

文章编号: 0529-6579 (2000) 04-0074-06

拟南芥 U59 snoRNA 基因簇的结构与功能研究^{*}

孟 清, 周 惠, 许 飞, 屈良鹄

(中山大学生物工程研究中心, 广州 510275)

摘 要: 在拟南芥 (*Arabidopsis thaliana*) 中发现和鉴定了具有两个反义序列的 snoRNA 组成的一个基因簇. 该基因簇由两个相同的 snoRNA 基因组成. 位于 box D' 上游的反义序列与人 U59 snoRNA 基因的反义序列相同, 因此命名为拟南芥 U59a 和 U59b. 位于 box D 上游的反义序列在所有哺乳动物和酵母 snoRNA 基因的反义序列中未找到同源拷贝. 经低浓度的 dNTP 逆转录检验证实此反义序列并不指导拟南芥 25S rRNA G1845 位点 2'-O-核糖甲基化. 这些结果表明: 拟南芥 U59a 和 U59b snoRNA 基因与动物 U59 snoRNA 在基因结构、组织和功能上有较大的差异. 拟南芥 U59 snoRNA 基因簇的发现为进一步研究植物 rRNA 的加工和修饰、植物 snoRNA 基因的组织与表达多样性以及两个反义序列 snoRNA 基因的起源提供了一个良好的研究体系.

关键词: 拟南芥 (*Arabidopsis thaliana*); U59; snoRNA; 基因簇

中图分类号: Q752 **文献标识码:** A

在核糖体生物合成中起重要作用的核仁小分子 RNA (small nucleolar RNA, snoRNA) 有 3 种不同的基因组织: 独立转录的 snoRNA 基因组织, 内含子编码的 snoRNA 基因组织和 snoRNA 基因簇^[1]. 而 snoRNA 基因簇由于基因组织的多样性和表达机制的复杂性以及在不同的生物基因组中存在有选择性, 备受从事分子生物学领域研究的科学家们的关注. 已成为当前分子生物学领域研究的热点. 迄今为止在动物基因组中尚未发现 snoRNA 基因簇; 从已报道的酿酒酵母 (*Saccharomyces cerevisiae*) 基因组全序列中只发现了 7 个 snoRNA 基因簇^[2]. 虽然在部分植物基因组中发现和鉴定了 snoRNA 基因簇, 例如玉米中发现独立转录的 snoRNA 基因簇, 水稻中则发现由蛋白质基因 Hsp70 内含子编码的 snoRNA 基因簇^[3], 但是植物 snoRNA 基因簇的表达过程和机制至今尚未被阐明. 拟南芥 (*Arabidopsis thaliana*) 基因组全序列的测定即将完成, 许多编码蛋白质的基因已被大量地发现, 但是编码 RNA 的基因却鲜有报道. 如我们最近发现的拟南芥 Z2 snoRNA 基因簇是拟南芥中首次报道的 snoRNA 基因簇 (周惠等, 待发表). 因此采用各种方法发现和鉴定其中新的 snoRNA 及其基因组织, 对于阐明植物 snoRNA 基因簇的表达过程和调控机制具有重要的意义.

通过对拟南芥基因组数据库分析并经过实验验证, 本文首次发现和鉴定了一个具有两个反义序列的 snoRNA 基因簇.

^{*} 基金项目: 国家自然科学基金重点项目资助项目 (39730300); 国家教育部重点项目资助项目

收稿日期: 2000-05-11; 作者简介: 孟清 (1955~), 男, 博士; 通讯联系人: 屈良鹄.

