

杨永恒,徐晓洋,孙玉明,等.甜菊葡萄糖苷转移酶基因 *SrUGT76G1* 启动子的克隆及其功能的瞬时表达分析[J].江苏农业科学,2018,46(6): 27-33.

doi:10.15889/j.issn.1002-1302.2018.06.007

甜菊葡萄糖苷转移酶基因 *SrUGT76G1* 启动子的克隆及其功能的瞬时表达分析

杨永恒,徐晓洋,孙玉明,原海燕,刘清泉,张永侠,王银杰,黄苏珍,佟海英

(江苏省中国科学院植物研究所,江苏南京 210014)

摘要:甜菊苷(stevioside, St)和莱鲍迪苷 A(rebaudioside A, R-A)是甜菊叶片中两大主要糖苷组分,甜菊糖苷生物合成途径的研究表明 *SrUGT76G1* 基因在 St 向 R-A 转化中起关键作用。为了进一步研究 *SrUGT76G1* 基因的表达调控,本试验采用 hiTAIL-PCR 的方法,经过 2 次步移从甜菊基因组中克隆到 *SrUGT76G1* 基因翻译起始位点上游 2 283 bp 的启动子序列。采用 PlantCARE、PLACE 等在线工具对序列进行分析,结果显示在这段序列上共有 475 个顺式调控元件,除了 TATA-box、CAAT-box、MYB 结合位点等典型的保守元件外,还包括光照、干旱、缺氧等环境因子响应元件,生长素、赤霉素等植物激素响应元件,芽、胚乳等组织特异性表达元件等。为了初步研究 *SrUGT76G1* 基因启动子的功能,本试验通过克隆 *SrUGT76G1* 转录起始位点上游 1 989 bp 带酶切位点的序列,将其替换植物表达载体 pCambia1301-220 中的 *CaMV35S* 启动子,连接下游的 *GUS* 报告基因,构建重组植物表达载体 pCambia1301-220-*SrUGT76G1P*,以 pCambia1301-220 载体作对照,通过农杆菌(EHA105)真空渗透法转入拟南芥和甜菊幼苗。瞬时表达结果表明,该 *SrUGT76G1* 启动子序列能驱动 *GUS* 基因在拟南芥和甜菊叶片中表达,具有启动子活性。

关键词:甜菊;*SrUGT76G1* 启动子;调控元件;植物表达载体;基因瞬时表达

中图分类号: S566.901 **文献标志码:** A **文章编号:** 1002-1302(2018)06-0027-07

甜菊(*Stevia rebaudiana* Bertoni)又称甜叶菊,是菊科(Asteraceae)甜菊属(*Stevia*)多年生草本植物,原产于南美洲巴拉圭等地。甜菊因其叶片可提取高甜度低热量的甜菊糖苷而从 20 世纪 70 年代开始受到人们的广泛关注^[1]。甜菊叶片所含甜菊糖苷有多达 30 种组分,各组分的结构相似,性质迥异。甜菊苷(stevioside, St)和莱鲍迪苷 A(rebaudioside A, R-A)是两大主要组分,占总糖苷含量的 60%~80%。St 苷的甜度为蔗糖的 250~300 倍,略带苦味,具有药用价值, R-A 苷的甜度为蔗糖的 350~450 倍,且具有更接近蔗糖的甜味^[2],因此在甜菊种质创新研究中,为满足不同应用目的的需要,提高 St 苷、R-A 苷等单一糖苷的生物积累及其在总苷中的比例成为甜菊育种学家竞相研究的热点。

在甜菊糖苷生物合成途径中,葡萄糖苷转移酶 *SrUGT76G1* 专一地对 St 苷的 C-13 位 C-3' 进行糖基化,将其转化为 R-A 苷^[3],所以 *SrUGT76G1* 是 St 苷和 R-A 苷合成的关键基因。研究该基因的表达调控对了解 R-A 苷的生物合成和今后通过基因工程等手段人为调控 R-A 苷的合成具有重要意义,但是目前对 *SrUGT76G1* 的表达调控模式还不清楚。启动子作为基因表达调控的重要区域,研究

SrUGT76G1 启动子对于了解这一关键基因的表达调控具有重要意义,但在甜菊研究中目前尚无相关报道。

本试验在 *SrUGT76G1* 基因研究的基础上,根据已克隆到的 *SrUGT76G1* 基因 DNA 序列设计嵌套引物,扩增基因上游未知序列,通过启动子序列分析初步探究序列上的功能元件。为了更深入地研究该基因的表达调控,进一步构建植物表达载体,用 1 989 bp 的 *SrUGT76G1* 启动子取代 pCambia1301-220 中的 *CaMV35S* 组成型启动子,连接 *GUS* 报告基因,构建 pCambia1301-220-*SrUGT76G1P* 载体,通过农杆菌介导的基因瞬时表达的方法初步验证 *SrUGT76G1* 启动子的活性,为今后深入研究 *SrUGT76G1* 基因的表达调控和甜菊分子育种奠定基础。

1 材料与方法

1.1 材料

试验于 2014 年 7 月至 2015 年 6 月在江苏省中国科学院植物研究所进行。用于 DNA 提取的甜菊叶片取自本实验室甜菊种质资源圃。植物表达载体 pCambia1301-220 由南京农业大学园艺学院菊花实验室惠赠;根癌农杆菌(*Agrobacterium tumefaciens*) EHA105、大肠杆菌(*Escherichia coli*)菌株 TOP10、琼脂糖凝胶回收试剂盒、DNA Marker、测序载体 pMD19-T Vector、Taq DNA polymerase、各种限制性核酸内切酶与试剂购自 TaKaRa 公司;氨苄青霉素(Amp)、卡那霉素(Kan)、IPTG、X-gal、Real Master Mix SYBR Green I、质粒提取试剂盒购自北京天根生化科技公司;GUS 活性检测试剂购自 Sigma 公司;其他常用化学试剂为国产分析纯。

