

棉花曲叶病毒互补链基因启动子 功能区的缺失分析*

谢迎秋^① 刘玉乐^② 朱 祯^{①**}

(^①中国科学院遗传研究所植物遗传操作实验室, 北京 100101; ^②中国科学院微生物研究所, 北京 100080)

摘要 棉花曲叶病毒(CLCuV)互补链基因启动子是一种新型的启动子, 它能驱动外源基因在植物体内高效表达. 为了研究其最佳启动子区域, 对启动子 5' 端进行了一系列缺失, 得到 5 种不同长度的启动子片段与 *gus* 基因融合的植物表达载体. 继而导入根癌农杆菌, 采用叶盘法转化烟草(*Nicotiana tabacum* L. cv. *Xanthi*), 并检测转基因植株的 GUS 活性. 实验结果表明, 自启动子 5' 端缺失至转译起始位点上游-287, -271 时启动子活性分别是全长启动子的 5 倍, 3 倍. -271~-176 元件对启动子在韧皮部的表达活性起重要作用. 自-176 缺失至-141 时, 启动子的活性降低至全长启动子的 1/30~1/20, 启动子在根中失去表达活性, 但在叶、茎中仍有微弱的活性. 对棉花曲叶病毒互补链基因启动子的功能区进行了分析比较, 发现缺失负调控元件的启动子比全长启动子具有更强的活性, 平均活性是 CaMV 35S 启动子的 12 倍, 暗示该启动子具有巨大的应用潜力. 研究结果也为进一步了解双生病毒基因表达调控机制及病毒-植物间的相互作用提供了新的线索.

关键词 双生病毒 启动子 顺式元件 根癌农杆菌 烟草

双生病毒是一类具有单链环状 DNA 基因组的植物病毒, 该类病毒通过滚环方式复制基因组 DNA, 其中产生的复制体(replication form, RF)双链 DNA, 分别称为病毒链和互补链. RF 可同时作为复制和转录的模板. 双生病毒最显著的特征是通过同一启动子区进行两个方向的转录, 分别调控病毒链基因(如外壳蛋白基因)和互补链基因(如复制蛋白基因)的表达, 因此两个方向的启动子分别称为病毒链基因启动子和互补链基因启动子. 在原生质体中的瞬时表达及转基因研究均表明, 互补链基因启动子的活性明显高于病毒链基因启动子的活性^[1], 但目前已报道的互补链基因启动子均比 CaMV 35S 启动子活性低, 如非洲木薯花叶病毒(ACMV)的互补链基因启动子活性是 CaMV 35S 启动子的 1/40^[1].

双生病毒病毒链和互补链基因启动子都具有典型的真核启动子特征, 转录起始位点上游均存在 TATA 盒序列^[2]. 研究表明, 互补链基因启动子除含有调控基因转录的顺式元件外, 还含有与病毒复制相关的序列, 如茎环结构序列, 这些调控元件与启动子的活性有关. 非洲木薯

1999-09-08 收稿, 2000-02-16 收修改稿

* 国家自然科学基金(批准号:39989001, 39880023)、中国科学院重大项目基金(批准号:KY951-A1-302-12-08)、国家“863”高技术计划(批准号:Z-17-01-01, BH-02-01-01)和国家转基因植物研究与产业化开发专项基金(批准号:ZJY-A-07)资助项目

** 联系人 (E-mail: pgmlzz@igtp.ac.cn)

花叶病毒(ACMV)互补链基因启动子的缺失分析表明, 启动子的远上游序列存在多个正调控元件, 其中与病毒复制有关的茎环结构及其 5' 端序列对启动子活性也有较大的影响, 该区序列的缺失导致启动子的活性降低至全长启动子的 1/2^[1]. 目前还未见互补链基因启动子顺式调控元件的组织化学定位方面的正式报道.

棉花曲叶病毒(cotton leaf curl virus, CLCuV)是新近发现并鉴定的一种双生病毒^[3,4], 我们从 CLCuV 基因组中分离了互补链基因启动子, 发现其为强启动子, 在根、茎、叶等器官中有非常高的活性, 这些特性使之成为可驱动外源基因高效表达的具有巨大应用价值的新型强启动子. 本文对一系列 5' 端缺失启动子驱动的 *gus* 基因表达结果进行了比较, 以探求双生病毒互补链基因启动子不同区域的功能.

