

文章编号: 1674-7054(2011)04-0291-06

棉花蔗糖合成酶3基因第1个内含子在基因表达中的作用

阮孟斌, 于晓玲, 彭 明

(中国热带农业科学院 热带生物技术研究所在海南 海口 571101)

摘 要: 为了研究棉花 *SUS3* 基因编码区的第1个内含子在基因表达调控中的作用, 分离该内含子并将其插入到 *GUS* 基因编码区中构建载体 *Susfig*: : 121, 将载体 *Susfig*: : 121 通过农杆菌介导转入拟南芥。对转基因拟南芥进行 *GUS* 活性分析, 发现该内含子对 *GUS* 基因在拟南芥中的表达具有负调控作用, 并且特异抑制报告基因在花粉中表达。其结果显示, 棉花 *SUS3* 基因编码区的第1个内含子对基因的表达具有一定的调控作用。

关键词: 棉花; *SUS3*; 内含子; 基因表达调控

中图分类号: Q 786

文献标志码: A

棉花蔗糖合成酶3基因(*SUS3*)在分解碳水化合物和胚珠发育中起着重要作用^[1]。它一方面可催化蔗糖分解为果糖和UDP-葡萄糖, 为纤维素的合成提供底物; 另一方面, 蔗糖的分解能提高细胞内的膨压, 且这一作用已被证明了与纤维细胞伸长密切相关^[2]。在开花后1天的胚珠中, 对 *SUS3* 蛋白进行原位杂交时, 发现其只在棉花胚珠表皮层表达, 且 mRNA 的丰度非常高, 这些特点说明 *SUS3* 基因与棉纤维细胞的分化有非常密切的关系^[3]。将一个无毛突变体与野生型的棉花进行对比分析发现, 野生型的胚珠中 *SUS3* 基因的 mRNA 丰度远远高于无毛突变体的胚珠^[4]。通过 RNA 干扰实验抑制 *SUS3* 的转录后, 出现了破坏棉花纤维细胞起始、抑制纤维伸长和影响种子发育, 这几个对棉花生长发育不利的性状^[5-6]。

因为决定细胞命运的基因功能与基因的表达模式密切相关, 所以基因的表达调控显得尤为重要。研究与棉纤维发育相关基因的表达, 有助于了解棉花纤维起始及发育的分子机制。棉花 *SUS3* 基因在棉纤维细胞起始阶段起着非常重要的作用, 因此, 研究该基因的表达调控有着重要的意义。近几年的相关研究表明, 内含子在基因表达调控中具有相当重要的作用^[7-12]。绝大多数真核生物基因都含有一个或多个内含子, 在 mRNA 的成熟过程中, 这些内含子序列的删除和外显子序列的连接的过程被称为前体 mRNA 的剪辑。内含子在前体 mRNA 中被剪辑的过程既是基因表达的一个重要环节, 也是基因表达调控的一个关键步骤。在本实验中, 笔者获得了 *SUS3* 基因的 DNA 全长序列, 并对其基因结构特征进行分析, 同时将基因编码区的第一个内含子分离出来后, 插入到 *GUS* 基因的编码区中, 构建相应的表达载体。通过农杆菌介导法将表达载体转入拟南芥中, 对内含子的功能进行分析, 探讨内含子在 *SUS3* 基因表达中的调控作用, 为进一步研究该基因的表达调控提供一定的基础。

1 材料与方法

1.1 棉花 *SUS3* 基因 DNA 全长克隆及结构分析 实验所用的棉花为陆地棉(*Gossypium hirsutum*) 种植于中国热带农业科学院热带生物技术研究所在实验基地。取棉花的嫩叶, 以改良 CTAB 法(2% CTAB, 4% PVP) 提取总 DNA。根据 *SUS3* 基因的 cDNA 序列(U73588) 合成用于 PCR 反应的 1 对引物 SSP1 和 SSP2

收稿日期: 2011-10-19

基金项目: 国家转基因生物新品种培育重大专项资助项目(2009ZX08005-027B)