收稿日期:2017-10-29

基金项目:国家自然科学基金青年科学基金(编号:31601371);江苏省自然科学基金青年基金(编号:BK20160600)。

作者简介:杨永恒(1985—),女,陕西洋县人,博士,助理研究员,主要从事甜菊育种方面的研究。E-mail:yyh8576@126.com。

通信作者:佟海英,高级实验师,主要从事甜菊栽培繁育方面的研究。

E-mail:njtonghy@163.com。

1.2 方法

1.2.1 *SrUGT76G1* 启动子的克隆 采用改良 CTAB 法提取甜菊基因组 DNA。以提取的甜菊基因组 DNA 为模板,根据已知的 *SrUGT76G1* 序列设计 3 个嵌套的特异性引物 SP1、SP2、SP3。hiTAIL-PCR 反应程序参照 Liu 等的方法^[4]进行。第 1 轮 PCR 用特异性引物 SP1 与简并引物 LAD1-1、LAD1-2、LAD1-3、LAD1-4 组合,建立 4 个平行的 25 μL PCR 反应体系进行扩增;第 2 轮 PCR 将第 1 轮的 PCR 产物稀

释 50 倍,取 1 μL 作为反应模板,用简并引物 AC1 与特异性引物 SP2 扩增;第 3 轮 PCR 将第 2 轮的 PCR 产物稀释 50 倍,取 1 μL 作为反应模板,用简并引物 AC1 与特异性引物 SP3 扩增,经过 3 轮 PCR 扩增得到 *SrUGT76G1* 的上游序列。根据第一次步移得到的序列设计特异性引物 SP4、SP5、SP6,以相同方法进行第二次步移,将 2 次步移得到的序列拼接,并设计全长引物 pro-1F、pro-1R,最终得到 *SrUGT76G1* 的启动子序列。试验所用引物序列及反应程序详见表 1、表 2。

表 1 甜菊 *SrUGT76G1I* 启动子步移及分析引物

引物名称	编号	序列 (5'→3')
基因组步移引物	LAD1-1	ACGATGGACTCCAGAGCGGCCGCVNVNNNGGAA
	LAD1-2	ACGATGGACTCCAGAGCGGCCGCBNBNNNGGTT
	LAD1-3	ACGATGGACTCCAGAGCGGCCGCVNVNNNNCCAA
	LAD1-4	ACGATGGACTCCAGAGCGGCCGCBDBNBNNCGGT
	AC1	ACGATGGACTCCAGAG
	SP1	GAAACCCACTCGCTTGTCTTCCA
	SP2	ACGATGGACTCCAGTCCGGCCAAACACGATACCTCTCCATCTTCTCAG
	SP3	TGAGTCGGTAGATTGGAAATGCGT
	SP4	TTGTTATTGCCGATCCTTGTCTTATT
	SP5	GCGGCGCAAGAGGTAAGGTAAC
	SP6	GGAAATAACCAGATATTTGATGTAGCTAAACTTTATATG
启动子全长引物	pro-1F	CGGTATCGGAGTTCTTGTA
	pro-1R	GGTTTGTGGAAGTTGGTGTCG
<i>SrUGT76G1P</i> 重组质粒构建引物	pro-2F	CCGAGCTCAAAATCCAATCC
	pro-2R	CATGCCATGGGTTTGACTGAC

表 2 hiTAIL-PCR 反应程序

步骤	预扩增		第一次 TAIL-PCR			第二次 TAIL-PCR		
	温度(℃)	时间(min)	步骤	温度(℃)	时间(min)	步骤	温度(℃)	时间(min)
1	93	2	1	94	3	1	94	3
2	95	1	2	94	0.5	2	94	0.5
3	94	0.5	3	59	1	3	65	1
4	60	1	4	72	3	4	72	3
5	72	3	5	转到步骤 2	2 次循环	5	94	0.5
6	转到步骤 3	10 次循环	6	94	0.5	6	65	1
7	94	0.5	7	61	1	7	72	3
8	25	2	8	72	3	8	94	0.5
9	升温至 72	0.5℃/s	9	94	0.5	9	50	1
10	72	3	10	62	1	10	72	3
11	94	0.5	11	72	3	11	转到步骤 2	12 次循环
12	58	1	12	94	0.5	12	72	5
13	72	3	13	50	1	13	4	保存
14	转到步骤 11	24 次循环	14	72	3			
15	72	5	15	转到步骤 6	13 次循环			
16	4	保存	16	72	5			
			17	4	保存			

1.2.2 *SrUGT76G1* 启动子作用元件分析 将甜菊 *SrUGT76G1* 启动子序列提交 PLACE 服务器 (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>),预测该启动子保守区域中潜在的顺式作用元件。

1.2.3 植物表达载体构建 植物表达载体 pCAMBIA1301-220 是含有 *CaMV35S* 启动子和 *GUS* 基因系统的二元载体。为了研究 *SrUGT76G1* 启动子的功能,本研究中用 *SrUGT76G1*

启动子取代 pCAMBIA1301-220 中的 *CaMV35S* 组成型启动子,连接 *GUS* 报告基因,构建植物表达载体 pCAMBIA1301-220-*SrUGT76G1P*。将引物 pro-2F 和 pro-2R(引物序列见表 1,划横线处分别为限制性内切酶 *Sac* I 和 *Nco* I 的酶切位点)扩增得到的含酶切位点的 *SrUGT76G1* 启动子片段连接到 pMD19-T(2 692 bp)载体上,得到重组质粒 pMD19-T-*SrUGT76G1P*。将重组质粒 pMD19-T-*SrUGT76G1P* 和质粒

pCambia1301-220 分别用限制性内切酶 *Sac* I 和 *Nco* I 进行双酶切,酶切产物用 1% 琼脂糖凝胶电泳分离,分别切胶回收 pMD19-T-*SrUGT76GIP* 小片段和 pCambia1301-220 大片段,用 T4 DNA 连接酶连接 2 个目的片段,连接产物转化至大肠杆菌 TOP10 菌株感受态细胞中。将经蓝白斑筛选、PCR 检验和 *Sac* I、*Nco* I 双酶切鉴定都正确的重组质粒命名为 pCambia1301-220-*SrUGT76GIP*。