1 材料与方法

1.1 材料

Escherichia coli DH5 α , *Agrobacterium tumefaciens* LBA4404, 质粒 pGEM7Zf(+), pBI121 均为本实验室保存. 烟草(*Nicotiana tabacum* L. cv. *Xanthi*)种子为本实验室保存. 感染 CLCuV 的烟草叶片由苏格兰作物研究所 Harrison 教授惠赠. 限制性内切酶、修饰酶购自华美公司, BRL 公司. CLCuV 互补链基因启动子由本室分离并克隆于质粒 pGEM7ZP1P2.

1.2 表达载体的构建

pDMC202(CAMBIA 公司)为无启动子的 *gus*-T nos 中间质粒, 首先将 pGEM7ZP1P2 的 *Nco* I / *Nsi* I 双酶解片段插入 pDMC202 的 *Nco* I / *Pst* I 位点, 得到 p7RPGUS202. 为了快速构建一系列 5' 端缺失的启动子突变体, 以 p7RPGUS202 为基础, 将其 *Hind* III / *Bgl* II 双酶解片段(包含全长启动子-*gus*-T nos)插入植物表达载体 pCAMBIA2300 的 *Hind* III / *Bam* H I 位点, 得到质粒 pA. 启动子突变体 pB 是通过 *Eco* R I 酶解 p7RPGUS202, 回收缺失启动子-*gus*-T nos 的片段, 插入 pCAMBIA2300 的 *Eco* R I 位点, 进一步筛选与 *Npt* II 基因表达盒方向一致的重组质粒获得的. pC 的获得是通过 *Bam* H I / *Eco* R I 双酶解 p7RPGUS202, 回收缺失启动子-*gus*-T nos 的片段, 插入 pCAMBIA2300 的 *Bam* H I / *Eco* R I 位点. pD 的构建是通过 *Apa* I 酶解 p7RPGUS202, 并用 DNA 聚合酶 I Klenow 补平末端, 再用 *Bam* H I 酶解 p7RPGUS202, 回收缺失启动子-*gus*-T nos 的片段, 插入 pCAMBIA2300 的 *Sma* I / *Bam* H I 位点. pE 的获得是先通过 *Ssp* I / *Nco* I 双酶解 p7RPGUS202, 得到 141 bp 的缺失启动子片段并插入 pDMC202 的 *Pvu* II / *Nco* I 双酶解位点以构建中间质粒 pSNRPGUS, 再将其用 *Xho* I / *Bgl* II 双酶解, 回收缺失启动子-*gus*-T nos 的片段, 插入 pCAMBIA2300 的 *Sal* I / *Bam* H I 位点.

将 pDMC202 用 *Eco* R I / *Hind* III 双酶解, 回收 *gus*-T nos 片段并分别插入 pBin19, pCAMBIA2300(CAMBIA 公司)载体的 *Eco* R I / *Hind* III 双酶解位点, 获得质粒 pBinGUS, pCAGUS, 作为无启动子的阴性对照. 阳性对照采用 pBI121, 它含有 35S 启动子-*gus*-T nos 的表达结构. CaMV 35S 启动子长 800 bp, 没有经过修饰及增加增强子.

1.3 电激法转化农杆菌

按照 BIO-RAD 公司电激仪的操作程序, 采用电激法使重组 DNA 质粒转化入受体菌.

1.4 叶盘法转化烟草及再生植株的筛选, PCR 检测

按文献[5]的方法进行烟草转化及再生植株的筛选. 对卡那霉素抗性植株进行基因组 DNA

的提取^[2], 最终 DNA 溶解于 TE 缓冲液中. 以提取的总 DNA 为模板, 进行 *gus* 基因的一部分片段的 PCR 扩增(441 bp), 引物序列为: 5' 端引物: 5' -CTGCGACGCTCACACCGATAACC-3'; 3' 端引物: 5' -TCACCGAAGTTCATGCCAGTCCAG-3'.

