

渡调控的秘密,有可能为找到细胞周期正常调控与异常调控的差别,以及为阐明癌变及治疗肿瘤奠定基础。

### 参 考 文 献

- [1] Lalande M., 1990, *Exp. Cell Res.*, 186: 332—339.
- [2] Resenblatt J. et al., 1992, *Proc. Natl. Acad. Sci. USA*, 89: 2824—2828.
- [3] Soloman MJ. et al., 1993, *The EMBO J.*, 12: 3133—3142.
- [4] Sebastian B. et al., 1993, *Pro. Natl. Acad. Sci. USA*, 90: 3521—3524.
- [5] Pagano M. et al., 1992, *The EMBO J.*, 11: 961—971.
- [6] Maridor G. et al., 1993, *The Journal of cell Science*, 106: 535—544.
- [7] Li J-L. et al., 1993, *Proc. Natl. Acad. Sci. USA*, 90: 3554—3558.
- [8] Marx J. 1994, *Science*, 263: 319—321.
- [9] Hunter, T. 1993, *Cell*, 75: 839—841.
- [10] Hollingsworth RE. Jr. et al., 1993, *current Opinion in Genetics and Development*, 3: 55—62.
- [11] Dyson N et al., 1989, *Science*, 243: 934—937.
- [12] Hinds PW. et al., 1992, *Cell*, 70: 993—1006.
- [13] Faha B., et al., 1992, *Science*, 255: 87—90.
- [14] Cao L. et al., 1992, *Nature*, 355: 176—179.
- [15] Kovcsdi I. et al., 1986, *CELL*, 45: 219—228.
- [16] Nevins JR., 1992, *Science*, 258: 424—429.
- [17] Helin K. et al., 1992, *Cell*, 70: 337—350.
- [18] Slansky JE. et al., 1993, *Mol. Cell. Biol.*, 13: 1610—1618.
- [19] Roehl HH, et al., 1993, *J. Virol.*, 67: 4964—4971.
- [20] Moberg KH. et al., 1992, *Oncogene*, 7: 411—421.
- [21] Sherr, C. J., 1994, *Trends in Cell Biology*, 4: 15—18.
- [22] Johnson DJ. et al., 1993, *Nature*, 365: 349—352.
- [23] Cardoso MC., et al., 1993, *Cell*, 74: 979—992.
- [24] Seto E. et al., 1992, *Proc. Natl. Acad. Sci. USA*, 89: 12028—12032.
- [25] Yuan J-N, et al., 1993, *Biochem. Biophys. Res. Commun.*, 191: 662—668.
- [26] Dutta A. et al., 1993, *Nature*, 365: 79—82.
- [27] Yamamoto M. et al., 1994, *Exp. Cell Res.*, 210: 94—101.

## 高等植物中外源基因表达的转录后调控

李 红 白永延

(中国科学院上海植物生理研究所 200032)

高等植物中基因表达与调控的分子机制是细胞和分子生物学的重点之一,随着植物基因克隆技术以及基因转化系统不断完善,在高等植物中,外源基因可以通过根癌农杆菌或直接基因转移导入高等植物细胞并再生新的植株,转入的基因存在于每个植物细胞中并能遗传。因此转基因植物成为进一步研究基因表达调控的重要手段和对象,这方面的突破与进展有助于人们将更多的外源基因引入到经济作

物中去并获得高表达、稳定遗传的转化植物。

在早期的遗传转化工作中,提高和稳定外源基因表达的研究大多集中在转录水平的调控(例如在外源基因的5'末端连接强启动子,在3'末端连接有效的终止信号)。这当然因为转录过程的调控是基因表达调控中最为重要的一环,转录能否开动是基因能否表达的决定性一步。但是,对于基因的整体表达过程,它包括转录、初级转录产物的加工、运输、mRNA的翻译

以及蛋白产物的后加工这一系列步骤,每个步骤上都存在着复杂而精细的调节,尤其在真核生物中,转录和翻译在时间和空间上都是分隔开的,转录后水平的调节对基因的表达调控也就十分重要了。在许多实验中,提高外源基因的转录活性并不能有效地提高细胞中相应的 mRNA 水平和蛋白含量,说明转录后水平的调控也影响着外源基因的表达。