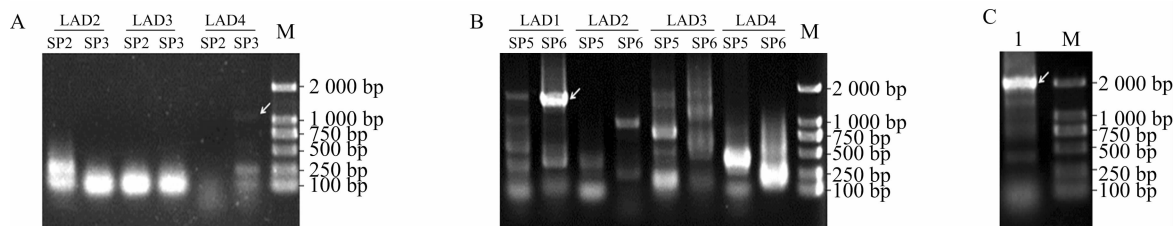
1.2.4 *SrUGT76G1* 启动子功能的瞬时表达分析 将重组质粒 pCambia1301-220-*SrUGT76GIP* 和质粒 pCambia1301-220 分别转化至农杆菌 EHA105 中^[5]。已转化的农杆菌在 5 mL 含利福平和卡那霉素的 YEP 液体培养基中 28 ℃ 振荡培养 24 h,取 2 mL 菌液转入 10 mL 含同样抗生素的新鲜培养液中,28 ℃ 振荡培养 36~48 h,离心培养物并用 5 mL MMA {1×MS,10 mmol/L MES [2-(*N*-morpholino)ethane sulfonic acid],200 μmol/L 乙酰丁香酮,pH 值 5.6} 重悬菌体,28 ℃ 保温 3 h,5 000 min 离心 10 min,沉淀用 2 mL 10 mmol/L 的 $MgCl_2$ 洗涤 1 次,菌体用 2 mL 10 mmol/L 的 $MgCl_2$ 悬起。取 14 d 左右的甜菊和拟南芥无菌幼苗分别浸没在含重组质粒 pCambia1301-220-*SrUGT76GIP* 和质粒 pCambia1301-220 的上述农杆菌菌液中,以 10 mmol/L $MgCl_2$ 溶液为空白对照,真空 (10 Pa) 处理 30 min,然后用无菌水将幼苗冲洗 3 次,放置在铺有湿润滤纸的培养皿中,

24℃、12 h/d 光照培养。分别于 2、4、6 d 后取样进行 *GUS* 染色。*GUS* 染色液的配制按照 Jefferson 等的方法^[6]进行,将待测材料和对照材料放入 1.5 mL 的离心管中,加入 *GUS* 染色液,37 ℃ 保温 24 h,将染色后的材料用 70%、80%、90%、100% 乙醇依次脱色,至空白对照为白色时用体视显微镜观察并拍照。

2 结果与分析

2.1 *SrUGT76G1* 启动子序列的获得

根据已知 *SrUGT76G1* 基因序列设计步移引物,采用 Liu 等的未知侧翼序列扩增的方法^[4],第一次步移克隆到了 1 107 bp 的片段,经测序和 NCBI blast 比对,结果表明所得序列的 3'端与数据库中 *SrUGT76G1* 基因 5'序列完全重叠,得到了 ATG 前 862 bp 的序列。根据克隆到的序列再设计 3 个嵌套引物继续向 5'端步移,得到 1 580 bp 片段,亚克隆后测序,比对结果显示其 3'端与第一次步移所得序列的 5'端重叠。将 2 次步移得到的序列进行拼接,并重新设计 1 对序列特异引物从甜菊基因组 DNA 中克隆完整的 *SrUGT76G1* 启动子区域,将扩增到的序列进行 NCBI blast,结果表明该序列 3'端与数据库中 *SrUGT76G1* 基因序列 5'端同源,最终得到 *SrUGT76G1* 基因 ATG 前 2 283 bp 的启动子序列 (图 1)。



A—第一次步移; B—第二次步移; C—启动子全长的扩增。箭头所指为获得的目的条带

图1 *SrUGT76G1* 启动子扩增电泳

2.2 启动子顺式作用元件预测和序列分析

将克隆得到的启动子全长序列提交 PLACE,进行顺式调控元件的分析,启动子序列正反向包含 475 个调控元件 (表 3),含有典型的启动子保守元件,有 21 个 TATA box;如: TATABOX2、TATABOX3、TATABOX4、TATABOX5 和 TATAPVTRNALEU;36 个 CAATBOX1 和 9 个 CCAATBOX1,它们是 CAAT box 的基序;2 个 ACGTTBOX,ACGTTBOX 是 1 个 ACGT 元件;有 2 个 AMYBOX1,它是 5'上游区域的保守序列;19 个 MYB 结合位点的同源序列,如: MYB1AT、MYB2AT、MYB2CONSENSUSAT、MYBATRD22、MYBCORE、MYBCORE ATCYCB1 和 MYBST1 等。有 12 个 MYCCONSensusAT 的同源序列,1 个 MYCATERD1 和 1 个 MYCATRD22,它们是 MYC 结合位点;发现有植物多腺苷酸化信号的保守序列,8 个 POLASIG1,3 个 POLASIG2 和 6 个 POLASIG3^[7]。

除了以上所述的保守元件,该启动子序列还含有多个植物激素相关的顺式调控元件,1 个 SURECOREATSULTR11,它是硫和生长素的响应元件^[8];ARFAT 是 ARF (植物生长素响应因子, auxin response factor) 的结合位点;2 个 ASF1MOTIFCAMV 同源序列,它是植物生长素和/或水杨酸响应元件,并参与光调控;3 个 WBOXATNPR1 的同源序列,参与