1 材料和方法

通过 Internet 采用 Blast 和 Fasta 程序对国际基因数据库中的核酸数据进行分析，再以 PC gene 6.0 软件包对所获数据作进一步序列鉴定和分析^[3]。

用于总 RNA 提取制备的拟南芥 新鲜幼苗为本实研室栽培。按酸性异硫氰酸胍一步法提取纯化总 RNA^[4]。拟南芥总 DNA 的制备按 CTAB 法进行^[5]。

根据拟南芥 U59 snoRNA 基因簇的保守序列设计出 Northern 杂交分析的寡核苷酸探针序列为：5'CCTCAGCTCAGTTCAGGTCTC 3'，可分别与 U59 snoRNA 基因 序列的 3'末端互补配对。探针的 5'端以 [γ-³²P] ATP（北京亚辉公司）标记。10 μg 总 RNA 以 8%变性聚丙烯酰胺凝胶电泳分离，转移至尼龙膜（Amersham 公司），尼龙膜经紫外交联后以标记探针进行杂交检测。

拟南芥 25S rRNA 部分序列的逆转录及序列分析。逆转录反应的体积为 40 μL，含 10 μg 总 RNA 和 20 ng [γ-³²P] ATP 标记的 Pz42 M1840 特异性引物（引物序列为：5' CATTGTTCCATCGACCAGAGGC 3'），按照不同的 dNTP 浓度要求，分别加入适量的 4 mmol/L 和 40 μmol/L 的 dNTP，制成终浓度分别为 1.5 mmol/L 和 4 μmol/L 的反应液。反应液经 65 ℃变性 5 min 后，冷却至 42 ℃。加入 200 U M-MLV 逆转录酶（Promega），于 42 ℃延伸反应 30 min，以 8 mol/L 尿素—8% 聚丙烯酰胺凝胶电泳检测逆转录结果。以 Z44M2116 和 At28S4F 为引物（序列分别为：5' CCGCAGGGACCATCGCAATGC 3'和 5'TATAGGGGCGAAAGACTAATCG 3'），对拟南芥总 DNA 进行 PCR 扩增，扩增产物经纯化后克隆至 PIZ18 质粒的 SmaI 位点，以 5'末端 [γ-³²P] ATP 标记的 Pz42M1840 为引物，测定重组质粒 DNA 的序列作为分子量标准。

2 结 果

通过 Internet 联网，我们用计算机对欧洲 EMBL 基因库和美国 Stanford 大学的拟南芥基因组数据进行分析，寻找具有 snoRNA 结构特征的核酸序列。经大量的分子数据处理，在拟南芥基因组中发现了一批具有 snoRNA 结构特征的 DNA 序列，其中的一段位于拟南芥第 5 号染色体上（克隆编号为 F20B22），被命名为拟南芥 U59 DNA（图 1）。该 DNA 序列中包含两个分别具有典型的 snoRNA 结构特征的序列，即具有 box C（TGATGA）和 box D

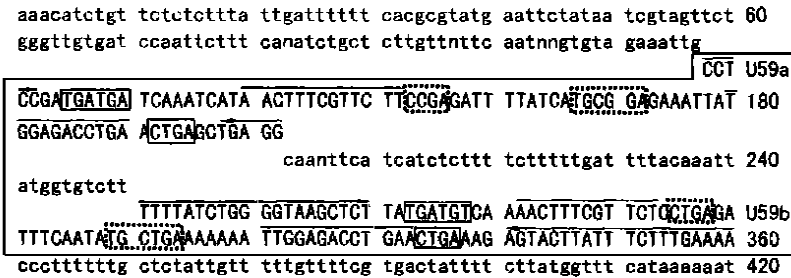


图 1 拟南芥 U59 snoRNA 基因簇的序列

Fig 1 The sequence of *A. thaliana* U59 snoRNA gene cluster

NA 基因以大写字母表示，box C/D 或 box C'/D' 结构元素分别以实线或虚线框示之，
线表示与 18S 和 25S rRNA 互补的反义序列，反向箭头表示可形成末端配对区的序列

(CTGA) 这两个保守的 snoRNA 结构元素. 位于 box C 和 box D 这两个结构元素之内还有 box C' (TGCT/CGA) 和 box D' (CT/CGA) 结构元素. 紧位于 box D' 上游分别具有一段长为 13 和 12 个核苷酸的反义 snoRNA 功能序列, 两段序列都可与拟南芥 18S rRNA 中部的一段保守核心序列互补. 位于 box D 上游同样分别具有一段长为 11 和 12 个核苷酸的反义 snoRNA 功能序列 (图 2), 两段序列可与拟南芥 25S rRNA 中部的一段保守核心序列互补. 在 box C 的上游和 box D 的下游分别有一段反向重复的核苷酸序列, 可使对应的 RNA 分子形成较为稳定的末端二级结构 (图 3).

