

# 果蝇 3 个新的小分子非编码 RNA 的鉴定

贺华良 周惠 肖振东 曾献芬 陈俊宇 郑涛 屈良鹄\*

(中山大学基因工程教育部重点实验室, 有害生物控制及资源利用国家重点实验室, 广州 510275.)

\* 联系人, E-mail: lsbr04@zsu.edu.cn

**摘要** 通过比较基因组和分子生物学方法分析了果蝇属中 5 种果蝇全基因组内含子区域的保守序列, 获得了 3 个新的非编码 RNA 基因. 其中一个为具有典型的 box C/D 家族保守元件及结构特征的核仁小分子 RNA 基因, 其功能序列可介导 28S rRNA 的 C2673 位点的核糖甲基化修饰. 另外两个为 miRNA 基因, 其转录序列可形成典型的 miRNA 前体茎环结构; 在果蝇发育的 4 个时期均可表达产生长度为 23 个核苷酸的成熟 RNA 分子. 结果还表明, 在长度为 100~500 bp 区间的黑腹果蝇基因内含子中存在 396 个多物种保守序列(MCIS), 这些序列除编码小分子 RNA 外, 还可能与影响基因转录或转录后加工的顺式元件有关.

**关键词** 果蝇 box C/D snoRNA miRNA 内含子 MCIS

近年来, 非编码 RNA (non-coding RNA, ncRNA) 日益引起人们的高度关注<sup>[1]</sup>. 根据 RNA 分子大小、序列或结构、结合蛋白质的类型以及亚细胞的定位将非编码 RNA 归入各种特定的 RNA 类型, 如核仁小分子 RNA (small nucleolar RNA, snoRNA)<sup>[2]</sup>和微 RNA (microRNA, miRNA) 等<sup>[3,4]</sup>. snoRNA 在核糖体的生物合成中起重要作用, 大多数 snoRNA 负责指导 rRNA 和 snRNA 等的转录后修饰, 少数 snoRNA 参与 rRNA 前体的剪切. 除 RNase MRP 外, snoRNA 主要可以分为 2 种类型<sup>[5]</sup>, 即 box C/D snoRNA 和 box H/ACA snoRNA. 其中 box C/D snoRNA 的 5' 端有保守序列元件 box C (RUGAUGA), 3' 端有保守序列元件 box D (CUGA). 在 box D 或另一个内部 box D' 的上游, 通常有一段能与 rRNA 互补的长度为 10~21 个核苷酸的反义序列, 该序列指导 box D (或 D') 上游第 5 个碱基对处对应目标 RNA 分子上 2'-O-核糖甲基化修饰<sup>[6]</sup>. miRNA 是细胞内另一类长约 22 个核苷酸的调控非编码 RNA 家族, 它们广泛存在于从植物、线虫到人类的基因组中. 目前尚未发现 miRNA 中含有保守序列元件. miRNA 在细胞中的生物合成十分复杂, 它的初级转录本可形成具有茎环结构的前体分子 (pre-miRNA), 再从茎环结构前体的一条臂上或两条臂上加工生成成熟分子<sup>[7]</sup>. miRNA 主要是通过抑制相关 mRNA 的翻译<sup>[8]</sup>或对靶 mRNA 分子的切割<sup>[9]</sup>等方式来调控靶基因的表达. 与蛋白质编码基因一样, 非编码 RNA 基因在进化上也具有一定的保守性. 各种非编码 RNA 在物种间的保守度差异很大, 可能与所执

行的功能紧密相关. 成熟的 miRNA 分子进化上高度保守, 许多分子从低等生物线虫到高等生物人的基因组中都可以找到同源分子<sup>[10,11]</sup>, 如人和鼠之间的同源性都在 90% 以上<sup>[12]</sup>. 对于 snoRNA, 人与小鼠之间的同源性为 80%~90%<sup>[12]</sup>. 在高等生物中, 许多非编码 RNA, 特别是小分子非编码 RNA, 来源于基因内含子区域的原初转录物<sup>[13~16]</sup>. 对两种亲缘关系较近的果蝇进行比较发现, 基因组中除了外显子和基因间隔区外, 内含子同样受到选择压力, 这种压力与适应性演化和纯化选择有关, 意味着内含子像基因的编码区一样具有重要的功能<sup>[17]</sup>. 目前有报道利用人、小鼠和大鼠之间的内含子保守序列成功地鉴定了一批 snoRNA<sup>[18]</sup>, 表明利用物种间序列保守性可以发现和鉴定内含子中隐藏的功能因子.