对高等植物而言,一般基因的 mRNA 的平均寿命为几个小时<sup>[1]</sup>。如果植物体仅仅通过调节基因的转录开关是难以及时适应环境条件的变化的,而转录后水平的调节则可能迅速有效地调节基因的表达适应环境的变化。转录后水平调控的分子机制十分复杂,目前的许多研究工作者认为, mRNA 代谢和翻译活性的调节在植物中广泛存在<sup>[2]</sup>。mRNA 代谢调控的分子机制还不清楚,但普遍认为 mRNA 代谢机制和代谢速率根本上取决于转录产物自身的结构<sup>[3]</sup>。高等植物中外源基因表达的转录后调控的研究也着重于分析和改造转录产物的结构,提高 mRNA 的稳定性和翻译活性。

### 一、5' 末端帽子结构

真核基因 mRNA 5' 末端转录出来不久即在鸟苷酸转移酶和甲基转移酶催化下戴上帽子结构,这是真核基因 mRNA 区别于原核基因 mRNA 的显著特征。帽子结构参与翻译起始,可以促进蛋白质生物合成起始复合物的形成。没有甲基化的帽子(如 GpppN-)或是采用化学或酶学的方法除去帽子,其 mRNA 的翻译活性显著下降。帽子结构的类似物如 m<sup>7</sup>GMP, m<sup>7</sup>GDP 等都强烈抑制戴帽 mRNA 的翻译,帽子结构不仅提高翻译起始频率还能保护 mRNA 免遭外切核酸酶的破坏,提高 mRNA 的稳定性<sup>[4]</sup>。

### 二、先导序列

从真核基因 mRNA 5' 末端帽子到起始密码子之间的不翻译核苷酸序列称为先导序列。

植物基因的先导序列富含 AU (平均约 60—70%)。先导序列的长度和碱基顺序在不同的物种和不同的基因中变化很大,甚至同一基因通过不同的转录起始得到不同长度的先导序列。先导序列的长度对翻译效率的影响是存在的,但不是绝对的。二级结构较多的先导序列对翻译起始是不利的。深入研究先导序列的结构、组成和长度对 mRNA 的加工、运输和蛋白翻译活性的影响是很有意义的问题。烟草花叶病毒 (TMV) 基因组 RNA 68 个碱基的先导序列和苜蓿花叶病毒 (AMV) RNA 436 个碱基的先导序列具有提高外源基因 mRNA 翻译活性的功能<sup>[5]</sup>。实验推测可能因为 AMV RNA 4 的先导序列不形成特定的二级结构,无需翻译起始因子消除二级结构从而提高了翻译效率<sup>[6]</sup>;而 TMV 基因组 RNA 的先导序列正相反,它具有 Ω 式的特定的二级结构作为某种 RNA 结合蛋白的特异的识别信号从而提高了翻译活性<sup>[7,8]</sup>。

### 三、3' 末端序列及 poly(A) 尾巴

在外源基因的 3' 末端连上某些植物基因的 3' 末端序列可以增强外源基因的表达。高等植物基因的 3' 末端区域常有多个加 poly(A) 的信号,在第一个加 poly(A) 信号的下流还有一段 (G)T 含量丰富的序列。在报告基因磷酸新霉素转移酶基因 II (nptII) 的 3' 端连上植物的章鱼碱合成酶基因的 3' 末端序列 (3'ocs), 可以显著提高报告基因在转化植物中的表达。嵌合基因中 3'ocs 的缺失实验表明,包含加 poly(A) 位点及其下游的 144 碱基对对 npt II 表达活性影响最大,其中又以包含主要加尾信号和主要及次要加尾位点在内的 35 碱基对最为重要,这 35 碱基对的缺失使 nptII 的表达降低到本底水平。此外,据统计结果在大部分哺乳动物基因 3' 末端加尾信号下游约 30 碱基对处含有 YGTGTTY 序列,这一序列与加尾信号之间距离的改变会降低 mRNA 3' 末端的加工效率,植物基因 3' 端也含有类似的结构元

件,不过距离加尾信号的位置不确定。3'ocs中类似 YGTGTTY 的序列位于加尾信号下游约 140 碱基对处,在 npt II 和 3'ocs 嵌合基因加尾信号下游 98-142 碱基对的缺失致使 npt II 活性下降 4 倍<sup>[9]</sup>。

