

编码叶绿体蛋白质的某些核基因

林标扬 白永延

(中国科学院上海植物生理研究所)

高等植物叶绿体中许多重要的蛋白质是由细胞核基因组编码。这些基因在核内转录出的信使RNA转运进细胞质，并在核糖体上翻译成蛋白质，这些蛋白质是无活性的前体，它们被运输进叶绿体时才转变为成熟的有功能的蛋白质，后者与叶绿体基因组编码的蛋白质共同组成有活性的酶和其他有功能的蛋白质。由于存在这样特殊而复杂的核质关系，这类基因的结构极为特殊，它们的表达受到精细的调节。近年来关于细胞核的这类基因的研究进展很快，尤其是关于1,5-二磷酸核酮糖羧化酶/加氧酶的小亚基和叶绿素a/b结合蛋白的基因研究最为详细而深入。本文着重介绍这一领域的最新进展。

一、1,5-二磷酸核酮糖羧化酶/加氧酶(Rubisco)小亚基基因(rbc S)

Rubisco是高等植物光合作用中同化CO₂的一个关键酶，占叶细胞总蛋白量的一半左右，该酶由大亚基和小亚基组成，大亚基基因位于叶绿体基因组上，而小亚基基因位于细胞核基因组上，核基因转录的mRNA在细胞质中合成小亚基前体分子，进入叶绿体后除去转运肽(transitpeptide)转变为成熟分子，然后与大亚基装配成全酶。

Rubisco小亚基基因是由多基因家族所编码。在矮牵牛^[1]、豌豆^[2]、小麦^[3]、大豆^[4]和膨胀浮萍^[5]等的细胞核DNA中均发现有多多个拷贝的小亚基基因，如矮牵牛的小亚基多基因家族至少有8个成员^[1]，膨胀浮萍有13个成员^[5]，根据成员之间的同源性而分成亚族，如矮牵牛的8个小亚基基因可分为3个亚族，其中两个亚族只有1个基因，第三个亚族则有

6个基因。

根据双子叶植物的豌豆^[2,7]、矮牵牛^[1]、大豆^[4]和单子叶植物的小麦^[3]、膨胀浮萍^[5]等的小亚基基因顺序分析可以得到以下几个特点：

1. 小亚基基因都有内含子(IVS)，数目因植物而异，小麦有1个，烟草有3个，豌豆则有2个(图1)。豌豆的5个小亚基基因(rbc S 3A, 3C, E9, 3.6, 8.0)的两个内含子位置相同，但rbc S 3A的IVS-1较大，并有高度变化^[6]。此外，矮牵牛小亚基SSU 11A有两个内含子，而SSU 8则有3个内含子，两个位置相同，第三个内含子位于蛋白质的高度保守区域。

2. 小亚基基因有很大的同源性。如矮牵牛的SSU 11A和SSU 8编码区的同源性达89%，成熟肽区域最为保守，两个小亚基基因间仅有1个氨基酸残基的区别，而转运肽却有5个氨基酸残基的差别^[8]。在不同植物的小亚基基因之间也存在很大的同源性，如矮牵牛SSU 8和烟草的rbc S之间有86%的碱基同源^[9]。图2表示不同植物的rbc S蛋白质中氨基酸顺序的同源性，双子叶植物的烟草、大豆、豌豆之间同源性大于与单子叶植物小麦的同源性。

rbc S基因5'端顺序也有很大的同源性，

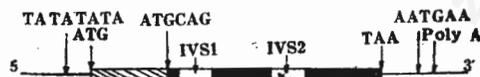
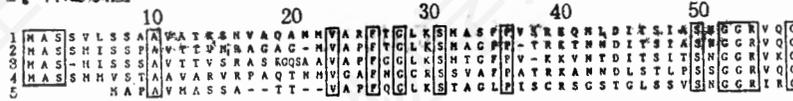


图1 豌豆的rbc S 3.6基因结构

IVS 1和IVS 2: 内含子; 斜线区: 转运肽区; 黑框区: 成熟肽区; TATA: 起动机; ATG: 起始密码; TAA: 终止密码; poly A: 添加信号。