1.5 GUS 活性测定

按 Jefferson^[6]的方法进行 GUS 活性的定量测定及组织化学染色. 农杆菌叶盘转化烟草 10 周后, 取长势较一致的再生植株提取叶片总蛋白, 进行 GUS 活性测定. 对 GUS 活性阳性植株的根、茎、叶徒手切成薄片后染色, 固定, 70% 乙醇脱去叶绿素后在 Olympus 解剖镜上进行观察照相.

2 结果与分析

2.1 CLCuV 互补链基因启动子序列分析及 5' 端缺失体的构建

从 CLCuV 基因组中分离了启动子片段, 序列分析表明, 克隆片段全长 436 bp, 包括与复制相关的茎环结构序列、互补链基因启动子的 TATA 盒及其他顺式作用元件(图 1), 如 GC 富集区^[7]、位于 TATA 盒附近的 6 个重复序列. 据推测, 这些重复序列很可能是病毒复制蛋白的识别及结合位点, 与病毒复制起始有关^[8]. 有趣的是, 这些重复序列中含有类似 MSA(M-specific activators)序列(AACGGT). 研究表明, MSA 序列与植物细胞周期调控有关^[9]. 已有报道, 双生病毒互补链基因编码产物 AC1 能诱导寄主细胞中 DNA 聚合酶 δ 辅助因子 PCNA (proliferating cell nuclear antigen)的大量积累, 进而激活寄主细胞复制系统^[10]. 推测 AC1 可能是通过启动子区中的类似 MSA 序列与细胞调控因子相互作用, 以参与寄主细胞的周期调控.

进一步构建了中间质粒 p7RPGUS202, 并以其为基础, 分别采用启动子内部的限制性内切酶酶解以获得全长及不同缺失程度的互补链基因启动子-*gus*-T nos 片段, 插入带有左右边界序列的植物表达载体, 最终获得了 5 种表达载体(图 2).

2.2 CLCuV 互补链基因启动子 5' 端缺失体在转基因烟草中的活性

为了进一步分析互补链基因启动子 5' 端顺式元件在真核生物中驱动外源基因表达的特性, 采用农杆菌介导的方法将缺失体转化烟草, 获得了卡那霉素抗性植株. 对抗性转基因植株进行了进一步的 PCR 检测(图 3). 叶片组织 GUS 活性测定结果表明(图 4), 就同一缺失体转化的植株而言, 转化个体之间存在一定的差异, 可能与整合基因的拷贝数及位置效应等有关. 就每种缺失体转化群体的 GUS 活性平均值而言, 自启动子 5' 端缺失至转译起始位点上游 -176 时, 启动子仍具有较强的 GUS 活性. -287 上游可能存在一个负调控元件, 该元件缺失后启动子活性是全长启动子的 5 倍左右, 平均值达 CaMV 35S 启动子的 12 倍. -271 上游可能也存在一个负调控元件, 该元件缺失后启动子活性是全长启动子的 3 倍左右. 图 1 的序列表明, 该区含有推测的病毒链基因转录的 TATA 盒序列. 大多数学者认为, 病毒链基因产物为晚期基因表达产物, 因此病毒链基因启动子不如早期基因即互补链基因启动子的活性强^[11], 但病毒链基因启动子顺式元件是否也影响互补链基因的启动子活性, 目前还不清楚. 启动子序列从 -176 缺失到 -141 时 GUS 活性下降约 1/30~1/20 倍, 其中缺失的调控元件包含与复制相关的茎环结构序列及推测的 GC 富集区(图 1), 暗示这些顺式元件对互补链基因转录至关重要. 目前已发现, 玉米核提取物中有反式因子与玉米条纹病毒启动子中的 GC 富集区结合^[7].

