



# 大豆 phyA 下游基因预测和表达分析及敲除载体构建

张 婷,马丽欣,刘 俊,林晓雅,刘宝辉

(广州大学 分子遗传与进化研究中心/广州市重点实验室,广东 广州 510006)

**摘 要:** 为了进一步研究大豆光周期调控通路中 *E3* 和 *E4* 的下游基因,为获得其突变体植株奠定基础,以拟南芥远红光响应受体 phyA 下游的重要信号传递因子 *PIF3*、*LAF1*、*FHY1* 和 *FHL* 的序列作为参考序列,在 Phytozome 12 数据库中查找其相似序列,进行序列比对和进化树分析,确定它们在大豆中的同源基因并采用 qRT-PCR 方法进行组织表达分析,使用 CRISPR/Cas 9 系统设计同源基因的敲除靶点并通过发根系统验证靶点的有效性。结果显示:*AtPIF3* 在大豆中有 6 个同源基因,*AtLAF1* 在大豆中有 4 个同源基因,*AtFHL* 在大豆中有 2 个同源基因,而 *AtFHY1* 在大豆中没有同源基因。4 个 *LAF1* 基因主要在顶端生长点表达,6 个 *PIF3* 和 2 个 *FHL* 基因主要在荚中表达。共鉴定出 12 个有效靶点,能够分别将大豆中 6 个 *PIF3*、4 个 *LAF1* 和 2 个 *FHL* 基因成功敲除。研究结果为进一步获得稳定的大豆突变体材料和大豆 phyA 下游基因的功能研究提供了理论基础。

**关键词:** 大豆;*PIF3*;*LAF1*;*FHY1*;*FHL*;组织表达分析;CRISPR/Cas9;基因敲除载体

## Soybean phyA Downstream Gene Prediction and Expression Analysis and Its Knockout Vector Construction

ZHANG Ting, MA Li-xin, LIU Jun, LIN Xiao-ya, LIU Bao-hui

(Innovation Research Center of Molecular Genetics and Evolution, Guangzhou University/Guangzhou Key Laboratory, Guangzhou 510006, China)

**Abstract:** To further study the downstream genes of *E3* and *E4* in the photoperiod regulatory pathway of soybean, and lay a foundation for obtaining its mutant soybean, we focused on *PIF3*, *LAF1*, and *FHY1/FHL*, which are important signaling factors downstream of *Arabidopsis* far-red light receptor phyA. We used genome sequences of *Arabidopsis PIF3*, *LAF1* and *FHY1/FHL* as references for searching soybean homologs in Phytozome 12 database, made sequence alignment and phylogenetic tree to analyze the homologs in soybean, and performed real-time quantitative PCR to detect the expression patterns of these soybean homologs in various tissues. We applied CRISPR/Cas 9 system to knockout these homologous genes, and verified the effectiveness of each target through the soybean hair root system. The results showed that: *AtPIF3*, *AtLAF1*, and *AtFHL* has six, four, and two homologs in soybean, respectively. And *AtFHY1* doesn't have homolog in soybean. Four *LAF1* homologs were mainly expressed in shoot apical meristem, six *PIF3* homologs and two *FHL* homologs were mainly expressed in pods. According to the CRISPR/Cas9 knockout results, a total of twelve effective targets were identified, which can successfully knock out six *PIF3*, four *LAF1*, and two *FHL* homologs in soybean. This research provides a theoretical basis for obtaining stable soybean knockout mutant, which paved a way for studying the functions of soybean phyA downstream genes.

**Keywords:** Soybean; *PIF3*; *LAF1*; *FHY1*; *FHL*; Tissue expression analysis; CRISPR/Cas 9; Gene knockout vector

光对植物的生长和发育具有重要作用,除了为植物提供主要能源外,还控制着植物生命周期中多个发育过程,包括种子萌发、幼苗去黄化、子叶扩张、气孔闭合、叶绿体运动、庇荫反应、昼夜节律和开花时间等<sup>[1-2]</sup>。植物体内有响应不同光质的光受体,这些光受体通过响应不同的光信号从而进行信号转导,调节下游基因的表达,并最终使整个植物的形态产生适应性变化<sup>[3]</sup>。植物体内的光信号受

体包括光敏色素、蓝光受体、向光素等<sup>[2]</sup>。光敏色素主要感受光波长为 600 ~ 730 nm,其中 phytochrome A (phyA)、phytochrome B (phyB) 两个光受体起主要作用,而 phyA 是唯一可以感受远红光的受体,可以调控植物在远红光条件下的光形态建成。

*FHY1/FHL*、*PIF3* 和 *LAF1* 在模式植物拟南芥 phyA 通路中发挥着重要的作用。拟南芥光敏色素

收稿日期:2020-12-22

基金项目:广州市科技计划(202002030180)。

第一作者:张婷(1992—),女,在读硕士,主要从事大豆光敏色素 A 基因功能研究。E-mail:809165378@qq.com。

通讯作者:刘宝辉(1964—),男,博士,教授,主要从事大豆光周期开花机制基础性研究。E-mail:liubh@gzhu.edu.cn。

受体有 5 个,分别是 phyA ~ E,其中 phyA 是植物体内唯一响应远红光的受体<sup>[4]</sup>。phyA 对光强的不同反应可分为 3 个等级: VLFR ( Very-Low-Fluence Response)、LFR ( Low-Fluence Response ) 和 HIR ( High-Irradiance Responses )<sup>[3]</sup>。phyA 可以响应 VLFR 和 HIR 反应,即在远红光照射下,phyA 由 Pr (非活性形式) 转化为 Pfr (活性形式),并由重要的 phyA 的结合蛋白 FHY1/FHL ( Far-Red Elongated Hypocotyl1/FHY1-Like) 协助 phyA 进入细胞核。活性形式的 phyA 进入到细胞核之后一方面可以抑制 PIF3 (Phytochrome Interacting Factor 3) 的表达从而促进下游基因的转录表达,进而响应 phyA 控制的 HIR 和 VLFR 反应<sup>[3]</sup>;另一方面活性形式的 phyA 进入细胞核后通过抑制 COP1 而促进 LAF1 (Long After Far-Red Light 1) 的表达从而促进 HIR 和 VLFR 反应。FHY1 和 FHL 是两个小蛋白,分别包含 202 和 201 个氨基酸,且在单子叶植物和双子叶植物中均具有同源性,每个蛋白包含 1 个入核信号 (NLS) 和 1 个出核信号 (NES),能够协助并结合 phyA 形成复合物进入核内发挥作用<sup>[3-5]</sup>。HY1/FHL 能够和两个特别的转录因子 HFR1 (Long Hypocotyl in Far-Red 1) 和 LAF1 相互作用,在 phyA 下游进行独立的信号传导,HFR1 还能与 LAF1 发生相互作用,通过抑制 COP1 的泛素化而相互稳定,从而增强 phyA 的光响应<sup>[6-7]</sup>。PIFs 蛋白属于拟南芥中 bHLH 转录因子超家族的 15 个亚家族成员之一,PIF3 是 PIFs 家族成员中的基础成员,拟南芥中一共有 8 个 PIFs 基因,但是只有 PIF1 和 PIF3 能够和 phyA 发生相互作用,因此其与 phyA 互作的结构域被称为 APA 结构域<sup>[8-10]</sup>。