作者简介: 阮孟斌(1979-) 男, 江西广丰人, 中国热带农业科学院热带生物技术研究所在助理研究员, 博士后。

通信作者: 彭明, 中国热带农业科学院热带生物技术研究所在研究员, E-mail: mmpeng_2000@yahoo.com

(见表1)。采用宝生物公司提供的 *LA Taq* 酶进行 PCR。反应后的产物经凝胶电泳回收后,与 pMD18-T 载体连接并转化大肠杆菌 Top10 感受态细胞,挑取经鉴定后的阳性克隆进行测序。用 Vector NTI 10 中的 AlignX 将所获得的棉花 *SUS3* 基因的 DNA 序列与 Genbank 中的 cDNA 序列进行比对,并分析基因的结构。

表 1 *SUS3* 基因 DNA 全长克隆和编码区第 1 个内含子相关载体构建的引物

引物	序列 (5'→3')
SSP1	AACGAGGTTTGGCCTTGCTC
SSP2	TTACTCCTCTGCCAATGGAA
121figP1	CCCTTATGTTACGTCCTGTAGTAATTTAAATATAACCCC
121figP2	CACGGGTTGGGGTTTCCTACAATAGAAAACAACAAAAC
121figP3	GTTTTGTTGTTTTCTATTGTAGGAAACCCCAACCCGTG
121figP4	CGAGCTCGGTAGCAATTCCCGAG (<i>Sac</i> I)

1.2 *SUS3* 基因编码区第 1 个内含子分离及载体构建 根据 *SUS3* 基因的 DNA 序列信息与结构分析的结果,合成 1 对引物 121figP1 和 121figP2(见表 1)用于 PCR 分离编码区的第 1 个内含子,在 121figP1 的 5' 端有 15 bp 的片段为 *GUS* 编码区的 5' 序列,121figP2 的 5' 端有 16 bp 的片段为 *GUS* 基因编码区中 16~31 bp 处的序列。121figP3 与 121figP2 为互补序列,121figP4 为 *GUS* 编码区 3' 端引物,引物的 5' 端带有 *Sac*I 酶切位点(见表 1)。采用宝生物公司的 *Pyrobest Taq* 酶,以 121figP1 和 121figP2、121figP3 和 121figP4 为引物,分别对含 *SUS3* 基因的 DNA 全长序列的质粒和 *pBI121* 质粒进行第 1 轮 PCR 反应。回收内含子和 *GUS* 基因的 PCR 产物后,将这 2 个片段按摩尔比 1:1 混合后作为第 2 轮 PCR 的模板,以 121figP1 和 121figP4 为引物进行 PCR 反应。回收反应产物,与 pMD18-T 载体连接,并转化大肠杆菌 Top10 感受态细胞,挑取经鉴定后的阳性克隆进行测序。以 *Bam*HI 和 *Sac*I 对带有内含子成功插入 *GUS* 编码区片段(121fig)的质粒进行双酶切,回收 121fig 片段,替换载体 *pBI121* 中的 *GUS* 基因,构建成载体 *Susfig::121*,测序验证构建是否正确。

1.3 拟南芥转化及筛选 将经测序表明构建正确的 *Susfig::121* 载体,通过农杆菌介导的 DIP 法转化野生型拟南芥 Col-0 (*Arabidopsis thaliana* Heyhn. ecotype Columbia)。转化用的野生型拟南芥为本实验室在培养房中培养的,培养条件:以蛭石为基质,温度控制在 20~22℃,光照时间为 12 h,光强为 4 000 lx。收集转化拟南芥的种子,消毒后在选择培养基(1/2MS,10 g 蔗糖,8 g 琼脂粉,25 mg·L⁻¹潮霉素 B,pH 5.7)上筛选抗性苗(T1)。待转基因 T1 代拟南芥结籽后,收取种子,并在选择培养基进行抗性分离比筛选,在符合分离比 3:1 的株系中随机选取 8 个株系用于随后的实验分析,以转化载体 *pBI121* 的转基因拟南芥作为对照。