模式生物果蝇一直在遗传和发育研究中具有举足轻重的地位. 在黑腹果蝇的全基因组序列完成后, 果蝇属其他 4 种果蝇的基因组序列也相继被测定, 为多物种的基因组比较提供了基础. 对黑腹果蝇的全基因组的注释发现基因内含子占整个基因组的 22%, 超过了蛋白编码区的 19%<sup>[19,20]</sup>. 研究果蝇内含子的结构及其功能对于揭示真核生物非编码区域参与的生命活动调控网络具有重要的意义. 本研究通过比较基因组学的方法分析了黑腹果蝇的 3 万多个内含子, 鉴定出 3 种新的小分子非编码 RNA 基因和一批多物种内含子保守序列, 为进一步研究果蝇内含子的结构与功能提供了重要线索.

## 1 材料和方法

( ) 基因组数据来源. 黑腹果蝇(*D. melanogaster*)基因组序列、内含子注释数据库以及果蝇属中其他4种果蝇的基因组序列从网站(<http://genome.ucsc.edu/>)获得. 几个果蝇种基因组的比对结果利用VISTA软件(<http://genome.lbl.gov/vista/>)分析获得. 全基因组比对分析涉及果蝇属中5个果蝇种, 分别是*D. melanogaster*, *D. ananassae*, *D. pseudoobscura*, *D. mojavensis*和*D. virilis*, 其中*D. melanogaster*分别与其他4个种进行全基因组比对分析.

( ) 内含子保守序列的获取. 在VISTA所提供的全基因比对数据基础上, 设计保守序列搜索平台CISa, 获取5个果蝇种内含子中都保守的序列, 即MCIS (multi-species conserved intronic sequences). 在CISa程序运算端口输入黑腹果蝇所有的蛋白质基因序列号或基因名. 对于发生选择性剪接的宿主基因, 本研究只选择最大的转录本; 而同一基因相同长度的不同转录本之间, 又只选择内含子数目最多的转录本. 本研究只分析了全基因组中3万多个小于500 bp的内含子. 初步将这些内含子再分为100 bp以下和100~500 bp两个区间, 其中小于100 bp的内含子主要由“minimal intron”组成. 果蝇“minimal intron”的大小分布范围为 $(61 \pm 10)$  bp<sup>[21,22]</sup>. 以上两个区间的内含子分别约为果蝇基因组内含子总数的50%和25%. 内含子中MCIS输出的参数设置: 序列保守性计算的移动窗口为50 bp, 保守阈值最小为70%, 保守序列的最小长度为50 bp. 为了避免宿主内含子长度对MCIS输出的影响, 本研究还对小于100 bp的内含子中的MCIS调整了参数重新输出. 参数调整后的设置为: 序列保守性计算的移动窗口为20 bp, 保守阈值最小为70%, 保守序列的最小长度为20 bp. 至于后期只针对miRNA候选分子的分析, 通过对已知的miRNA进行测试后, 本研究再调整了对应miRNA候选分子的MCIS输出的最佳参数: 序列保守性计算的移动窗口最小为20 bp, 保守阈值为100%, 保守序列的最小长度为20 bp. 最终实验鉴定所用的保守序列来源于黑腹果蝇基因组, 所取序列区间为上述4组基因组比对产生的保守序列映射在黑腹果蝇基因组上的共同区间. 保守序列的命名由宿主基因的序列号或基因名、保守序列所在的内含子的次序编号和CISa软件对这个保守序列的编号组成.

