 *TM6* 均是由 *paleoAP3* 复制、演化而来的,在功能分化上可能会出现非功能化与亚功能化。大豆中的 *TM6* 基因 *GmTM601* 和 *GmTM611* 在 Williams 82 花器官中有较高表达量,推测其与其他 *euAP3* 系基因 *GmAP304* 和 *GmAP306* 存在功能部分重叠,即复制后的亚功能化。

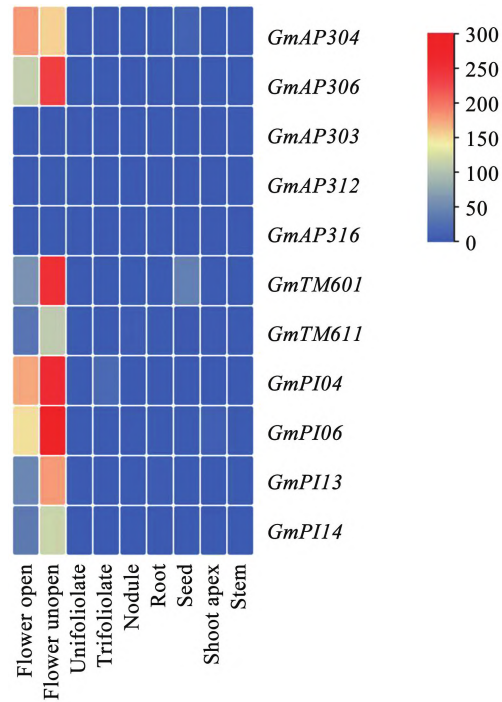


图 4 大豆 B 类基因表达量热图

Fig. 4 Expression heatmap of B-class genes in soybean

2.5 大豆 B 类花器官特征基因的共线性分析

大豆的 11 个 B 类花器官特征基因分别分布在 9 条染色体上,其中,4 号、6 号染色体上各含有 2 个 B 类基因,1、3、11、12、13、14、16 号染色体上各有 1 个 B 类基因,2、5、7、8、9、10、15、17、18、19、20 号染色体上未发现 B 类基因。由于 3 个 AP3 基因可能不具有功能,故本研究只对其它 8 个功能性 B 类基因进行共线性分析,结果如图 5 所示,*GmTM611* 分

别与 *GmAP306*、*GmAP304* 和 *GmTM601* 存在大片段重复,*GmPII3* 分别与 *GmPI06*、*GmPI04* 和 *GmPII4* 存在大片段重复,*GmPII4* 分别与 *GmPI04* 和 *GmPI06* 存在大片段重复,*GmTM601* 分别与 *GmPI06* 和 *GmAP304* 存在大片段重复,*GmAP304* 与 *GmAP306* 存在大片段重复,*GmPI04* 与 *GmPI06* 存在大片段重复。上述结果说明大豆 B 类基因的数量扩增主要是由于片段重复造成的。