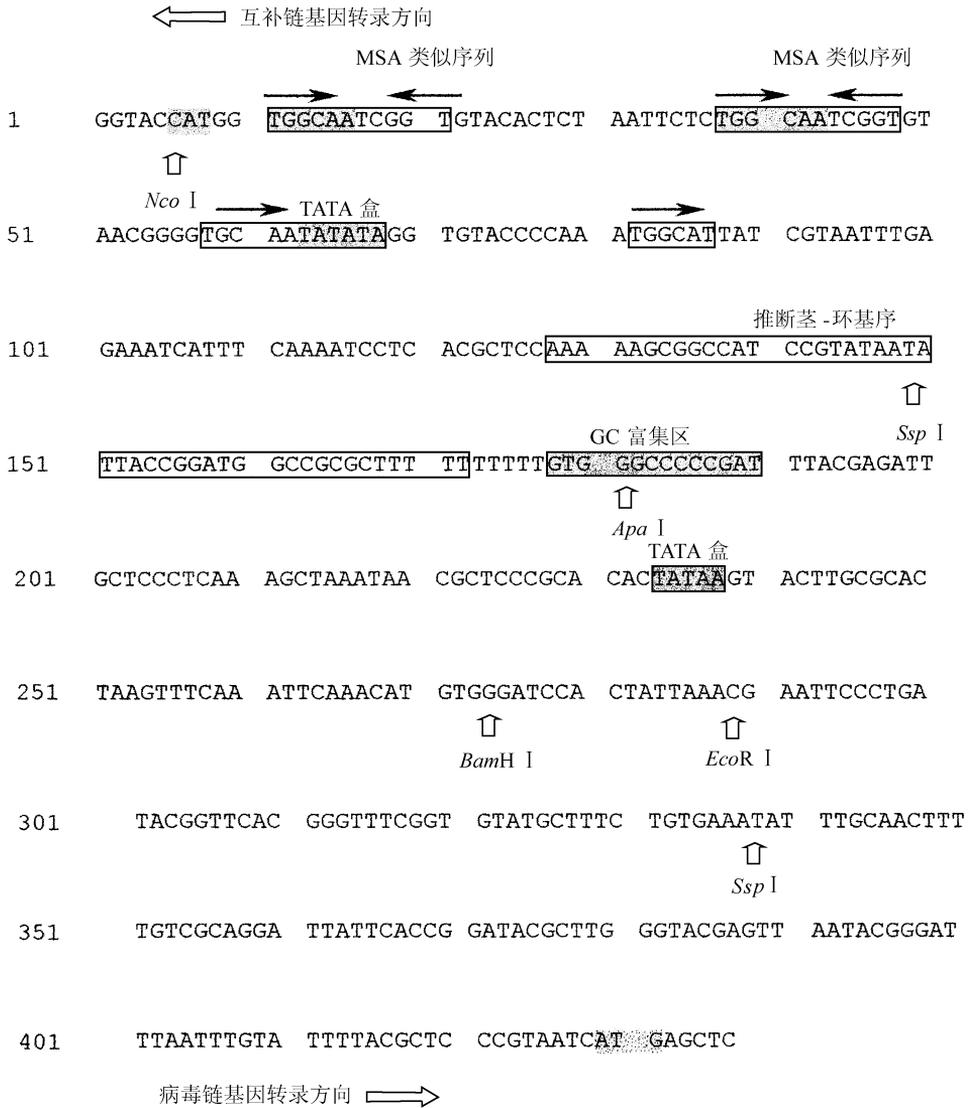


图 1 CLCuV 启动子序列及结构

转译起始密码子以阴影表示; TATA 盒, MSA 类似序列, GC 富集区以方框加阴影表示, 重复序列以箭头表示; 互补链及病毒链基因的转录方向分别以不同方向空心箭头表示. 图左数字表示启动子片段大小

2.3 互补链基因启动子 5' 端缺失体转基因烟草的 GUS 活性组织化学定位

我们对各缺失启动子控制下的 GUS 基因表达载体转化植株进行了根、茎、叶不同部位的组织化学染色(图 2, 图版 I, 附本刊后). 根部染色结果表明, 全长启动子及缺失至-176 时的启动子在根尖分生区及根中部维管束中均有较高的活性, 但缺失至-141 时启动子在根尖分生区及根中部维管束中未能检出 GUS 活性. 与 35S 启动子不同的是, 全长及各缺失启动子在根冠细胞中活性均极低, 在根尖分生区及根中部维管束中活性较高. 而 35S 启动子在根冠、根分生区均有较强的活性, 根中部维管束中活性较低. 研究表明, 35S 启动子在细胞周期的 S 期表达活性较高, 这很可能与 CaMV 的 35S 转录本在病毒复制中的功能有关^[12]. 如前文所述,

CLCuV 启动子的重复序列中含有与细胞周期调控有关的类似 MSA 序列, 我们推测, 该启动子在根尖分生区的高表达活性与 MSA 序列有关, 但奇怪的是, 缺失至-141 时启动子在根尖分生区失去活性, 可能是因为除类似 MSA 序列外, 茎环结构周边也含有与细胞分裂有关的序列, 它们共同决定了启动子在分生区的活性.