水杨酸 (SA) 信号转导;3 个 DPBFCOREDCDC3,是 bZIP 转录因子结合位点和受 ABA 诱导的^[9];含有 CPBCSPOR 同源序列,是细胞分裂素增强蛋白 (cytokinin-enhanced protein) 结合的关键序列,最早在黄瓜 POR (NADPH-protochlorophyllide reductase) 基因启动子序列中被发现^[10];以及乙烯响应元件 ERELEE4、GCCCORE、LECPLEACS2 和赤霉素响应元件 GARE1OSREPI、GAREAT、PYRIMIDINEBOXHVEPB1、PYRIMIDINEBOXOSRAMY1A 等。

该启动子序列还含有多个环境因子响应的顺式调控元件:有 2 个 CBFHV 同源序列,分别为 ABRELATERD1 和 DRECRTCOREAT,它们是脱水响应元件;6 个 ACGTATERD1 的同源序列,是 erd1 (early responsive to dehydration) 基因^[11]黄化诱导表达必需的。LTRECOREATCOR15 是在拟南芥 *cor15a* 基因启动子中发现的低温响应元件 LTRE 的核心基序^[12]。还有厌氧响应元件 ANAERO1CONSENSUS,以及一些光调控元件,如:GT1CONSENSUS、IBOXCORE、INRNTPSADB、REALPHALGLHCB21、SORLIP1AT 和 SORLIP2AT 等。其中 GT1CONSENSUS 是许多光调控基因保守的 GT-1 结合位点^[13-14]。顺式调控元件 10PEHVPSBD、CIACADIANLELHC 和 EVENINGAT 参与生理节律调控^[15-17]。

此外,还发现 *SrUGT76G1* 启动子序列含有一些组织特异性表达元件:有 22 个 GATABOX 的同源序列,它们参与光调控和组织特异性表达^[18];ACGTOSGLUB1 元件参与胚乳特异性表达;DPBFCOREDCDC3 参与胚特异性表达;NTBBF1ARROLB 参与组织特异性表达和植物生长素诱

导^[19];OSE1ROOTNODEULE 和 OSE2ROOTNODEULE 为器官特异性表达元件,在根瘤感染细胞中被激活;POLLEN1LELAT52 是 2 个相互依赖的花粉特异性表达的元件之一;RYREPEATBNNAPA 是种子特异性表达元件;TAAAGSTKST1 是保卫细胞特异表达调控元件。

表 3 *SrUGT76G1* 启动子序列上部分顺式调控元件

转录因子或位点名称	数目 (个)	功能描述
AMYBOX1	2	5'上游区保守序列
CAATBOX1	36	CAAT 启动子保守序列
CCAATBOX1	9	真核生物基因 5'非编码区共有序列
TATABOX4	6	TATA 框
TATABOX5	9	TATA 框
ACGTATERD1	6	黄化诱导 <i>erd1</i> (脱水的早期响应)表达的必需元件
CURECORECR	6	铜响应元件;氧响应元件
GT1CONSENSUS	24	许多光调控基因共有的 GT-1 结合位点.
LTRECOREATCOR15	2	低温响应元件;ABA 响应元件
MYB1AT	5	MYB 识别位点;脱水响应元件
MYB2CONSENSUSAT	3	脱水应答基因启动子中的 MYB 识别位点
MYCCONSSENSUSAT	12	MYC 识别位点;冷调控元件
SORLIP2AT	6	光诱导的启动子元件
SREATMSD	1	糖抑制元件(SRE);主茎截断后下调基因
ARFAT	1	生长素响应因子(ARF)结合位点
ASF1MOTIFCAMV	2	生长素和/或水杨酸响应元件;光调控元件
CPBCSPOR	2	细胞分裂素增强蛋白(cytokinin-enhanced Protein)结合位点
ERELEE4	3	乙烯响应元件
GARE1OSREPI	2	赤霉素响应元件
GAREAT	1	赤霉素响应元件
PYRIMIDINEBOXHVEPBI	1	受赤霉素诱导的
PYRIMIDINEBOXOSRAMY1A	3	赤霉素响应元件;糖抑制
SURECOREATSULTR11	2	硫响应元件;生长素响应元件
WBOXATNPRI	3	受水杨酸诱导的
WRKY71OS	11	赤霉素信号途径中的转录抑制子
DPBFCOREDCDC3	3	bZIP 转录因子的结合位点;胚胎特异,受 ABA 诱导的
NTBBF1ARROLB	2	组织特异性表达和生长素诱导所必需的
OSE1ROOTNODEULE	4	器官特异性表达的,在根瘤感染细胞中被激活
POLLEN1LELAT52	7	激活花粉特异性表达的 2 个互作调控因子之一
RYREPEATBNNAPA	2	种子特性性表达必需元件
TAAAGSTKST1	8	保卫细胞特性性表达必需元件
CIACADIANLELHC	2	节律性表达必需区域
EVENINGAT	1	夜晚响应元件;节律调控
GATABOX	22	高度,光调控和组织特异性表达必需元件

2.3 含双酶切位点 *SrUGT76G1* 基因启动子片段的获得

以含酶切位点的引物 pro-2F 和 pro-2R 克隆得到甜菊 *SrUGT76G1* 基因的启动子序列,其长度为 1 989 bp,提交至 NCBI GenBank,获得登录号为 KM206772。序列分析结果见图 2,除了 TATA box 等保守元件外,还有 MYB、MYC 识别位点,病原体和盐响应元件 GT1GMSCAM4,钙响应元件 ABRERATCAL,水杨酸响应元件 WBOXATNPRI,糖响应元件 WBOXHVISO1,赤霉素响应元件 GAREAT 和增强子 CTRMCAMV35S 等。

2.4 *SrUGT76G1* 基因启动子植物表达载体的构建

植物表达载体 pCAMBIA1301-220 是含有 *CaMV35S* 启动子和 *GUS* 基因系统的双元载体。为了验证所克隆的

SrUGT76G1 启动子序列是否具有活性,设计引物引入 *Sac* I 和 *Nco* I 酶切位点,用 *SrUGT76G1* 启动子取代 pCAMBIA1301-220 中的 *CaMV35S* 组成型启动子,连接 *GUS* 报告基因,构建植物表达载体 pCAMBIA1301-220-*SrUGT76G1P*(图 3)。根据启动子全长序列设计质粒构建引物 pro-2F(5'端带 *Sac* I 酶切位点)和 pro-2R(3'端带 *Nco* I 酶切位点),*Nco* I 酶切位点中的 ATG 是 *SrUGT76G1* 基因的起始密码子,同时也是载体上 *GUS* 基因起始密码子。

