

## 特约综述

## 长链非编码 RNA 的功能及其研究

余良河 李楠 程树群\*

(第二军医大学东方肝胆外科医院, 上海 200438)

**摘要** 长链非编码 RNA (Long noncoding RNAs, lncRNAs) 是一类无或少有蛋白编码能力的复杂长链非编码 RNA, 在大部分真核生物基因组被转录。目前大多数 lncRNAs 已分类, 但其功能有待深入研究。已有研究表明 lncRNAs 在调节生长发育、细胞定向分化、亚细胞结构分布、进化选择和人类疾病的关系等方面有重要作用。本文就对 lncRNAs 的功能及其在医学方面的研究进展作一综述。

**关键词** 长链非编码 RNA; 转录; lncRNAs;

长链非编码 RNA (Long noncoding RNAs, lncRNAs) 首先是在大鼠全长 cDNA 序列文库中被描述的。最初分离出长链非编码 RNA 接近蛋白编码基因的非编码 RNA, 包括重叠、顺反义、双向或内含子非编码 RNA。lncRNAs 不单独表现任何已知 RNA 的特性, 而是表现出这些特性的综合作用 (图 1)。lncRNAs 异常丰富, 最初被认为是由于转录过程中 RNA 聚合酶的低保真度所致<sup>[1]</sup>。但是有些 lncRNAs 的表达仅限于特定的生长发育阶段<sup>[2]</sup>, 如大鼠 lncRNAs 在胚胎干细胞分化过程中<sup>[3]</sup>, 以及在大脑中多表现出精确的亚细胞分布<sup>[4]</sup>。不同于小分子 RNA 或蛋白质, 目前尚无法根据非编码 RNA 序列或结构推断其功能。本文就对 lncRNAs 的功能及其在医学方面的研究进展作一综述。

## 1 lncRNAs 对转录的影响

lncRNAs 可以通过许多机制调节 RNA 聚合酶 II 的活性, 包括通过干扰起始复合物的形成, 影响启动子的选择性。在人 DHFR 位点上, 一条长链非编码链起源于主要 DHFR 启动子上游区, 可以抑制下游蛋白编码基因表达<sup>[5]</sup>。PcG 蛋白已被证实与 1 000 多种哺乳动物基因的表达与沉默相关<sup>[6]</sup>, 但 PcG 蛋白是如何作用于特定的靶点上尚未明确。有研究提示 PcG 蛋白和特定基因位点结合可能是通过 lncRNAs 完成的<sup>[7]</sup>。有研究发现 Ezh2 直接结合一长度为 1.6 Kb 的 lncRNAs 成为 RepA<sup>[8]</sup>。当稳定结合的亲嗜性 RepA 基因进行转录时, PcG 蛋白结合于特定的基因位点, 从而使 PRC2 结合在染色质上。内源性 RepA 非编码

RNA 链是由 X 基因的重复 A 区转录的, 对哺乳动物早期 X-染色体非活化起着关键作用。TrxG 蛋白和激动型 PcG 蛋白相互作用维持转录的状态, 可能是通过 lncRNAs 使二者的靶点接近<sup>[9]</sup>。源自 Hox 基因的特定非编码 RNA 在体内直接与组蛋白甲基转移酶 Ash1 结合, 其目的是使 TrxG 蛋白特异结合染色质<sup>[10]</sup>。虽然没有报道发现 lncRNAs 的特异性表达和转录激活间的伴随作用, 但也有文献报道非编码 RNA 的转录可能抑制相邻 Hox 基因的表达<sup>[10,11]</sup>。虽然尚有争议, 但已证实特定非编码 RNA 通过结合 PcG 和 TrxG 蛋白靶点, 对维持基因表达的活性或非活性状态起关键作用。