1.4 *GUS* 活性分析 *GUS* 组织化学染色及荧光定量分析按照 Jefferson 等^[13]的方法进行。将材料浸泡在染色液(50 mmol·L⁻¹磷酸钠缓冲溶液 pH7.0,0.1 mmol·L⁻¹ K₃[Fe(CN)₆],0.1 mmol·L⁻¹ K₄[Fe(CN)₆],10 mmol·L⁻¹ Na₂EDTA,φ=0.001% Triton X-100,φ=20%的甲醇,0.5 g·L⁻¹ X-Gluc)中,于 25~37℃保温 1 h 至过夜,然后转入 φ=70%的乙醇中脱色 2~3 次至阴性对照材料呈白色后,在显微镜下观察,使用 Olympus SZX12 体视显微镜和 Nikon Eclipse 80i 荧光显微镜照相。

2 结果与分析

2.1 棉花 *SUS3* 基因 DNA 全长克隆 根据 Genbank 中的棉花 *SUS3* 基因的 cDNA 序列(U73588)合成了引物 SSP1 和 SSP2 作为 PCR 引物,以获取 *SUS3* 基因的 DNA 全长片段和 PCR 反应产物,以 *w*=0.8% 的琼脂糖凝胶进行电泳分析,结果显示获得了大小约为 3.5 kb 的 DNA 片段(见图 1)

2.2 棉花 *SUS3* 基因结构及第 1 个内含子序列分析 测序结果表明,棉花 *SUS3* 基因的 Genomic DNA 全长为 3 458 bp,将测序结果与 *SUS3* 的 cDNA 序列进行比对分析,分析软件为 Vector NTI Suite 10 软件包中的 AlignX。比对结果显示,所获得的片段即为棉花 *SUS3* 的基因组 DNA 序列(具体比对数据未显示)。在不考虑其 5' 非翻译区的情况下,该基因具有 12 个外显子,其中有 2 个在 400 bp 以上(见图 2)。蔗糖合成酶的蛋白

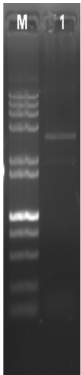


图 1 PCR 获取棉花 *SUS3* 基因的 DNA 全长
M 表示 DL15 000 bp +2 000 bp marker,泳道 1 为 PCR 产物

结构可分为两部分,即蔗糖合成酶和糖基转移酶。根据这一特点,棉花 *SUS3* 基因从结构上可清楚地划分出这 2 个不同的部分,其中第 1~9 个外显子编码蔗糖合成酶,而第 10 个外显子单独编码糖基转移酶。

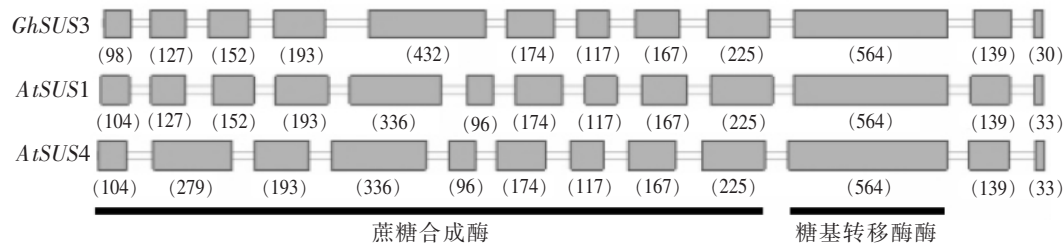


图 2 棉花 *SUS3* 基因与拟南芥中的 2 个蔗糖合成酶基因结构比较
(深色条形框表示外显子,中间连线为内含子,数字表示长度)

将棉花 *SUS3* 的基因结构与拟南芥蔗糖合成酶家族基因结构进行对比分析,其基因结构在外显子的数目和长度上与 *AtSUS1*、*AtSUS4* 具有非常高的相似性,具体分析结果如图 2 所示。棉花 *SUS3* 基因中,第 5 个外显子长度为 432 bp,而在拟南芥中,这个外显子被内含子分割成 336 bp 和 96 bp 的 2 个外显子。