大豆是一种古四倍体的短日照作物,是人类植物蛋白的重要来源。大豆在全世界范围被种植,因此不同地区控制大豆光周期开花和成熟的数量性状基因座有所不同,目前已经发现的大豆中的主要控制开花和成熟的 QTL 有 *E1* ~ *E4*、*E6* ~ *E10* 和 *J*<sup>[11]</sup>。其中,*E3* 和 *E4* 被证实是拟南芥中 *PHYA* 的同源基因<sup>[12]</sup>,研究表明,*E3* 和 *E4* 基因是大豆开花期前后控制光周期反应的主要因子,对它们的研究可以更好地解释大豆中光周期不敏感的遗传多样性和对开花后光周期的响应机制<sup>[11]</sup>。初期研究表明,*E3* 和 *E4* 虽然主要在光周期调控开花中起重要作用,但对该基因的研究显示,*E4* 可能参与了部分光形态建成,而 *E3* 的表现更弱一些<sup>[11,13]</sup>,甚至可能

不参与远红光形态建成,而 *E3* 更主要的功能是参与大豆开花期的调控<sup>[11]</sup>。大豆基因组在进化过程中发生过两次复制,而导致产生多个等位基因,使得对大豆 *E3* 和 *E4* 基因的研究较为缓慢。因此,虽然在大豆中 *E3* 和 *E4* 是重要的影响开花的生育期基因,但是其下游基因对大豆开花的影响尚且知之甚少。另一方面大豆的转化周期相对较长,非常需要设计用于根毛检测的有效靶点从而保证转化的高效性。而且对于多倍体植物而言,基因组中的同源基因大多有功能冗余情况,所以需要找到其所有的同源基因,并同时基因敲除才能有效避免功能冗余情况。

本研究通过查找大豆中与拟南芥 *PIF3*、*LAF1* 和 *FHY1/FHL* 的相似基因,进行序列进化树分析,确定它们在大豆中的同源基因,检测这些基因的组织特异性表达情况,采用高效准确的 CRISPR/Cas 9 技术构建大豆 *PIF3*、*LAF1* 和 *FHY1/FHL* 基因敲除载体,并在大豆发根系统中检验 phyA 下游基因敲除靶点的有效性。本研究以大豆中的 phyA 下游基因的鉴定和敲除为基础,旨在为今后研究大豆的 phyA 下游基因对花期的影响提供参考,也为培育适应不同纬度大豆新品种奠定理论基础。

### 1 材料与方法

#### 1.1 材料

大豆发根转化材料为大豆品种 Williams 82,由本实验室保存。大肠杆菌菌株 DH5 $\alpha$ 、发根农杆菌菌株 K599,由本实验室保存。用于基因敲除载体构建的 AtU3d、AtU3b、AtU6-1、AtU6-29、Cas9 载体,由华南农业大学刘耀光院士提供。

#### 1.2 方法

##### 1.2.1 大豆中 phyA 下游基因的鉴定和进化树分析

分别利用拟南芥 phyA 下游 4 个基因 *AtPIF3* ( AT5G59240 )、*AtLAF1* ( AT4G25560 )、*AtFHY1* ( AT1G79790 ) 和 *AtFHL* ( AT5G02200 ) (拟南芥基因号参考网站: <https://www.arabidopsis.org/>) 的蛋白序列在 Phytozome 12 ( <https://phytozome.jgi.doe.gov/pz/portal.html> ) 数据库进行 BLAST,根据注释搜索并下载与拟南芥 phyA 下游基因 *AtPIF3*、*AtLAF1*、*AtFHY1* 和 *AtFHL* 同源的大豆基因序列。使用 MEGA7 软件,利用 NJ 方法建立蛋白序列进化树,并根据 PIFs 家族保守的 bHLH 结构域,进一步确认大豆 *PIF3* 基因。利用 SMART 网站 ( <http://smart.embl-heidelberg.de/> ) 进行结构预测,分析和绘制

LAF1 和 PIF3 蛋白结构图。

1. 2. 2  组织特异性表达分析  将 Williams 82 大豆种子播种于钵中(土:蛭石 = 2:1),每钵 3 粒,在温度 25 ℃,光照 16 h/8 h(光照/黑暗)的温室内培养。在大豆 V1 期时取大豆的种子、根、茎、初生叶、侧生分生组织和生长点于液氮中速冻。R1 期时取大豆的花于液氮中速冻。R8 期时取大豆的荚于液氮中速冻。使用 RNA 提取试剂盒(康为)提取以上组织 RNA,进行 qRT-PCR 试验<sup>[14-15]</sup>,3 次生物学重复。设计大豆中 *GmPIF3*、*GmLAF1* 和 *GmFHL1* 及其同源基因的特异 qRT-PCR 引物(表 1)。

1. 2. 3  CRISPR/Cas 9 载体构建  利用 CRISPR/Cas9 在线网站(<http://skl.scau.edu.cn/>),根据刘耀光实验室网站(<http://skl.scau.edu.cn/>)方法设计靶点引物(表 2)。