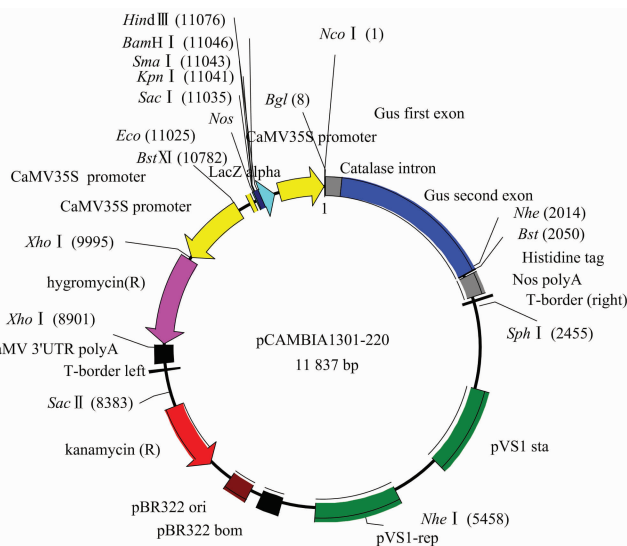
2.5 基因瞬时表达分析

采用农杆菌渗透法分别转化拟南芥和甜菊幼苗,被侵染的甜菊和拟南芥幼苗分别在不同的处理时间取出(1 d、2 d、3 d 和 4 d)进行 *GUS* 染色,结果用含质粒 pCAMBIA1301-

1 CCGAGCTCAAAATCCAATCCACTAAACCGTACTTGCTACATTGCGATTTCCCATTTAC
61 TATTAGATTTTGATAAGGAAGTTCGATTGGGTTTTCTCTCTCTCTCTGAATGATCT
121 TCGTTTTGTTCAATTCATCAGTGTATTGCAITTCATTAGCATTTTATGGGTTTTTG
181 AAGCGTAAGTAGTGCTTAGGATTATTCGACTAATCGATCGAACATAAATAGACCAG
241 TGTCAGGGTGAACAAGATGAATAAGAAATGAAGCATGATTTTGAGAAAAAGAATAACA
301 GAAGCATGAAATGAAGCAAAGATCAATCTGATAGTTTAAATCAATGGAAGCATATCTC
361 TCTTTGATTAGTTAGCATTTTCATTGTGTGTGCACTTCATCGCTGACTCTAAATTA
421 GGGTCGAATTGATTGAAATCTTTTCATAATCATCAGTGTTTCCAATTTATTTGAAAT
481 AGGGTCGAATTGATTGAAATGTTTCATGAGATAGTGATTATCTATGCCCAATATTTTC
541 TTTTCGAGTTTGTCAGTCGATAGAAAGAAATGGAATAAACCTAAGATAATCATTCTGTC
601 GACAGATCCACATTCATTTTTTTCGATACAGATTCAGTAGATTTTATGCATATATAG
661 TTGTACCGCTTGCAATTAATGAGATAACCACTCTGTTGAATTTGCTAAAATAATAGGATA
721 TATTCGAAGCGAGTCTCGACTACCTTAAAAATACGGGGCCCAAGCGAGAAAAA
781 TGGGCCCCCTAAAAAGCTCAATAGTTTATATCTAAAGATTAAATACCATATTAGTT
841 ATTAAAAATAAACAAACCAATGATTATAGCAAAATGATAAATTAACGCTTAAGC
901 TAATTGTGAGGCTCTCTATCATTTTTTGAAGCAAAATTCGATCAACTCTTTATAGTC
961 CATCTGTTTAAAGATAGCGTTCTCAATATATAGATATCATTTATCAATCGTTATGCTG
1021 CTGTGACATTTGCGATTGAAGTAAAGTAAAGATTTAATTTGAAAACTCTCTTCTGCAAT
1081 GCCACGCTTATTGGAATACTCAACAACTCTATATGCAATTAACGCATTAGGAATATAC
1141 CCAGCTCGCTGACAGCGTTTATGACATCAAGAGAGTGAACAATCTCATTAGATAAAAT
1201 GTCTCAACAACTTCACTCCATATAAAGTTTATGCTACATCAATATGTTTATTTCCCA
1261 TACTTCAATGCATCTTGAAGAAGATAACAACAGCATTAAGATCATTTTTCATTAAGTTG
1321 CTCAACTTATAGGAAAAAGAACCAATAGTTTTTCATATTATTGGTATTGTCGAATC
1381 TTGATTAAGGAGAAGCAATGTTTGTATCAACAATACATAAAAAAGGGGCTGTATAGAAG
1441 TGGGGCTTAAGCACAAAGCTTGGGTTACCTTACCTCTGGGCGCCCTAATCGAGTTCAA
1501 CAATTAAGCTCATCGATGATTATCTATAATCTACTTAACATAAACAAGAATGAAT
1561 GCTAGTTAACGTTTATCTTTTTTCCAACTAAGGTAGTAAGGCAATGGTTATGTAATA
1621 AACTAAATAAATAGAGTACAACTAAATAAATAAACCTTTATATAAATAATCTTTGCA
1681 TACAATAATATATAAATAAATAAAGATTATGGGAACTTTTTGTCATGGTTCAGAA
1741 TTATTGGAATCTTTGCTTTTATTTATTTTGAACAATACATAGATACACACAAGATC
1801 TTCAGGACCAATCATCTGTATAAAGAGCAAGGATCGGCAATAACAACCTGGACCACAC
1861 AAGATGTGCAATTTCCACATCGAGATCTCAGCTACCTAACCACTGTAAATTAATGAAT
1921 GCTTGCCGACACTATTTATTGTTATATATTTTAATAATTTCTGCGGTAAACGTCAGTC
1981 AAACCCATCGCATG