在嵌合基因中,不同的植物基因的 3' 末端序列对外源基因表达活性的影响也不相同。将一系列嵌合基因导入烟草原生质体做瞬间表达,结果表明 3' 端连接 1,5-核糖二磷酸羧化酶小亚基基因 3' 端序列的 npt II 活性比连接 3'ocs 的活性高出 3 倍,后者又分别比连接 2 s 种子贮藏蛋白基因和伸展素基因 3' 端序列的 npt II 活性高 5—10 倍。而连有查尔酮合成酶基因 3' 端序列的 npt II 活性比连接 3'ocs 的要低 20 倍<sup>[9]</sup>。由此看来,细胞中高产稳定组成型表达的基因的 3' 端序列有利于外源基因的表达,而由环境诱导短期表达的基因的 3' 端序列无益于外源基因的表达。

poly(A)尾巴在有些基因的表达中具有稳定 mRNA 的作用,在另一些基因中则具有调节 mRNA 翻译活性的功能或是两种功能兼有。poly(A)尾巴和某些蛋白因子的结合能保护 mRNA 免遭外切核酸酶的降解。poly(A)的降解很可能是 mRNA 降解的第一步,较长的 poly(A)尾巴可能有利于提高外源基因 mRNA 的稳定性<sup>[10]</sup>。

#### 四、内含子

由于内含子序列存在于原初转录物中而在 RNA 的剪切反应中被去除,过去很长一段时间内含子被认为是基因中多余的、无用的部分。其实不然,现在的研究已经表明内含子在基因的表达调控中起着重要的作用。目前研究得比较多的是玉米 Adh 1 基因第一内含子(Adh 1 intron 1),玉米 Shrunke n-1 基因第一内含子(Shl intron 1)和水稻 actin 基因第一内含子(actin intron 1)等。通常是将内含子插入 35 s 启动子或其它启动子和报告基因之间构建重组质粒来转化植物细胞,测报告基因的表达活性的变

化。在 35 s 启动子与  $\beta$ -葡萄糖醛苷酶基因编码区(uidA)之间插入 actin intron 1 转化水稻和玉米原生质体,uidA 活性分别提高 40 倍和 56 倍<sup>[11]</sup>。在 35 s 启动子和胭脂碱合成酶基因启动子(NOP)与氯霉素酰基转移酶基因之间插入 Adh1 intron1 转化玉米原生质体,cat 表达活性分别提高了 8—20 倍和 170 倍<sup>[12]</sup>。同样在转化的大麦原生质体中,Adh1 intron1 使 cat 表达活性提高了 30 倍<sup>[18]</sup>。玉米 Sh1 基因 5' 端不翻译引导序列中的第一外显子(Sh1 exon1)与第一内含子(Sh1 intron 1)分别使 cat 基因表达活性提高 10 倍和 100 倍,两者共同使用可使报告基因活性增强 1000 倍之多<sup>[14]</sup>,其中 exon1 与 intron1 的作用机理是不同的。exon1 的作用与启动子上游元件有关,在双子叶植物与单子叶植物中具有相同的增强表达作用;Sh1 intron1 只在转化的单子叶植物中有效,在双子叶植物中反而抑制了报告基因的表达<sup>[14]</sup>。玉米 Adh1 intron1 在双子叶植物中也没有增强基因表达的作用<sup>[12]</sup>。实验分析认为 intron 的作用与 RNA 的剪接有关,它很可能是通过提高细胞质中成熟 mRNA 的含量而不是通过提高 mRNA 的稳定性或翻译活性起到增强表达作用的。内含子的增强作用与其两侧核苷酸顺序密切相关,其作用机理尚不清楚<sup>[14]</sup>。核基因 mRNA 的内含子剪接位点很保守,均为 5' $\downarrow$ GU.....AG $\downarrow$ 3'。一般而言,双子叶植物内含子 AU 含量较单子叶植物丰富,双子叶植物只剪接 AU 丰富的内含子,内含子中的茎环结构抑制剪接。而在单子叶植物中剪接条件要宽松得多<sup>[15,16]</sup>,上述玉米 Adh 1 intron 1、Sh1 intron 1 等内含子在转化的烟草、大豆等双子叶植物中不能有效剪接,因而不影响或抑制报告基因的表达<sup>[16]</sup>。目前在单子叶植物中为增强目的基因的表达常使用 Adh1 intron 1 等内含子<sup>[17,18]</sup>。

#### 五、mRNA 不稳定序列

在高等真核细胞中,有些基因的 mRNA

半衰期很短(<60 min),如植物的光敏色素基因, SAUR 基因和几种动物细胞中的原癌基因。通过分析这些不稳定的 mRNA 序列,发现了一些与 mRNA 降解有密切关系的所谓不稳定序列。