根据棉花 *SUS3* 基因的 DNA 序列与其 cDNA 序列的比对结果,可以确定第 1 个内含子的序列(见图 3)。如图 3 所示,灰色框内为第 1 个内含子序列,从 GT 开始到 AG 结束,符合真核生物内含子序列的特点。在内含子的 GT 之前,是 98 个碱基的第 1 个外显子,AG 之后是 127 个碱基的第 2 个外显子。

ATGGCAGAGCGTGTGATCACCCGAGTACACAGCCTCCGGGAGCGTCTGGATGACACCC
TTATTGCCCATAGAAACGAGGTTTTGGCCTTGCTCACAAG GTAATTTAAATATAACCCCTT
CAACGATTGCTTTTCCTTTTGGAAATGCTTAGTTTTGTTGTTTTCTATTGTAGGATCGAGG
GTAAGGGAAAAGGGATTCCC

图 3 棉花 *SUS3* 基因第 1 个内含子序列(下划线为第 1 个内含子序列)

2.3 棉花 *SUS3* 基因编码区第 1 个内含子分离及载体构建 根据第 1 个内含子在棉花 *SUS3* 基因编码区中所处的相对位置,按图 4A 所示,将第 1 个内含子(F_1)插入到 *GUS* 基因编码区的 +15 bp 处。先以 PCR 的方式获得 *GUS* 基因片段和棉花 *SUS3* 第 1 个内含子片段后,再用套叠 PCR 的方法将 2 个片段拼接起来,结果见图 4B。图 4B 所示 *Bam*HI 和 *Sac*I 双酶切结果显示,带有内含子的 *GUS* 基因片段成功替换了 *pBI121* 载体上的 *GUS* 基因,成功构建了 *Susfig::121* 载体。随后的测序结果进一步验证了载体构建的正确性。

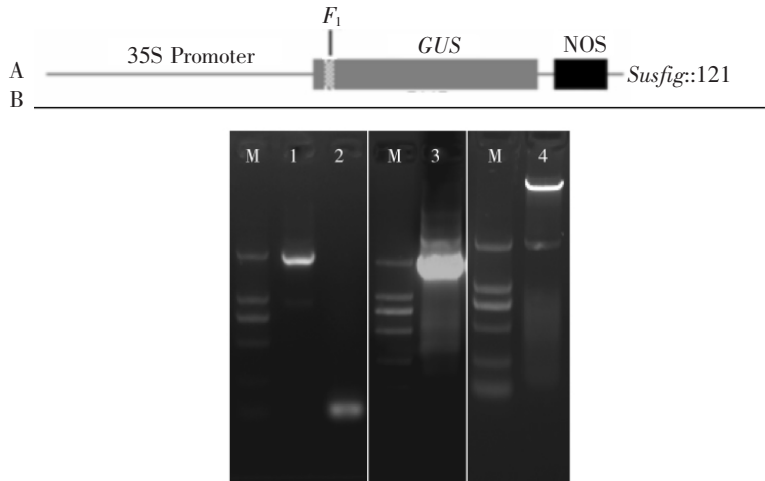


图 4 棉花 *SUS3* 基因编码区第 1 个内含子分离及载体构建

(A: 载体构建示意图, F_1 代表棉花 *SUS3* 基因的第 1 个内含子; B: 载体构建, M 表示 D 2 000 marker, 泳道 1 是以 121figP3 和 121figP4 对 *GUS* 基因进行 PCR 的产物, 泳道 2 是 PCR 分离第 1 个内含子的结果, 泳道 3 是内含子与 *GUS* 基因大片段连接结果, 泳道 4 是以 *Bam*HI 和 *Sac*I 对 *Susfig::121* 进行双酶切的结果)

2.4 转基因拟南芥的 GUS 活性分析 分别对 *Susfig::121* 和 *pBI121* 的转基因拟南芥的 14 d 幼苗、20 d 叶片及花器进行组织化学染色,完全脱色后在显微镜下观察。比较 *Susfig::121* 和 *pBI121* 转基因拟南芥的染色情况时发现, *GUS* 在维管组织、叶片、毛刺以及根部的表达都没有组织上的差异(见图 5-a, b, e, f),但在花器上的表达却有明显的区别,从图 5-c、图 5-d 和图 5-i、图 5-g 的对比中可以看出, *Susfig::121* 转基因拟南芥中 *GUS* 活性在花粉中的表达被完全抑制,因此,笔者认为棉花 *SUS3* 基因的第 1 个内含子具有特异抑制基因在拟南芥花粉中表达的功能。