表 1  qRT-PCR 引物

Table 1  Primers of qRT-PCR

引物名称 Primer name	引物序列(5'-3') Primer sequence (5'-3')	目的基因 Target gene
PIF3-RT T1F	AACGACGTGTCCAAGGTGA	<i>PIF3a</i>
PIF3-RT T1R	TCTCGTGGCTCCAAATCCTC	<i>PIF3a</i>
PIF3-RT T2F	TTTCCTTCCCGCCTCAAACA	<i>PIF3b</i>
PIF3-RT T2R	GACGCATCCTGGTTTGCATT	<i>PIF3b</i>
PIF3-RT T3F	TCCTGGACAAGGCTATCCCA	<i>PIF3c</i>
PIF3-RT T3R	GACTTGTCGAGTCACAGCCA	<i>PIF3c</i>
PIF3-RT T4F	GTGGAGCTAGCTGTGCTCTC	<i>PIF3d</i>
PIF3-RT T4R	GTTCTGTAAACGCACGCATC	<i>PIF3d</i>
PIF3-RT T5F	GGGACTTGGTATGGGGTTCA	<i>PIF3e</i>
PIF3-RT T5R	GTGCTAGTTGCAAGAGGGGT	<i>PIF3e</i>
PIF3-RT T6F	GTGGAGCTAGCTGTGCTCTC	<i>PIF3f</i>
PIF3-RT T6R	GTTCTGTAAACGCACGCATC	<i>PIF3f</i>
LAF1-RT T1F	CAACATGCCGGCTCAAGTTC	<i>LAF1a</i>
LAF1-RT T1R	TCCACTGATAGCCACTCAGAA	<i>LAF1a</i>
LAF1-RT T2F	CTCACCCCTTCACGACATGCT	<i>LAF1b</i>
LAF1-RT T2R	GATCTGTTGGCCTTGAGTTGC	<i>LAF1b</i>
LAF1-RT T3F	CAGCATGCTAGCTCAAGTTCA	<i>LAF1c</i>
LAF1-RT T3R	TCCACTGAAAGCCACTCAGA	<i>LAF1c</i>
LAF1-RT T4F	TCACCACATGCTAGGCAACA	<i>LAF1d</i>
LAF1-RT T4R	GTGGGCATTGAGTTGCAAGATT	<i>LAF1d</i>
FHL-RT-T1F	AGGAAGCATTTTGTGATAGCCA	<i>FHLa</i>
FHL-RT-T1R	CAGGGTCCTCATTGCTGTCTT	<i>FHLa</i>
FHL-RT-T2F	TGATGATGCCAATGCACAAGC	<i>FHLb</i>
FHL-RT-T2R	TTCGCCCCACTGGAGTAAAG	<i>FHLb</i>

表 2 基因靶点引物  
Table 2 Primers of gene targets

靶点引物 Target primer	引物序列(5'-3') Primer sequence (5'-3')	目的基因 Target gene
PIF3-T1F	gtcaTTAGAAGGCTAGTCCCATC	PIF3a
PIF3-T1R	aaacGATGGGACTAGCCTTCTAA	PIF3a
PIF3-T2F	gtcaCATCTCACAGTCCCAAGGCT	PIF3b
PIF3-T2R	aaacAGCCTTGGGACTGTGAGATG	PIF3b
PIF3-T3F	gtcaTGCCTTTGTATGAGTTATAC	PIF3d
PIF3-T3R	aaacGTATAACTCATACAAAGGCA	PIF3d
PIF3-T4F	gtcaCAAGCTTCCCTTTAGCCACA	PIF3c
PIF3-T4R	aaacTGTGGCTAAAGGGAAGCTTG	PIF3c
PIF3-T5F	gtcaCAAGCTTCCCTTTAGCCACA	PIF3f
PIF3-T5R	aaacTGTGGCTAAAGGGAAGCTTG	PIF3f
PIF3-T6F	gtcaTGCCTGATTCCACCAAACC	PIF3e
PIF3-T6R	aaacGGTTTGGTGGAATCAGGCA	PIF3e
PIF3-T1F	gtcaTTAGAAGGCTAGTCCCATC	PIF3a
PIF3-T1R	aaacGATGGGACTAGCCTTCTAA	PIF3a
PIF3-T2F	gtcaCATCTCACAGTCCCAAGGCT	PIF3b
PIF3-T2R	aaacAGCCTTGGGACTGTGAGATG	PIF3b
LAF1-T1F	gtcaTAGCACAGCATTTGCCAGGA	LAF1a、LAF1c
LAF1-T1R	aaacTCCTGGCAAATGCTGTGCTA	LAF1a、LAF1c
LAF1-T2F	gtcaGAAACCACATCCTTAAACA	LAF1b、LAF1d
LAF1-T2R	aaacTGTTTAAGGATGTGTTTC	LAF1b、LAF1d
FHL-T1F	gtcaGATCCAGATGGAGGAAAGC	FHLa
FHL-T1R	aaacGCTTTCCTCCATCTGGATC	FHLa
FHL-T2F	gtcaCAACCCCTTCTCAAGTTGAC	FHLa
FHL-T2R	aaacGTCAACTTGAGAAGGGTTG	FHLa
FHL-T3F	attgCCAAACACGAGACAGATTG	FHLb
FHL-T3R	aaacCAATCTGTCTCGTGTTCG	FHLb
FHL-T4F	attgCTTTTCAGATTCCAGGATGA	FHLb
FHL-T4R	aaacTCATCCTGGAATCTGAAAAG	FHLb

参照甘卓然等<sup>[16]</sup>和侯智红等<sup>[17]</sup>和张峰的  
CRISPR-Cas 9 技术方法 ( [https://www. ncbi..nlm.  
nih. gov/pmc/articles/PMC3969860/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969860/)) 构建 CRISPR/

Cas 9 载体,利用 PCR 将 gRNA 与靶点引物<sup>[16]</sup>连接,  
每个基因设计多个靶点,并连接在同一个载体上;  
将 Cas 9 载体与连接好靶点的 gRNA 连接;将得到

的产物转化至大肠杆菌感受态并过夜培养;利用载体引物 SP1 和 SP3(表 4)鉴定转化大肠杆菌的阳性单克隆;将鉴定的阳性克隆送至广州天一辉远测序公司进行测序,并对结果序列比对和分析。

1.2.4 大豆根毛转化试验 参照 Chen 等<sup>[15]</sup>方法进行根毛转化试验,具体步骤为:将良好 Williams 82 种子放在干净的培养皿中,中间放 1 个 200 mL 的烧杯,烧杯中加入 96 mL 次氯酸钠和 4 mL 浓盐酸,灭菌 12 ~ 16 h;根据表 3 配制大豆萌发培养基,在超净