I. 转运肽区



II. 成熟肽区

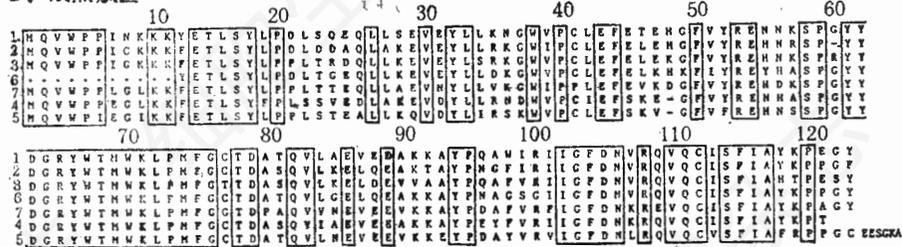


图 2 几种植物 rbcS 转运肽、成熟肽顺序比较

1. 烟草; 2. 大豆; 3. 豌豆; 4. 膨胀浮萍; 5. 小麦; 6. 矮牵牛; 7. 菠菜。线框表示保守区^[9]。

表 1 几种植物 rbc S 基因 5' 端的起动力结构

克隆名称	CAAT 盒	TATA 盒	文献
矮牵牛 ssu 11 A	CAAT	TATATATA	[2]
烟草 Ntss 7	CCAAT	TATATATA	[9]
Ntss 23	CCAATG	TATATATA	[9]
豌豆 rbc S-3 A		TATATATA	[7]
rbc S-3 C		TATATACCTA	[7]
pPS-2.4	GGCCAACT	TATATA	[10]
矮牵牛 SSU301(SSU 8)	CCAATG	TATATATAAAGG	[2]
SSU 611	GCAAAG	TATATATAGG	[12]
SSU 491	GCAAAT	TATAAATAGA	[12]
SSU 112	CCAAT	TATAAATAGA	[12]
SSU 911	CCAAAT	TATAAATAGA	[12]
豌豆 rbc S 3.6		TATATATAGC	[20]
rbc S 8.0		TATATAAAGC	[19]
大豆 SRS 1, 2, 3,	CAAT	TATATATA	[4]

在 -40 bp 左右有类似真核生物 TATA 盒区, 在 -80 bp 左右有类似真核生物 CAAT 盒区 (表 1)。此外, 小亚基基因 5' 端还有类似病毒或哺乳动物增强子核心顺序 GTGC AAA G 或 TTT AAA CCAC 的构造。矮牵牛的 SSU 8 基因有 5 个类似的顺序^[2], 烟草的 rbcS 基因有 4 个这样的顺序, 并在第 1 和第 3 的内含子中也各有一个^[9]。此外, 在小麦、大豆和豌豆的 rbc S 基因中都发现了这种顺序^[3,4,10] (图 3)。

在烟草 rbc S 基因中, 含有 4 个类似增强子顺序的 5' 端能形成茎环结构, 其中的两个信号

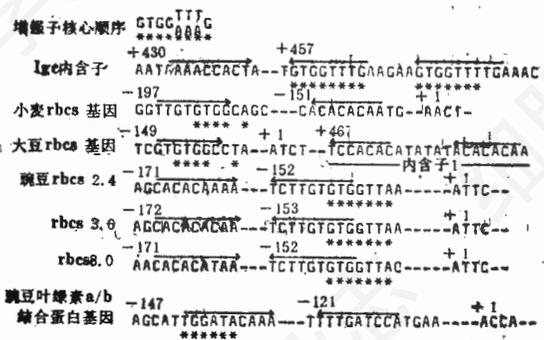


图 3 基因中类似增强子元件的顺序^[3,4,10]

通过一个能形成 67 bp 茎环结构的倒置重复而



图4 烟草 *rbc S* 基因 5' 端可能形成茎环结构的示意图

在 67 bp 中有 54 bp 完全配对, 6bp G:T 配对, 5bp 错配, 2bp 不配对。划线部分为类似增强子顺序^[8]。

连接(图4)。

小亚基基因 3' 端有 polyA 添加信号, 但与一般真核生物的 polyA 信号 AATAAA 有些不同。烟草的 *rbcS* 基因 3' 端的 polyA 添加信号是 TAATAATAAT^[9], 大豆中是 AATG-AA^[4], 在膨胀浮萍中是 AATTA^[11], 矮牵牛中是 AATAAT^[12], 豌豆中是 AATAAA (*rbcS 3A*) 或 AATGAA (*rbc S 3C, E 9, 3.6, 8.0*)。

