

酵母线粒体 DNA 与表达

周婉冰

(华南理工大学生物工程系 广州 510641)

线粒体有它独特的 DNA、rRNA、tRNA 和核糖体。线粒体 DNA(Mitochondria DNA, 简称 mtDNA)是真核细胞中分子量较小而又较易纯化的单位,是研究 DNA 结构、基因表达与调节的良好材料。与其它真核生物相比,酿酒酵母(*S. cerevisiae*)能在线粒体呼吸活性缺陷情况下存活,特别适合于线粒体基因表达的研究。本文就酵母线粒体基因组、线粒体内含子的结构与产物、基因表达等问题作一综合介绍。

1 酵母线粒体 DNA 的结构与产物^[1~3]

酵母 mtDNA 是个双链环状分子,大小介乎 73 ~ 82kb 之间,因菌株而异。现今统计酿酒酵母有 46 种线粒体基因功能产物(表 1),包括 2 个 rRNA、26 个 tRNA、细胞色素 C 氧化酶的三个亚基,细胞色素 b 脱辅基蛋白,ATP 酶的三个亚基,线粒体核糖体 37S 亚基中的 Var1 蛋白等的基因(遗传图谱见文献[1])。其中七种位于线粒体内膜的蛋白(即 CoX 1、2、3、ATP6、8、9 和 CoB)直接参与呼吸作用,28 种 RNA 和 Var1 蛋白在线粒体翻译过程中起作用;还有 10 种线粒体产物非常特殊,由内含子编码,并非线粒体呼吸所必需,其功能见下述。

线粒体基因组有多个转录起始位点;有大量富含 A、T 短序列的非编码区,含有多个 mtDNA 复制点。

2 酵母线粒体内含子及其编码的蛋白质^[1,3,4]

与哺乳动物线粒体基因显著不同,酵母菌有些线粒体基因存在内含子,如酿酒酵母的 CoX1、CoB 和 LSU 基因是断裂基因,内含子数目因菌种而异。根据内含子的保守序列及二级结构,内含子可分为若干类型。已鉴定出有 80 多种 I 类内含子和 70 多种 II 类内含子。酵母线粒体内含子大多属于 I 类,少数是 II 类。

对酵母 mtDNA 编码区域序列研究发现,许多内含子都含有相当长的可译读框(oper reading frames, ORF)或未定译读框(unassigned reading frames, URF)具有编码蛋白质的潜力,其中有些还能与其上

游或下游的外显子(exon)编码序列连续,其产物为融合蛋白或称杂蛋白,功能上与断裂基因初级转录物的拼接有关。这种蛋白因子称为 mRNA 成熟酶。

2.1 由 I 类内含子编码的蛋白质:对酿酒酵母 I 类内含子的七个未定译读框(URF)的分析发现,其产物形成了一个蛋白质家族,都含有命名为 P1 和 P2 或 LAGLI-DADG 的十肽单位。最先证明这一家族成员与酵母线粒体拼接有关的是对 CoB 基因的研究。该长基因有 3 个 URF。第一个位于长基因的非必需内含子 bI2I 中,其序列已被测定。翻译从细胞色素 b 脱辅基蛋白的起始密码子(AUG)开始,通过 BE1 和 BE2 两外显子,延伸 954 bp,产生一个融合蛋白,其 N-端部分来自外显子序列(具高度疏水性和非极性),C-端部分则来自内含子序列(具强碱性)的 318 个氨基酸的蛋白质。推测此融合蛋白是介导 bI2I 切除的成熟酶(或称反式激活因子, TAF),把 BE1+BE2 与 BE3 连接起来,从而消除了产生成熟酶的 mRNA 本身。现证明 I 类内含子 bI3I^{6I} 和 bI1s /bI4I 的翻译产物也具有成熟酶活性。由 bI4I 表达的成熟酶不仅协助 CoB 的 bI4I 切除,而且也是 CoX1 的 aI4 切除所必需,因此其突变不仅抑制细胞色素 b 脱辅基蛋白的表达,还能抑制 CoX1 的表达^[7]。

