



研究报告 Original Papers

# 大豆根系特异表达基因全基因组水平挖掘及其启动子活性鉴定

薛冬<sup>1,2</sup>, 闫强<sup>1</sup>, 胡亚群<sup>1</sup>, 周琰琰<sup>1</sup>, 韦雅雯<sup>1</sup>, 袁星星<sup>1</sup>, 王学军<sup>2</sup>, 陈新<sup>1,\*</sup>

<sup>1</sup>江苏省农业科学院经济作物研究所, 江苏省高效园艺作物遗传改良重点实验室, 南京210014

<sup>2</sup>江苏沿江地区农业科学研究所, 江苏南通226500

\*通信作者(cx@jaas.ac.cn)

**摘要:** 通过改变根系形态建成和提高根系对逆境的抗性可以有效提高作物产量。鉴定根特异性启动子是研究根系发育和对根部性状进行遗传操作的前提。大豆作为一种重要的作物, 目前对其根部特异表达基因及启动子尚缺乏系统性的研究。基于大豆幼苗期根、茎、叶片组织的基因转录组数据, 本文对大豆根特异或偏好表达的基因启动子进行了挖掘, 并通过RT-PCR方法对其中105个候选基因进行了验证, 获得33个根特异表达基因, 并克隆了其中11个基因启动子(pro1~pro11)。对转基因大豆发状根组织和稳定转化本氏烟草幼苗植株进行GUS染色, 结果表明11个启动子均具有显著的根部特异或偏好表达活性。在大豆发状根中, 由pro1、pro2、pro8和pro9启动的GUS基因表达活性达到35S的2倍以上。本研究为大豆及其他作物中组织特异性基因及启动子的研究提供了参考。此外, 本研究提供的启动子可用于驱动抗性基因在大豆根系组织中特异表达, 从而提高大豆对根部病害的抵抗能力, 实现品种改良的目的。

**关键词:** 大豆; 启动子; 根特异性; GUS酶活性

## Genome-scale mining of root-preferential genes from soybean and characterization of their promoter activity

XUE Dong<sup>1,2</sup>, YAN Qiang<sup>1</sup>, HU Yaqun<sup>1</sup>, ZHOU Yanyan<sup>1</sup>, WEI Yawen<sup>1</sup>, YUAN Xingxing<sup>1</sup>,  
WANG Xuejun<sup>2</sup>, CHEN Xin<sup>1,\*</sup>

<sup>1</sup>Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences/Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Nanjing 210014, China

<sup>2</sup>Jiangsu Yanjiang Institute of Agricultural Sciences, Nantong, Jiangsu 226500, China

\*Corresponding author (cx@jaas.ac.cn)

**Abstract:** Crop yield can be effectively improved by changing root architecture and improving root resistance to stress. Identification of root-specific promoters is prerequisite for research into the regulation of root development and genetic manipulation of root traits. As an important crop species, researches on the genes and promoters specifically expressed in soybean roots is still devoid. In our study, genomic-scale

收稿 2022-02-26 修定 2022-09-08

资助 江苏省农业科技自主创新资金项目[CX(20)3119]、国家自然科学基金(32101668)、江苏省特粮特经产业技术体系集成创新中心([2021]423)和江苏沿江地区农业科学研究所青年科技基金项目[YJ(2020)003]。

mining of root-specific genes were performed based on the transcriptome data of soybean root, stem and leaf tissues in the seedling stage, among which 105 putative genes were further verified by RT-PCR and 33 root-specific genes were identified. Then, 11 promoters (pro1 to pro 11) were cloned from soybean. GUS staining of transgenic soybean hairy roots and *Nicotiana benthamiana* seedlings showed that the 11 promoters had high root-specific or preferential expression activities. Besides, the GUS enzyme activities driven by pro 1, pro 2, pro 8 and pro 9 were all two folds higher than that of 35S. Our research provided a reference for the identification of tissue specific promoters in soybean and other species. In addition, the promoters identified in this study can be used to drive the specific expression of resistance genes in soybean root tissues, so as to improve the resistance of soybean to root diseases and achieve the purpose of variety improvement.

**Key words:** soybean; promoter; root-preferential; GUS activity

转录的起始是基因表达的关键阶段,而启动子是调控基因表达起始的关键位点。启动子通常指位于基因编码区5'端上游的一段序列,包含了RNA聚合酶、转录因子识别及结合位点(Hernandez-Garcia和Finer 2014)。启动子通过调控基因组不同基因在时间和空间上的有序表达,从而保证了植物正常的生长发育和生命周期(Li等2019)。同时,作为研究基因分子生物学功能的一个重要工具,启动子在植物基因工程领域也有着广泛的用途。

启动子按照其驱动基因的表达方式可以分为组成型启动子、诱导型启动子和组织特异性启动子三类(Hernandez-Garcia和Finer 2014)。组成型启动子能在几乎所有组织中启动基因转录,目前在商业化转基因作物中使用最为广泛,用于驱动转基因作物中目的基因如工业酶、抗虫和筛选标记基因的超表达(Jiang等2018)。目前在植物基因工程中使用的来自花椰菜花叶病毒的35S (CaMV35s)启动子、水稻的Actin启动子和玉米的Ubiquitin启动子都属于组成型启动子(Odell等1985; Christensen等1992; Mcelroy等1991)。但是,外源基因在植株全方位的高效表达会大量消耗植株体内的基础物质,会对转基因植株的正常生理代谢过程及生长发育带来不利影响(Pino等2007)。

诱导型启动子在没有诱导因子存在的情况下不启动基因转录或转录水平很低,而在特定信号刺激下可以大幅度提高基因的转录水平(于壮等2009; 邓泽宜等2021)。植物诱导型启动子依据诱导因子可划分为内源信号(植物激素)响应型、外部物理刺激(生物非生物胁迫)响应型和外部化学信号刺激

响应型(Hernandez-Garcia和Finer 2014)。诱导型启动子使目的基因在可控的条件下表达,因此能够减少对作物生长发育和产量的影响。利用逆境诱导型启动子rd29A驱动*DREB1A*基因在烟草中过表达能够增强转基因植物对于干旱和低温胁迫的抗性,同时不显著影响转基因植物的正常生长(Kasuga等2004)。同样,利用rd29A驱动*AtCBF*基因能够使转基因马铃薯植株获得与35S启动子相同水平的耐冷能力,但是却显著改善了由35S启动子导致的转基因植株矮小、花期延迟及产量降低等不利性状(Pino等2007)。