小亚基多基因家族在不同器官中表达水平不一样, 豌豆 *rbc S* 多肽和 mRNA 在绿色组织中含量最高, 在果皮、花瓣和种子中含量次之, 黄化叶、茎和根中含量极低。光刺激不仅能提高叶中 *rbc S* mRNA 水平, 也能提高茎和根中的 *rbc S* mRNA 水平, 但叶的效果最显著, 说明光与组织类型有协同效应^[10]。小亚基基因的光诱导区可能位于 TATA box 或类似于增强子的顺序中^[13, 14]。

不同的小亚基基因在同一组织中表达水平不同。在豌豆叶组织中高表达的 *rbc S* 基因 (*rbc S3a, 3c, 8.0*) 转录子与低表达的 *rbc S* 基因 (*rbc S, E9, 3.6*) 的转录子水平相差 10 倍^[10], 在矮牵牛的 8 个 *rbc S* 基因中, 其中的一个 (SSU 301) 在叶中的表达占总的表达水平的 47%。在矮牵牛的 5 个基因中都发现有 mRNA 有效翻译的起始顺序 A/GXXATGG, 但只有在表达最多的 SSu 301 基因中相对于 ATG 的 -3 位置是 A 残基。体外实验表明: A 残基位于 ATG 的 -3 位置最有利于 mRNA 的翻译^[15]。

二、叶绿素 a/b 结合蛋白基因 (Cab)

叶绿素 a/b 结合蛋白是叶绿体类囊体膜的主要蛋白, 与光系统 II 联系在一起, 能把吸收

的光能从辅助色素转移到 PS II 反应中心。此外, 叶绿素 a/b 结合蛋白还参与类囊体膜形成基粒的过程。叶绿素 a/b 结合蛋白是由核基因编码, 在细胞质中合成前体分子, 转移到叶绿体中, 加工为成熟蛋白质, 与叶绿素 a, b 和类胡萝卜素非共价结合, 最后嵌入类囊体膜。

叶绿素 a/b 结合蛋白是由多基因家族编码, 矮牵牛至少有 16 个, 豌豆至少有 8 个, 膨胀浮萍至少有 10—12 个, 小麦至少有 7 个。目前已从小麦^[16]、膨胀浮萍^[17]、矮牵牛^[18]、豌豆^[19]、拟南芥菜^[20]分离出一些基因组克隆并进行了顺序分析。它们的 5' 端都有 TATA box 和 CAAT box (表 2), 但它们的 3' 端没有一般真核生物基因 3' 端所具有的 AATAA 的 polyA 添加信号。在矮牵牛的 5 个 cDNA 克隆终止密码子 3' 端下游 90 bp 有 TTTGTTT 顺序。在其基因组克隆 Cab 25, Cab 22 R, Cab 91 R 中也有这样的顺序^[18]。在拟南芥菜基因组克隆 AB 140 3' 端终止密码子下游 64 bp 也发现有 TTGTTT 顺序^[20]。小麦 whAB 1.6 基因 3' 端终止密码子下游 88 bp 也有 TTGTTT 顺序, 而豌豆基因组克隆 AB 66, AB 80 的 3' 端却有 AATACAA^[19]。膨胀浮萍 Cab 基因 3' 端却是 AATGAA 或 CACAAA^[21]。这些顺序可能代替一般的 polyA 添加信号 AATAAA 而起着 Cab 基因 3' 端的 polyA 添加信号作用。

在矮牵牛 Cab 基因翻译起始点上游约 50 bp 有类似真核转录起始顺序 TCAT^[18]。在拟南芥菜 AB 180, AB 165 翻译起始点上游 52 bp 有 TCACTC 重复顺序, 在 AB 140 上游 65 bp 有 ACCAC^[21]。在豌豆、小麦 Cab 基因翻译起始位置上游 68 bp, 70 bp 有 ACCAT^[22]。这些顺序作为 Cab 基因的转录起始顺序, RNA 聚