现已证实包括 Airn 和 Kcnqlot1 在内的大量 lncRNAs 表达印记基因位点在等位基因的表达过程中起关键作用<sup>[12,13]</sup>, 最新研究证实单簇的印记基因能通过不同机制沉默。大鼠胎盘中, 以顺式方式沉默亲链特定含有 Slc22a3、Slc22a2 和 Igf2r 基因<sup>[14]</sup>在内长 400 kb 的一区域, 需要不到 108kb 的非编码 RNA 链 Airn。类似于 Xist, Airn 停留在细胞核内<sup>[15]</sup>, 并表达在子代染色体的印记基因外侧<sup>[16]</sup>。Airn 并非是一成不变的分布在印记结构域, 而是首先聚集在 Slc22a3 启动子, 然后与组蛋白 H3-赖氨酸-9-甲基-转移酶 G9a 相互作用, 导致 Slc22a3 启动子的甲基化和沉默。去除 G9a 导致 Slc22a3 的缺失, 但对 Igf2r 无影

国家自然科学基金(No.30873352), 国家科技部专项基金(2008ZX10002-025), 上海优秀学科带头人项目资助(No.10XD1405800)上海市慈善基金会资助项目(2009)

\* 通讯作者。Tel: 021-81875251, Fax: 021-65562400 E-mail: chengshuqun@yahoo.com.cn

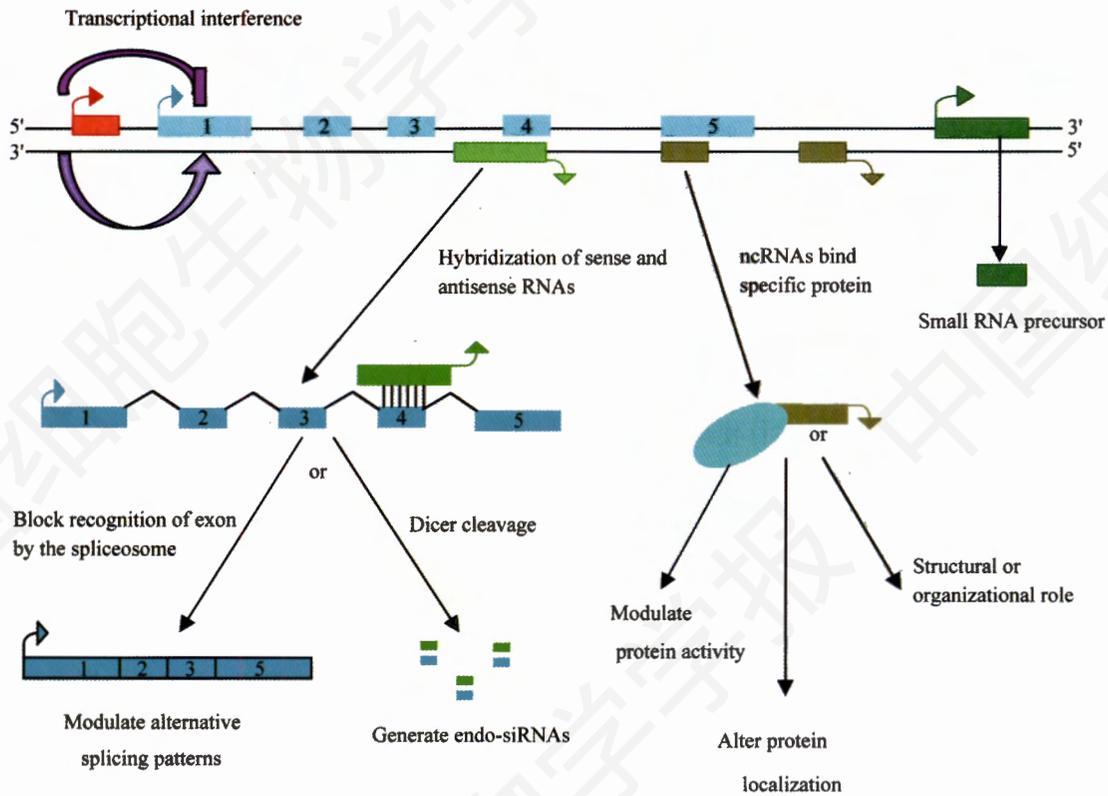


Fig.1 Paradigms for how long ncRNAs function

响, *Igf2r* 仍保持单个等位基因表达。因此, 即使子代染色体上的 *Slc22a3* 和 *Igf2r* 的基因沉默都需要 *Airn*, 也必定是通过不同机制完成的。