I 类内含子编码的蛋白质属于单一蛋白质家族,但并不意味着它们只有单一活性。一些 I 类内含子具有移动性,能从原位切离(称内含子丧失)或插入无内含子的等位基因中繁殖自身(intron-homing)。通过内含子繁殖缺陷突变体的分析,揭示内含子可读框(RF)的重要作用。1986 年报道^[8],内含子 rI 的翻译产物是一种双链 DNA 内切酶,专一性地切割无内含子基因的“建家”(homing)位点。酵母的 aI3、aI4、aI5 的翻译产物(表 1)及其它生物有几种 I 类内含子产物都是核酸内切酶,经它们切割后可产生四种碱基的游

1995-08-14 收稿

表 1 酵母线粒体基因和产物

基 因	图谱位置	产 物
CoX1	41 ~ 58	细胞色素 C 氧化酶亚基 1
内含子 a11		a11 成熟酶
内含子 a12		a12 成熟酶
内含子 a13		ds DNA 核酸内切酶 I Sce III
内含子 a14		ds DNA 核酸内切酶 I Sce II
内含子 a15a		ds DNA 核酸内切酶 I Sce IV
内含子 a15b		未知
CoX2	14.2 ~ 15.3	细胞色素 C 氧化酶亚基 2
CoX3	19.9 ~ 21.0	细胞色素 C 氧化酶亚基 3
ATP6	61.5 ~ 62.6	ATP 酶亚基 6
ATP8	~ 61	ATP 酶亚基 8
ATP9	~ 81	ATP 酶亚基 9
CoB	71 ~ 76	细胞色素 b 脱辅基蛋白
内含子 b12		b12 成熟酶
内含子 b13		b13 成熟酶
内含子 b14		b14 成熟酶
VAR1	~ 86	核糖体蛋白 Var1
ENS2	~ 65	ds DNA 核酸内切酶亚基 2
内含子 r1	在 LSV 基因中	ds DNA 核酸内切酶 I Sce I
RF1		CoX2 下游的 URF
RF2		CoX3 下游的 URF
其它 RFs		两个推断的 URF
25tRNA	19 个在 0 ~ 20 4 个在 20 ~ 25 范围内	25 种 tRNA
LSU	94.2 ~ 99.4	21S rRNA
SSU	36 ~ 38.3	15S rRNA
tRNA _{syn}		mtRNA _{sep} RNA 组分

离 3'-端。双链 DNA 的破坏可能启动了修复过程，此过程使用一个含有内含子的 DNA 序列作模板。CoB 内含子不能翻译为核酸内切酶，表明 CoB 内含子没有“建家”的能力。

酵母 mtDNA 中有三个非内含子可读框，其产物也象内含子编码的蛋白质那样具有 LAGLI-DADG 单元。RF1 和 RF2 的作用仍不清楚。RF3 位于 ATP6 下游，编码与基因重组有关的线粒体内切酶 *Secl* 杂二聚体中的 ENS2 亚基 (ENS1 亚基由核编码)，是一个分

子量为 70000u 的热休克蛋白。

LAGLI-DADG 家族的两个成员已在酵母的核基因组中鉴定出来，产物也是核酸内切酶：由 HO 基因编码的位点专一性核酸内切酶(切割 MAT 区域的一个位点来启动接合型的开关控制)和 VDE 核酸内切酶(可能是通过对初级翻译产物进行拼接而成，这一拼接是受这种蛋白质前体本身所催化)。因此，LAGLI-DADG 家族既包括 RNA 成熟酶，也包括 DNA 核酸内切酶。通过简单的点突变(min2-1)可将

a14 核酸内切酶转变为能切割 a14 和 b14 两种内含子的成熟酶。这一结果表明, 几乎完全一样的蛋白质能具有很不相同的催化功能。可以把这些编码蛋白质的内含子看作是一种寄生分子, 能将它们自身传播到新的位置中去。这对于内含子通过自我催化方式切割拼接自身的作用至关重要。

2.2 由 II 类内含子编码的蛋白质: 酵母线粒体 CoX1 基因的 a11 和 a12 属 II 类内含子, 它们的可读框编码两个功能与结构相近的 RNA 成熟酶。其表达方式与 b14 成熟酶相似, 即先被翻译为杂蛋白质再经加工后形成成熟酶。在缺少 a12 成熟酶时可通过提高 a11 成熟酶含量来补偿, 说明两者功能相近。II 型翻译产物与 I 型相应组分很不相同。“内含子建家”现象在 a11 和 a12 中都被发现。