( ) 小分子RNA的提取和富集. 野生型黑腹果

蝇原种由中山大学生命科学学院遗传学教研室提供. 总RNA的提取来源于果蝇4个发育时期的材料: 卵、幼虫、蛹和成虫. 总RNA的提取方法采用酸性异硫氰酸胍法<sup>[23]</sup>. 总RNA提取后小分子RNA富集方法如下<sup>[24]</sup>: 取1~2 mg总RNA, 加入50  $\mu$ L 50% PEG8000和50  $\mu$ L 5 mol/L NaCl, 混匀. 冰上放置2 h后以15000  $\times$  g离心10 min. 取上清, 加1/10体积3 mol/L NaAc (pH 5.2), 2倍体积无水乙醇, -20  $^{\circ}$ C放置2 h. 再以15000  $\times$  g离心10 min, 弃上清. 沉淀用75%乙醇洗涤2次, 干燥, 溶于DEPC水中备用.

( ) RNA分析. Northern杂交实验的步骤<sup>[25]</sup>: 30  $\mu$ g富集后的小分子RNA经8 mol/L尿素-8%聚丙烯酰胺(针对snoRNA分子)或8 mol/L尿素-12%聚丙烯酰胺(针对miRNA分子)变性胶电泳分离, 然后转移至尼龙膜(Hybrid N<sup>+</sup>, Amersham)或Zeta膜(Zeta-probe GT, Bio-Rad), 紫外交联4 min, 再用5'末端标记的探针杂交. 杂交膜压磷屏后用Typhoon 8600 Variable Mode Imager成像分析. 逆转录实验的步骤<sup>[25]</sup>: 20  $\mu$ L反应体系中含20  $\mu$ g富集小分子RNA, 8 ng 5'末端以<sup>32</sup>P标记的引物和50 mmol/L dNTPs. 反应液充分混合后于65  $^{\circ}$ C变性5 min, 冷至42  $^{\circ}$ C后, 加入10 U AMV或100 U MLV逆转录酶(Promega公司), 在42  $^{\circ}$ C延伸反应60 min后终止反应, 经8 mol/L尿素-8%变性聚丙烯酰胺凝胶电泳分离后压磷屏, 用Typhoon 8600 Variable Mode Image 成像分析.

( ) 寡核苷酸探针和引物. 寡核苷酸探针由上海生工公司合成, 探针序列如下: nop5-1-1/probe-1, 5'-CAGCCAAGGTTTATCGTCATCT-3'; nop5-1-1/probe-2, 5'-TCGCTCAGACAGCCAAGGTTTAT-3'; tomosyn-3-1, 5'-ATAGCCGCGTGTGTCTTATGATA-3'; CG8877-8-1, 5'-CGTGAGGTTTGCAACAAGGGGATG-3'. 用于Northern杂交的寡核苷酸探针5'末端以 $[\gamma\text{-}^{32}\text{P}]$ ATP标记后可直接使用<sup>[25]</sup>.

## 2 结果

本研究在100~500 bp大小的内含子区域获得396个在果蝇属5个果蝇种中保守性为70%以上的多物种保守序列, 即MCIS. 对这些序列的进一步分析表明, 其中含有38个已知的小分子非编码RNA序列, 包括21个box H/ACA snoRNA, 13个box C/D snoRNA和4个miRNA. 同时, 分析还发现了3个新的小分子非编码RNA.

2.1 一个新的 box C/D snoRNA 的发现和鉴定

研究发现 *nop5* 基因的第一个内含子中含有一个 MCIS, 命名为 *nop5-1-1*. 在 *nop5-1-1* 序列中含有“CTGA”, 有可能是 snoRNA 家族的 box D 元件. 通过分析 *nop5* 基因的第一个内含子的全部序列, 发现在 MCIS 序列上游还存在“GTGATGA”, 与 snoRNA 家族中的 box C 元件特征吻合. 同时对其宿主基因 *nop5* 的功能分析表明, *nop5* 是一种核仁蛋白, 与大部分已知 snoRNA 的宿主基因一样参与核糖体复合物的形成. 因此推测 *nop5-1-1* 可能是一个参与 rRNA 转录后修饰的小分子向导 RNA, 并且判断这个新的 snoRNA 基因的长度应该为 74 个核苷酸左右. *nop5-1-1* 的序列分析见图 1.