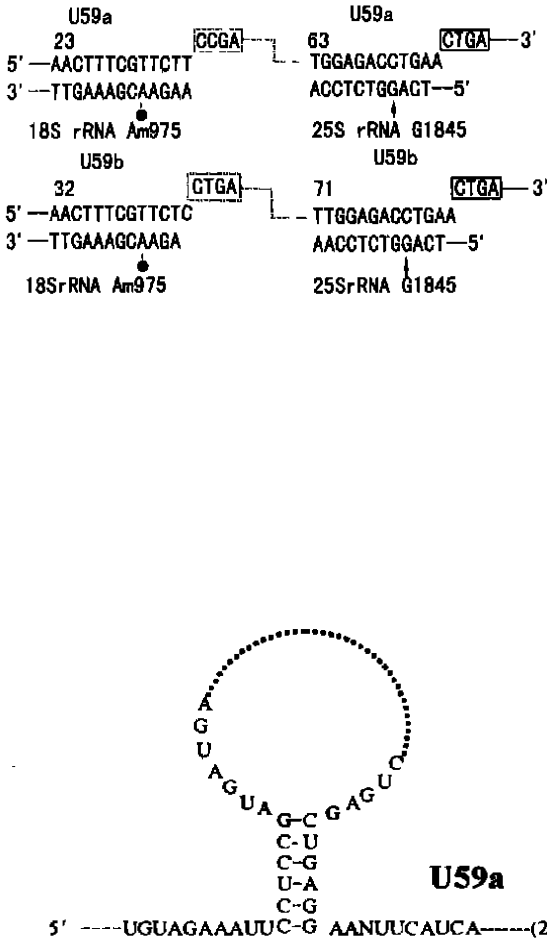


图 2 拟南芥 U59 snoRNA 的反义序列与 18S 和 25S rRNA 2'-O-核糖甲基化位点关系

Fig. 2 Methylation guide duplexes between U59 snoRNAs and *A. thaliana* rRNA 2'-O-ribose methylation sites

根据理论计算的甲基化位点以箭头表示, 与人 rRNA 同源的甲基化位点以实心圆示之

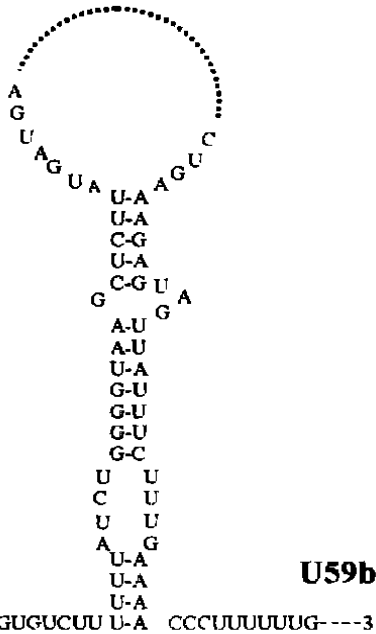


图 3 拟南芥 U59 snoRNA 的末端 (包括部分上下游序列) 可形成的二级结构

Fig. 3 Prediction of the secondary structures of U59 snoRNA terminal region including part flanking sequences

比较上述 2 个可能编码 snoRNA 的 DNA 序列, 发现两者之间存在着很大的差异, 但是功能序列几乎完全相同. 按照 snoRNA 结构与功能的关系计算^[6], box D' 上游的反义序列可指导拟南芥 18S rRNA A 975 位点 2'-O-核糖甲基化, 而位于 box D 上游的反义序列可能指导拟南芥 25S rRNA G1845 位点 2'-O-核糖甲基化 (图 2), 因此它们是同一种 snoRNA 基因的不同拷贝. 将该基因与各种生物中已鉴定的 snoRNA 基因进行比较, 发现该基因位于 box D' 上游的反义序列与鉴定自人 U59 snoRNA 基因中的反义序列相同. 所以命名为拟南芥 U59a

和 U59b snoRNA 基因. 但是将位于 box D 上游的反义序列与各种生物中已鉴定的 snoRNA 基因的反义序列进行比较未发现具有同源功能序列的 snoRNA 基因拷贝, 是一种新的反义序列.