红色: 起始密码子; 蓝色框: TATA box; 1—MYB 识别位点; 2—MYC识别位点; 3—GT1GMSCAM4, 病原体和盐响应元件; 4—ABRERATCAL, 钙响应元件; 5—WBOXATNPRI, 水杨酸响应元件; 6—WBOXHVISOI, 糖响应元件; 7—GAREAT, 赤霉素响应元件; 8—CTRMCMAMV35S, 增强子

图2 甜菊 *SrUGT76G1* 基因启动子序列

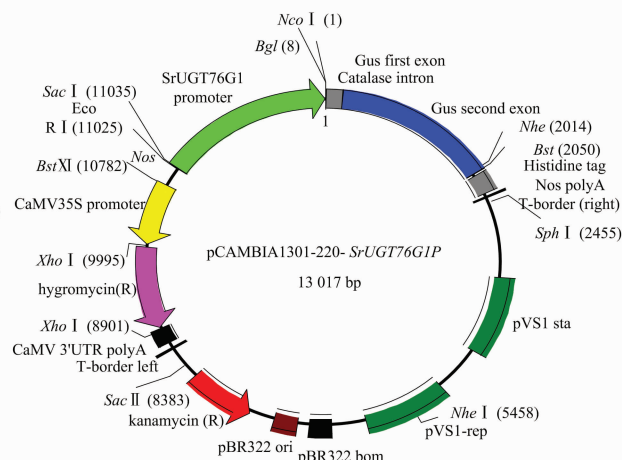


220 和 pCambia1301-220-*SrUGT76G1P* 的农杆菌侵染的幼苗都表现出 *GUS* 基因的有效表达(图4),说明所克隆到的 *SrUGT76G1* 基因启动子具有启动活性,可以用于下一步的研究。同时试验中发现,被侵染的甜菊幼苗在第3天 *GUS* 的表达达到高峰,随后组织褐化;而在拟南芥中24 h后逐渐增强并在第3、第4天持续表达。

3 讨论与结论

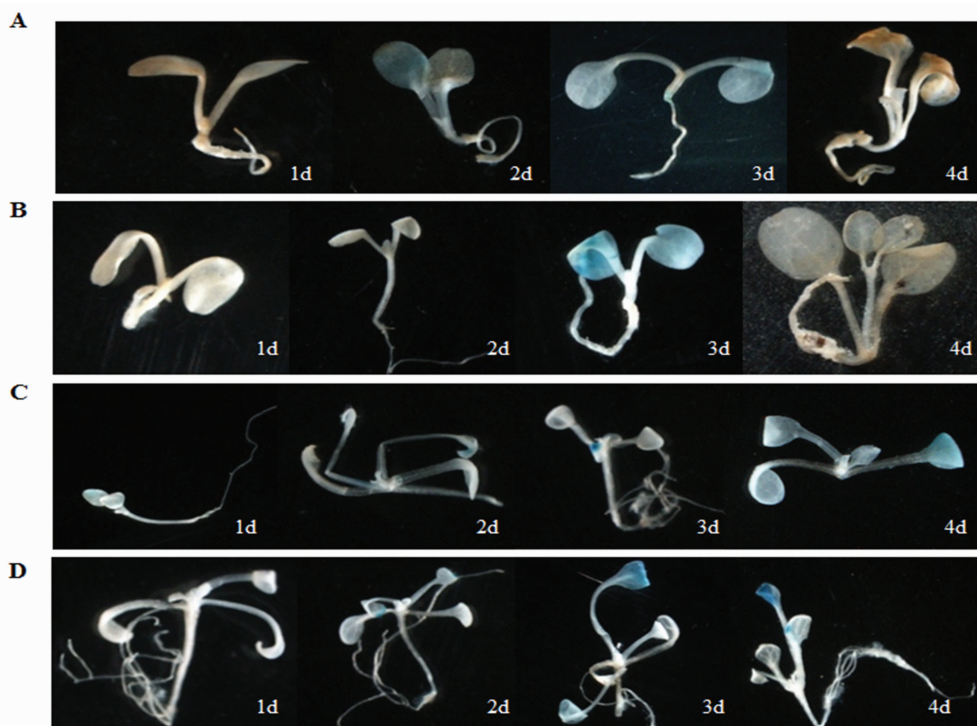
基因的表达调控是基因功能的基本部分,基因启动子上关键元件是其表达调控的概况^[20-21]。了解植物启动子非常有意义,并且在很多方面通过生物技术控制基因表达提供可能^[22-23]。尽管如此,目前甜菊基因启动子研究的报道还很少,仅有关于 *SrHDR* 基因启动子的研究^[24]。甜菊 *SrUGT76G1* 基因是 RA 昔合成的关键基因^[25-26],研究 *SrUGT76G1* 基因的表达调控对于了解植物体内甜菊糖苷积累的规律和今后通过生物技术手段进行甜菊育种都有重要的意义,但是目前尚无相关报道。本试验在研究 *SrUGT76G1* 基因的基础上,通过染色体步移的方法克隆了该基因的5'端侧翼序列,采用在线的启动子分析工具预测该 *SrUGT76G1* 启动子序列上的调控元件,结果显示双向序列包含475个调控元件,除了典型的启动子元件,如:TATA-box、CAAT-box、MYB结合位点等外,还有生长素、赤霉素、细胞分裂素、乙烯、ABA、水杨酸和茉莉酮酸酯等植物激素响应元件,和光、脱水、低温、厌氧等环境因子响应元件,以及芽、胚乳等组织特异性表达元件,可见 *SrUGT76G1* 表达调控的复杂性。