## 2 LncRNAs调节蛋白结合模型

蛋白质通过形成不同基序与 RNA 结合, 起调节加工、分布及稳定 RNA 的功能<sup>[17]</sup>。相反地, RNA 也可以影响与其结合的蛋白的活性和分布。如 lncRNAs 在转录调节中可作为关键蛋白的共激活因子。据报道表明只有在 *Evf-z* 非编码 RNA 链存在的条件下, *DLX2* 基因才起转录增强子作用<sup>[18]</sup>。同样地, 在细胞热休克应激中, 非编码 RNA 链 *HSR1* (热休克 RNA-1) 与 *HSF1* (热休克转录因子 1) 和一种有转录共激活因子的激素受体功能的非编码 RNA 异构体 *SRA* (激素受体 RNA 激动剂) 共同形成一复合物, 使转录因子介导热休克蛋白表达<sup>[19]</sup>。相反地, 在热休克反应中, 由 *SINEs* (短交杂元件) 产生的非编码链结合 RNA 聚合酶 II, 阻止其他 mRNA 的转录, 如抑制肌动蛋白的 mRNA 转录<sup>[20,21]</sup>。

最近已证实从细胞周期蛋白 *D1* (*CCND1*) 启动子区域产生的非编码 RNA 脂肪肉瘤易位 -*TLS*<sup>[22]</sup>, 有

RNA 结合蛋白变构效应器功能。*TLS* 蛋白与非编码 RNA 链结合后, 发生蛋白构象变化, 从抑制状态变为激活状态, 这样它结合并抑制组蛋白乙酰转移酶 *CBP* 和 *P300* 的活性, 从而沉默 *CCND1* 基因的转录。*LncRNAs* 能通过调节亚细胞定位来调整蛋白质的活性。转录因子 *NFAT* (活化 T 细胞的核因子), 在钙依赖信号作用下, 从胞质转入胞核里, 在核内 *NFAT* 激活靶基因启动转录<sup>[23]</sup>。*NFAT* 发生核转运的一个关键调节子是 *NRON* (即 *NFAT* 的非编码抑制子)<sup>[24]</sup>, 它是长为 0.8~3.7 Kb 的非编码 RNA。*NROR* 可以特异性抑制核中 *NFAT* 的表达。

## 3 LncRNAs作为小RNA的前体

最近全基因组研究表明, *LncRNAs* 链的功能之一可能是作为长度小于 200 Kb 的小 RNA 的前体<sup>[25,26]</sup>。另外, 通常是由 *Drosha* 酶和 *Dicer* 酶通过序列剪切一 *LncRNAs* 链产生 *microRNA*<sup>[27,28]</sup>, 而且有可能通过加工 *LncRNAs* 产生 *Piwi*-干扰 RNA (*piRNAs*)<sup>[29]</sup>。如发现在小 RNA 链 5'-端和 3'-端有基因簇<sup>[30-35]</sup>。转染的 RNA 类似物启动相关小 RNA, 可减少 mRNA 增强子重叠部分的基因表达, 这表明这些新发现的小 RNA 影响

基因表达<sup>[26]</sup>。更有研究发现许多蛋白编码 RNA 和 lncRNAs 可能转录后加工成 5'-端帽状结构的小 RNA。利用第二代测序技术发现许多小 RNA 有重复的 CAGE 标记(帽状结构的基因表达分析), 这种帽状结构戴在长 RNA 链的 5'-端。虽然许多 CAGE 标记在转录起始点, 但也发现在外显子区, 通过剪接部分加工成 mRNA。因此, 有人认为成熟长链(包括蛋白编码 mRNA 和 lncRNC), 在转录后可加工生成小 RNA, 然后通过戴帽结构修饰。