组织特异性启动子驱动基因只在特定的器官或组织部位表达,此类启动子能够有效避免植物遗传转化中应用组成型启动子造成的营养与能量的浪费,又能够实现目的基因在特定组织器官表达的目的。组织特异性启动子的应用还能有效缓解对转基因食品安全性的恐慌及对环境的不良影响。因此,许多组织特异性启动子被鉴定并用于玉米、水稻、大豆、马铃薯、番茄和烟草等作物的遗传转化(Kummari等2020; Jeong和Jung 2015)。

根作为植物的重要器官,根部特异性启动子的分离及鉴定,对于利用工程技术改良作物根系形态建成、养分吸收、盐分和干旱等非生物胁迫、根部病虫害等性状的改良具有重要意义(Koprivova等2019; Werner等2010)。目前为止,已经从水稻、烟草、马铃薯、番茄、大豆等作物中分离出许多根部特异性启动子(Li等2019; Kummari等2020; Jeong和Jung 2015)。这些启动子的获得往往是利用模式植物中根部特异或偏好表达基因的信息进行同源

搜索并克隆获得其启动子序列的, 通过转录组数据进行全基因组水平根部特异性启动子分析的研究还较少。

大豆作为重要的粮食作物, 已有5 000年栽培历史并广泛栽培于世界各地。大豆种子含有丰富的植物蛋白质、脂肪和多种营养元素, 是人类重要的油脂和蛋白质来源, 同时还是畜牧业重要的饲料来源(Zhang等2022)。由土传病原菌引起的大豆根部病害是严重危害大豆生产的世界性病害。由于根部侵染的特性, 在田间已经发生肉眼可辨病症的时候往往已经处于发病后期, 此时进行农药防治难以取得预期效果(Whitham等2016)。因此, 通过基因工程手段构建仅在根中特异表达抗病基因的转基因品种改良是一种有前途的策略。基于作物不同器官转录组数据的基因差异表达生物信息学分析为快速鉴定组织特异性启动子提供了强有力的工具(Li等2019)。

本研究中, 我们通过分析大豆幼苗期根、茎、叶片组织的基因转录组数据, 挖掘在根中特异或偏好表达的基因, 并通过RT-PCR方法对其中105个候选基因进行了验证。为了进一步验证所选候选基因的准确性, 我们克隆了其中11个候选基因的启动子序列并融合GUS报告基因。利用转基因大豆发状根组织对其GUS活性进行检测, 同时对稳定转化本氏烟植株GUS表达特异性进行组织化学染色分析, 结果表明选择的11个候选基因均具有强烈的根部特异或偏好表达活性。本研究为大豆根部性状的遗传操作提供了有利资源, 为其他作物组织特异性启动子分离鉴定提供了参考。

## 1 材料与方法

### 1.1 实验材料

大豆(*Glycine max* L.)栽培种‘Williams82’与本氏烟草(*Nicotiana benthamiana* L.)植物种子均由本实验室保存。温室温度保持在25°C, 光周期为16 h光照/8 h黑暗, 烟草在相同的生长条件下于温室中培养。

### 1.2 大豆转录组测序

分别收集生长3周大豆幼苗植株叶片、茎和根样品, 在液氮中迅速研磨成粉末, 取约100 mg组织样品利用FastPure® Plant Total RNA Isolation Kit试

剂盒(Vazyme, 南京)提取样品总RNA, 每个样品3个重复。RNA样品送北京擎科生物公司进行转录组测序及进行基因表达量分析, 测序采用华大MGIS-EQ-2000测序平台。下机的raw reads利用fastp软件进行质控, 过滤低质量数据后获得clean reads。利用Bowtie2软件将reads比对核糖体RNA数据库, 并删除比对到数据库中的clean reads。剩余的clean reads用于进一步的组装和基因丰度计算。利用HISAT2软件并设置“-rna-strandness RF”和其他参数为默认值开展基于参考基因组的比对分析。根据HISAT2的比对结果, 利用StringTie重构转录本, 并利用RSEM计算每个样本中基因的FPKM值作为基因表达量水平。用R语言计算每两个样品之间的皮尔森相关系数( $R^2$ ), 并据此评价重复样本之间的重复性。利用DESeq2软件进行2个样本间的差异表达分析, 差异基因筛选条件设置为:  $FDR < 0.01$ ,  $|\log_2 Ratio| > 1$ 。

### 1.3 RT-PCR检测

收集与转录组测序同时期根、茎、叶样品并提取RNA, 基因组去除和反转录采用HiScript II 1st Strand cDNA Synthesis Kit (+gDNA wiper)试剂盒(Vazyme, 南京), 按照说明书操作获得各组织cDNA并调整浓度至100 ng· $\mu\text{L}^{-1}$ 。利用Primer 6软件设计扩增长度400~600 bp的候选基因引物用于检测各候选基因的组织表达特异性。选择*GmACT11*基因(GenBank: NM\_001252731.2)作为内参基因, 设计扩增片段大小为246 bp的引物GmACT11-F: 5'-GG-TGGTTCTATCTTGGCATC-3', GmACT11-R: 5'-CTTTCGCTTCAATAACCCTA-3'。在25  $\mu\text{L}$  PCR体系中加入100 ng cDNA模板, 并同时加入候选基因和内参基因引物在同一反应体系中, 由于扩增片段大小不同, 电泳时同时观察到内参基因和候选基因目标条带。

### 1.4 启动子克隆和载体构建

根据pBIN-GFP载体上GFP表达盒序列设计引物(P3301-GFP-F: 5'-tatgaccatgattacgaattcGTCAC-TGGATTTTGGTTTGTAGGAAT-3'; P3301-GFP-R: 5'-ccgggtaccgagctcgaattcATGGTGGAGCACGACACACTT-3'), 上下游引物5'端分别包含pCAMBIA-3301载体EcoRI酶切位点同源臂序列。利用高保真酶(2×Phanta Master, Vazyme, 南京)扩增获得长度

为1 829 bp的GFP表达盒序列, 切胶并纯化回收后与 *EcoRI*酶切线性化处理后的pCAMBIA3301质粒进行同源重组反应(ClonExpress II One Step Cloning Kit, Vazyme, 南京)。反应产物转化大肠杆菌DH5 $\alpha$ 感受态细胞, 通过PCR鉴定阳性克隆, 并经测序验证, 将获得的新载体命名为pCAMBIA3301-GFP。