因为获得稳定的转基因植株耗时费力,为了快速验证克隆的启动子片段是否有活性,我们采用了农杆菌真空渗透的瞬时转化法^[27]。结果表明 *SrUGT76G1* 启动子可以驱动 *GUS* 报告基因的表达,证明克隆到的5'侧翼序列具有启动子活性。启动子克隆、序列分析和启动子活性初步验证只是 *SrUGT76G1* 基因表达调控研究的开始,这些顺式元件的功能和对 *SrUGT76G1* 基因表达的调控作用等还有待深入研究,如进一步构建一系列5'端缺失启动子的表达载体,通过对拟南



左图: 植物表达载体 pCambia1301-220; 右图: 重组载体 pCambia1301-220-*SrUGT76G1P*; 绿色部分代表取代 *CaMV35S* 的 *SrUGT76G1* 启动子

图3 试验所用植物表达载体的构建



A—被含质粒 pCambia1301-220 的农杆菌侵染的甜菊幼苗；B—被含质粒 pCambia1301-220-*SrUGT76GIP* 的农杆菌侵染的甜菊幼苗；C—被含质粒 pCambia1301-220 的农杆菌侵染的拟南芥幼苗；D—被含质粒 pCambia1301-220-*SrUGT76GIP* 的农杆菌侵染的拟南芥幼苗

图4 *GUS* 基因在甜菊和拟南芥幼苗中的瞬时表达

芥及甜菊中的遗传转化获得稳定的转基因植株,进而研究各顺式调控元件的功能及调控因子等。

甜菊糖苷的生物合成与赤霉素有着密切的关系,它们共用生物合成途径的起始阶段,直至共同的前体物质贝壳杉烯酸的合成^[28]。因此我们曾推测 *SrUGT76G1* 的表达受赤霉素调控,正如我们所期望的,启动子顺式调控元件预测的结果表明该序列上有 7 个赤霉素响应元件,包括 2 个 GARE1OSRE1、1 个 GAREAT、1 个 PYRIMIDINEBOXHVEPB1 和 3 个 PYRIMIDINEBOXOSRAMY1A。GARE1OSRE1 (TAACAGA) 是在水稻一个半胱氨酸蛋白酶 (*REP-1*) 基因启动子上发现的,该基因的表达受赤霉素的诱导^[29]。GAREAT 则是在拟南芥中发现的赤霉素响应元件^[30]；PYRIMIDINEBOXHVEPB1 是一个嘧啶盒 (Pyrimidine box),在大麦 EPB-1 (半胱氨酸蛋白酶) 基因启动子上发现,并且确定为赤霉素诱导必需的元件^[31]；PYRIMIDINEBOXOSRAMY1A 存在于水稻和大麦的 α -淀粉酶基因启动子上,是转录因子 GARE 的赤霉素响应顺式调控元件,并部分参与糖抑制^[32-33]。尽管在 Kumar 等的研究中 *SrUGT76G1* 基因的转录水平在 GA_3 处理下并没有显著变化^[24],但是鉴于甜菊糖苷与赤霉素共用很长一段生物合成途径,赤霉素对 *SrUGT76G1* 基因的表达调控还需深入研究。

在 Kumar 等的研究中发现甜菊糖苷主要在叶片中积累,尤其是顶部向下第三对叶,花和茎中含量次之,根中几乎没有,甜菊糖苷合成途径相关基因的转录水平也有相似的结果,表现出很强的组织相关性^[24]。通过对 *SrUGT76G1* 启动子顺式元件的分析,确实发现该序列上存在一些组织特异性表达元件,且大多是芽特异性表达相关的。由于甜菊叶片中糖苷

含量在现蕾期达到最高,开花后逐渐下降^[34-35],我们推测一些关键基因受生理节律的调控。在 *SrUGT76G1* 启动子序列中也确实发现有生理节律调控相关元件,如: -10PEHVPSBD、CIACADIANLHC 和 EVENINGAT。

目前,植物基因工程中常用的启动子仍以 CaMV35S 为主,CaMV35S 是组成型启动子,不表现出时间或空间的特异性。组成型启动子驱动外源基因持续、稳定地表达,可能会使植物消耗过多的能量和物质,而且在研究中发现用相同的组成型启动子同时驱动 2 个或 2 个以上外源基因的表达可能引起基因沉默或共抑制^[36]；所以克隆甜菊自身基因的启动子,并研究其启动子的功能对未来甜菊的分子育种具有重要参考价值。

参考文献:

- [1] 舒世珍. 甜菊引种三十年[J]. 中国种业,2010(6):21-23.
- [2] Yadav S K, Guleria P. Steviol glycosides from *Stevia*: biosynthesis pathway review and their application in foods and medicine [J]. Critical Reviews in Food Science and Nutrition,2012,52(11):988-998.
- [3] Richman A, Swanson A, Humphrey T, et al. Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana* [J]. Plant Journal,2005,41(1):56-67.
- [4] Liu Y G, Chen Y L. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences [J]. Biotechniques,2007,43(5):649-654.
- [5] Holsters M, de Waele D, Depicker A, et al. Transfection and transformation of *Agrobacterium tumefaciens* [J]. Molecular & General