根据 U59 DNA 中 snoRNA 基因编码区序列, 设计合成了一个 21 个核苷酸长的 U59 snoRNA 特异性探针 Pu59, 对 U59 snoRNA 基因簇的表达产物进行实验验证. 以 ^{32}P 标记 Pu59 探针对拟南芥总 RNA 进行的 Northern 杂交的结果证实 (图 4), 在严格的条件下, Pu59 探针分别与大小为 85 和 110 个核苷酸的 RNA (即 U59a 和 U59b snoRNA) 稳定地杂交, 这些长度与预期的结果相吻合. 虽然大部分具有 box C 和 box D 的 snoRNA 的 5' 和 3' 末端都在其 box C 上游和 box D 下游

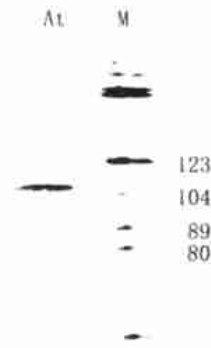


图 4 以 Pu59 为探针对拟南芥总 RNA 进行 Northern 杂交的结果

Fig. 4 Northern analysis of *A. thaliana* total RNA
M: pBR322 质粒用 *Hae* III 和 *Tag* I 酶切后以 ($\gamma\text{-}^{32}\text{P}$) ATP 标记的分子量标准

10 个核苷酸以内, 但也有个别例外. 本实验中鉴定的 U59b snoRNA 即是如此 (图 1). 原因可能是对 snoRNA 的加工不完全造成的, 也可能是 snoRNA 稳定自身结构所需要的. 如果能量平衡对二级结构的稳定性产生影响, 那么长末端反向互补序列更有利于 U59b snoRNA 二级结构的稳定存在 (图 3).

与哺乳动物 U59 不同, 在拟南芥 U59a 和 U59b snoRNA 中包含了一段与 18S rRNA 保守区互补的核苷酸序列, 该序列是否作为指导一个 25S rRNA (G1845) 新的甲基化位点向导序列, 还是有其它功能? 由于拟南芥中 rRNA 的甲基化核苷酸尚未被准确定位, 所以有必要用实验的方法对此加以验证. 在逆转录反应中, 逆转录酶是否能正常通过 RNA 模板合成预计的 cDNA, 除其它因素影响之外, 直接与反应液中 dNTP 浓度的大小有关. 如果 dNTP 的浓度较低, RNA 模板中的 2'-O-核糖甲基化核苷将阻碍逆转录酶的正常通过, 一部分 cDNA 的合成将在该甲基化核苷前停顿; 提高反应液中 dNTP 的浓度, 可以有效地减弱这一影响. 因此, 可以用不同浓度的 dNTP 进行逆转录的方法来鉴别 RNA 分子中的 2'-O-核糖甲基化核苷的存在^[3]. 鉴别拟南芥 25S rRNA 分子中的 2'-O-核糖甲基化核苷的逆转录反应所选用的 dNTP 的浓度为 $4\text{ }\mu\text{mol/L}$. 为了对上述逆转录反应中与 dNTP 的浓度相关的 2'-O-核糖甲基化核苷进行精确的定位, 以 Z44M2116 和 At28SC4F 为引物, 扩增了拟南芥 25S rRNA 包含该 25S rRNA G1845 核苷互补的序列, 并将该序列克隆到质粒中. 以逆转录反应中所用的同一引物对克隆的 25S rRNA 进行序列测定, 作为分子量对照.