下载大豆‘Williams82’参考基因组数据([ftp://ftp.ensemblgenomes.org/pub/plants/release-48/fasta/glycine\\_max/dna/Glycine\\_max.Glycine\\_max\\_v2.1.dna.toplevel.fa.gz](ftp://ftp.ensemblgenomes.org/pub/plants/release-48/fasta/glycine_max/dna/Glycine_max.Glycine_max_v2.1.dna.toplevel.fa.gz)), 利用软件TBtools提取候选基因起始密码子ATG上游1 500 bp序列片段(Chen等2020)。利用Vazyme CE Design在线引物设计软件(Vazyme, 南京)设计扩增引物, 上、下游引物5'端分别含有pCAMBIA3301-GFP载体 *HindIII*和 *NcoI*酶切位点同源臂序列。收集大豆幼嫩叶片组织样品提取基因组(FastPure Plant DNA Isolation Mini Kit, Vazyme, 南京)作为模板, PCR扩增获得启动子片段切胶纯化回收, 与 *HindIII*和 *NcoI*双酶切线性化处理后的pCAMBIA3301-GFP进行同源重组反应, 反应产物通过热激法转化大肠杆菌DH5 $\alpha$ 感受态细胞, 获得的单菌落经过菌落PCR验证后并送南京擎科生物公司进一步测序验证。测序正确的质粒通过电击转化法分别转化发根农杆菌K599和根癌农杆菌EHA105, 并进行菌落PCR验证。

### 1.5 大豆发状根转化

挑选饱满且种皮无瑕疵的大豆‘Williams82’种子进行氯气灭菌处理, 灭菌后种子经无菌蒸馏水浸泡过夜后种植到含有1/2MS培养基的组培瓶里, 然后放置到16 h光照/8 h黑暗、25°C条件培养箱培养。生长5 d待子叶变为深绿色时, 用手术刀将子叶切下, 并小心切除子叶节区域。处理完毕的子叶在近子叶节端切一个直径约4 mm伤口, 然后伤口朝上放置到含有White培养基的培养皿中。在伤口处滴加20  $\mu$ L含有目的质粒的K599菌液, 侵染处理过的子叶放到25°C黑暗条件下继续培养3周。待发状根从伤口处长出约3 cm, 利用荧光体视镜筛选GFP阳性转化体并从子叶上切下转接到新的White培养基上继续培养。

### 1.6 烟草叶片瞬时表达和稳定转化

将28°C、200 r·min<sup>-1</sup>条件下震荡培养36 h含有

pCAMBIA3301-GFP质粒的土壤农杆菌GV3101菌液用10 mmol·L<sup>-1</sup> MgCl<sub>2</sub>沉淀洗涤2次, 最后用10 mmol·L<sup>-1</sup> MgCl<sub>2</sub>重悬浮并调节OD值至0.6。将处理好的菌液利用1 mL注射器注射进幼嫩烟草叶片, 并在16 h光照/8 h黑暗、25°C条件下继续培养48 h。荧光检测利用LUYOR-3415荧光激发光源灯和配套滤镜进行拍照(上海路阳, 上海)。

利用根癌农杆菌介导的叶盘转化法获得转基因烟草, 具体方法按照Yan等(2019)报导。从T<sub>0</sub>代植株收获T<sub>1</sub>代种子后, 播种于含有50 mg·L<sup>-1</sup>草铵膦的MS培养基上, 种子萌发后经荧光和草铵膦抗性筛选后移栽到营养土中继续生长并收集T<sub>2</sub>代转基因种子用于启动子表达特异性分析。

### 1.7 GUS染色及活性检测

收集在含有50 mg·L<sup>-1</sup>草铵膦的MS培养基上生长10 d并经荧光筛选后的T<sub>2</sub>代转基因烟草幼苗, 整株浸没在GUS染色液中(pH值为7.0的50 mmol·L<sup>-1</sup>磷酸缓冲液、体积分数为0.1% Triton·X-100、2 mmol·L<sup>-1</sup> K<sub>4</sub>[Fe(CN)<sub>6</sub>]·3H<sub>2</sub>O、2 mmol·L<sup>-1</sup> K<sub>3</sub>[Fe(CN)<sub>6</sub>]、10 mmol·L<sup>-1</sup> EDTA、2 mmol·L<sup>-1</sup> X-Gluc), 在37°C黑暗条件下处理4~8 h。染色后的植株经70%酒精隔水加热脱色处理直至背景无色, 白色背景下的蓝色即为GUS表达位点, 染色结果拍照记录。切取White培养基上继代培养1周的发状根末端组织, 浸没在GUS染色液中按上述条件染色并拍照观察。

大豆根毛中GUS基因表达活性检测利用GUS基因定量检测试剂盒(coolaber, 北京), 并按照说明书操作。

### 1.8 启动子序列调控元件的生物信息学分析

将提取的启动子序列提交PLACE在线数据库(<https://www.dna.affrc.go.jp/PLACE/?action=new-place>)进行预测分析。利用软件TBtools对根部特异表达相关元件在启动子上的分布进行可视化分析(Chen等2020)。

## 2 实验结果

### 2.1 大豆根特异表达基因筛选

分别提取生长3周大豆幼苗植株根、茎、叶片样品RNA进行转录组测序, 测序数据经过质量控

表1 转录组测序数据统计  
Table 1 Transcriptome sequencing data statistics

| 样品    | 过滤得到的测序数据  | 过滤得到的数据的总碱基数  | GC含量/% | 测序错误率≤0.1%的碱基数目比例/% |
|-------|------------|---------------|--------|---------------------|
| leaf1 | 22 949 020 | 6 878 592 948 | 44.81  | 93.21               |
| leaf2 | 23 307 719 | 6 986 614 552 | 45.10  | 92.94               |
| leaf3 | 21 531 994 | 6 454 691 604 | 45.10  | 93.05               |
| root1 | 23 017 125 | 6 897 413 072 | 43.95  | 93.09               |
| root2 | 22 916 058 | 6 866 347 912 | 44.29  | 93.23               |
| root3 | 23 235 424 | 6 963 433 422 | 44.05  | 93.15               |
| stem1 | 23 536 123 | 7 055 618 744 | 44.34  | 93.07               |
| stem2 | 23 481 044 | 7 038 894 328 | 44.39  | 93.07               |
| stem3 | 23 319 398 | 6 989 911 696 | 44.71  | 92.77               |

制共得到约62.14 Gb Clean Data, 各样品的Q30碱基百分比均在90%以上(表1)。将获得的Clean reads与大豆参考基因组进行序列比对, 各样品的Reads与参考基因组的比对效率在94.26%~97.57%之间。各生物学重复样品间的皮尔森相关系数( $R^2$ ) $\geq 0.803$ , 说明重复样品间重复性较好(图1)。通过序列比对及新基因预测, 并经基因表达量统计, 共计获得56 942个基因表达数据(FPKM值)。