- Genetics, 1978, 163(2): 181–187.
- [6] Jefferson R A, Kavanagh T A, Bevan M W. Gus fusions – beta – glucuronidase as a sensitive and versatile gene fusion marker in higher – plants[J]. The EMBO Journal, 1987, 6(13): 3901–3907.
 - [7] Heidecker A G, Messing J. Structural analysis of plant genes[J]. Annual Review of Plant Physiology, 1986, 37(1): 439–466.
 - [8] Maruyama – Nakashita A, Nakamura Y, Watanabe – Takahashi A A, et al. Identification of a novel cis – acting element conferring sulfur deficiency response in *Arabidopsis* roots[J]. The Plant Journal, 2005, 42(3): 305–314.
 - [9] Kim S Y, Chung H J, Thomas T L. Isolation of a novel class of bZIP transcription factors that interact with ABA – responsive and embryo – specification elements in the Dc3 promoter using a modified yeast one – hybrid system[J]. The Plant Journal, 1997, 11(6): 1237–1251.
 - [10] Fusada N, Masuda T, Kuroda H, et al. Identification of a novel cis – element exhibiting cytokinin – dependent protein binding *in vitro* in the 5' – region of NADPH – protochlorophyllide oxidoreductase gene in cucumber[J]. Plant Molecular Biology, 2005, 59(4): 631–645.
 - [11] Simpson S D, Nakashima Kazuo, Narusaka Y, et al. Two different novel cis – acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark – induced senescence[J]. The Plant Journal, 2003, 33(2): 259–270.
 - [12] Baker S S, Wilhelm K S, Thomashow M F. The 5' – region of *Arabidopsis thaliana cor15a* has cis – acting elements that confer cold – regulated, drought – regulated and ABA – regulated gene – expression[J]. Plant Molecular Biology, 1994, 24(5): 701–713.
 - [13] Terzaghi W B, Cashmore A R. Light – regulated transcription[J]. Annual Review of Plant Physiology and Plant Molecular Biology, 1995, 46: 445–474.
 - [14] Zhou D X. Regulatory mechanism of plant gene transcription by GT – elements and GT – factors[J]. Trends in Plant Science, 1999, 4(6): 210–214.
 - [15] Piechulla B, Merforth N, Rudolph B. Identification of tomato Lhc promoter regions necessary for circadian expression [J]. Plant Molecular Biology, 1998, 38(4): 655–662.
 - [16] Harmer S L, Hogenesch L B, Straume M, et al. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock [J]. Science, 2000, 290(5499): 2110–2113.
 - [17] Thum K E, Kim M, Morishige D T, et al. Analysis of barley chloroplast *psbD* light – responsive promoter elements in transplastomic tobacco[J]. Plant Molecular Biology, 2001, 47(3): 353–366.
 - [18] Reyes J C, Muro – Pastor M I, Florencio F J. The GATA family of transcription factors in *Arabidopsis* and rice[J]. Plant Physiology, 2004, 134(4): 1718–1732.
 - [19] Baumann K, de Paolis A, Costantino P, et al. The DNA binding site of the Dof protein NtBBF1 is essential for tissue – specific and auxin – regulated expression of the *roIB* oncogene in plants[J]. Plant Cell, 1999, 11(3): 323–334.
 - [20] 张春晓, 王文棋, 蒋湘宁, 等. 植物基因启动子研究进展[J]. 遗传学报, 2004, 31(12): 1455–1464.
 - [21] 雒雅婧, 李 杰, 张 爽, 等. 植物启动子研究进展[J]. 北方园艺, 2015(22): 186–189.
 - [22] Lescot M, Déhais P, Thijs G, et al. PlantCARE, a database of plant cis – acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences[J]. Nucleic Acids Research, 2002, 30(1): 325–327.
 - [23] Shahmuradov I A, Gammerman A J, Hancock J M, et al. PlantProm: a database of plant promoter sequences [J]. Nucleic Acids Research, 2003, 31(1): 114–117.
 - [24] Kumar H, Kumar S. A functional (E) – 4 – hydroxy – 3 – methylbut – 2 – enyl diphosphate reductase exhibits diurnal regulation of expression in *Stevia rebaudiana* (Bertoni) [J]. Gene, 2013, 527(1): 332–338.
 - [25] 郭书巧, 杨郁文, 倪万潮. 甜叶菊葡萄糖基转移酶基因 *UGT76G2* 的克隆及生物信息学分析[J]. 基因组学与应用生物学, 2009, 28(3): 422–428.
 - [26] Yang Y H, Huang S Z, Han Y L, et al. Base substitution mutations in uridinediphosphate – dependent glycosyltransferase *76G1* gene of *Stevia rebaudiana* causes the low levels of rebaudioside A; mutations in *UGT76G1*, a key gene of steviol glycosides synthesis [J]. Plant Physiology and Biochemistry, 2014, 80: 220–225.
 - [27] Cook D, Dreyer D, Bonnet D, et al. Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti* [J]. Plant Cell, 1995, 7(1): 43–55.
 - [28] Humphrey T V, Richman A S, Menassa R, et al. Spatial organisation of four enzymes from *Stevia rebaudiana* that are involved in steviol glycoside synthesis [J]. Plant Molecular Biology, 2006, 61(1/2): 47–62.
 - [29] Sutoh K, Yamauchi D. Two cis – acting elements necessary and sufficient for gibberellin – upregulated proteinase expression in rice seeds [J]. Plant Journal, 2003, 34(5): 635–645.
 - [30] Ogawa M, Hanada A, Yamauchi Y, et al. Gibberellin biosynthesis and response during *Arabidopsis* seed germination [J]. Plant Cell, 2003, 15(7): 1591–1604.
 - [31] Cercós M, Gómezcadenas A, Ho T H. Hormonal regulation of a cysteine proteinase gene, *EPB – I*, in barley aleurone layers: cis – and trans – acting elements involved in the co – ordinated gene expression regulated by gibberellins and abscisic acid [J]. Plant Journal, 1999, 19(2): 107–118.
 - [32] Morita A, Umemura T A, Kuroyanagi M, et al. Functional dissection of a sugar – repressed α – amylase gene (*RAmy1A*) promoter in rice embryos [J]. FEBS Letters, 1998, 423(1): 81–85.
 - [33] Mena M, Cejudo F J, Isabellamonedá I, et al. A role for the DOF transcription factor BPBF in the regulation of gibberellin – responsive genes in barley aleurone [J]. Plant Physiology, 2002, 130(1): 111–119.
 - [34] Bondarev N I, Sukhanova M A, Reshetnyak O V, et al. Steviol glycoside content in different organs of *Stevia rebaudiana* and its dynamics during ontogeny [J]. Biologia Plantarum, 2004, 47(2): 261–264.
 - [35] Ceunen S, Geuns J M. Spatio – temporal variation of the diterpene steviol in *Stevia rebaudiana* grown under different photoperiods [J]. Phytochemistry, 2013, 89: 32–38.
 - [36] Kumpatla S P, Hall T C. Recurrent onset of epigenetic silencing in rice harboring a multi – copy transgene [J]. Plant Journal, 1998, 14(1): 129–135.