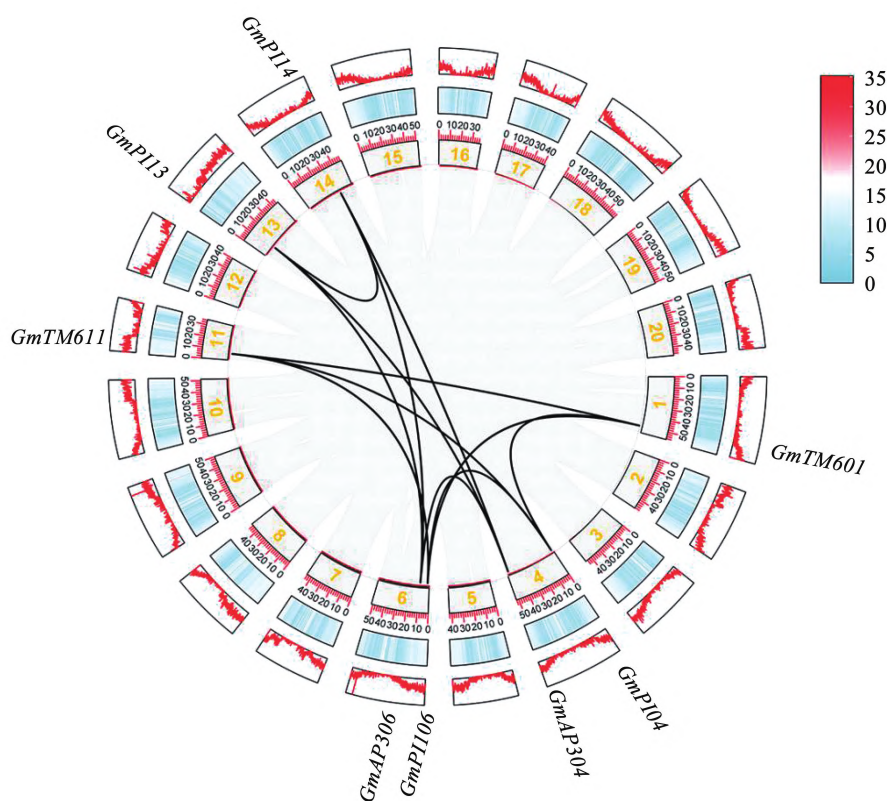


图 5 大豆 B 类基因共线性分析

Fig. 5 Collinearity of B-class genes in soybean

2.6 大豆 B 类花器官特征基因的单倍型分析

利用 644 份大豆材料的基因组重测序数据,对 B 类花器官特征基因的单倍型和分布频率进行分析,结果如图 6 所示:AP3 同源基因 *GmAP304*、*GmAP306* 和 *GmAP303* 分别有 4、14 和 11 种单倍型,*GmAP312* 和 *GmAP316* 分别有 6 和 5 种单倍型;TM6 同源基因 *GmTM601* 和 *GmTM611* 分别有 3 和 5 种单倍型;PI 同源基因 *GmPI04*、*GmPI06*、*GmPII3* 和 *GmPII4* 分别有 4、6、8 和 4 种单倍型。

AP3 同源基因的单倍型:*GmAP304* 在栽培大豆中只包含 *GmAP304*^{H01} 单倍型,在野生大豆中则包含 *GmAP304*^{H01}、*GmAP304*^{H02}、*GmAP304*^{H03} 和 *GmAP304*^{H04} 等 4 种单倍型;*GmAP306* 在栽培大豆中含有 *GmAP306*^{H01} 和 *GmAP306*^{H02} 两种单倍型,在野生大豆

中主要含有 *GmAP306*^{H01}、*GmAP306*^{H02}、*GmAP306*^{H03}、*GmAP306*^{H04} 和 *GmAP306*^{H05} 等 5 种单倍型;在 *GmAP303* 的众多单倍型中,*GmAP303*^{H01} 在栽培大豆中占比最高,野生大豆中主要包含 *GmAP303*^{H01}、*GmAP303*^{H03} 和 *GmAP303*^{H05};在栽培大豆与野生大豆中,*GmAP312*^{H01} 均为占比最高的 *GmAP312* 单倍型;*GmAP316* 在栽培种中主要包含 *GmAP316*^{H01}、*GmAP316*^{H02} 和 *GmAP316*^{H04} 等单倍型,在野生大豆中主要包括 *GmAP316*^{H01}、*GmAP316*^{H02}、*GmAP316*^{H03} 和 *GmAP316*^{H04} 等 4 种单倍型(图 7)。