为了筛选大豆根特异表达基因, 将根中基因表达水平分别与茎和叶片样品比较, 各获得6 160和

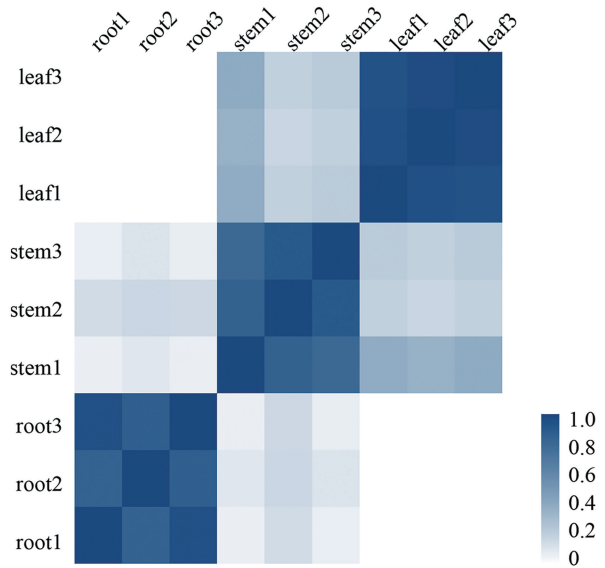


图1 样品间皮尔森相关系数  
Fig. 1 Person correlation between samples  
 $R^2$ : 皮尔森相关系数的平方。

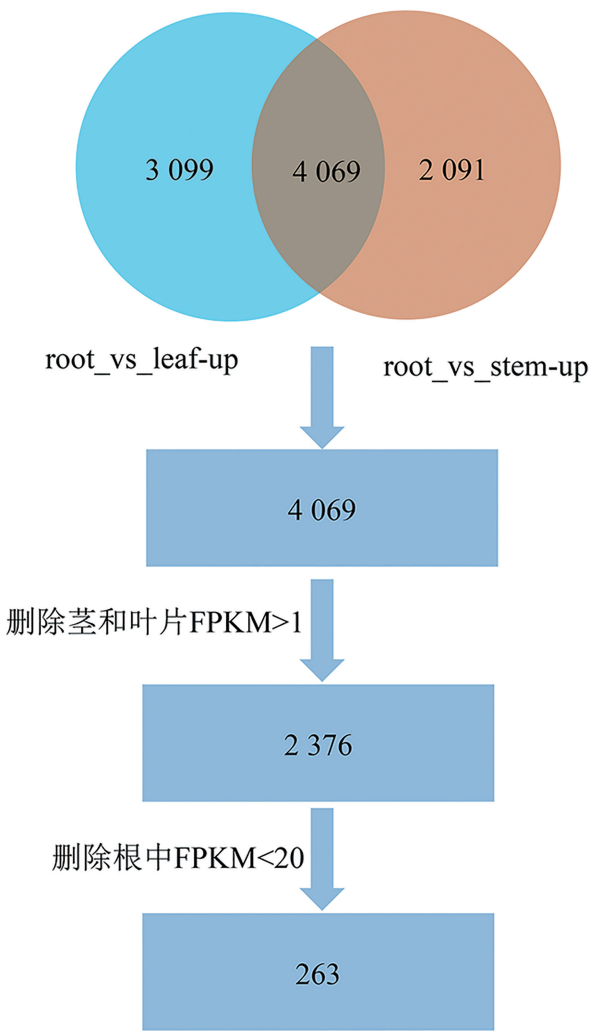


图2 根特异性表达基因筛选策略示意图  
Fig. 2 Schematic diagram of the strategy for identifying root-specific genes

7 168个在根中显著上调表达的基因( $FDR < 0.01$ ,  $\text{Log}_2\text{-Ratio} > 1$ ), 其中包含4 096个共同基因(图2)。下一步为保证候选基因在茎与叶片中的低表达水平, 删除在茎和叶片 $\text{FPKM} > 1$ 的差异基因, 获得2 376个候选基因。为保证候选基因在根中的高表达水平, 进一步删除根中 $\text{FPKM} < 20$ 的差异基因, 最终获得263个在根中高水平表达同时在茎和叶片中低水平表达的候选基因(图2)。

2.2 候选基因组织特异性表达验证

为了证明候选基因的根特异表达特性, 从263

个候选基因中选择了76个在根中具有高表达水平( $\text{FPKM} > 80$ )的差异基因; 同时为了全面了解该分析方法的可靠性, 又选择了29个基因 $\text{FPKM} < 20$ 的差异基因, 对选择的105个候选基因利用RT-PCR检测了其分别在大豆根、茎和叶片中的表达水平。结果表明其中33个基因表现出强烈的根部特异表达特性(图3)。其中有7个基因 $\text{FPKM} < 20$  (第1、9、13、24、39、40、46号), RT-PCR结果也表明这7个基因表达量也低于选自高 $\text{FPKM}$ 值的基因(图3和4)。通过以上研究, 我们从选择的105个候选基



图3 候选根特异表达基因组织特异性表达检测

Fig. 3 Expression of candidate genes with root-specificity

M: 标准分子量; R: 根; S: 茎; L: 叶片; 1~105表示候选基因编号。白色方框标记的候选基因为后续用于启动子克隆的候选基因。

| 编号  | 基因ID            | 根       | 茎   | 叶   |
|-----|-----------------|---------|-----|-----|
| 1   | GLYMA_16G212400 | 19.5    | 0.0 | 0.0 |
| 9   | GLYMA_12G054000 | 4.8     | 0.0 | 0.0 |
| 13  | GLYMA_03G038600 | 6.3     | 0.0 | 0.0 |
| 14  | GLYMA_08G341400 | 844.5   | 0.0 | 0.2 |
| 15  | GLYMA_13G251600 | 998.6   | 0.1 | 0.1 |
| 16  | GLYMA_13G252300 | 814.4   | 0.0 | 0.1 |
| 17  | GLYMA_13G251700 | 1 209.4 | 0.1 | 0.1 |
| 18  | GLYMA_16G060800 | 379.2   | 0.1 | 0.1 |
| 19  | GLYMA_13G252400 | 1 595.8 | 0.1 | 0.1 |
| 23  | GLYMA_15G062800 | 727.6   | 0.1 | 0.1 |
| 24  | GLYMA_09G239900 | 18.3    | 0.0 | 0.0 |
| 36  | GLYMA_16G060900 | 49.6    | 0.0 | 0.0 |
| 39  | GLYMA_12G176900 | 9.9     | 0.0 | 0.0 |
| 40  | GLYMA_07G124400 | 5.6     | 0.0 | 0.0 |
| 44  | GLYMA_16G060700 | 584.7   | 0.1 | 0.1 |
| 45  | GLYMA_11G185000 | 23.1    | 0.0 | 0.0 |
| 46  | GLYMA_04G035600 | 6.9     | 0.0 | 0.0 |
| 51  | GLYMA_17G053000 | 113.7   | 0.0 | 0.0 |
| 56  | GLYMA_13G252000 | 1 728.6 | 0.1 | 0.1 |
| 58  | GLYMA_17G019300 | 814.3   | 0.0 | 0.1 |
| 59  | GLYMA_12G053800 | 654.0   | 0.0 | 0.1 |
| 60  | GLYMA_20G220800 | 431.6   | 0.0 | 0.0 |
| 61  | GLYMA_10G168900 | 378.4   | 0.0 | 0.0 |
| 62  | GLYMA_03G061300 | 340.4   | 0.1 | 0.0 |
| 75  | GLYMA_01G083300 | 218.2   | 0.3 | 0.5 |
| 83  | GLYMA_09G022500 | 184.9   | 0.1 | 0.0 |
| 85  | GLYMA_11G035300 | 174.0   | 0.1 | 0.1 |
| 89  | GLYMA_10G122300 | 166.6   | 0.0 | 0.0 |
| 90  | GLYMA_16G170000 | 157.6   | 0.2 | 0.2 |
| 93  | GLYMA_11G195200 | 148.8   | 0.0 | 0.0 |
| 96  | GLYMA_19G263100 | 141.7   | 0.2 | 0.6 |
| 98  | GLYMA_06G286600 | 132.0   | 0.3 | 0.0 |
| 100 | GLYMA_08G341300 | 128.0   | 0.0 | 0.9 |