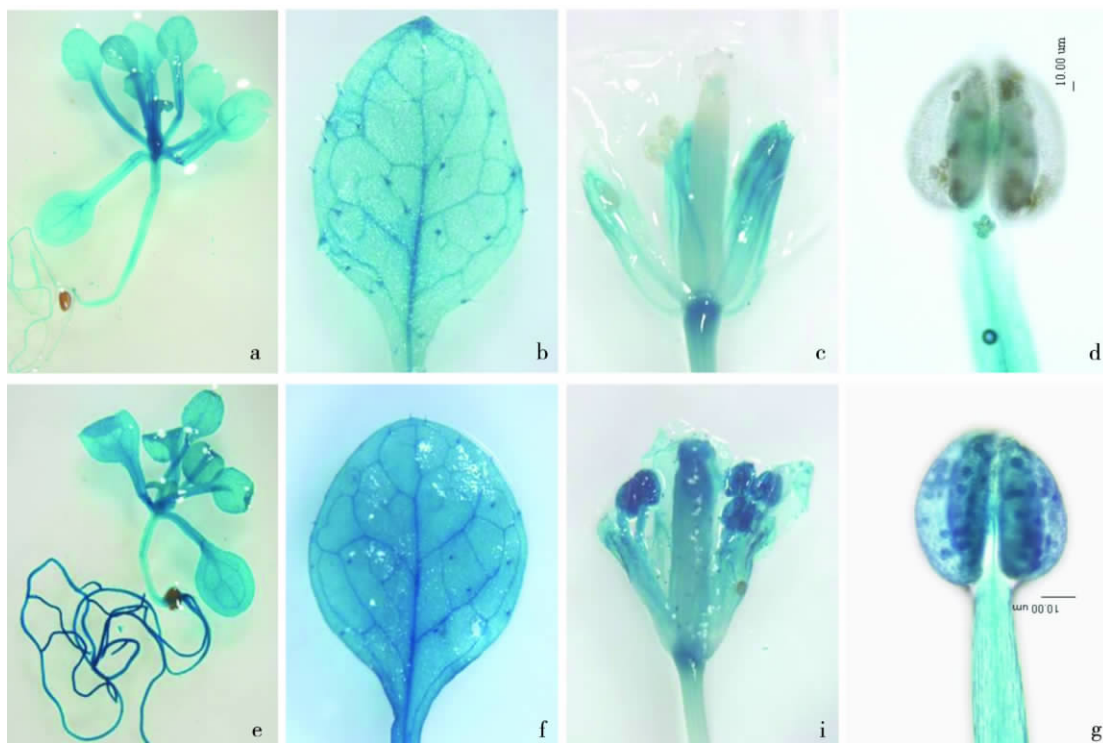


图 5 转基因拟南芥 *GUS* 组织化学染色分析

a, b, c, d 分别为 *Susfig::121* 转基因拟南芥的幼苗、叶片、花、雄蕊;
e, f, i, g 分别为 *pBI121* 转基因拟南芥的幼苗、叶片、花、雄蕊