工作台中将已灭菌的大豆摆放到萌发培养基中;大豆种子于温室中萌发 5 d 左右,根据表 3 配制发根培养基,在超净工作台中利用手术刀在大豆的子叶上面切 1 个小洞,将切好的大豆子叶放置到发根培养基上,取 10 μL 含有目的载体的农杆菌 K599 放到大豆子叶的小洞中;将培养基密封,放于无菌培养室中培养 15 d 左右,收集生长出来的根毛并快速置于液氮中备用。

表 3 萌发培养基与根诱导培养基配方

Table 3 The germination medium and root medium formulation

培养基类型 Medium type	试剂 Reagent	用量 Dosage
萌发培养基 Germination medium	B5 盐 B5 salt mixture	1 ×
	蔗糖 Sucrose/(g·L <sup>-1</sup> )	20
	琼脂 Agar/(g·L <sup>-1</sup> )	8
发根培养基 Rooting medium	MS 合成盐 MS salt mixture	1 ×
	蔗糖 Sucrose/(g·L <sup>-1</sup> )	30
	2 - (4 - 吗啉) 乙磺酸 MES/(g·L <sup>-1</sup> )	0.6
	琼脂 Agar/(g·L <sup>-1</sup> )	8
	头孢霉素 Cephalosporin/(mg·L <sup>-1</sup> )	250
	羧苄青霉素 Carbenicillin/(mg·L <sup>-1</sup> )	250

1.2.5 基因靶点编辑有效性检测 使用 DNA 提取试剂盒(诺唯赞生物科技股份有限公司)提取大豆根毛 DNA,利用 CRISPR/Cas 9 载体通用引物 SP1 和 SP3(表 4)检测大豆根毛是否为转基因根毛,目的片段的长度约为靶点数 × 500 bp 左右,每个载体检测 20 个转基因事件,统计靶点的编辑效率。根据每个基因的 DNA 序列,分别设计目的基因的靶点 PCR 检测特异引物(表 4),将 PCR 产物送至广州天一辉远生物科技有限公司进行测序分析。测序峰图结果如果出现双峰则表示靶点有效,可用于后续试验。靶点有效性检测试验 PCR 反应体系和程序参考刘耀光实验室试验方法和体系(<http://skl.scau.edu.cn/>)。

1.3 数据分析

使用 Bioedit 软件进行 DNA 序列和蛋白序列比对;使用 MEGA7 软件制作蛋白进化树;使用 Tbttool 软件进行蛋白结构域结果展示;使用 Graphpad 软件进行定量 PCR 结果展示。

表 4 靶点检测引物

Table 4 Detection of target primer

靶点检测引物 Primer for target detection	引物序列(5' - 3') Primer sequence (5' - 3')
SP3	GTCGTGCTCCACATGTTGACCGG
SP1	GAAGTTATTGCATCTATGTCGGG
PIF3a-F	TGGCAGACCCAACAATGATT
PIF3a-R	TGGATCTAACGGATGCAA
PIF3b-F	ATGATGTGCAGGCCAGAGAATG
PIF3b-R	ATCAGAACCATACTCTTGTGTGC
PIF3c-F	GTTTGACCCTGATGCTCTTTG
PIF3c-R	ACTTCACCTGAAGGGACAAG
PIF3d-F	TGTTGCTGTGATGCAATGTTA
PIF3d-R	ACCTTGGGACAGTGAGATGG
PIF3e-F	TTGTCCACTTCCATCCCCTCT

续表 4

靶点检测引物	引物序列(5'-3')
Primer for target detection	Primer sequence (5'-3')
PIF3e-R	CCCGGTGATCCTTTAGGTGAG
PIF3f-F	GTTTCCGTGTTTGGGCTGTC
PIF3f-R	ACGCTGGATCTAACAGGTGC
LAF1a-F	ATGAACATATACTGCACAAT
LAF1a-R	TGAAGCATGCACACATTTC
LAF1b-F	AGTTTGCTCTGTGTATTCTGTG
LAF1b-R	GATGACCAATTAGATGTCTCCA
LAF1c-F	ATCACAACTATCTATTGCAT
LAF1c-R	TGATTACAGCCAAATAATGT
LAF1d-F	ATAAGTCTGTTTGTGCACTGTTT
LAF1d-R	TCAAATAAATAGTAGGCGTACT
FHLa-F	AGAACAAATGAGCAACTATG
FHLa-R	GTTGATACATCCAAAGGGCACT
FHLb-F	ATGCACGCGCCATGGGTATAT
FHLb-R	AACGGTAGCATCCAATCCAT