叶片横截面染色结果表明, 全长启动子及缺失至-141 时的启动子在叶肉及维管组织均有较强的活性, 但染色强度存在差异. 上述各转化体的 GUS 活性组织化学定位结果与 35S 启动

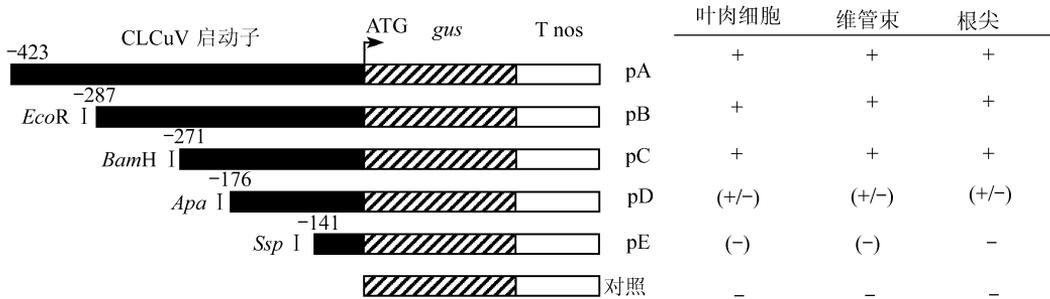


图 2 CLCuV 启动子缺失体示意图及组织化学定位比较

pA 为全长启动子载体, pB, pC, pD, pE 分别为不同的启动子缺失体. 对照为无启动子的质粒载体 pCAGUS. 各缺失体的 GUS 活性组织化学定位比较分为 4 个等级: + 示不低于全长启动子的活性; (+/-) 示比全长启动子活性略低; (-) 示活性极低; - 示未能检出活性

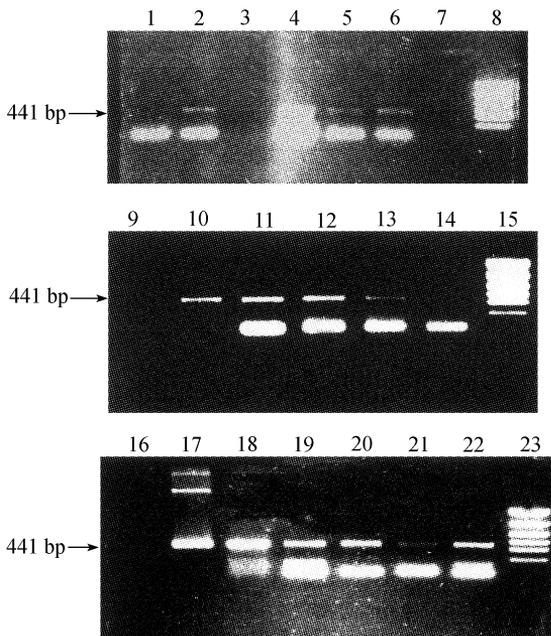


图 3 转基因烟草的 PCR 检测

1~4 示全长启动子 pA, 5~6 示 pB, 7, 9, 16 示阴性对照 pCambia2300, 10~11 示 pC, 12~14 示 pD, 17 示质粒 p7RPGUS202, 18~21 示 pE, 22 示 pCAGUS 转化植株, 8, 15, 23 示 PCR Marker. 箭头示 gus 基因片段的 PCR 扩增产物

子相似, 均在叶脉维管组织显示了较强的蓝色. 已知双生病毒在韧皮部复制, 但至今还不清楚复制蛋白基因的组织定位, 更不清楚决定其组织定位的顺式元件. 根据实验结果, 我们认为, 与复制蛋白基因定位有关的韧皮部特异性元件很可能延伸至互补链基因启动子-141 下游.