3 线粒体基因的表达与调节^[1,2,4]

线粒体只能合成有限的几种 RNA 和蛋白质。90% 以上线粒体多肽由核基因编码, 线粒体的 DNA 聚合酶、RNA 聚合酶以及绝大多数 γ -蛋白也都由核基因编码。因此线粒体基因的复制与表达高度依赖核基因。

3.1 线粒体基因的转录加工: 线粒体基因组并不是转录成全长的前体 RNA 分子。已证明酵母 mtDNA 包含 19 ~ 20 个转录起始位点, 转录起始序列(A/T) TATAAGTA, 最后的 A 为“+1”。没有发现相关的加帽转录序列。强弱启动子与 RNA 聚合酶结合的能力可相差 20 倍之多, 强启动子的 +2、+3 位点上大多是 A 和 T。由于许多 mtDNA 基因缺乏转录起始位点, 以 RNA 转录起点作为前体 RNA_s 的分隔, 形成 LSU /tRNA^m、CoX2 /RF1、CoX1 /ATPase 等较大的转录单位, 经加工形成成熟的 RNA。

至于转录终止的标志还不很清楚。已知许多 mRNA_s 的 3'-端具有 12 碱基序列(AAUAUAUUCUU), 这是内切核酸酶裂解主要位点, 转录遇到这一序列往往会减慢或停止。这已在 CoX1 /ATPase 转录单位证实, CoX1 转录的 3'-端直接与 ATPase 前体 RNA 的 5'-末端邻近, 并都在这段序列上。由于在此段裂解而产生 RNA 有一个稳定的 3'-末端和进一步处理的下游 5'-末端。发现 CoB-mRNA 也在这 12 碱基序列处终止。事实上这段序列可作为转录衰减和处理加工的双功能信号, 还作为 mRNA 稳定性或可翻译性

的决定体。下面两个例子说明, 这 12 碱基的裂解控制是调节基因表达的一种手段。例一, 12 碱基存在于编码蛋白序列的 3'-末端, 其缺失突变体不能执行 RNA 裂解功能, 且内含子交叉传递有缺陷。12 碱基裂解可能是 DNA 内切酶未拼接的前体 RNA 产生功能性 mRNA 的途径之一; 例二, 说明 12 碱基的裂解控制对 RF1 表达是必须的, RF1 位于 CoX2 基因下游, 其起始点与 CoX2 的 C-末端部分重叠。12 碱基裂解发生在 CoX2 转录产物 3'-末端位置时将导致 RF1 表达的丧失, RF1 的合成依赖于这点的裂解或使裂解发生在此点的上游端。

Var 1 基因与 ATPase 9 和 tRNA^m 序列的前体组成一个转录单位。加工对于产生功能性 mRNA 很重要, 而裂解主要发生在 t 碱基序列上(AUAUAA)。此序列常发现于基因内部, 现在还不清楚如何保护这点不被裂解, 可能涉及其它序列或其二级结构。

3.2 RNA 拼接与内含子切除: 如上所述, 酵母 mtDNA 中有 3 个基因是拼接的。I、II 类内含子二级结构不同, 在体外自我拼接机制不同, 但都要经过两步连续的转酯反应, 导致分别与 5'-及 3'-拼接点断裂。两类内含子拼接的主要区别是激发第一步反应的羟基特性, I 类以辅助因子 G3'-OH 作为亲核攻击基团, 而 II 类则以分子内部的 A-2'-OH 为攻击基团。第二步反应都以 5'外显子 3'-OH 作为攻击基团。拼接结果除产生成熟的 RNA 外, I 类还有线状内含子(经再次转酯反应后为环状), II 类有一个以分子内 A 为分支点的套索状内含子。虽然有些内含子在体外试验始终不表现自我拼接作用, 但这种拼接机理对大多数内含子是适用的。