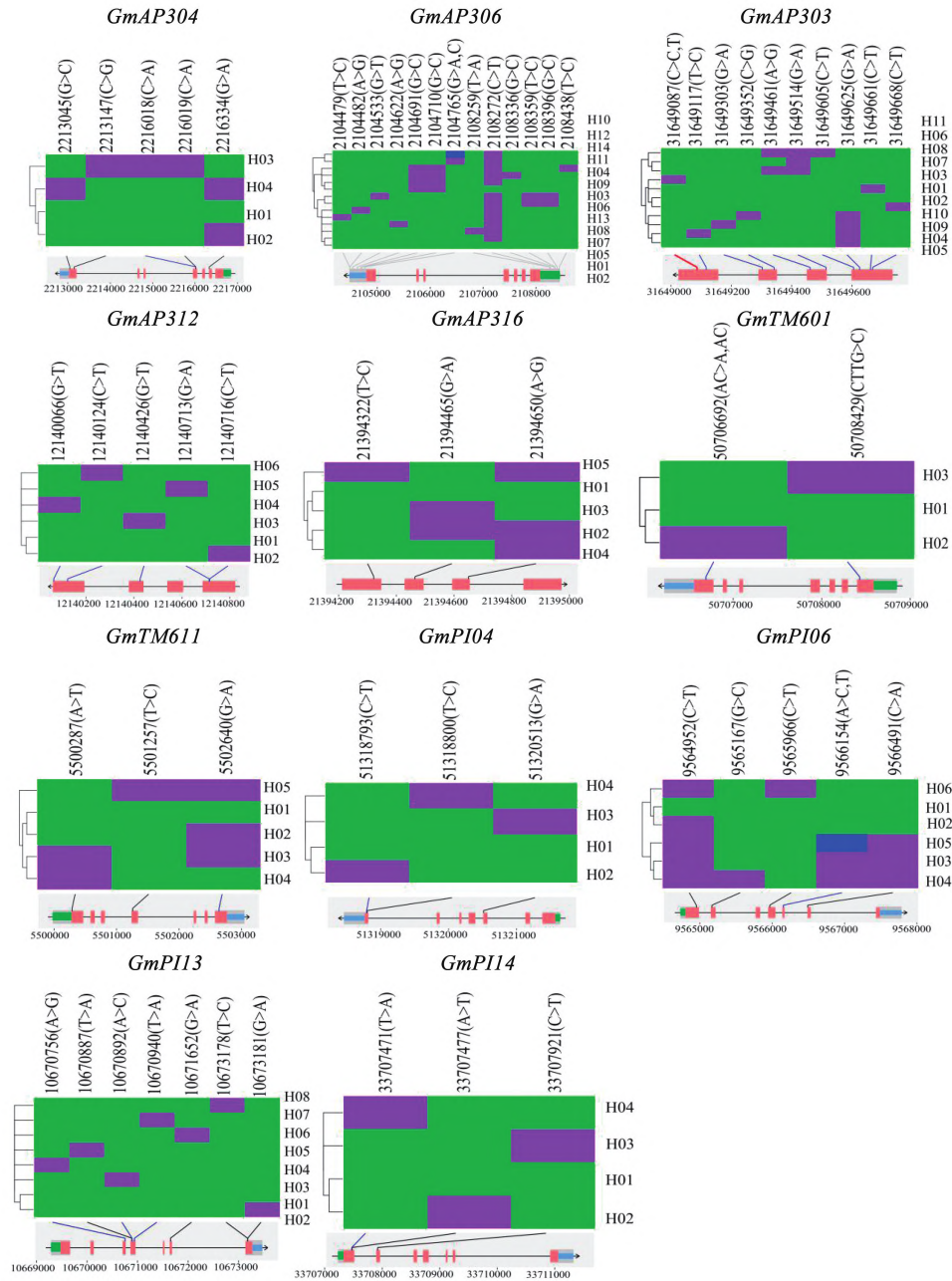
TM6 同源基因的单倍型:*GmTM601*^{H01} 和 *GmTM601*^{H02} 在栽培大豆与野生大豆中均为主要单倍型;*GmTM611* 在栽培大豆中的单倍型为 *GmTM611*^{H01} 和 *GmTM611*^{H02},在野生大豆中的单倍型为 *GmTM611*^{H01}、