对 *nop5-1-1* 设计了 2 个不同的探针进行 Northern 杂交, 实验表明, 杂交信号都为 74 个核苷酸左右, 与预期结果吻合(图 2(a)). 逆转录分析进一步证实了 snoRNA 基因 5'末端在 C box 上游的 6 个核苷酸处(图 2(b)). 测序分析表明, cDNA 序列与预测的 snoRNA 基因的序列相符. 对该 RNA 功能的计算机分析表明,

在 snoRNA 的 D box 上游有一段与 28S rRNA 互补、长度为 10 个核苷酸的反义序列, 预测其可指导 28S rRNA 的 C2673 位的甲基化修饰(图 2(c)). 目前, 在果蝇中尚没有报道指导修饰该位点的 snoRNA 分子, 因此将这个新的 snoRNA 基因命名为 Me28S-C2673.

2.2 两个新的 miRNA 的发现和鉴定

根据果蝇中已知的 miRNA 基因在多个近缘物种中保持 100%保守度的特点, 将 396 个在 5 种果蝇之间保守度为 70%的 MCIS 全部调整为 100%保守度的参数后重新输出序列, 发现了 8 个 MCIS 都具有一段 23 bp 左右的序列可以在 5 个近缘物种中达到 100%的保守. 而相应的 70%保守度输出的序列可以被折叠成典型的茎环结构. 通过与果蝇 miRNA 数据库比较表明, 其中 4 个 MCIS 编码果蝇中已知的 4 个 miRNA, 另外 4 个 MCIS 在果蝇 miRNA 数据库中没有发现对应的序列, 推测可能是新的 miRNA 候选分子, 分别暂命名为 CG8877-miRNA, *cpx*-miRNA, *tomosyn*-miRNA 和 CG6831-miRNA.

通过与其他物种的 miRNA 数据库 [13,14] 比对分析,

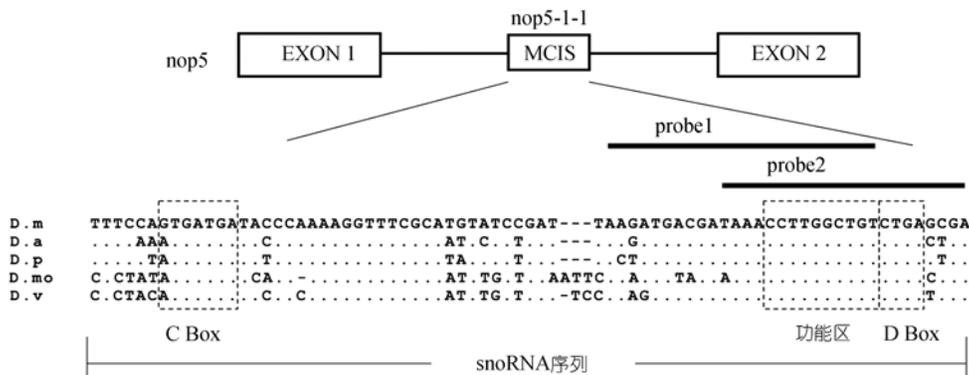


图 1 MCIS *nop5-1-1* 及对应的 snoRNA 的序列分析

物种间相同的碱基用点表示, 缺失碱基用“-”表示. 探针的位置用“probe”表示. D.m, D.a, D.p, D.mo 和 D.v 分别表示 *D. melanogaster*, *D. ananassae*, *D. pseudoobscura*, *D. mojavensis* 和 *D. virilis*