表2 Cab基因的5'端顺序

植 物	克 隆	CAAT盒	TATA盒	起始顺序	参考文献
拟南芥菜	pAB 140	CCAAT	TATTATATATA	ACCAC	[16]
	pAB 165	CCAAT	TATATAAT	TCACTC	[16]
	pAB 180	CCAAT	TATATAAT	TCACTC	[16]
小 麦	whAB 1.6	CCAAC	TTTAAATA	ACCAT	[19]
	AB 66	CCAAC	TATAAATA	ACCAT	[3]
豌 豆	AB 80	CCAAC	TATAAATA	ACCAT	[3]
	AB 19	CCAAT	TATTAA		[20]
膨胀浮萍	AB 30	CAAT	TATTA	ACCAC	[17]
	Cab 91 R	CCAAT	TATATATA		[18]
矮 牵 牛	Cab 13	CCAAT	TATAA		[18]
	Cab 25	CCAAT	TATATA		[18]
	Cab 22 L	CCAAT	TATATATA		[18]
	Cab 22 R	CCAAT	TAAATAAA		[18]

合酶可能从这里开始转录。不同的 Cab 基因也有些不同的特点。在拟南芥菜 AB 165, AB 180 5' 端有一个 23 bp 和 22 bp polyA 区, 其间被两个 C 残基所打断, 而 AB 140 无这个 polyA 区。在矮牵牛 Cab 91 R, 22 R, 25 转录起始点上游 - 300 bp 有一个 48 bp 的同源区, 这个区域可能与转录的调节有关^[18]。小麦 whAB 1.6 克隆 5' 端有 3 个重复区域, 第一个富含 GT, 第二个富含 CA, 第三个重复区有 3 个拷贝的 CTCAAG, 其中一个紧接在 TATA 盒前。它的 3' 端有 4 拷贝的 AATTG, 3 拷贝的 C/GAACATG, 在膨胀浮萍 Cab 克隆 AB 19 5' 端有长的 GT 重复顺序 (255—270 bp) 紧接着有 GA 重复顺序 (273—312 bp)。这样的顺序与 Z-DNA 的形成及一系列剪切连接过程有关。在豌豆 AB 80 5' 端发现有两个 TATA 盒。第一个位于加帽位点上游 31 bp 位点, 第二个位于起始密码子 5' 端 7 bp 位点, 而且两个 TATA 盒附近的顺序也很类似, 它们的启动子均可能起作用。此外, 在 AB 80 和 AB 65 5' 端 - 143 处有 TTGGATAC 顺序, 类似病毒增强子的核心顺序 (GTGG^{TTT}_{AAA}G)。

Cab 基因中一般没有插入顺序, 但在膨胀浮萍的两个基因组的一个克隆 (AB 19) 中发现有 1 个 84 bp 的插入顺序, 在插入顺序与外显子的边缘处的顺序为顺向重复 (Direct repeat), 与之紧连的是反向重复顺序, 说明该插入顺序

可能是起源于转座子。Cab 基因的编码区可分为转运肽区和成熟肽区。Cab 基因表达受光敏色素的调控。在黄化大麦、膨胀浮萍、豌豆中, Cab mRNA 水平很低, 但可被短暂的红光照射显著提高, 该效应并可被远红光逆转^[22]。Lamppa 等利用 Ti 质粒载体将小麦的 Cab 1 基因 (whAB 1.6) 转入烟草基因组, 发现在烟草中小麦 Cab 1 基因的表达仍受光的调节^[23]。Nagy 等的研究表明, Cab 基因 5' 端 1.8 kb (- 1816 到 + 31) 片段使 Cab 基因接受光调节, 并在叶组织中特异地表达^[24]。

此外, 还有两个核基因编码的叶绿体蛋白基因已分离出来, 它们是麦瓶草 (*Silene pratensis*) 的铁氧还蛋白 (ferredoxin)、cDNA 克隆

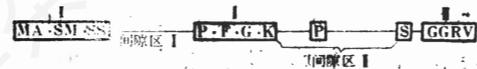


图5 在核基因编码的叶绿体蛋白转运肽中的同源顺序^[27]
框内大写字母表示共同的氨基酸残基。

pFD 1 和质蓝素 (plastocyanin) cDNA 克隆 pPC 1, pPC 2^[25,26]。最近从豌豆中也分离得到铁氧还蛋白的 cDNA 克隆^[27]。以上 4 个核基因编码的叶绿体蛋白的转运肽有 3 个同源区, 它们分别位于转运肽的 N-端、中部和 C-端, 并形成两个间隙区 (图 5)。间隙间中有两