图4 33个根特异表达基因在大豆不同组织表达热图  
Fig. 4 Heat map of the expression profiles of the 33 root-specific genes

热图上数据为转录组3个样品平均FPKM值。

因中选择出33个根特异表达基因。

2.3 候选基因启动子克隆和载体构建

为便于进行阳性转化体筛选, 根据图5-A设计, 以本实验室保存的pBIN-GFP载体为模板扩增含有pCAMBIA3301载体EcoRI酶切位点两侧同源臂序列的35S-GFP-nos表达盒序列。将获得的目的序列与EcoRI酶切线性化的植物表达载体pCAMBIA3301连接并转化大肠杆菌感受态细胞, 单克隆经菌落PCR和测序验证后获得pCAMBIA3301-GFP载体(图5-A)。利用农杆菌介导的烟草叶片瞬时转化技术将pCAMBIA3301-GFP在本氏烟草叶片中过表达,

培养48 h后叶片注射区域表现出强烈的GFP荧光信号(图5-B)。说明插入到载体中的GFP表达盒能够在植物体内正常表达。

为了验证候选基因启动子在大豆根中的启动活性, 将启动子插入到植物表达载体pCAMBIA3301-GFP中融合GUS报告基因并转化到大豆发状根组织。从33个根特异表达基因中随机选择了11个基因(GLYMA\_13G251600、GLYMA\_13G252300、GLYMA\_16G060800、GLYMA\_15G062800、GLYMA\_16G060700、GLYMA\_11G185000、GLYMA\_04G035600、GLYMA\_13G252000、GLYMA\_17G019300、GLYMA\_12G053800、GLYMA\_06G286600), 提取其起始密码子ATG上游1 500 bp序列作为启动子序列, 依照上面基因名称顺序分别命名为pro1~pro11。PCR扩增获得5'端和3'端分别含有pCAMBIA3301-GFP载体HindIII和NcoI酶切位点同源臂序列的11个候选基因启动子序列(图5-C), 插入到pCAMBIA3301-GFP载体与GUS报告基因融合。获得的载体分别命名为pCAMBIA3301-GFP-pro1:GUS至pCAMBIA3301-GFP-pro11:GUS, 通过电击转化法转入发根农杆菌K599和根癌农杆菌EHA105, 用于后续遗传转化操作。

2.4 大豆发状根中启动子活性分析

获得的发状根在荧光显微镜下经过荧光筛选后转接到White培养基中继续培养1周, 切取根尖部分进行GUS组织化学染色分析, 结果表明pro1~pro11候选启动子在大豆发状根组织中表现出强烈的GUS信号(图6)。进一步利用转基因组织中GUS报告基因酶活来量化表达活性, 结果表明在空载体阳性对照转基因组织中, 组成型表达CaMV35S启动子能够驱动GUS报告基因的高水平表达。pro1、pro2、pro8和pro9均表现出显著高于35S的表达活性, pro3、pro5、pro10和pro11表现出与35S相当的表达活性, pro7转基因发状根组织中检测到的GUS酶活最低, 但是与35S对照没有表现出显著性差异(图7)。以上结果说明, 11个候选启动子在大豆发状根组织中均具有较强的启动活性。

2.5 转基因烟草中启动子组织特异性分析

T<sub>2</sub>代转基因烟草种子经草铵膦和GFP荧光筛选后(图8), 选择在培养基上生长10 d的阳性植株进

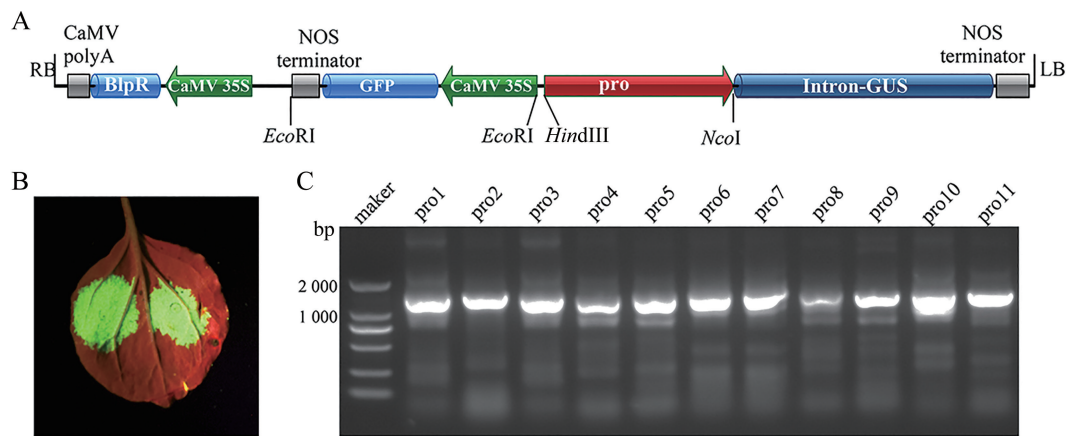


图5 植物表达载体构建及启动子克隆

Fig. 5 The construction of plant expression vector and promoter clone

A: pCambia3301-GFP-pro::Intron-GUS表达载体示意图; B: pCambia3301-GFP烟草叶片瞬时表达荧光检测; C: 候选启动子序列PCR扩增。