2 结果与分析

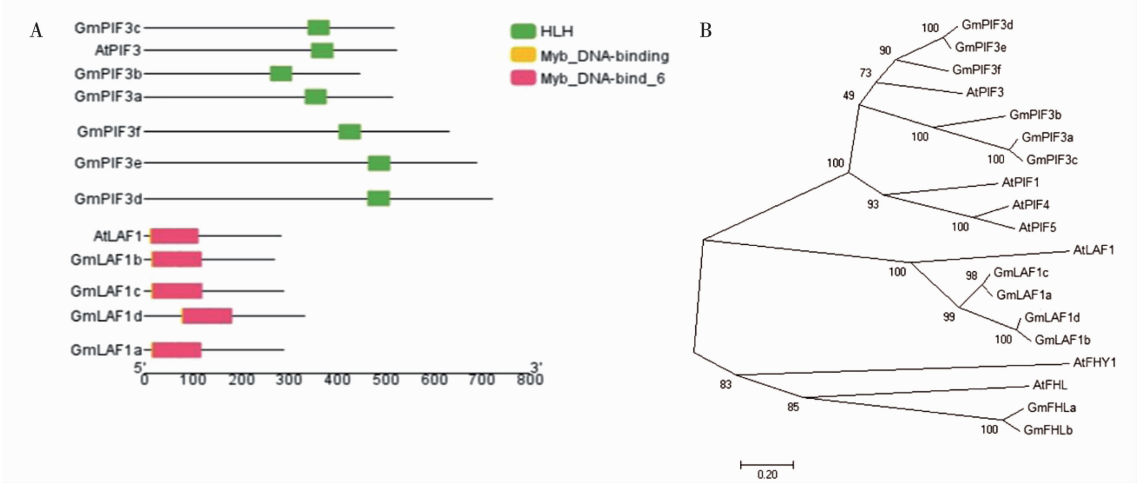
2.1 大豆同源基因预测与进化分析

分析结果显示,大豆中有 6 个 *AtPIF3* 同源拷贝,分别命名为 *GmPIF3a* (Glyma. 03G225000. 1)、

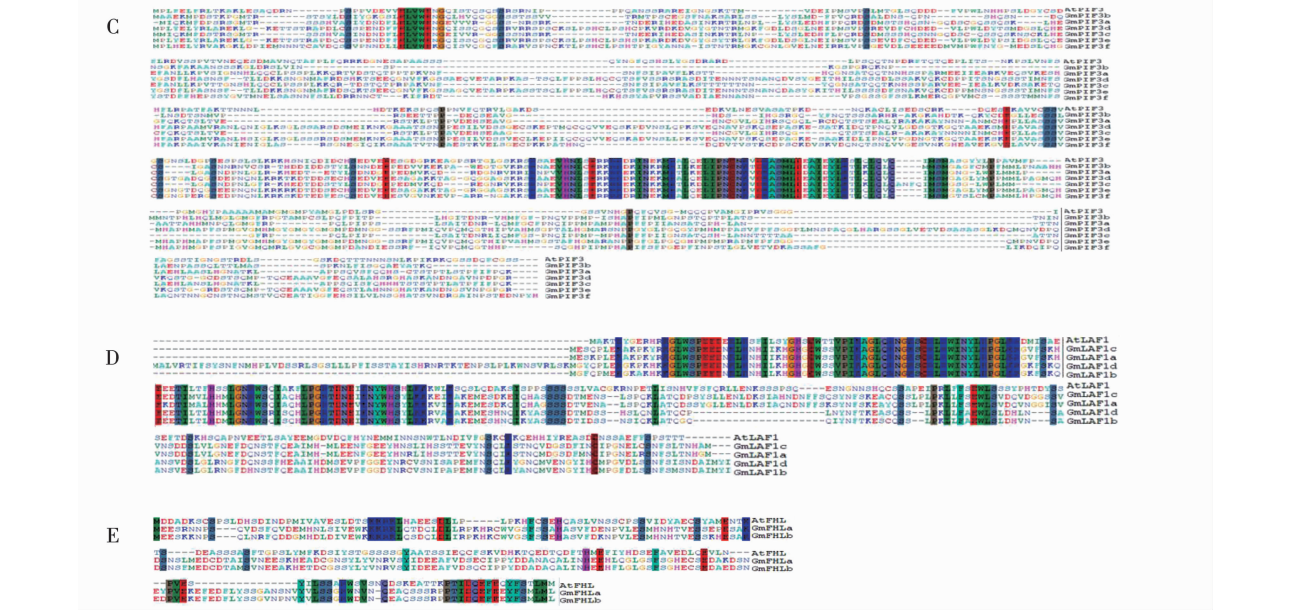
*GmPIF3b* ( Glyma. 10G138800. 1)、*GmPIF3e* ( Glyma. 10G142600. 1)、*GmPIF3c* ( Glyma. 19G222000. 1)、*GmPIF3f* (Glyma. 19G224700. 1)和 *GmPIF3d* (Glyma. 20G091200. 1)。AtPIF3 和 GmPIF3 中都含有 1 个保守的 bHLH 结构域,该结构域也是 PIFs 蛋白家族属于 bHLH 转录因子的最明显的特征(图 1A)。蛋白进化树分析结果显示 GmPIF3a/b/c 与 AtPIF3 更接近(图 1B 和图 1C)。

大豆中有 4 个 *AtLAF1* 同源拷贝,分别命名为 *GmLAF1a* (Glyma. 09G139000. 1)、*GmLAF1b* (Glyma. 10G236600. 1)、*GmLAF1c* (Glyma. 16G189400. 1)和 *GmLAF1d* (Glyma. 20G157900. 1)。AtLAF1 和 GmLAF1a/b/c/d 都含有 myb-DNA-binding 和 myb-DNA-bing-6 结构域(图 1A),且该区域能够和相应 DNA 直接结合,从而影响下游基因的表达。序列比对结果显示 GmLAF1a/b/c/d 的蛋白序列相似性都为 60% 以上,而 GmLAF1a 和 GmLAF1c 的序列相似性高达 85% (图 1B 和图 1D)。

利用拟南芥 *AtFHY1* 和 *AtFHL* 基因在大豆中寻找其同源基因,结果表明大豆中没有检测到 *AtFHY1* 的同源基因,检测到 2 个 *AtFHL* 同源基因(图 1B 和图 1E),编码蛋白都属于小蛋白家族,将这两个基因分别命名为 *GmFHLa* (Glyma. 18G183600. 1)和 *GmFHLb* (Glyma. 07G134000. 1), GmFHLa 和 GmFHLb 蛋白序列的相似性高达 80%。由于大豆中的 GmFHLa 和 GmFHLb 以及拟南芥中的 *AtFHY1* 和 *AtFHL* 蛋白较小并且没有特殊的蛋白结构元件,所以没有进行蛋白结构域分析。







A: *GmPIF3*/a/b/c/d/e/f 和 *GmLAF1*a/b/c/d 的结构域分析;B: 大豆和拟南芥 *FHY1*/*FHL*、*PIF3*、*LAF1* 蛋白序列进化树分析;C: *GmPIF3*/a/b/c/d/e/f 和 *AtPIF3* 氨基酸序列比对;D: *GmLAF1*a/b/c/d 和 *AtLAF1* 氨基酸序列比对;E: *GmFHL*a/b 和 *AtFHL* 氨基酸序列比

A: The domain analysis of *GmPIF3*/a/b/c/d/e/f and *GmLAF1*a/b/c/d in soybean; B: Evolutionary tree of *FHY1*/*FHL*, *PIF3* and *LAF1* protein of soybean and *Arabidopsis*; C: The protein sequence alignment of *GmPIF3*/a/b/c/d/e/f and *AtPIF3*; D: The protein sequence alignment of *GmLAF1*a/b/c/d and *AtLAF1*; E: The protein sequence alignment of *GmFHL*a/b and *AtFHL*

图 1 GmLAF1、GmPIF3 和 GmFHL 蛋白结构域、序列和进化树分析

Fig. 1 Analysis of protein domain, sequence alignment and evolutionary tree of *GmLAF1*, *GmPIF3* and *GmFHL*