茎的横截面染色结果表明, 全长启动子及缺失至-271 时的启动子在内、外韧皮部、木质部周围薄壁细胞及皮层薄壁细胞均有活性. 然而, 缺失至-176 时的启动子仅在维管束中有活性, 薄壁细胞检测不到 GUS 活性. 较为奇特的是, 该缺失启动子在外韧皮部未显示活性, 仅在内韧皮部显示活性, 而且活性相当强, 但缺失至-141 时的启动子在内、外韧皮部中却又显示了活性. 我们推测, -176~-141 很可能含有韧皮部特异性的调控元件, 植物基因组中可能存在与之相互作用的因子.

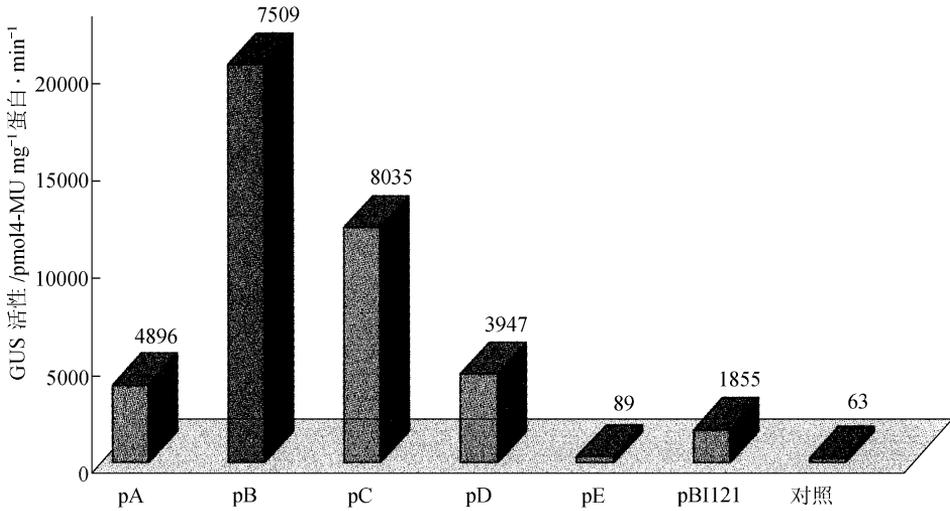


图 4 CLCuV 互补链基因启动子缺失体在转基因烟草中的活性

pA 为全长启动子载体, pB, pC, pD, pE 分别为不同的启动子缺失体, pBI121 为 CaMV 35S 启动子-gus 转化植株; 柱型图上方的数据表示标准差; 对照为无启动子的质粒载体 pCAGUS 转化植株; 分析的转基因植株数目从 pA 至对照依次为 23, 30, 26, 18, 30, 29, 8

3 讨论

实验结果表明 CLCuV 互补链基因启动子包括许多顺式元件. 图 5 列出了在植物不同器官中对 CLCuV 互补链基因启动子起作用的可能区域, 我们发现同一顺式元件在不同器官中的功能有所不同, 启动子组织定位可能由多个不同区域的顺式元件共同决定. 除了上文所述的正、负调控元件外, CLCuV 互补链基因启动子可能还含有其他具有不同功能的顺式元件. 值得提出的是, 自 5' 端缺失至 -141 时的启动子在叶肉及叶脉、茎的维管束仍旧具有启动子活性, 表明启动子 -141 的下游序列还包括有功能元件. 进一步对该区启动子进行缺失分析将会增进对互补链基因启动子功能的了解. 另外, 我们在转基因烟草的研究中还发现, 全长启动子在根、茎、叶、花、果实、种子不同器官中的表达活性有差异, 表明启动子存在时空调控(结果将另文发表). 我们推测, 不同器官中可能存在不同的反式作用因子与顺式元件相互作用. 因此, 分离这些反式作用因子以揭示植物与病毒间的相互作用的方式, 并利用这些原理以调控 CLCuV 启动子的活性, 使之更加有效地应用于植物遗传操作, 将具有十分重要的意义.