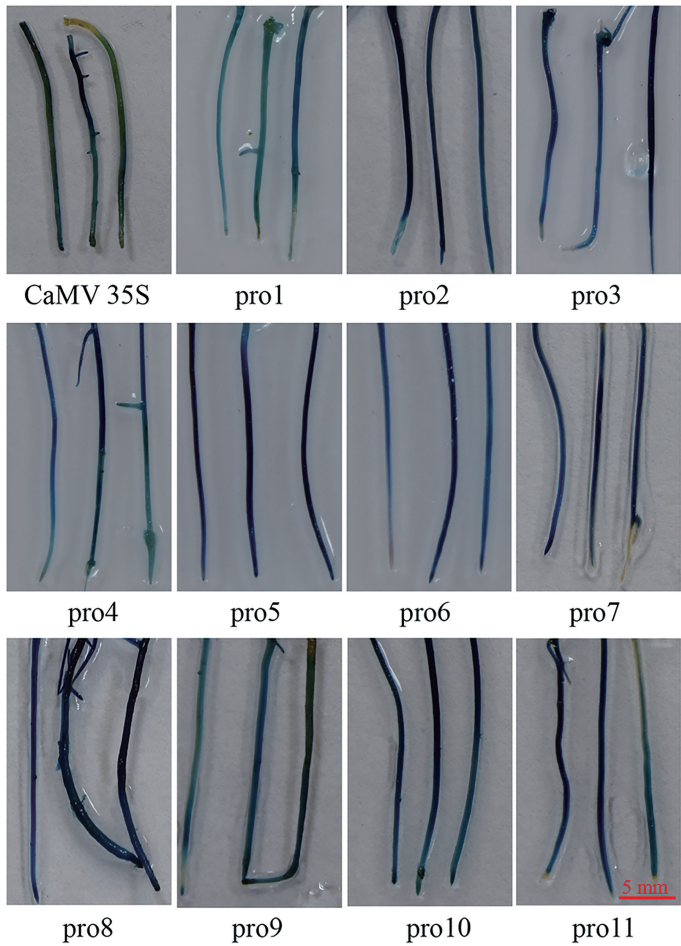


图6 不同转基因大豆发状根GUS组织化学染色

Fig. 6 GUS histochemical assays in transgenic soybean hairy roots

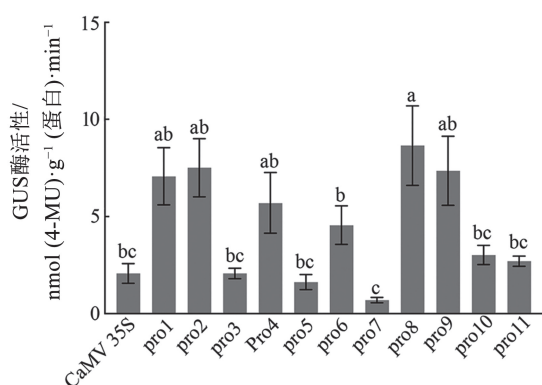


图7 不同启动子转基因大豆发状根中GUS酶活性定量分析  
Fig. 7 Quantitative analysis of GUS activity of transgenic soybean hairy roots with different promoter constructs

酶活性数据展示为3个样品平均值±SD, 柱形图上端小写字母表示显著性差异( $P < 0.05$ )。

行GUS组织化学染色, 分析候选启动子组织特异性。转化空载体阳性对照转基因烟草叶片、根和茎均检测到强烈的GUS表达活性; 11个启动子都在根中检测到强烈的GUS表达活性, 表现出根表达特异性和偏好性。其中pro5、pro6和pro10在叶片中也表现出低水平的GUS表达活性; pro11在叶片和茎中均表现出低水平的GUS表达活性, 但是具有明显的根部表达偏好性(图9)。以上结果说明选择的11个启动子均具有启动活性, 可以驱动下游GUS报告基因在转基因本氏烟草中表达, 且表现出强烈的根部特异表达特性。

## 2.6 11个候选启动子序列调控元件分析

为了分析选择的11个根特异性/偏好表达启动子序列中的调控元件分布特征, 使用PLACE在线数据库对1 500 bp启动子中已报导的MYCCONSENSUSAT、OSE1ROOTNODULE、OSE2ROOTNODULE、RAV1AAT、ROOTMOTIFTAPOX1、SP8B-FIBSP8AIB和SP8BFIBSP8BIB 7个根部特异表达相关的顺式作用元件进行分析(李濯雪和陈信波2015)。结果表明11个启动子序列上均含有大量的根部特异表达相关的作用元件, 其中包含最多的是MYCCONSENSUSAT和ROOTMOTIFTAPOX1, 而SP8BFIBSP8AIB元件仅在pro1序列中预测到1个(图10和表2)。

## 3 讨论

由于组织特异性启动子可以克服许多组成型启动子的缺点, 因而在遗传转化操作中受到青睐。最近, 利用已公开的组学数据, 建立了一个在线分析玉米、大豆和高粱根部特异表达基因及其启动子的数据库, 为研究这三种作物的根特异启动子提供了便利(Moisseyev等2020)。随着转录组测序成本的降低和越来越多作物全基因组测序的完成, 使得通过不同组织样本转录组数据鉴定组织特异表达基因和启动子成为一个便捷的手段。例如通过利用基因芯片数据, 研究者成功鉴定到几个水稻根特异启动子和玉米胚特异性启动子(Huang等2015;