## 2.2 组织特异性表达分析

如图 2A 所示, *GmPIF3* 在各大豆组织中都有表达。 *GmPIF3c/d* 在荚中表达量很高, 可能与植株的生长发育有关; *GmPIFa* 在茎中的表达量最高, 可能与株高有关; *GmPIF3f* 在籽粒中表达量最高, 而在其他组织中表达量相对较低, 推测其参与调控籽粒的生长发育。

如图 2B 所示, *GmLAF1a/b/c/d* 这 4 个基因在各大豆组织中均有表达, 但它们在茎、花和侧生分生组织中表达量都相对较低, 而 *GmLAF1b/d* 这两个基因在顶端生长点表达量较高, 最有可能调控生长点的生长。

如图 2C 所示, *GmFHL1b* 在各大豆组织中的表达量都相对较低, 推测该基因可能只需要少量表达就会发挥一定的功能或者该基因在 *GmFHL1a* 存在时不发挥功能, 而 *GmFHL1a* 在荚中表达量比较高,

预示其可能在荚中发挥功能。

## 2.3 基因敲除载体构建

靶点接头与 gRNA 载体酶切连接产物、第二轮扩增和菌落 PCR 检测结果如图 3 所示。最终构建载体的测序及分析结果显示, 构建出 2 个载体, 可通过 8 个靶点同时敲除 *PIF3a/b/c/d/e/f* 基因; 构建出 1 个载体, 可通过 4 个靶点同时敲除 *LAF1a/b/c/d* 基因; 构建出 1 个载体, 可通过 2 个靶点同时敲除 *FHLa/b* 基因。

## 2.4 靶点有效性检测

对被侵染的大豆毛根 DNA 的特异检测结果如图 4 所示, *PIF3a/b/c/d/e/f*、*LAF1a/b/c/d* 和 *FHLa/b* 均出现了双峰, 显示基因敲除载体编辑成功。本研究共鉴定出 12 个有效靶点, 能够分别将大豆中 6 个 *PIF3*、4 个 *LAF1* 和 2 个 *FHL* 成功敲除。

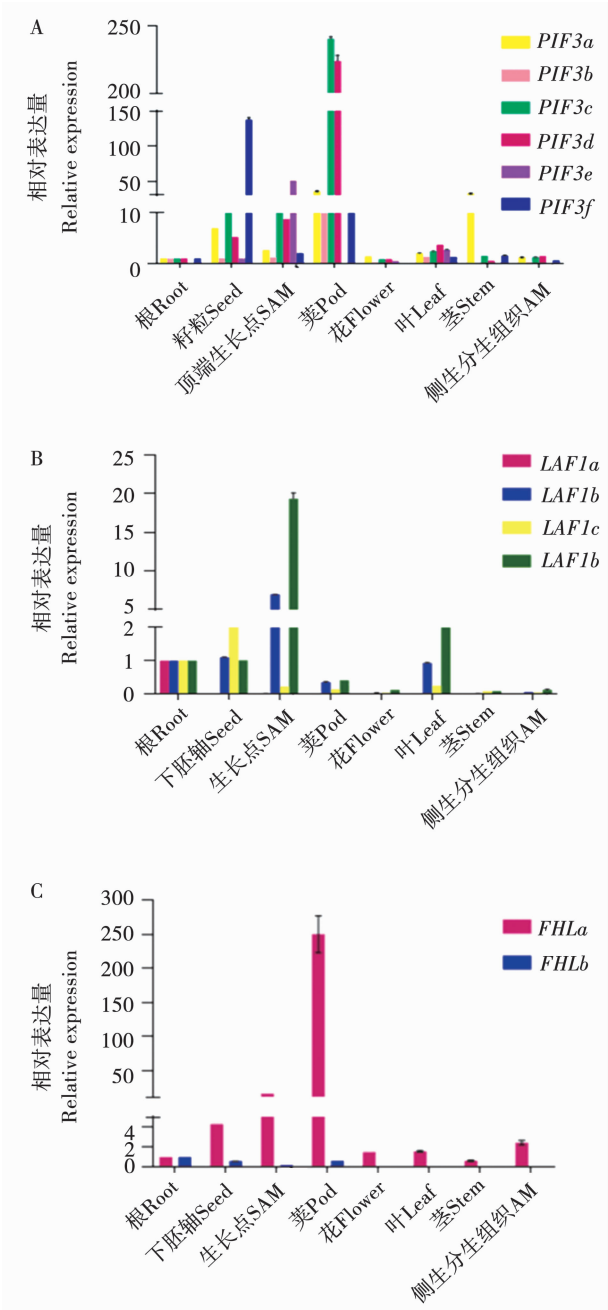
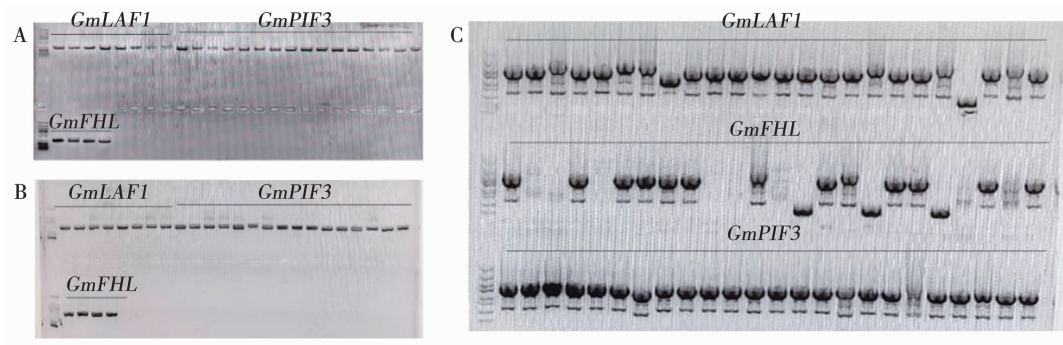