最新研究提示: lncRNAs 的单一链加工成许多小 RNA 后可分布于不同亚细胞结构, 并行使其独特的功能。长度不到 7 Kb 非编码 RNA 链 MALAT1 (肺腺癌相关转移转录 1), 即 NEAT2, 在许多人类癌症中发生表达改变。原先研究提示 MALAT1 是在细胞核里的核散斑, 行使组装、修饰和 / 或储存前 mRNA 的作用<sup>[36]</sup>。最近研究提示: 通过小 RNA 探针与 MALAT1 位点匹配, 发现 — 长度 61-nt 高度保守的 tRNA 样小 RNA<sup>[37]</sup>。

RNA 干扰的机制是 miRNA 和 siRNA 调节转录后水平的基因表达<sup>[38,39]</sup>。然而, 在大鼠一段未剪接的长度为 2.4 Kb 的多聚 A 尾非编码 RNA 链即 mrh1, 由 Drosha 酶加工生成成长为 80-nt 小 RNA<sup>[40]</sup>。有趣的是, 80-nt 小 RNA 链在体内通过 Dicer 酶不再进一步加工, 可能是因为它与染色质一起留在核内。最近发现, 两条 lncRNAs 链通过甲基化和去甲基化之间的转换调节哺乳动物 X-染色体失活, 其中可能也产生长为 25~42nt 小 RNA<sup>[41]</sup>, 进一步调节甲基化和去甲基化基因位点配对。因为在体内甲基化和去甲基化时形成一小 RNA 杂交双链, 在 Dicer 酶去除时小 RNA 表达减少, 所以这表明双链 RNA 加工产生小 RNA<sup>[42]</sup>, 但是目前产生长度为 25~42nt 的小 RNA 链的机制还不清楚。另外, 最近有文章报道 X 染色体失活不依赖 Dicer 酶。lncRNAs 链加工生成许多其他类别的小 RNA, 此类小 RNA 可能有独特的功能。

#### 4 lncRNAs 影响其它 RNA 的加工

lncRNAs 能加工生成小 RNA, 但它们是如何影响其它链的加工? 例如通过调整切割小 RNA 或 pre-mRNA 剪接位点。线虫 RNA 非编码链 rncs-1, 以反式方式抑制另外一链, 从而抑制小 RNA 生成<sup>[43]</sup>。在体内, 发现随着 rncs-1 基因过分表达或缺失, 特定 siRNA 的表达水平呈减少或增加, 并伴随着 mRNA 水平相应的靶基因表达发生变化。因此, 有人认为

rncs-1 结合 Dicer 酶, 或在同其他双链 RNA 竞争辅助双链 RNA- 结合蛋白参与基因沉默。

有研究发现长链非编码链 RNA 通过与 miRNA 相互作用, 能够竞争性抑制 miRNAs 结合其靶位点, 类似于人工合成 miRNA 的功能<sup>[44]</sup>。这个特定模仿机制用的是拟南芥的 IPS1 (磷饥饿诱导 1)<sup>[45]</sup>, 它是一 lncRNAs 链, 长度不到 550-nt, 在进化时保守性很差。非编码 RNA 链 IPS1 有一短 23-nt 基序与 miR-399 十分配对, 可通过在 miRNA 的预期切割点上错配进行干扰。这种碱基配对使 lncRNAs 链 IPS1 不可以切割, 插入 miR-399 序列后, 导致 miR-399 靶基因表达增加。