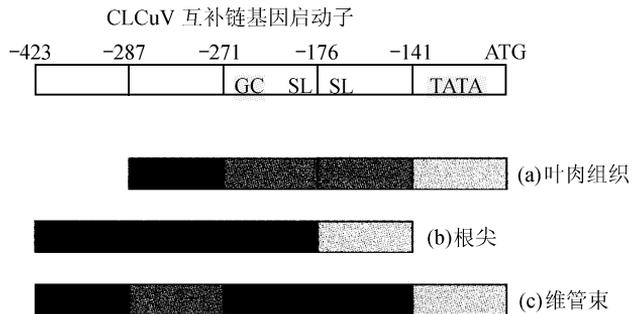


图 5 CLCuV 启动子顺式元件的功能分析

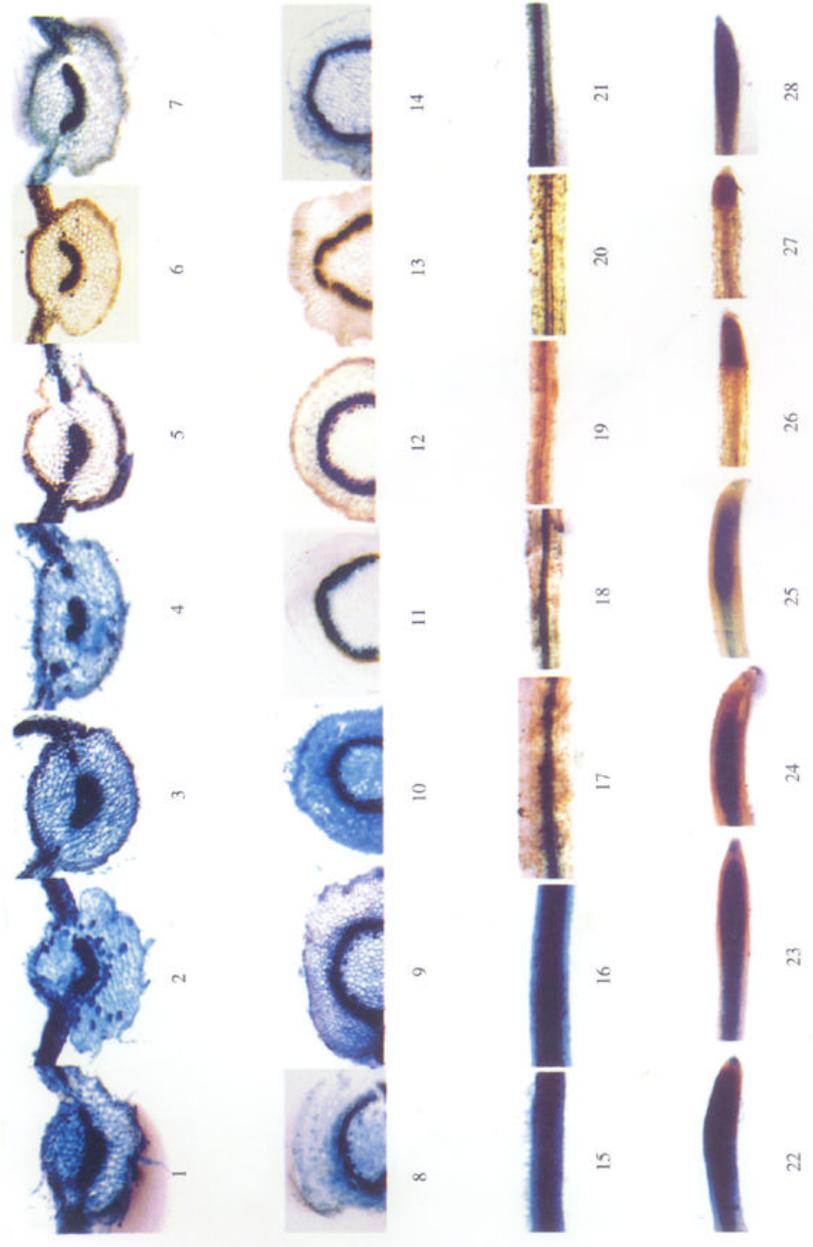
ATG 为起始密码子, 方框中颜色的深浅为顺式元件对启动子组织特异性的影响程度. SL 为茎环结构(stem-loop)序列; GC 为 GC 富集序列; TATA 为 TATA 盒序列

CLCuV 是新近发现的一种双生病毒, 其基因表达调控机理尚未知. 目前已知双生病毒采用类似原核生物的滚环方式复制其基因组, 可独立于染色体复制^[13], 并且能在原核生物中独立复制, 但是尚未见该类病毒启动子在原核生物中驱动外源基因表达的报道. 我们发现 CLCuV 互补链基因启动子在根癌农杆菌中也具有超强活性, 而且-287 上游的负调控元件在根癌农杆菌中也有抑制启动子活性的功能(另文发表). 图 1 表明启动子区存在真核生物的 TATA 盒序列, 在转录与转译起始位点之间不具有典型的 SD 序列, 却存在多个重复序列. 这些重复序列是否与启动子的活性调控有关, 还需要开展更加深入的研究.