GmTM611^{H02}、*GmTM611*^{H03}和 *GmTM611*^{H04} (图 7)。

PI 同源基因的单倍型:栽培大豆中主要包含 *GmPI04*^{H01} 和 *GmPI04*^{H02}, 野生大豆主要包含 *GmPI04*^{H01}; *GmPI06* 同源基因在栽培大豆中的单倍型主要为 *GmPI06*^{H01}, 而在野生大豆中有 *GmPI06*^{H01}、*GmPI06*^{H02}、*GmPI06*^{H03} 和 *GmPI06*^{H04} 等 4 种单倍型; 在栽培大豆与野生大豆中, *GmPI13* 基因的主要单倍

型均为 *GmPI13*^{H01}, *GmPI14* 基因的主要的单倍型均为 *GmPI14*^{H01} (图 7)。

在栽培大豆中, *AP3* 同源基因 *GmAP304* 和 *GmAP312*, *PI* 基因 *GmPI06*、*GmPI13* 和 *GmPI14* 均只包含 1 种单倍型; 在野生大豆中, *GmPI14* 也只包含 1 种单倍型。说明这些基因的保守性很强, 推测或在大豆花器官发育中行使重要功能。



注:绿色表示与参考基因组一致的等位变异,紫色和蓝色表示与参考基因组不同的等位变异。

Note: The green block represents alleles that are consistent with the reference genome, while purple and blue blocks represent alleles that are different from the reference genome.