图 2 *GmPIF3*、*GmLAF1* 和 *GmFHL* 基因组织特异性表达分析

Fig. 2 Tissue expression of *GmPIF3*, *GmLAF1* and *GmFHL* genes in soybean



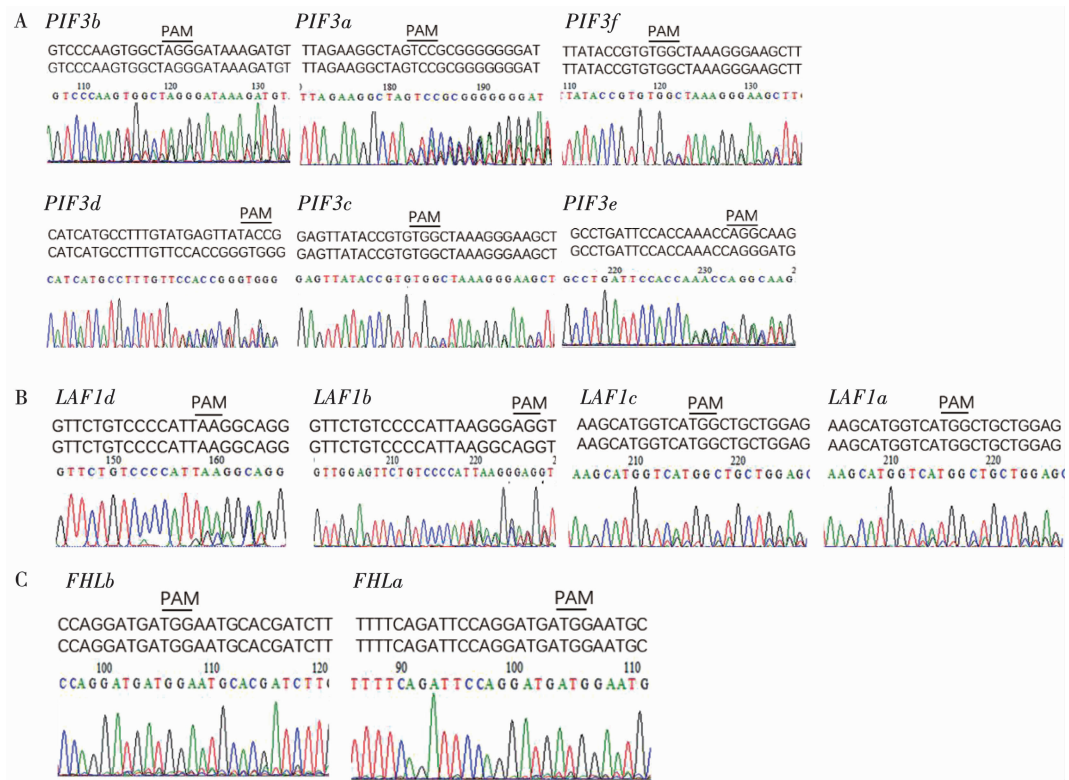
A: 第一轮 PCR; B: 第二轮 PCR; C: 菌落 PCR

A: PCR of the first round; B: PCR of the second round; C: Colony PCR

图 3 *GmPIF3*、*GmLAF1* 和 *GmFHL* 的 CRISPR/Cas 9 载体构建检测

Fig. 3 PCR detection of *GmLAF1*, *GmPIF3* and *GmFHL* CRISPR/Cas 9 vectors





A: *GmPIF3a/b/c/d/e/f* 靶点的突变情况; B: *GmLAF1a/b/c/d* 靶点的突变情况; C: *GmFHLa/b* 靶点的突变情况; PAM: CRISPR-Cas 9 载体识别位点  
A: The mutation of *GmPIF3a/b/c/d/e/f*; B: The mutation of *GmLAF1a/b/c/d*; C: The mutation of *GmFHLa/b*; PAM: The recognition site of CRISPR-Cas 9 vector

图 4 基因敲除靶点的测序峰图

Fig. 4 Detailed sequence of the knock-out targets site

3 讨 论

在拟南芥中, *PIFs* 在光形态建成和光信号途径中起作用, 参与植物体内多重信号转导途径, 能够参与黄化幼苗生长转化过程。 *PIFs* 属于转录因子 basic Helix-Loop-Helix (bHLH) 的第 15 亚族, 它们都能够和光敏色素相互作用, 其中 *PIF3* 既能够和光敏色素 A 相互作用也可以和光敏色素 B 相互作用; 在光照条件下 phy 由 Pr 转变为 Pfr, 然后转移到细胞核与 *PIF3* 相互作用, 从而调控靶基因的表达, 诱导光形态建成<sup>[2, 18-19]</sup>。 *AtPIF3* 主要在拟南芥的幼苗期表达量最高, 其次在莲座叶和顶端也有表达<sup>[20]</sup>, 而本研究结果显示大豆中预测到 6 个 *PIF3* 基因, *GmPIF3e* 的表达模式更接近拟南芥中 *AtPIF3* 的表达模式, 因此可以推测 *GmPIF3e* 的功能可能会更接近 *AtPIF3* 的功能; 而 *GmPIF3c* 和 *GmPIF3d* 两者在莢中的表达量很高。大豆中的多个 *PIF3* 基因可能出现了功能分化或者具有新的功能。 *AtLAF1* 是 R2R3-MYB 转录因子, 可以直接介导 phyA 下游基因的表达, 其主要在幼苗初期、花和叶片中表达<sup>[22-23]</sup>, 而大豆 *GmLAF1a/b/c/d* 4 个基因中 *GmLAF1b/d* 更

接近拟南芥 *AtLAF1* 的表达模式, 推测其功能更接近 *AtLAF1* 的功能, 而 *GmLAF1d* 在顶端生长点的表达量很高, 推测该基因可能与植株的生长相关。拟南芥中的相关研究表明, *AtFHY1/FHL* 是能够直接和 phyA 相互作用的转录因子, 其可以通过二聚化来调节自身稳定性, 从而协调不同信号分支的信号强度<sup>[21-23]</sup>。 *AtFHY1/FHL* 主要在幼苗期表达<sup>[24]</sup>, *GmFHLb* 在各个时期的表达量都较低, 而 *GmFHLa* 主要在莢表达, 其次在顶端分生组织和籽粒中的表达量也很高, 因此 *GmFHLa* 可能会影响莢的成熟。

基因编辑技术已经被广泛的应用于各个物种的研究之中, 但是依然存在脱靶的问题<sup>[23]</sup>。在大豆这个古四倍体作物中的应用则更需要先鉴定靶点的有效性, 从而使得研究更有效地开展<sup>[15-17]</sup>。本研究成果的有效靶点是经过多次试验而得出, 所以脱靶问题依然是一个有待解决的技术性难题<sup>[25]</sup>。 *E3* 和 *E4* 对大豆生育期十分重要, 后续研究将获得 *E3* 和 *E4* 下游基因的有效突变体, 并进一步根据突变体的表型研究目的基因的功能, 研究结果有助于发现影响大豆开花的新基因, 促进适应不同地区的大豆新品种的培育。