除成熟酶外, 由核编码的拼接因子与线粒体内含子加工有关。这些因子大致可分三类: (1) 可能直接与拼接有关; (2) 可能是 RNA 成熟酶合成所需; (3) 作为线粒体拼接缺陷的多拷贝校正因子。目前, 对于成熟酶及拼接因子等蛋白质如何控制拼接了解尚少, 推测可能: (1) 调整和稳定前体 RNA, 使之成为有利于拼接的结构; (2) 阻止异常的副反应; (3) 通过静电屏蔽和反应基团的活性来加强 RNA 的催化效率。

4 线粒体翻译系统的特点^[5]

线粒体翻译系统以酿酒酵母研究得最深入。在抑制翻译的抗生素图谱上, 在翻译系统的主要组分以及

核糖体组成等方面, 线粒体与原核生物比与该真核生物细胞质的更为相似。这里只概括线粒体翻译系统的一些独特之处。

(1) 不同物种的线粒体遗传系统使用的遗传密码本与通用的遗传密码本不同, 物种之间也有差异。

(2) 线粒体中 tRNA 数量有限, tRNA 辨认密码有所拓展。如酿酒酵母, 若反密码子的第一个碱基位点被未经修饰的尿苷占据, 那么单个 tRNA_s 能解读四码氨基酸家族的所有四个三联密码子。在反密码子的摆动位点为 5'-羧甲基氨基甲基尿苷占据的单个 tRNA_s 能辨认以嘌呤结尾的两码家族的所有两个三联密码子。

(3) 核糖体与原核生物极为相似, 但不同物种的线粒体核糖体也很不一样。酵母线粒体核糖体单体为 74S, 可分解为 50S 和 37S 两个亚基, 分别含有 21S rRNA 和 15S rRNA。核糖核蛋白含量及种类比原核生物及同种真核细胞质的核糖体更大。除与 37S 小亚基有关的 Var1 基因产物外, 其它 r-蛋白由核基因编码, 在线粒体外合成再运载到体内组装。

(4) 酵母线粒体的翻译需要一个或多个由核编码的特异性翻译激活因子的参与。例如 CoX3 mRNA 的翻译需要三个专一性核基因产物 PET494、PETS4 和 PET122。这三个蛋白因子与线粒体内膜相连, 通过与 CoX3-mRNA 的 5'-端不被翻译的前导序列相互作用而起激活作用; CoB-mRNA 的翻译与 CBS1 和 CBS2 因子有关; PET111 特异地与 CoX2-mRNA 正确翻译有关。这些特异性翻译激活因子与各自的目标 mRNA_s 的 5'-端前导序列间相互作用, 执行不同的功能: ① 可能把 mRNA 和核糖体小亚基带到线粒体

内膜上, 在内膜上发生翻译的起始, 因而使编码的蛋白质在多亚基呼吸酶系装配位点附近进行合成; ② 在选择翻译的起始位点上起作用(上述的激活因子都参与了翻译的起始); ③ 可能调节特异 mRNA_s 的翻译, 即激活因子的表达组成了部分的调节网络, 随细胞内外环境而调整线粒体合成的水平。此外, MSS51、CBP6、AEP1 和 ATP13 四个基因与专一性的线粒体 mRNA 翻译有关。

致谢 本文得到林影、潘力和刘淑文等同事的帮助, 特此致谢。

参 考 文 献

- [1] Herman J Pel, Leslie A Grivell. *Molecular Biology Reports*, 1993, **18**: 1 ~ 13.
- [2] Leslie A Grivell. *Eur J Biochem*, 1989, **182**: 477 ~ 493.
- [3] 陈士怡, 徐洪基. 酵母遗传学. 北京: 科学出版社, 1989. 147 ~ 170.
- [4] 孙乃恩. 分子遗传学. 南京: 南京大学出版社, 1993. 217 ~ 233.
- [5] Herman J Pel, Leslie A Grivell. *Molecular Biology Reports*, 1994, **19**: 183 ~ 194.
- [6] Lazowska J, Claisse M, Gargoani A, *et al.* *J Mol Biol*, 1989, **205**: 275 ~ 289.
- [7] Bunroques J, Perea J, Jacq C. *EMBO J*, 1987, **6**: 1085 ~ 1091.
- [8] Collear L, d'Auriol L, Betermier M, *et al.* *Cell*, 1986, **44**: 521 ~ 533.