3 讨 论

近年来,有许多基因的内含子被发现具有调控基因表达的功能。Kim 等^[10]在康乃馨的 S-腺苷甲硫氨酸脱羧酶(*SAMDC*)基因中发现其 5'非翻译区的内含子缺失后,该基因启动子驱动报告基因在花粉中的表达特性消失。拟南芥线粒体细胞色素 c 氧化酶第 5 亚基(*AtCOX5c*) 5'非翻译区中的内含子被移除后,该基因的启动子具有了在花粉特异启动报告基因表达的特性^[15]。对拟南芥 *SVUH3* 基因的启动子及 5'非翻译区中内含子的研究也表明了内含子具有决定基因表达组织特异性的功能^[16]。将本实验获得的 *SUS3* 基因进行基因结构分析,并与拟南芥中的 *SUS* 基因家族在基因结构上进行比较时,发现棉花 *SUS3* 与拟南芥的 *SUS1* 和 *SUS4* 在内含子数目与长度上有非常高的相似性,因此,推测棉花 *SUS3* 与拟南芥的 *SUS1* 和 *SUS4* 属于同源基因。同时还发现,棉花 *SUS3* 基因编码区的第 1 内含子与 *AtCOX5c* 的内含子一样,可特异抑制基因在花粉中表达,具有调控基因表达组织特异性的功能,与 *AtCOX5c* 和 *AtSVUH3* 基因的 5'非翻译区的内含子具有相似功能。

大部分内含子除决定组织表达特异性外,还可以提高基因表达的水平。但在本实验中,转基因拟南芥 *GUS* 染色的结果(见图 5)表明,棉花 *SUS3* 基因的第 1 个内含子会降低报告基因在除花粉外的其他组织中的表达水平。这说明内含子在调控基因表达的机制上有可能是一致的,但在调控方向上有一定的区别。内含子这种对基因表达水平的调控作用是通过影响转录或者翻译的效率来实现的^[7-12],但具体的机

制尚不清楚。同样, 内含子在基因表达组织特异性中的作用机制也还是未知的, 推测可能是内含子影响前体 RNA 的剪切, 而这种剪切又是具有组织相关性的^[9], 但目前还没有发现直接的证据可以证明这一点。

有些基因只有当其编码区中特定的内含子存在时, 该基因才会有相应的功能。WANG 等^[12]在对棉花 *GaMYB2* 基因以及拟南芥中的 *GL1* 和 *WER* 基因进行表达分析时发现, 以这 3 个基因的 cDNA 转化拟南芥相应的无毛刺突变体时, 它们均不能恢复转基因拟南芥的毛刺表型, 但当基因编码区的第 1 个内含子存在时, 这 3 个基因就可以发挥相应的功能, 从而使拟南芥无毛刺突变体的毛刺表型得到恢复。这 3 个内含子的序列有 1 个共同的特征, 它们都存在着 1 个可以被 MYB 类转录因子特异识别并结合的保守序列 (CAGTTA), 这似乎意味着可能有相应的转录因子与内含子结合, 从而调控基因的表达, 但研究中并没有证实这一点。在棉花 *SUS3* 基因编码区的第 1 个内含子存在着 1 个 P-box (CCTTTTG), 这一元件被认为是 DOF (DNA binding with one finger) 类转录因子识别并结合的位点, 且该位点受赤霉素 (GA) 信号的调控^[17], 但该内含子是否受 GA 的调控还需实验证明。

总而言之, 棉花 *SUS3* 基因编码区的第 1 个内含子可特异抑制报告基因在转基因拟南芥花粉中表达, 同时对报告基因在其他组织的表达具有负调控作用。这表明该内含子在棉花 *SUS3* 基因的表达中起一定的调控作用, 为研究该基因在棉纤维起始阶段中的表达调控提供新的线索。

参考文献:

- [1] HENDRIX D L. Carbohydrates and carbohydrate-enzymes in developing cotton ovules [J]. *Physiologia plantarum*, 1990, 78: 85–92.
- [2] RUAN Y L. Rapid cell expansion and cellulose synthesis regulated by plasmodesmata and sugar: insights from the single-celled cotton fibre [J]. *Functional Plant Biology*, 2007, 34: 1–10.
- [3] NOTLE K D, HENDRIX D L, RADIN J W, et al. Sucrose synthase localization during initiation of seed development and trichome differentiation in cotton ovules [J]. *Plant physiology*, 1995, 109: 1285–1293.
- [4] RUAN Y L, CHOUREY P S. A fiberless seed mutation in cotton is associated with lack of fiber cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds [J]. *Plant Physiology*, 1998, 118: 399–406.
- [5] RUAN Y L, LEWELLYN D, FURBANK R T. The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and *k(t)* transporters and expansin [J]. *The Plant Cell*, 2001, 13: 47–60.
- [6] RUAN Y L, LEWELLYN D J, FURBANK R T. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development [J]. *The Plant Cell*, 2003, 15: 952–964.
- [7] FU H Y, KIM S Y, PARK W D. A Potato *SUS3* sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects [J]. *The Plant Cell*, 1995, 7: 1395–1403.
- [8] JEON J S, LEE S, JUNG K H, et al. Tissue-preferential expression of a Rice α -Tubulin gene, *OsTubA1*, mediated by the first intron [J]. *Plant Physiology*, 2000, 123: 1005–1014.
- [9] LEHIR H, NOTT A, MOORE M J. How introns influence and enhance gene expression [J]. *Trends in the Biochemical Sciences*, 2003, 28: 215–220.
- [10] KIM Y J, LEE S H, PARK K Y. A leader intron and 115 bp promoter region necessary for expression of the carnation S-adenosylmethionine decarboxylase gene in the pollen of transgenic tobacco [J]. *FEBS Lett*, 2004, 578: 229–235.
- [11] MORELLO L, BARDINI M, SALA F, et al. A long leader intron of the *Ostub16* rice β -tubulin gene is required for high-level gene expression and can autonomously promote transcription both in vivo and in vitro [J]. *The Plant Journal*, 2002, 29: 33–44.
- [12] WANG S, WANG J W, YU N, et al. Control of plant trichome development by a cotton fiber *MYB* gene [J]. *The Plant Cell*, 2004, 16: 2323–2334.
- [13] JEFFERSON R A. Assaying chimaeric genes in plants: the *gus* gene fusion system [J]. *Plant Molecular Biology Reporter*, 1987, 5: 387–405.
- [14] BAUD S, VAULTIER M N, ROCHAT C. Structure and expression profile of the sucrose synthase multigene family in Arabi-

- dopsis [J]. Journal of Experimental Botany , 2004 , 396: 397 – 409.
- [15] CURI G C , CHAN R L , GONZALEZ D H. The leader intron of *Arabidopsis thaliana* genes encoding cytochrome c oxidase subunit 5c promotes high-level expression by increasing transcript abundance and translation efficiency [J]. Journal of Experimental Botany , 2005 , 56: 2563 – 2571.
- [16] JUAN A , LAO N T , KAVANAGH T A. Intron-regulated expression of *SUVH3* , an *Arabidopsis* Su(var) 3 – 9 homologue [J]. Journal of Experimental Botany , 2006 , 57: 3301 – 3311.
- [17] MENA M , CEJUDO F J , ISABEL L 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: 111 – 119.

Function of the First Intron of Cotton *Sucrose Synthase 3* for Expression Regulation

RUAN Meng-bin , YU Xiao-ling , PENG ming

(Chinese Academy of Tropical Agricultural Sciences , Institute of Tropical Bioscience and Biotechnology , Haikou 571101 , China)

Abstract: Cotton *sucrose synthase 3* (*SUS3*) plays an important role during the cotton fiber development. Compared to *Arabidopsis SUS1/SUS4* , the gene structure of cotton *SUS3* has a high similarity with that of those two genes. To determine the function of the first intron of cotton *SUS3* , the intron was isolated and inserted into the *GUS* coding region to construct vector *Susfig*: : 121 , and which was transformed into *Arabidopsis thaliana* mediated by agrobacterium. The results of *GUS* activity analysis of transgenic *Arabidopsis* plants indicated that the intron down-regulated *GUS* gene expression in non-pollen tissues and specifically repressed *GUS* expression in pollen , which suggested that the first intron played a role in gene expression.

Key words: cotton; *SUS3*; intron; gene expression regulation

更正启示

本刊于 2010 年第 3 期刊登的《小菜蛾对阿维菌素的抗性风险评估及交互抗性的室内测定》原文中的表 1 (230 页) 中的数据“世代 F0 的 LC_{50} 为 $0.2 \text{ mg} \cdot \text{L}^{-1}$ ”有误 , 更正为“世代 F0 的 LC_{50} 为 $0.12 \text{ mg} \cdot \text{L}^{-1}$ ”。