一种是多拷贝的串联 ATTTA 序列,这种重复序列普遍存在于许多哺乳动物基因的不稳定转录物中,位于基因的 3' 端不编码区,如哺乳动物基因 *c-fos*, GM-CSF (granulocyte-macrophage colony-stimulating factor) [19,20]。将人工合成的 11 拷贝的 ATTTA 序列插入报告基因 3' 端不编码区转化烟草,测得报告基因的 mRNA 迅速降解了,而作为对照实验的另外两种同样含有 A 和 T 的顺序没有引起 mRNA 降解的现象,说明引起 mRNA 不稳定的因素是特定的碱基序列 [21]。在有些植物基因中也发现存在 ATTTA 顺序,但还没有实验证明它是否对 mRNA 的稳定性具有调节功能,另外,在转化植物中表达水平很低的苏云金杆菌杀虫蛋白基因中也含有 ATTTA 顺序 [3]。

第二种是从大豆及其它几种植物的 SAUR 基因 mRNA 编码区下游发现的一约 40 碱基对的高度保守的顺序,称为 DST 顺序 [18]。对几种植物中不同基因的 DST 顺序进行分析比较,概括出其一般结构模式为 GGAg-avg. 5-cAT A G ATTa-avg. 6- (A/C) (T/A) (A/T) TuG TA (T/C) (加点及大小写表示保守性强弱),将人工合成的 DST 顺序分别插入 GUS 基因和  $\beta$ -珠蛋白基因的 3' 端不编码区转化烟草,结果表明两个串联的 DST 顺序足以使两种报告基因的 mRNA 迅速降解 [11]。DST 顺序是目前植物中发现的一种导致 mRNA 不稳定的顺序。

DST 顺序及多拷贝的 AUUUA 顺序引起 mRNA 降解的机制有两种模式:一是 poly(A) 尾巴降解模式。不稳定序列可能使 poly(A) 尾巴不能与 poly(A) 结合蛋白结合,使得 mRNA 3' 末端容易遭受外切核酸酶的降解 [21];二是核酸内切酶剪切模式。不稳定序列本身或是激活

mRNA 上另一段序列,成为特异的核酸内切酶的攻击位点,经内切酶作用后的 mRNA 碎片很快被外切酶降解 [22]。在植物细胞中寻找与 DST 顺序相作用的 RNA 结合蛋白或是 RNAase 将会进一步阐明 DST 顺序的功能和高等植物中 mRNA 的降解机制。

## 六、密码子的使用

外源基因的 G+C 含量及使用的密码子与植物基因的 G+C 含量和偏爱密码子的差别也是造成 mRNA 不稳定,翻译活性低的重要原因。

大多数植物基因以 AUG 作为起始密码子,起始密码子的最佳前后核苷酸顺序是 AACAAUGGC [23] 或 UAAACAAUGGCU [24],其中-3 位的 A 和 +4 位的 G (AUG 的 A 为第一位) 最为重要。有些实验表明在植物中使用动物基因中更常见的 CCACCAUGG 顺序能增强 mRNA 的翻译起始,这可能是因为 CCACCA 有利于与核糖体 RNA 的互补识别 [25]。植物基因的终止密码子多用 UGA,终止密码子下游第一个碱基多为 A。在双子叶植物中, UAA 的使用频率比 UGA 高 [26]。双子叶植物偏爱的密码子有 44 个,很少用 XCG (X 代表四种碱基中的任一种) 类的密码子。单子叶植物偏爱的密码子有 38 个,极少使用 XUA 类的密码子,且有一部分单子叶植物基因在密码子的第三位使用 G 或 C 的频率很高 [10]。

每种生物中常用的密码子与相应的 tRNA 丰度具有共进化的关系,使用对应的 tRNA 丰度高的密码子无疑会加快翻译的速度。在苏云金杆菌的杀虫晶体蛋白基因的改造实验中,利用密码子的简并性,不改变编码的氨基酸顺序,去掉植物中不常用的密码子,换之以植物偏爱的密码子,并去掉其中的 ATTTA 顺序和加 poly(A) 的信号序列 AATAAA 和 AAATAA,增加了 G+C 含量同时避免 mRNA 产生稳定的二级结构如茎环结构等。经过部分改造(3%的碱基改变)和全部改造(21%的碱基改变)的鳞翅目昆虫毒素蛋白基因 CryIA(b) 在转