个不变的氨基酸残基(脯氨酸P和丝氨酸S)。这些同源顺序可能与它们的蛋白前体进入叶绿体、加工以及定位有关。同源区Ⅲ的GRV可能与转运肽的切除有关。

研究核编码的叶绿体蛋白质基因结构和表达调节,对了解核质关系、核基因对细胞器功能的调控等具有重要的意义。

摘 要

高等植物叶绿体的一些重要蛋白质基因位于细胞核的基因组内。这些基因的细微结构已研究得较为清楚,并对不同植物的编码叶绿体蛋白质的核基因进行了比较研究。

参 考 文 献

- [1] Dean, C., et al., 1985, *Proc. Natl. Acad. Sci. USA.*, 82: 4964—4968.
- [2] Coruzzi, G. R., et al., 1983, *J. Biol. Chem.*, 258: 1399—1402.
- [3] Broglie, R., et al., 1983, *Biotechnology* 1, 55—61.
- [4] Berry-lowe, S. L. et al., 1982, *J. Mol. Appl. Genet.*, 1, 483—498.
- [5] Wimpee, C., et al., 1983, In *Plant Molecular Biology*, R. B. Goldberg (ed.) pp. 391—401.
- [6] Fluhr, R., et al., 1986, *The EMBO J.* 5: 2063—2071.
- [7] Polans, N. O., et al., 1985, *Proc. Natl. Acad. Sci. USA.*, 82: 5083—5087.
- [8] Tumer, N. -E., et al., 1986, *Nucl. Acids Res.*, 14: 3325—3342.
- [9] Mazur, B. J., & Chui, C. F., 1985, *Nucl. Acids Res.*, 13: 2373—2386.
- [10] Coruzzi, G. R., et al., 1984, *The EMBO J.*, 3: 1671—1679.
- [11] Stiekema, W. J. et al., 1983, *Nucl. Acids Res.*, 11: 8051—8061.
- [12] Dean C., et al., 1985, *The EMBO J.*, 4: 3055—3061.
- [13] Fluhr R. & Chua N. -H., 1986, *Proc. Natl. Acad. Sci. USA.*, 83: 2358—2362.
- [14] Fluhr R., 1986, *Science* 232: 1106—1112.
- [15] Kozak M., 1984, *Nature*, 308: 241—246.
- [16] Broglie R., et al., 1981, *Proc. Natl. Acad. Sci. USA.*, 78: 7304—7308.
- [17] Kohorn B. D., et al., 1986, *J. Cell Biol.*, 102: 972—981.
- [18] Dunsmuir P., 1985, *Nucl. Acids Res.*, 13: 2503—2518.
- [19] Timko M. P. et al., 1985, In *Molecular Biology of the Photosynthesis Apparatus*, Steinbeck K. E. et al., (eds) Cold Spring Harbor Laboratory Press. pp. 381—396 N. Y.
- [20] Karlin-Neumann G. A. et al., 1985, *J. Mol. Appl. Genet.*, 3: 45—61.
- [21] Leutwiler L. S. et al., 1986, *Nucl. Acids Res.*, 14: 4051.
- [22] Lamppa G. K., et al., 1985, *Mol. Cell Biol.*, 5: 1370—1378.
- [23] Lamppa, G. K. et al., 1985, *Nature*, 316:750—752.
- [24] Nagy, F. et al., 1986, *EMBO J.*, 5: 1119—1124.
- [25] Smeekens S., et al., 1986, *Nucl. Acids Res.*, 13: 3179—3184.
- [26] Smeekens S., et al., 1985, *Nature*, 317: 456—458.
- [27] Dobers, M. S. et al., 1987, *Plant Mol. Biol.*, 8: 77—85.

DNA 烷化损伤及修复的分子基础

章 扬 培

(军事医学科学院放射医学研究所)

韩 文 智

(中国农科院品种资源研究所)

DNA 是生命活动中最重要的遗传物质,也是环境中物理和化学因素攻击的主要靶分

子。DNA 损伤与修复的研究一直是生命科学的前沿课题。近年来,不少学者从分子水平上