实验结果清楚地表明: 在 dNTP 的浓度为 $4\text{ }\mu\text{mol/L}$ 时, 拟南芥 25S rRNA 的第 1846 位点的鸟嘌呤核苷处没有形成明显的逆转录停顿点 (图 5), 与在 1.5 mmol/L 高浓度 dNTP 下转录的情况一致. 即 25S rRNA 1845 位点的鸟嘌呤核苷上的 2'-O-核糖未被甲基化. 这意味着位于拟南芥 U59a、U59b snoRNA box D 上游的反义序列并不指导拟南芥 25S rRNA G1845 位点 2'-O-核糖甲基化 (图 2).

3 讨 论

U59 是一种保守的 snoRNA, 它已被发现存在于酿酒酵母 (snR54) 和哺乳动物中. 但是在

人和酵母中 U59 都是由内含子编码的单拷贝的 snoRNA, 且只具有一个反义序列. 拟南芥基因组中编码紧密相连多拷贝的且具有两个反义序列的 U59 snoRNA, 说明在不同的生物类群中 snoRNA 基因组织的多样性. 这些基因组织的表达 (如转录和转录产物的加工成熟的方式) 也有很大的差异. 拟南芥 U59 基因簇的发现反映出植物 snoRNA 基因组织的一个特点.

在人的 U21、U26 snoRNA 以及酿酒酵母 snR38 snoRNA 分子结构中, 都有 box D (CTGA) 和反义序列结构元素. 根据 snoRNA 结构与功能的关系推知: 上述的 snoRNA 应分别指导人类以及酿酒酵母 28S rRNA G1297、A329 和 G2805 位点 2'-O-核糖甲基化. 但是到目前为止还没有经实验证实上述基因组中的 rRNA 的相应位点被甲基化以及上述 snoRNA 具有指导甲基化的功能^[6].

鉴定自人的 U14 snoRNA 也具有 2 个与 rRNA 互补的反义序列. 其中一个可指导 18S rRNA Am462 位点 2'-O-核糖甲基化, 而另一个反义序列后面没有 boxD (CTGA) 结构元素. 现已证实: U14 snoRNA 不仅具有指导 rRNA 甲基化的功能, 而且参与 rRNA 前体的加工与剪切. 已证实拟南芥 U59a、U59b snoRNA 结构中位于 boxD 上游的序列元素不指导 rRNA 相应位点核苷的甲基化. 这意味着拟南芥 U59a、U59b snoRNA 除行使指导拟南芥 18S rRNA Am975 位点 2'-O-核糖甲基化功能之外, 还具有其它的生物学功能, 即很有可能如同存在于人的 U14 一样, 参与 rRNA 前体的加工与剪接; 或者如同存在于人的 U21、U26 snoRNA 以及酿酒酵母 snR38 一样, 具有其它还未揭示出的生物学功能. 值得指出拟南芥 U59 具有分别与两种不同的 rRNA (18S 和 25S rRNA) 配对的反义序列, 它是同时还是先后与这些 rRNA 相反应? 进一步探讨这一问题, 对于研究 snoRNA 与 rRNA 反应的过程和机制有较大的价值.

虽然拟南芥 U59a、U59b snoRNA 与哺乳动物 U59 snoRNA 为同源分子, 但却有很大的差异. 在基因的大小上, 鉴定自哺乳动物 U59 snoRNA 基因仅为 76 bp; 而拟南芥 U59a、U59b snoRNA 基因编码序列长度分别达到 85 和 110 bp, 而且相互的序列之间同源率也很低. 这些事实表明 snoRNA 分子中功能序列增加, 导致基因的编码区扩大, 与其它同源分子的同源率就降低. 拟南芥 U59a、U59b snoRNA 基因 boxD 上游与 25S rRNA 互补的反义序列在已鉴定的酵母、哺乳动物以及植物中的 snoRNA 分子中未发现有与其同源的拷贝. 从它们的结构元素与基因大小可以认为拟南芥 U59a、U59b snoRNA 基因在基因进化过程中发生了重组, 即由一个类似于哺乳动物的 U59 snoRNA 分子分别与另外一个植物特有的 snoRNA 基因进行了重组融合, 形成现存于拟南芥中的 U59a、U59b snoRNA 分子. snoRNA 基因的这一起源推测在分析鉴定自粟酒裂殖酵母 Z15 snoRNA 基因结构时, 能得到一些佐证. 最新研究结果表明, 酿酒酵母 Z2 snoRNA 基因具有移动性 (屈良鸽等, 待发表). 这种移