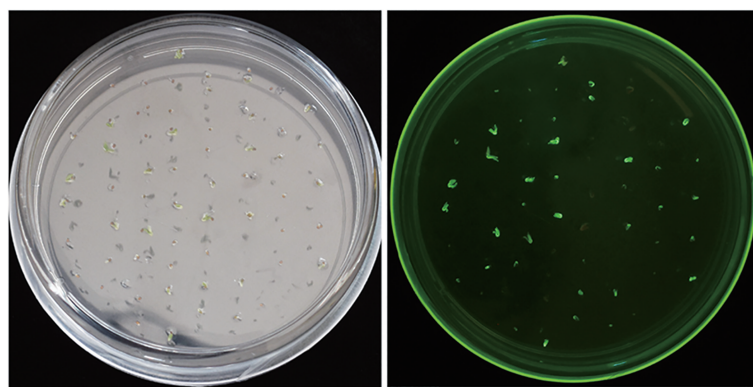


图8 转基因烟草荧光筛选

Fig. 8 Fluorescent screening of transgenic *N. benthamiana* seedlings

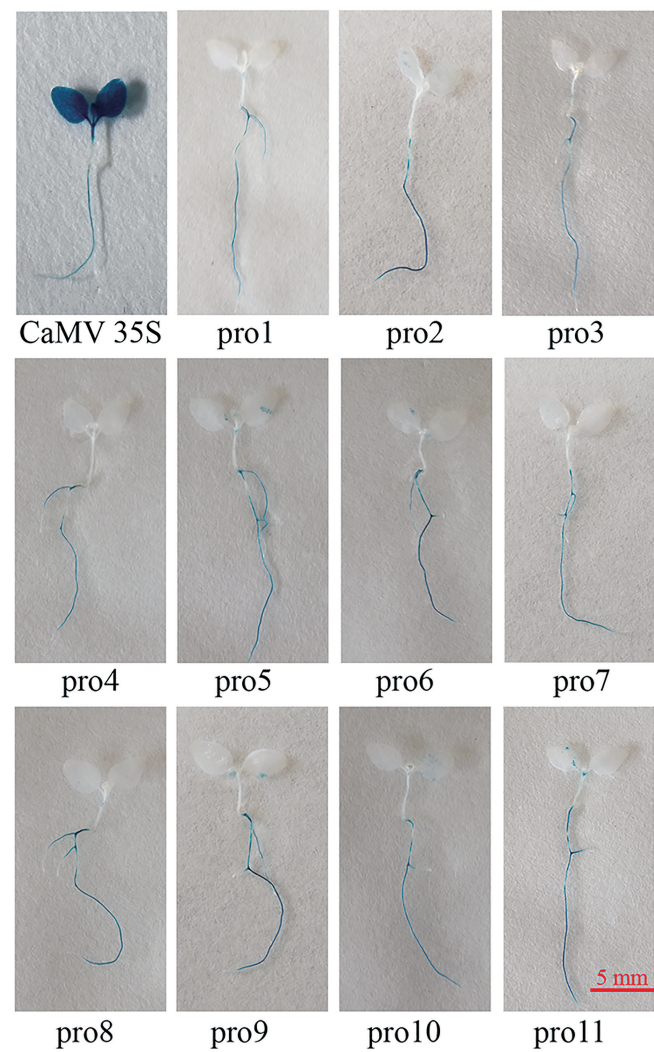


图9 不同启动子转基因本氏烟草幼苗GUS组织化学染色

Fig. 9 GUS histochemical assays of transgenic *N. benthamiana* seedlings with different promoter constructs

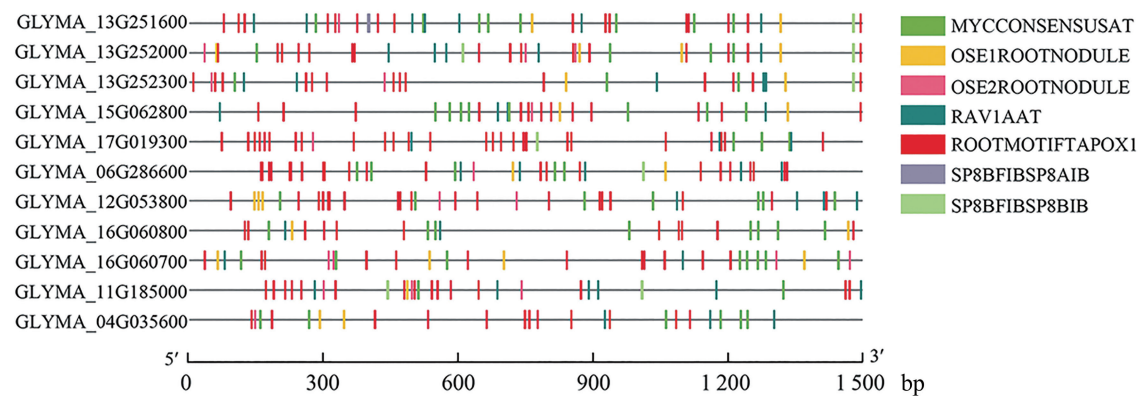


图10 pro1~pro11序列根特异表达相关顺式元件预测分析

Fig. 10 Prediction of the pro1 to pro11 *cis*-elements related to root-specific

表2 启动子pro1~pro11序列上根特异表达相关顺式作用元件数目统计  
Table 2 Statistics of the number of cis-elements related to root-specific expression on promoter pro1~pro11 sequence

| 编号    | 基因ID            | MYCCONSE<br>NSUSAT | OSEIROOT<br>NODULE | OSE2ROOT<br>NODULE | RAV1AAT | ROOTMOTI<br>FTAPOX1 | SP8BFIBSP8<br>AIB | SP8BFIBSP8<br>BIB |
|-------|-----------------|--------------------|--------------------|--------------------|---------|---------------------|-------------------|-------------------|
| pro1  | GLYMA_13G251600 | 12                 | 2                  | 1                  | 7       | 21                  | 1                 | 1                 |
| pro2  | GLYMA_13G252000 | 8                  | 4                  | 3                  | 5       | 19                  | 0                 | 2                 |
| pro3  | GLYMA_13G252300 | 6                  | 2                  | 2                  | 6       | 20                  | 0                 | 1                 |
| pro4  | GLYMA_15G062800 | 16                 | 2                  | 1                  | 4       | 18                  | 0                 | 0                 |
| pro5  | GLYMA_17G019300 | 6                  | 0                  | 1                  | 3       | 32                  | 0                 | 1                 |
| pro6  | GLYMA_06G286600 | 10                 | 3                  | 0                  | 5       | 27                  | 0                 | 1                 |
| pro7  | GLYMA_12G053800 | 14                 | 5                  | 0                  | 4       | 25                  | 0                 | 0                 |
| pro8  | GLYMA_16G060800 | 16                 | 2                  | 0                  | 2       | 14                  | 0                 | 0                 |
| pro9  | GLYMA_16G060700 | 16                 | 4                  | 4                  | 2       | 18                  | 0                 | 0                 |
| pro10 | GLYMA_11G185000 | 4                  | 4                  | 3                  | 6       | 22                  | 0                 | 2                 |
| pro11 | GLYMA_04G035600 | 12                 | 2                  | 1                  | 3       | 16                  | 0                 | 0                 |

Liu等2014)。最近Li等(2019)综合利用基因芯片和转录组数据,在全基因水平对玉米根特异性启动子进行了鉴定,并通过转基因验证获得玉米5个根特异性启动子。

目前对于大豆根特异性启动子的研究还较少,仅通过同源克隆的方法报道了大豆*GmPRP1*和*Gm-TIP*基因启动子具有根特异性(Chen等2014, 2015),对从全基因组水平分析大豆根特异表达基因及启动子还尚未有报道。在本研究中,利用大豆幼苗根、茎和叶三个器官的基因转录组数据,挖掘了根部特异性表达基因,选择其中105个候选基因利用RT-PCR对其组织表达特异性进行了验证,从中获得了33个根特异性或偏好表达的基因。对其中11个候选基因启动子的根部启动活性和组织表达偏好性进行分析,结果表明11个启动子在根中均表现出强烈的表达活性。本研究从全基因组水平对大豆根特异性表达基因及启动子进行了分析,并最终获得11个根特异性启动子。