图 6 大豆 B 类花器官特征基因单倍型的分类

Fig. 6 Haplotypes of B-class genes in soybean

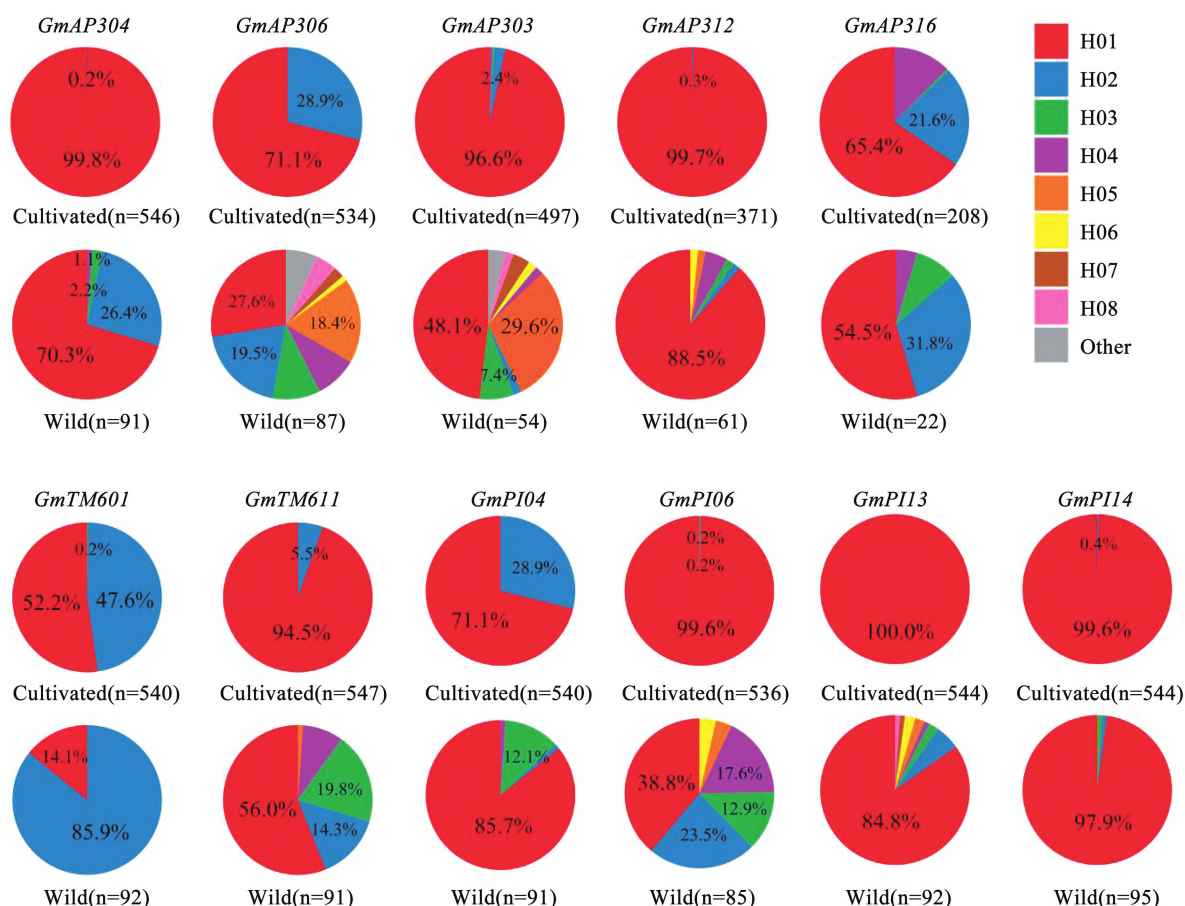


图7 大豆 B 类基因不同单倍型在野生种与栽培种中分布频率的比较

Fig. 7 Comparison of haplotype distributional frequency of B-class genes between the cultivated and wild soybeans

3 讨论

花器官特征基因控制花器官不同结构的产生,其中 B 类基因(*AP3/PI*)与花瓣和雄蕊的生成有关。*AP3/PI* 基因作为编码 MADS-box 蛋白的转录因子,在植物信号转导与生长发育过程中发挥重要作用,尤其是 B 类花器官特征基因在花器官形态发生中对雄蕊和花瓣的作用,具有重要的研究和利用价值。在水稻中,*PI* 同源基因 *OsMADS2* 和 *OsMADS4* 被沉默后,会导致其雄蕊组织转化为心皮样器官^[39]; *cfol-3* (*PI* 基因)突变体会产生小叶瓣特性丧失与雄蕊数量明显减少的缺陷花^[40]。

在以前的研究中,发现大豆中存在 12 个 B 类基因^[24-25]。本研究根据 *GmMADS175* 蛋白 (KAG5054864.1) 序列在 Phytozome 数据库中的 Blast 结果 (*Glyma.03G111400*),发现基因注释为 X-box 相关转录因子,不属于 B 类基因,故未将其列入研究范围。前人把已知的大豆 B 类基因分成 *AP3* 和 *PI* 两个亚组^[25],本研究根据系统进化树结构,将

B 类基因进行了细分,划分为 3 个亚组,其中, *GmAP304*、*GmAP306*、*GmAP303* 和 *GmAP312*、*GmAP316* 属于同一组,即 *euAP3* 系; *GmTM601* 和 *GmTM611* 属于 *TM6* 系; *GmPI04*、*GmPI06*、*GmPI13* 和 *GmPI14* 属于 *PI* 系。*euAP3* 系与 *TM6* 系亲缘关系较近,推测 *GmAP304*、*GmAP306*、*GmAP303*、*GmAP312*、*GmAP316*、*GmTM601* 和 *GmTM611* 基因具有较相似的功能。大豆中 B 类花器官特征基因的数量明显多于模式植物拟南芥,这可能是由于大豆为四倍体,多经历了一次染色体复制,导致大豆中 B 类基因的功能与作用机制较拟南芥中复杂。基因结构分析表明,绝大多数大豆 B 类花器官特征基因都含 7 个外显子与 6 个内含子。大豆 B 蛋白共含有 10 个 motif,其中 *GmAP303*、*GmAP312* 和 *GmAP316* 只含有 motif 1~4, *GmAP304* 和 *GmAP306* 不含 motif 7 和 motif 10, *GmTM601* 和 *GmTM611* 不含 motif 7、motif 9 和 motif 10, *GmPI04* 和 *GmPI06* 不含 motif 6、motif 8 和 motif 9, *GmPI13* 和 *GmPI14* 不含 motif 8~10。其中 motif 1、motif 4 和 motif 5 与 B 类基因的