化植物(烟草, 西红柿, 棉花)中的表达效率分别提高了10倍和100倍之多, 经改造的CryIA(c)基因的表达活性也大大提高。对部分和全部改造的CryIA(b)基因的mRNA丰度和蛋白产量的分析比较表明, 全部改造的CryIA(b)基因使用植物偏爱的密码子, 从整体水平上提高了翻译效率<sup>[27]</sup>。

除了上述mRNA自身结构因素对基因表达活性的影响之外, 还有一种抑制外源基因表达的现象称为共抑制现象, 即当引入的外源基因多拷贝插入到植物染色体DNA中或与植物中原有的基因有一定的同源性, 则此外源基因和植物中原有的基因的表达都可能受到抑制。这与反义RNA的抑制现象相类似, 同源顺序引起共抑制的分子机理有待进一步的研究。

在植物体的生命活动中, 基因的表达与调控还与植物体的生长发育, 细胞内的激素水平, 光照以及环境因素的变化密切相关, 深入地研究和阐明植物基因表达调控的分子机制, 将有助于解决植物基因工程中出现的许多问题, 推动植物基因工程的前进。

## 摘 要

转基因植物是研究高等植物中基因表达调控的重要手段和对象, 高等真核生物中基因表达的灵活性多样性与转录后调控密切相关。本文着重阐述RNA的自身结构与代谢调控之间的关系, 分别从5'端帽子结构、先导序列、3'末端序列及poly(A)尾巴、内含子序列、不稳定序列、使用的密码子等六个方面说明RNA的特定结构对其稳定性及翻译活性的影响。

## 参 考 文 献

- [1] Thomas, C. Newan et al., 1993, *The Plant Cell*, Vol 5: 701—714.
- [2] Pamela, J. Green, 1993, *Plant Physiol*, 102: 1065—1070.
- [3] EE. Murray et al., 1991, *Plant Molecular Biology*, 16: 1035—1050.
- [4] Gallie et al., 1991, *Gene & Devel.*, 5:

- 2108—2116.
- [5] Gallie et al., 1987, *Nucleic Acids Res* 15: 3257—3273.
- [6] Jobling et al., 1987, *Nature*, 325: 622—625.
- [7] Gallie et al., 1989, *Plant Cell*, 1: 301—311.
- [8] Gallie et al., 1987, *Nucleic Acids Res.*, 15: 8693—8711.
- [9] Ivan, L. W. Ingelbrecht et al., 1989, *The Plant Cell*, Vol 1: 671—680.
- [10] Daniel, R. Gallie, 1993, *Annu. Rev. Plant Physiol Plant Mol. Biol.*, 44:77—105.
- [11] McElroy, D. et al., 1990, *Plant Cell*, 2: 163—171.
- [12] Callis, J. et al., 1987, *Genes Devel.*, 1: 1183—1200.
- [13] Oard, JH. et al., 1989, *Plant Cell Rep.*, 8: 156—160.
- [14] Mass, C. et al., 1991, *Plant Mol Biol.*, 16: 199—207.
- [15] Goodall, G. J. et al., 1989, *Cell*, 58: 473—483.
- [16] Keith, B., et al., 1986, *EMBO J.*, 5: 2419—2425.
- [17] Vasil, V., et al., 1992, *Bio Technology.*, 10: 667—674.
- [18] Walters, D. A. et al., 1992, *Plant Mol. Bio.*, 18: 189—200.
- [19] McClure, B. A., et al., 1989, *Plant Cell*, 1: 229—239.
- [20] Shyu, A. -B. et al., 1989, *Genes Dev.*, 3: 60—72.
- [21] Bernstein, P. Peltz SW. et al., 1989, *Mol Cell Biol.*, 9: 659—670.
- [22] Peltz, S. W., et al., 1991, In GS Stein JL Stein, JB Lian, eds critical Review. *Eukaryotic Gene Expression* OCR Press, Boca Raton. FL, pp 99—126.
- [23] Lutcke, H. A., et al., 1987, *EMBO J.*, 6: 43—48.
- [24] Joshi, C. P., 1987, *Nucleic Acids Res.*, 16: 6643—6653.
- [25] Kozak, M. 1989, *Mol Cell Biol.*, 9: 5073—5080.
- [26] Angenon, G., et al., 1990, *FEBS Lett.*, 271: 144—146.
- [27] Frederick, J. Perlak, et al., 1991, *Proc. Natl. Acad. Sci., USA* Vol 88: 3324—3328.