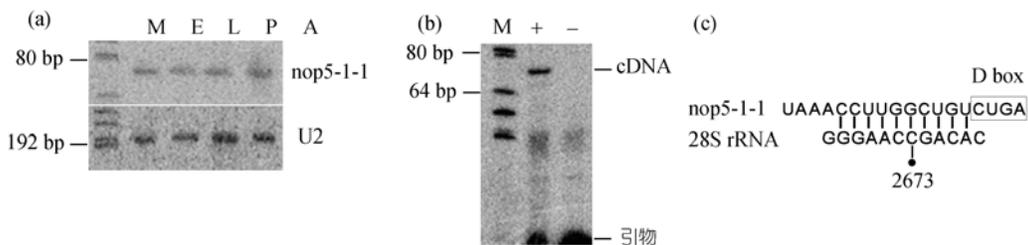


图 2 *nop5-1-1* 的鉴定和功能分析

(a) *nop5-1-1* 的 Northern 杂交鉴定, 探针为 probe 1. 采用 probe 2 杂交结果与 probe 1 相同, 故未显示. M 为标准分子量(pBR322 以 *Hae* 酶切, 产物以 [ $\gamma$ - $^{32}$ P]ATP 标记 5'末端). E, L, P 和 A 分别表示来自果蝇卵、幼虫、蛹和成虫的富集小分子 RNA. U2 为正对照. (b) *nop5-1-1* 的逆转录反应, 以 probe 2 为引物进行. (c) *nop5-1-1* 的功能分析. (c) 中的黑点表示甲基化修饰

CG6831-miRNA 与人和小鼠中 miR-190 为同源分子,而在所有模式生物中都没有发现另外 3 种新的 miRNA 候选分子的同源基因. 根据预测的成熟 miRNA 分子(图 3 和 4)设计探针进行杂交鉴定,结果表明, tomosyn-miRNA 和 CG8877-miRNA 均可产生为约 23 个核苷酸的稳定表达的成熟 RNA 分子,而 cpx-miRNA 没有得到杂交验证. 杂交图谱表明,以上两个 miRNA 基因在果蝇发育的 4 个时期的表达水平基本一致,与 let-7 组织特异性表达的特征不同(图 5). 对 tomosyn-miRNA 和 CG8877-miRNA 作用的 mRNA 的 3'UTR 区进行了分析,发现它们各有上百个可能的靶位点. 这两个 miRNA 的调控功能还有待进一步的研究.

### 3 讨论

目前果蝇全基因组中已鉴定编码 box C/D snoRNA, box H/ACA snoRNA 和 miRNA 的内含子分别有 105, 101 和 21 个 [26~30], 反映果蝇内含子中编码了大量的反式作用因子. 绝大部分已知的非编码 RNA 序列都位于黑腹果蝇与上述 4 种果蝇中的任意

一种果蝇基因组序列比对后的内含子保守序列中. 而在 5 种果蝇基因组内含子中都超过 70% 保守度的 MCIS 中则包含 65% 的 box C/D snoRNA, 50% 的 box H/ACA snoRNA 和 100% 的 miRNA. 由此可见, 小分子非编码 RNA 主要集中在高度保守的 MCIS 中, 而多物种基因组比较分析是一种鉴定内含子编码功能的有效方法. 本研究采用此方法在果蝇基因内含子中成功鉴定了 snoRNA 和 miRNA 两类不同的小分子非编码 RNA. 本研究所鉴定的 Me28S-C2673 snoRNA 在前人的实验 RNA 组学研究中未被克隆 [26]. 在利用大规模的 snoscan 软件搜索果蝇基因组的过程中也因为阈值较低被排除在 snoRNA 候选基因之外 [27]. 同样, 在前人应用计算机搜索软件 miRseeker 对果蝇进行全基因组水平搜索过程中, 本研究所鉴定的两个 miRNA 也曾被列入了果蝇 miRNA 的候选分子, 但一直没有被进一步证实 [30]. 本研究通过多物种比较基因组学搜索 MCIS 的方法, 从另一个角度获得了这两个 miRNA 分子的序列, 并通过实验证实了这些 miRNA 成熟分子的稳定表达. 基于模式识别的