MADS 结构域重叠。本研究将 B 类基因的 MADS 结构域做了细致的划分,进一步明确了大豆 B 基因在基因结构上的差异,为对该类基因进行遗传操作提供了有价值的信息。

对 Phytozome 数据库中相关基因表达量的分析发现,*GmAP303*、*GmAP312* 和 *GmAP316* 在大豆的任何部位均不表达,推测这可能是由复制后基因沉默所致,其余 8 个 B 类基因均在花器官中有较高表达。单倍型分析结构表明,在大面积种植品种中,*GmAP304*、*GmAP312*、*GmPI06*、*GmPI13* 和 *GmPI14* 几乎只包含 1 种单倍型,说明这 5 个 B 类基因具有很强的保守性,可作为进一步研究和利用的重点基因。如结合第三代基因编辑技术 CRISPR/CAS9,可获得敲除 B 类基因的突变体,根据 B 类功能基因对大豆花瓣及雄蕊产生的影响,将有可能得到大豆花瓣与雄蕊缺失的开放授粉新种质。创制和应用 B 基因缺失突变体,可望代替传统不育系,解决人工杂交成本过高、效率低下的问题,为提升大豆天然异交率、拓宽大豆遗传基础提供理论、技术和材料支撑。

4 结论

本研究探究了大豆中的 11 个 B 类花器官特征基因,通过系统发育树构建将这 11 个基因分为 *euAP3*、*TM6*、*PI* 系 3 组。上述基因不均匀地分布在大豆 9 条染色体上,大多数在花器官中有较高表达,但 *GmAP303*、*GmAP312* 和 *GmAP316* 例外,在大豆的任何部位均未见表达。单倍型分析表明,*GmAP304*、*GmAP312*、*GmPI06*、*GmPI13* 和 *GmPI14* 保守性较强,可作为大豆花器官结构遗传改良的重要分子靶点。