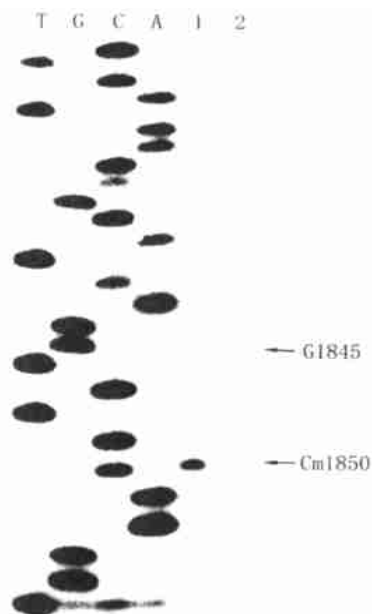


图 5 拟南芥 25S rRNA 第 1845 位点 2'-O-核糖甲基化核苷的鉴别

Fig. 5 The identification of 2'-O-ribose methylation of *A. thaliana* 25S rRNA G1845
1.4 μmol/L dNTP; 2. 1.5 mmol/L dNTP

动性可能导致了分散存在于生物基因组中的 snoRNA 基因发生重组. snoRNA 基因组织的多样性及其机制的研究将为阐明真核生物 snoRNA 基因的起源和进化提供新的证据.

拟南芥 U59a 和 U59b snoRNA 基因序列已被国际基因数据库 EMBL 收录, 收录号分别为 AJ240064, AJ240065.

参考文献:

- [1] MAXWELL E S, FOURNIER M J. The small nucleolar RNAs [J]. Annu Rev Biochem, 1995, 35: 897 ~ 934.
- [2] TODD M L, SEAN R E. A computational screen for methylation guide snoRNAs in yeast [J]. Science, 1999, 283: 1168 ~ 1171.
- [3] QU L H, ANTHONY H, et al. Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in Yeast [J]. Mol cell Biol, 1999, 19 (2): 1144 ~ 1158.
- [4] CHOMCZYNSKI P, SACCHI N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction [J]. Anal Biochem, 1987, 162: 732 ~ 735.
- [5] SAMBROOK, FRITSCH E F, MANIATIS T. Molecular cloning: a laboratory manual [M]. 2nd ed. New York: CSH Press, 1989.
- [6] NICOLOSO M, QU L H, MICHOT B, et al. Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-O-ribose methylation of rRNAs [J]. J Mol Biol, 1996, 260: 178 ~ 195.

Study on the Structure and Function of U59 snoRNA Gene Cluster from *A. thaliana*

MENG Qing^{*}, ZHOU Hui, XU Fei, QU Liang-hu

Abstract: U59 snoRNA gene cluster, consisted of two same snoRNA genes possessing two antisense elements each, was identified and characterized from *A. thaliana*. The antisense element located upstream of box D' is identical to that of U59 snoRNA identified from human, so, they were termed as *A. thaliana* U59a and U59b snoRNA respectively. The antisense element located in the upstream of box D is different from all of those of snoRNA identified from mammalian and yeast genome. It was verified that the antisense element located upstream of box D does not guide *A. thaliana* 25S rRNA G1845 2'-O-ribose methylation revealed by lower concentration dNTP reverse transcription analysis. The results showed the important difference between *A. thaliana* U59 snoRNA and human U59 snoRNA, suggesting *A. thaliana* U59 snoRNA gene cluster as a good system for further studying diversities of plant snoRNA gene organization and expression, and the origin of snoRNA genes.

Keywords: *A. thaliana*; U59; snoRNA; gene cluster

^{*} Biotechnology Research Center, Zhongshan University, Guangzhou 510275, China