4 结 论

本研究挖掘拟南芥 phyA 下游重要的 3 个转录因子编码基因 *PIF3*、*LAF1* 和 *FHL* 的大豆同源基因,不同基因的主要表达部位存在明显差异,可能出现了功能分化。获得具有 CRISPR/Cas 9 编辑作用的靶点序列 12 个;构建出 4 个有效的基因敲除载体,1 个可敲除 *LAF1a/b/c/d* 基因,1 个可敲除 *FHLa/b* 基因,另外 2 个同时使用可敲除 *PIF3a/b/c/d/e/f* 基因。研究结果为大豆 *PIF3*、*LAF1* 和 *FHL* 同源基因突变体的研究提供技术支持,也为大豆 phyA 下游开花基因的研究奠定基础。

参考文献

[1] Jiao Y L, Lau O S, Deng X W. Light-regulated transcriptional networks in higher plants[J]. Nature Reviews Genetics, 2007, 8 (3): 217-230.

[2] Wang H Y, Deng X W. Dissecting the phytochrome A-dependent signaling network in higher plants [J]. Trends Plant Science, 2003, 8(4): 172-178.

[3] Li J, Li G, Wang H, et al. Phytochrome signaling mechanisms [M]. Arabidopsis Book, 2011, 9: e0148.

[4] Zeidler M, Zhou Q W, Sarda X, et al. The nuclear localization signal and the C-terminal region of FHY1 are required for transmission of phytochrome A signals[J]. Plant Journal, 2004, 40(3): 355-365.

[5] Genoud T, Schweizer F, Tscheuschler A, et al. FHY1 mediates nuclear import of the light-activated phytochrome A photoreceptor [J]. PLoS Genetics, 2008, 4(8): e1000143.

[6] Yang S W, Jang I C, Henriques R, et al. FAR-RED ELONGATED HYPOCOTYLI and FHY1-LIKE associate with the Arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome A signals for inhibition of hypocotyl elongation[J]. Plant Cell, 2009, 21(5): 1341-1359.

[7] Jang I C, Yang S W, Chua N H, et al. Independent and interdependent functions of LAF1 and HFR1 in phytochrome A signaling [J]. Genes & Development, 2007, 21 ( 16 ): 2100-2111.

[8] Shin J, Kim K, Kang H, et al. Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors[J]. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106 ( 18 ): 7660-7665.

[9] Ni M, Tepperman J M, Quail P H. PIF3, a phytochrome interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein[J]. Cell, 1998,95(5): 657-667.

[10] Leivar P, Quail P H. PIFs: Pivotal components in a cellular signaling hub[J]. Trends Plant Science, 2011,16(1): 19-28.

[11] Xu M L, Xu Z H, Liu B H, et al. Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-

regulated post-flowering responses of soybean [J]. BMC Plant Biology, 2013,13(1): 91.

[12] Xia Z J, Watanabe S, Yamada T, et al. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering[J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109: E2155-E2164.

[13] Liu B H, Kanazawa A, Matsumura H, et al. Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene[J]. Genetics, 2008,180(2): 995-1007.

[14] Wang J F, Zhang C M, Liu W, et al. Construction of *tgw 6* mutants in rice based on CRISPR/Cas 9 technology [J]. Acta Agronomica Sinica, 2016, 42(8): 1160-1167.

[15] Chen K L, Wang Y P, Zhang R, et al. CRISPR/Cas genome editing and precision plant breeding in agriculture [J]. Annual Review of Plant Biology, 2019, 70(1): 667-697.

[16] 甘卓然, 石文茜, 黎永力, 等. 大豆生物钟基因 *GmLNK1/2*、*GmRVE4/8* 和 *GmTOC1* CRISPR/Cas 9 组织表达分析及敲除靶点的鉴定[J]. 作物学报, 2020, 46(8): 1291-1300. ( Gan Z R, Shi W Q, Li Y L, et al. Identification of CRISPR/Cas 9 knockout targets and tissue expression analysis of circadian clock genes *GmLNK1/2*, *GmRVE4/8*, and *GmTOC1* in soybean [J]. Acta Agronomica Sinica, 2020, 46(8): 1291-1300. )

[17] 侯智红, 吴艳, 程群, 等. 利用 CRISPR/Cas 9 技术创制大豆高油酸突变系[J]. 作物学报, 2019, 45(6): 839-847. ( Hou Z H, Wu Y, Cheng Q, et al. Creation of high oleic acid soybean mutation plants by CRISPR/Cas 9 [J]. Acta Agronomica Sinica, 2019, 45(6): 839-847. )

[18] 张媛媛. 光敏色素的结构及其信号调控机制[J]. 湖北农业科学, 2020, 59 ( 4 ): 5. ( Zhang Y Y. Structure and signal mechanism of phytochrome [J]. Hubei Agricultural Sciences, 2020, 59(4):5. )

[19] Quail P H. Phytochrome-interacting factors[J]. Seminars in Cell & Developmental Biology, 2000, 11(6): 457-466.

[20] Pham V N, Kathare P K, Huq E. Phytochromes and phytochrome interacting factors [J]. Plant Physiology, 2018, 176 ( 2 ): 1025-1038.

[21] Wang H, Studer A J, Zhao Q, et al. Evidence that the origin of naked kernels during maize domestication was caused by a single amino acid substitution in *tgal* [J]. Genetics, 2015, 200(3): 965-974.

[22] Castillon A, Shen H, Huq E. Phytochrome Interacting Factors: Central players in phytochrome-mediated light signaling networks [J]. Trends Plant Science,2007,12(11):514-521.

[23] Jang I C, Henriques R, Chua N H. Three transcription factors, HFR1, LAF1 and HY5, regulate largely independent signaling pathways downstream of phytochrome A [J]. Plant Cell Physiology, 2013, 54(6): 907-916.

[24] Zhu H, Li C, Gao C. Publisher correction: Applications of CRISPR-Cas in agriculture and plant biotechnology [J]. Nature Reviews Molecular Cell Biology, 2020, 21(11): 661-677.

[25] Zhang B. CRISPR/Cas gene therapy [J]. Journal of Cellular Physiology, 2021, 236(4): 2459-2481.