最近发现来自 RNA 链的假基因能引起 mRNA 的功能蛋白编码基因加工成小 RNA<sup>[46,47]</sup>。这是因为假基因产生的长反义链能够杂交相应的剪切 mRNA, 形成 dsRNA, 通过 Dicer 酶分解成内源性小干扰 RNA, 即 endo-siRNAs。编码 mRNA 产生内源性小干扰 RNA, 可能导致 RISC (诱导 RNA 沉默的复合物) 分解其他 mRNA 链, 进一步导致编码基因下调。因此假基因是 lncRNAs 链转录时基因表达的关键调节子。

一些天然的反义链(NATs)和假基因一样, 也能杂交重叠基因, 并产生内源性小干扰 RNA<sup>[47~50]</sup>。NATs 有许多调节重叠基因的可变剪切位点的例子, 例如在 Zeb2/sip1 基因位点上, Zeb2/sip1 是一种转录 E-钙粘连蛋白的抑制子, E-钙粘连蛋白的转录表达是受到严密调控的<sup>[51]</sup>。在上皮细胞中, IRES 区是从成熟 mRNA 剪切的。NTA 的加工是与 5'-端剪切位点的内含子互补配对的, 所以, 阻断剪切小体从成熟 mRNA 移至 IRES, 可使 Zeb2/sip1 蛋白表达增加。

#### 5 lncRNAs 可能是 RNAs 的结构

哺乳动物细胞核不仅与细胞质分隔开, 而且含有作为特定功能的胞膜结构成分<sup>[52]</sup>。虽然核内 RNA-蛋白复合体的确切功能尚不清楚, 但是认为它可以储存核内 RNA<sup>[53]</sup>。最近有报道称 MENε/β lncRNAs 是细胞核内 RNA-蛋白复合体的核心物质<sup>[54~56]</sup>。MENε 也称为 NEAT1, 和 MEBβ 分布在 3'-端的不同部位, 他们都是同一 RNA 聚合酶 II 启动子启动转录的, 都存在核内 RNA-蛋白复合体上<sup>[54,55]</sup>。不同于抑制 CTN-RNA 表达一样<sup>[53]</sup>, 抑制 MENε/β 表达会导致核内 RNA-蛋白复合体的分解, 故有人认为核内 RNA-蛋白复合体的出现和维持需要这些 lncRNAs<sup>[54~56]</sup>。

已发现 RNA 在细胞结构和有丝分裂的纺锤体上

具有保持有序性和稳定性的作用。在非洲蟾蜍卵细胞中, 十分有序的细胞角蛋白的结构依赖于两种 RNA, 即 X1sirts 非编码 RNA 和 Veg TmRNA, 它们整合在细胞骨架里<sup>[57,58]</sup>。虽然 Veg T 是一种蛋白编码 mRNA, 阻断它的翻译对细胞角蛋白网状结构无影响, 但是有人认为 Veg TRNA 的功能是维持细胞结构。同样地, 许多 RNA 特别是核糖体 RNA 和许多未分类出的转录链一样, 发现它们一直伴随有丝分裂的纺锤体存在<sup>[59]</sup>。虽然 RNA 酶干扰纺锤体组装, 并引起纺锤体的崩解, 但对翻译抑制子进行处理无影响。所以有人认为在 M 阶段纺锤体组装步骤中, RNA 起独立翻译作用。

鉴于 mRNA 定位模式繁多, 在观察果蝇的早期胚胎中发现<sup>[60]</sup>, 更多的 RNA (特别是 ncRNA) 在细胞中起维持结构高度有序的作用。如在细胞周期 S 阶段, 分布在核周部位的 ncRNA Xist<sup>[61]</sup> 和 kcnqlot1<sup>[62]</sup> 都能使染色质沉默。