参考文献

- [1] YOSHIMURA Y. Wind tunnel and field assessment of pollen dispersal in soybean [*Glycine max* (L.) Merr] [J]. *Journal of Plant Research*, 2011, 124(1): 109-114.
- [2] 闫昊, 张井勇, 彭宝, 等. 大豆异交率性状研究概述 [J]. *安徽农业科学*, 2017, 45(36): 14-16. (YAN H, ZHANG J Y, PENG B, et al. Overview of outcrossing rate traits in soybeans [J]. *Journal of Anhui Agricultural Sciences*, 2017, 45(36): 14-16.)
- [3] BEARD B H, KNOWLES P F. Frequency of cross-pollination of soybeans after seed irradiation [J]. *Crop Science*, 1971, 11(4): 489-492.
- [4] JARNE P, DAVID P. Quantifying inbreeding in natural populations of hermaphroditic organisms [J]. *Heredity*, 2008, 100(4): 431-439.
- [5] WEBER C R, HANSON W D. Natural hybridization with and without ionizing radiation in soybeans [J]. *Crop Science*, 1961, 1(6): 389-392.
- [6] 李继存, 黄新阳, 王妙, 等. 大豆杂交中伪杂种产生的原因分析 [J]. *大豆科学*, 2012, 31(3): 492-494. (LI J C, HUANG X Y, WANG M, et al. Analysis on reasons for false hybrid seeds of soybean [J]. *Soybean Science*, 2012, 31(3): 492-494.)
- [7] 曹永强, 宋书宏, 王文斌, 等. 拓宽大豆育种遗传基础研究进展 [J]. *辽宁农业科学*, 2005(6): 34-36. (CAO Y Q, SONG S H, WANG W B, et al. Research advance on broadening genetic base of soybean breeding [J]. *Liaoning Agricultural Sciences*, 2005(6): 34-36.)
- [8] HUANG W, HOU J, HU Q, et al. Pedigree-based genetic dissection of quantitative loci for seed quality and yield characters in improved soybean [J]. *Molecular Breeding*, 2021, 41(2): 14.
- [9] 谢甫锦. 大豆雄性不育及杂种优势利用研究进展 [J]. *沈阳农业大学学报*, 2008, 39(2): 131-136. (XIE F T. Review of soybean male sterility and heterosis utilization [J]. *Journal of Shenyang Agricultural University*, 2008, 39(2): 131-136.)
- [10] KAY S. The molecular and genetic control of ovule development [J]. *Current Opinion in Plant Biology*, 1999, 2(1): 13-17.
- [11] ANGENENT G C, FRANKEN J, BUSSCHER M, et al. A novel class of MADS box genes is involved in ovule development in petunia [J]. *The Plant Cell*, 1995, 7(10): 1569-1582.
- [12] FANG Z W, QI R, LI X F, et al. Ectopic expression of *FaesAP3*, a *Fagopyrum esculentum* (Polygonaceae) *AP3* orthologous gene rescues stamen development in an *Arabidopsis ap3* mutant [J]. *Gene*, 2014, 550(2): 200-206.
- [13] THEIBEN G, MELZER R, RÜMLER F. MADS-domain transcription factors and the floral quartet model of flower development: Linking plant development and evolution [J]. *Development*, 2016, 143(18): 3259-3271.
- [14] 丛楠, 程治军, 万建民. 控制花器官发育的 ABCDE 模型 [J]. *中国农学通报*, 2007, 23(7): 124-128. (CONG N, CHENG Z J, WAN J M. The ABCDE model of floral organ development [J]. *Chinese Agricultural Science Bulletin*, 2007, 23(7): 124-128.)
- [15] 黄方, 迟英俊, 喻德跃. 植物 MADS-box 基因研究进展 [J]. *南京农业大学学报*, 2012, 35(5): 9-18. (HUANG F, CHI Y J, YU D Y. Research advances of MADS-box genes in plants [J]. *Journal of Nanjing Agricultural University*, 2012, 35(5): 9-18.)
- [16] HIBINO Y, KITAHARA K, HIRAI S, et al. Structural and functional analysis of rose class B MADS-box genes 'MASAKO BP', euB3, and B3': Paleo-type *AP3* homologue 'MASAKO B3' association with petal development [J]. *Plant Science*, 2006, 170(4): 778-785.
- [17] KRAMER E M, DORIT R L, IRISH V F. Molecular evolution of genes controlling petal and stamen development: Duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages [J]. *Genetics*, 1998, 149(2): 765-783.
- [18] LIU Y, NAKAYAMA N, SCHIFF M, et al. Virus induced gene silencing of a DEFICIENS ortholog in *Nicotiana benthamiana* [J].