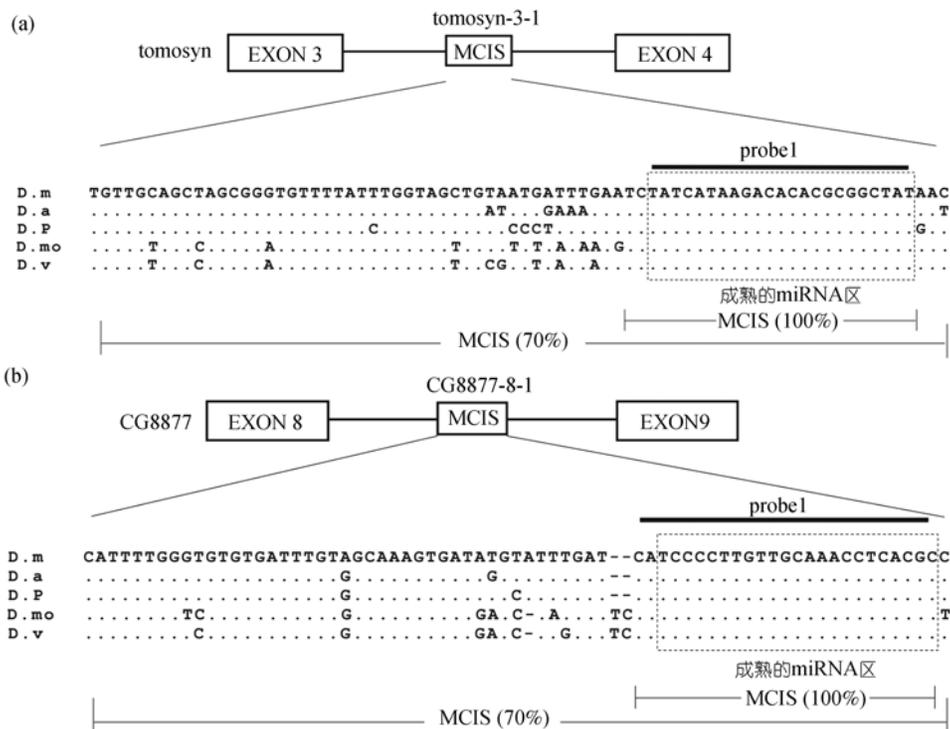


图3 MCIS 及对应的 miRNA 的序列分析

(a) tomosyn-3-1 的序列分析; (b) CG8877-8-1 的序列分析. 物种间相同的碱基用点表示, 缺失碱基用“-”表示. 探针的位置用“probe”表示. MCIS (70%)表示保守度参数为 70%时的 MCIS 输出. MCIS (100%)表示保守度参数为 100%时的 MCIS 输出. D.m, D.a, D.p, D.mo 和 D.v 分别代表 *D. melanogaster*, *D. ananassae*, *D. pseudoobscura*, *D. mojavensis* 和 *D. virilis*

计算机RNA组学分析需要对搜索对象有较全面的了解 [31]，在搜索新的非编码RNA基因时有一定的局限性，而多物种比较基因组学的方法可以为实验RNA组学和计算机RNA组学分析提供较好的补充。