启动子pro1、pro2、pro8和pro9在转基因大豆发状根组织中表现出显著的表达活性,同比之下,其表达活性能够达到35S启动子表达活性的2倍以上。但根特异性启动子驱动GUS基因在本氏烟草根部的表达并非绝对的。在本实验中,pro5、pro6、pro10及pro11转基因植株在叶片中也检测到低水平的GUS表达活性;此外pro11转基因植株还在茎中表现出低水平的GUS表达活性。基因的组织特异性表达受多种因素影响,其中一个主要因素是启动子上的顺式表达元件(Li等2019)。最后我们对11个启动子序列的根特异表达元件进行了分析,以期分析其元件数目和位置对启动子的活性是否存在影响。虽然每个启动子上均存在多个根特异性表达相关顺式元件,但是经过比较不同启动子转基因大豆发状根组织中的GUS酶活性,我们并没有发现其与已报道的8个根特异性表达相关元件数目之间存在相关性,因此还需要进行进一步的启动子截短实验来验证调控根特异性启动子中核心调控组织特异性及表达活性的关键调控元件。

由于缺乏合适的启动子,因此很难调控外源基因在转基因作物中按预期模式进行表达。本研究提供了系列大豆根特异性表达启动子,这些启

动子可作为有力的工具广泛应用于大豆遗传转化,为后续通过遗传转化改良大豆根部抗病性提供了具有价值的材料。

### 参考文献(References)

- Chen CJ, Chen H, Zhang Y, et al (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*, 13 (8): 1194–1202
- Chen L, Jiang BJ, Wu CX, et al (2014). *GmPRP2* promoter drives root-preferential expression in transgenic *Arabidopsis* and soybean hairy roots. *BMC Plant Biol*, 14: 245
- Chen L, Jiang BJ, Wu CX, et al (2015). The characterization of *GmTIP*, a root-specific gene from soybean, and the expression analysis of its promoter. *Plant Cell Tiss Org*, 121 (2): 259–274
- Christensen AH, Sharrock RA, Quail PH (1992). Maize polyubiquitin genes-structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol*, 18 (4): 675–689
- Deng ZY, Song X, Hong Y, et al (2021). Applications of promoters in the genetic engineering of ornamental plants: a review. *Acta Horti Sin*, 48 (6): 1250–1264 (in Chinese with English abstract) [邓泽宜, 宋想, 洪艳等(2021). 启动子在观赏植物基因工程中的应用综述. *园艺学报*, 48 (6): 1250–1264]
- Hernandez-Garcia CM, Finer JJ (2014). Identification and validation of promoters and *cis*-acting regulatory elements. *Plant Sci*, 217–218: 109–119
- Huang LY, Zhang F, Qin Q, et al (2015). Identification and validation of root-specific promoters in rice. *J Integr Agr*, 14 (1): 1–10
- Jeong HJ, Jung KH (2015). Rice tissue-specific promoters and condition-dependent promoters for effective translational application. *J Integr Plant Biol*, 57 (11): 913–924
- Jiang P, Zhang K, Ding Z, et al (2018). Characterization of a strong and constitutive promoter from the *Arabidopsis* serine carboxypeptidase-like gene *AtSCPL30* as a potential tool for crop transgenic breeding. *BMC Biotechnol*, 18 (1): 59
- Kasuga M, Miura S, Shinozaki K, et al (2004). A combination of the *Arabidopsis DREB1A* gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol*, 45 (3): 346–350
- Koprivova A, Schuck S, Jacoby RP, et al (2019). Root-specific camalexin biosynthesis controls the plant growth-promoting effects of multiple bacterial strains. *Proc Natl Acad Sci USA*, 116 (31): 15735–15744
- Kummari D, Palakolanu SR, Kishor PBK, et al (2020). An update and perspectives on the use of promoters in plant genetic engineering. *J Biosci*, 45: 119
- Li Y, Liu X, Chen R, et al (2019). Genome-scale mining of root-preferential genes from maize and characterization of their promoter activity. *BMC Plant Biol*, 19: 584
- Li ZX, Chen XB (2015). Research advances in plant tissue specific promoters and related *cis*-acting elements. *J Biol*, 32 (6): 91–95 (in Chinese with English abstract) [李濯雪, 陈信波(2015). 植物组织特异性启动子及相关顺式作用元件研究进展. *生物学杂志*, 32 (6): 91–95]
- Liu X, Jian T, Zhou X, et al (2014). Identification and characterization of promoters specifically and strongly expressed in maize embryos. *Plant Biotechnol J*, 12 (9): 1286–1296
- Mcelroy D, Blowers AD, Jeness B, et al (1991). Construction of expression vectors based on the rice *Actin-1 (Act1)* 5' region for use in monocot transformation. *Mol Gen Genet*, 231 (1): 150–160
- Moiseyev G, Park K, Cui A, et al (2020). RGPDB: database of root-associated genes and promoters in maize, soybean, and sorghum. Database (Oxford), 2020
- Odell JT, Nagy F, Chua NH (1985). Identification of DNA-sequences required for activity of the cauliflower mosaic virus-35s promoter. *Nature*, 313 (6005): 810–812
- Pino MT, Skinner JS, Park EJ, et al (2007). Use of a stress inducible promoter to drive ectopic *AtCBF* expression improves potato freezing tolerance while minimizing negative effects on tuber yield. *Plant Biotechnol J*, 5 (5): 591–604
- Werner T, Nehnevajova E, Kollmer I, et al (2010). Root-specific reduction of cytokinin causes enhanced root growth, drought tolerance, and leaf mineral enrichment in *Arabidopsis* and tobacco. *Plant Cell*, 22 (12): 3905–3920
- Whitham SA, Qi M, Innes RW, et al (2016). Molecular soybean-pathogen interactions. *Annu Rev Phytopathol*, 54: 443–468.
- Yan Q, Si JR, Cui XX, et al (2019). *GmDAD1*, a conserved *defender against cell death 1 (DAD1)* from soybean, positively regulates plant resistance against *Phytophthora* pathogens. *Front Plant Sci*, 10: 107
- Yu Z, Zhu LP, Zou CX, et al (2009). Application of stress-induced promoter in plant genetic engineering. *Plant Physiol Commun*, 45 (11): 1141–1145 (in Chinese with English abstract) [于壮, 朱丽萍, 邹翠霞等(2009). 胁迫诱导型启动子在植物抗逆基因工程中的应用. *植物生理学通讯*, 45 (11): 1141–1145]
- Zhang M, Liu S, Wang Z, et al (2022). Progress in soybean functional genomics over the past decade. *Plant Biotechnol J*, 20 (2): 256–282