我们先前在转基因烟草(*Nicotiana tabacum* L. cv. NC89)中发现 CLCuV 互补链基因启动子属强启动子. 本文发现在烟草品种 *Xanthi* 中, 缺失了-287 上游负调控元件的互补链基因启动子比全长启动子具有更强的活性. 实验结果进一步表明, CLCuV 互补链基因启动子具有巨大的潜力成为超强启动子而取代如今应用最为广泛的 CaMV 35S 启动子. 同时, 对 CLCuV 互补链基因启动子功能区的更深入的分析可增强对该启动子驱动外源基因高效表达机理的认识, 为植物遗传操作提供了有益的调控元件, 也为进一步研究双生病毒基因表达调控及病毒与寄主植物间相互作用的分子机制提供了新的线索.

参 考 文 献

- 1 Zhan X, Haley A, Richardson K, et al. Analysis of the potential promoter sequence of African cassava mosaic virus by transient expression of the *b* - glucuronidase gene. *J Gen Virol*, 1991, 72: 2 849~2 852
- 2 谢迎秋, 朱 祯. 植物双生病毒的复制及转录调控. *病毒学报*, 1999, 15: 377~386
- 3 Zhou X, Liu Y, Robinson D J, et al. Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. *J Gen Virol*, 1998, 79: 915~923
- 4 Harrison B D, Liu Y L, Khalid S, et al. Detection and relationships of cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. *Ann Appl Biol*, 1997, 130: 61~75
- 5 刘春明, 朱 祯, 周兆澜, 等. 豇豆胰蛋白酶抑制剂抗虫转基因烟草的获得. *科学通报*, 1992, 37: 405~411
- 6 Jefferson R A. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Rep*, 1987, 5: 387~405
- 7 Fenoll C, Schwarz J J, Black D M. The intergenic region of maize streak virus contains a GC-rich element that activates rightward transcription and binds maize nuclear factors. *Plant Mol Biol*, 1990, 15: 865~877
- 8 Argüello-Astorga G, Herrera-Estrella L, Rivera-Bustamante R. Experimental and theoretical definition of geminivirus origin of replication. *Plant Mol Biol*, 1994, 26: 553~556
- 9 Ito M, Iwase M, Kodama H, et al. A novel *cis*-acting element in promoter of plant B-type cyclin genes activates M phase-specific transcription. *The Plant Cell*, 1998, 10: 331~341
- 10 Negar S, Pedersen T J, Carrick K M, et al. A geminivirus induces expression of a host DNA synthesis protein in terminally differentiated plant cells. *The Plant Cell*, 1995, 7: 705~719
- 11 Eagle P A, Crozco B M, Hanley-Bowdoin L. A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *The Plant Cell*, 1994, 6: 1 157~1 170
- 12 Jefferson R A, Kavanagh T A, Bevan M W, et al. GUS fusions: *b* -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO J*, 1987, 6: 3 901~3 907
- 13 Timmermans M C P, Dason O P, Messing J. Geminiviruses and their uses as extrachromosomal replicons. *Ann Rev Plant Physiol Mol Biol*, 1994, 45: 79~112



CLCuV 互补链基因启动子及其缺失体驱动的 GUS 表达活性组织化学定位比较

转化植物的叶片横截面(1~7)、茎横截面(8~14)、根尖(22~28) GUS 染色结果. 1, 8, 5, 22 示 pA, 为全长启动子载体; 2, 9, 16, 23 示 pB; 3, 10, 17, 24 示 pC; 4, 11, 18, 25 示 pD; 5, 12, 19, 26 示 pE; 6, 13, 20, 27

示阴性对照 pCAGUS; 7, 14, 21, 28 示阳性对照 pB II 121