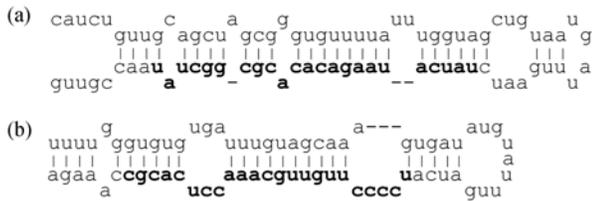


图 4 miRNA 前体序列折叠图

(a) tomosyn-miRNA 前体折叠图。(b)CG8877-miRNA 前体折叠图。  
粗体序列部分为成熟 miRNA 分子

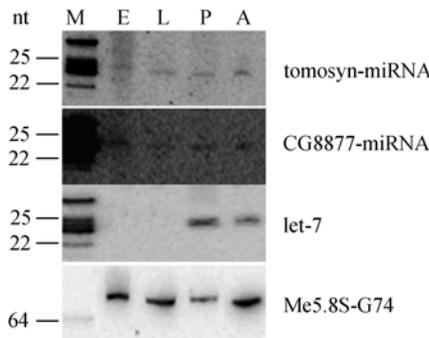


图 5 tomosyn-miRNA 和 CG8877-miRNA 的杂交鉴定

M 为 DNA Marker，凝胶上 RNA 分子的实际大小为 DNA 标准分子量减去 2~3 个核苷酸。E、L、P 和 A 分别表示来自果蝇卵、幼虫、蛹和成虫的富集小分子 RNA。let-7 miRNA 和 Me5.8S-G74 snoRNA 为对照

本研究分析了 3 万多个果蝇内含子，其中 2 万多个为“minimal intron”。我们发现仅有 0.2% 的“minimal intron”含有MCIS；在“minimal intron”中也没有发现任何已知的小分子RNA编码区。这一结果进一步支持绝大部分“minimal intron”只是在长度上而不是在序列上承受选择压力 [22]。与“minimal intron”不同，我们发现 3.9% 的长度为 100~500 bp 的内含子含有MCIS。目前，在这 396 个MCIS中已发现 41 种小分子RNA编码区，表明这类内含子序列具有多种功能意义。剩下的MCIS的功能将会有那些表现形式呢？本研究对随机抽取的 40 个功能未知的MCIS进行初步Northern 杂交实验，发现仅有 2 个MCIS可以产生稳定表达的RNA分子(结果未显示)。以上结果表明，在长度为 100~500 bp 区间的内含子中，除了少部分的MCIS编码小分子RNA外，大部分MCIS可能起顺式元件的作用，与基因的转录或转录后加工有密切关系 [32,33]。对果蝇基因内含子保守序列的功能意

义还有待进一步研究。

致谢 感谢本实验室陈晓红老师给予的技术支持和帮助。本工作为国家自然科学基金(批准号：30230200, 30370328)、教育部和广东省创新团队基金(批准号：IRT0447, NSF-05200303)以及国家重点基础研究发展计划(批准号：2005CB724600)资助项目。

参 考 文 献

- Eddy S R. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*, 2001, 2(12): 919—929[DOI]
- Maxwell E S, Fournier M J. The small nucleolar RNAs. *Annu Rev Biochem*, 1995, 64: 897—934[DOI]
- Lee R C, Feinbaum R L, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 1993, 75(5): 843—854[DOI]
- Reinhart B J, Slack F J, Basson M, et al. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 2000, 403(6772): 901—906[DOI]
- Balakin A G, Smith L, Fournier M J. The RNA world of the nucleolus: Two major families of small RNAs defined by different box elements with related functions. *Cell*, 1996, 86(5): 823—834[DOI]
- Kiss-Laszlo Z, Henry Y, Bachelierie J P, et al. Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell*, 1996, 85(7): 1077—1088[DOI]
- Lagos-Quintana M, Rauhut R, Lendeckel W, et al. Identification of novel genes coding for small expressed RNAs. *Science*, 2001, 294(5543): 853—858[DOI]
- Zeng Y, Wagner E J, Cullen B R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell*, 2002, 9(6): 1327—1333[DOI]
- Llave C, Xie Z, Kasschau K D, et al. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science*, 2002, 297(5589): 2053—2056[DOI]
- Lim L P, Glasner M E, Yekta S, et al. Vertebrate microRNA genes. *Science*, 2003, 299(5612): 1540[DOI]
- Lim L P, Lau N C, Weinstein E G, et al. The microRNAs of *Caenorhabditis elegans*. *Genes Dev*, 2003, 17(8): 991—1008[DOI]
- Pang K C, Frith M C, Mattick J S. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet*, 2006, 22(1): 1—5[DOI]
- Griffiths-Jones S. The microRNA Registry. *Nucleic Acids Res*, 2004, 32(Database issue): D109—111[DOI]
- Griffiths-Jones S, Grocock R J, van Dongen S, et al. miRBase microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, 2006, 34(Database issue): D140—144[DOI]
- Lestrade L, Weber M J. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res*, 2006, 34(Database issue): D158—162[DOI]
- Brown J W, Echeverria M, Qu L H, et al. Plant snoRNA database. *Nucleic Acids Res*, 2003, 31(1): 432—435[DOI]
- Andolfatto P. Adaptive evolution of non-coding DNA in *Drosophila*