- Plant Molecular Biology, 2004, 54(5): 701-711.
- [19] VANDENBUSSCHE M, ZETHOF J, ROYAERT S, et al. The duplicated B-class heterodimer model: Whorl-specific effects and complex genetic interactions in *Petunia* hybrida flower development [J]. The Plant Cell, 2004, 16(3): 741-754.
- [20] ROQUE E, SERWATOWSKA J, CRUZ ROCHINA M, et al. Functional specialization of duplicated AP3-like genes in *Medicago truncatula* [J]. The Plant Journal: for Cell and Molecular Biology, 2013, 73(4): 663-675.
- [21] SEVERIN A J, CANNON S B, GRAHAM M M, et al. Changes in twelve homoeologous genomic regions in soybean following three rounds of polyploidy [J]. The Plant Cell, 2011, 23(9): 3129-3136.
- [22] SCHMUTZ J, CANNON S B, SCHLUETER J, et al. Genome sequence of the palaeopolyploid soybean [J]. Nature, 2010, 463: 178-183.
- [23] ROULIN A, AUER P L, LIBAULT M, et al. The fate of duplicated genes in a polyploid plant genome [J]. The Plant Journal: for Cell and Molecular Biology, 2013, 73(1): 143-153.
- [24] 胡瑞波, 范成明, 李宏宇, 等. 大豆 MIKC 型 MADS-box 基因家族分析 [J]. 分子植物育种, 2009, 7(3): 429-436. (HU R B, FAN C M, LI H Y, et al. Analysis of MIKC-type MADS-box genes in soybean (*Glycine max*) [J]. Molecular Plant Breeding, 2009, 7(3): 429-436.)
- [25] SHEN G, JIA Y, WANG W L. Evolutionary divergence of motifs in B-class MADS-box proteins of seed plants [J]. Journal of Biological Research, 2021, 28(1): 12.
- [26] ZHANG A, HE H, LI Y, et al. MADS-box subfamily gene *GmAP3* from *Glycine max* regulates early flowering and flower development [J]. International Journal of Molecular Sciences, 2023, 24(3): 2751.
- [27] MA W Y, LIU W, HOU W S, et al. GmNMH7, a MADS-box transcription factor, inhibits root development and nodulation of soybean (*Glycine max* [L.] Merr.) [J]. Journal of Integrative Agriculture, 2019, 18(3): 553-562.
- [28] LAM H M, XU X, LIU X, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection [J]. Nature Genetics, 2010, 42(12): 1053-1059.
- [29] ZHOU Z, JIANG Y, WANG Z, et al. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean [J]. Nature Biotechnology, 2015, 33: 408-414.
- [30] QI X, JIANG B, WU T, et al. Genomic dissection of widely planted soybean cultivars leads to a new breeding strategy of crops in the post-genomic era [J]. The Crop Journal, 2021, 9(5): 1079-1087.
- [31] ZHANG T, WU T, WANG L, et al. A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content [J]. International Journal of Molecular Sciences, 2019, 20(23): 5915.
- [32] YUE Y, SUN S, LI J, et al. GmFULa improves soybean yield by enhancing carbon assimilation without altering flowering time or maturity [J]. Plant Cell Reports, 2021, 40(10): 1875-1888.
- [33] JIANG B, ZHANG S, SONG W, et al. Natural variations of FT family genes in soybean varieties covering a wide range of maturity groups [J]. BMC Genomics, 2019, 20(1): 230.
- [34] DUARTE J M, CUI L, WALL P K, et al. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis* [J]. Molecular Biology and Evolution, 2006, 23(2): 469-478.
- [35] MOORE R C, PURUGGANAN M D. The evolutionary dynamics of plant duplicate genes [J]. Current Opinion in Plant Biology, 2005, 8(2): 122-128.
- [36] KIM S, YOO M J, ALBERT V A, et al. Phylogeny and diversification of B-function MADS-box genes in angiosperms: Evolutionary and functional implications of a 260-million-year-old duplication [J]. American Journal of Botany, 2004, 91(12): 2102-2118.
- [37] VANDENBUSSCHE M, THEISSEN G, VAN DE PEER Y, et al. Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutations [J]. Nucleic Acids Research, 2003, 31(15): 4401-4409.
- [38] 黄和喜, 王岩, 许玉超, 等. 不结球白菜雄蕊瓣化相关 *AP3* 基因的克隆和表达分析 [J]. 南京农业大学学报, 2015, 38(5): 748-756. (HUANG H X, WANG Y, XU Y C, et al. Clone and expression analysis of stamen patelody-associated *AP3* genes from non-heading Chinese cabbage [J]. Journal of Nanjing Agricultural University, 2015, 38(5): 748-756.)
- [39] YAO S G, OHMORI S, KIMIZU M, et al. Unequal genetic redundancy of rice *PISTILLATA* orthologs, *OsMADS2* and *OsMADS4*, in lodicule and stamen development [J]. Plant and Cell Physiology, 2008, 49(5): 853-857.
- [40] WANG H, ZHANG L, CAI Q, et al. *OsMADS32* interacts with PI-like proteins and regulates rice flower development [J]. Journal of Integrative Plant Biology, 2015, 57(5): 504-513.