- phila*. *Nature*, 2005, 437(7062): 1149—1152[DOI]
- 18 Fedorov A, Stombaugh J, Harr M W, et al. Computer identification of snoRNA genes using a Mammalian Orthologous Intron Database. *Nucleic Acids Res*, 2005, 33(14): 4578—4583[DOI]
- 19 Adams M D, Celniker S E, Holt R A, et al. The genome sequence of *Drosophila melanogaster*. *Science*, 2000, 287(5461): 2185—2195[DOI]
- 20 Holt R A, Subramanian G M, Halpern A, et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*, 2002, 298(5591): 129—149[DOI]
- 21 Deutsch M, Long M. Intron-exon structures of eukaryotic model organisms. *Nucleic Acids Res*, 1999, 27(15): 3219—3228[DOI]
- 22 Yu J, Yang Z, Kibukawa M, et al. Minimal introns are not “junk”. *Genome Res*, 2002, 12(8): 1185—1189[DOI]
- 23 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, 1987, 162(1): 156—159[DOI]
- 24 Park W, Li J, Song R, et al. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol*, 2002, 12(17): 1484—1495[DOI]
- 25 Zhou H, Chen Y Q, Du Y P, et al. The *Schizosaccharomyces pombe* mgU6-47 gene is required for 2'-O-methylation of U6 snRNA at A41. *Nucleic Acids Res*, 2002, 30(4): 894—902[DOI]
- 26 Yuan G, Klambt C, Bachellerie J P, et al. RNomics in *Drosophila melanogaster*: Identification of 66 candidates for novel non-messenger RNAs. *Nucleic Acids Res*, 2003, 31(10): 2495—2507[DOI]
- 27 Huang Z P, Zhou H, He H L, et al. Genome-wide analyses of two families of snoRNA genes from *Drosophila melanogaster*, demonstrating the extensive utilization of introns for coding of snoRNAs. *RNA*, 2005, 11(8): 1303—1316[DOI]
- 28 Huang Z P, Zhou H, Liang D, et al. Different expression strategy: Multiple intronic gene clusters of box H/ACA snoRNA in *Drosophila melanogaster*. *J Mol Biol*, 2004, 341(3): 669—683[DOI]
- 29 Aravin A A, Lagos-Quintana M, Yalcin A, et al. The small RNA profile during *Drosophila melanogaster* development. *Dev Cell*, 2003, 5(2): 337—350[DOI]
- 30 Lai E C, Tomancak P, Williams R W, et al. Computational identification of *Drosophila* microRNA genes. *Genome Biol*, 2003, 4(7): R42[DOI]
- 31 Schattner P, Brooks A N, Lowe T M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*, 2005, 33(Web Server issue): W686—689[DOI]
- 32 Standiford D M, Sun W T, Davis M B, et al. Positive and negative intronic regulatory elements control muscle-specific alternative exon splicing of *Drosophila* myosin heavy chain transcripts. *Genetics*, 2001, 157: 259—271
- 33 Glazov E A, Pheasant M, McGraw E A, et al. Ultraconserved elements in insect genomes: A highly conserved intronic sequence implicated in the control of homothorax mRNA splicing. *Genome Res*, 2005, 15(6): 800—808[DOI]

(2006-05-07 收稿, 2006-09-06 接受)