## 6 医学方面的意义

在分化和发育过程中, 由于非编码 RNA 功能的异常往往导致一些疾病的发生。许多非编码 RNA 涉及疾病的病因, 现已引起研究者越来越大的兴趣。许多 lncRNAs 在疾病中发生错调, 尤其是癌症<sup>[63]</sup>。研究发现有些 lncRNAs 在肿瘤中非常敏感, 是肿瘤特异标志物, 如 DD3 (也称 PCA3) 在前列腺肿瘤中表达<sup>[64]</sup>。这些 lncRNAs 链通过什么机制影响肿瘤形成和 / 或进展, 目前尚不清楚。因此 lncRNAs 在疾病研究中仍是一个相对未开发领域, 可能为我们提供新的治疗靶点。最近对阿尔茨默病的研究发现  $\beta$ -分泌酶基因(BACE1) 的一非编码 RNA 反义链, 可产生不溶性淀粉样蛋白  $\beta$  (A $\beta$ ), 可能导致疾病的进展<sup>[65]</sup>。此长度不超过 2 kb 非编码 RNA 链介导许多细胞应激反应, 通过增加 BACE1 mRNA 的稳定性, 产生更多的 A $\beta$  肽, 从而促进疾病进一步恶化。用 siRNA 干扰非编码 RNA 链, 可使 A $\beta$  肽表达减少。所以有人认为这种非编码链可能成为治疗阿尔茨默病的十分有前景的靶向药物。

非编码 RNA 错误调节相关的蛋白编码基因的表达能够导致疾病发生, 如果对疾病有临床意义的基因下调, 这可能有助于疾病治疗。例如, 由 p15 肿瘤抑制基因转录的一反义非编码 RNA 链, 使局部染色质和 DNA 甲基化状态发生变化, 从而调节 p15 基因的表达<sup>[66]</sup>。在癌症中通过表观遗传机制, 许多抑癌基因

常被反义非编码 RNA 沉默。通过疾病相关基因多态性和染色体在非编码区的改变, 可能重新评估 lncRNAs 的功能。例如, 通过易位和反义 lncRNAs 的表达, 导致相邻的  $\alpha$ -珠蛋白基因表观遗传沉默, 导致  $\alpha$ -地中海贫血<sup>[67]</sup>。

## 7 展望

目前已越来越清楚的知道 lncRNAs 的许多分子功能, 包括调节转录模式、调节蛋白活性、作为小 RNA 的前体和改变 RNA 加工、维持细胞结构和保持其有序性等。目前只有很小的一部分 lncRNAs 被详细地确定分类, 但今后可能发现更多具有功能模式的 lncRNAs。随着对 lncRNAs 认识的加深, 将逐步建立有共同特征的非编码 RNA 文库, 它将有助于确认和预测非编码 RNA 的功能特征。目前面临的主要问题是, lncRNAs 的分子功能如何影响有机体。例如, lncRNAs 在许多生长发育过程中发挥作用<sup>[2]</sup>, 如 lncRNA 在视网膜的感光细胞及乳腺的发育过程中起到调节细胞周期和细胞存活的作用<sup>[69]</sup>。显而易见, RNA 不仅有编码蛋白质的功能, 还有内在信息传递的功能。这表明基因组远远比预期更为复杂, 甚至可能包含一种以 RNA 为基础的信息组件。假如大多数的非编码 RNA 被证明是具有功能性的, 那么它们的分类将有助于我们了解复杂有机体的遗传程序, 并将给进化、生长发育和疾病的认识, 带来新的见解。

## 参考文献(References)

1. Struhl K. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat Struct Mol Biol* 2007; 14: 103-5.
2. Amaral PP, Mattick JS. Noncoding RNA in development. *Mamm Genome* 2008; 19: 454-92.
3. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* 2008; 18: 1433-45.
4. Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res* 2006; 16: 11-9.
5. Blume SW, Meng Z, Shrestha K, Snyder RC, Emanuel PD. The 59-untranslated RNA of the human dhfr minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter. *J Cell Biochem* 2003; 88: 165-80.
6. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 2006; 441: 349-53.
7. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA,

- Wang H, *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* 2003; 300: 131-5.
8. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 2008; 322: 750-6.
  9. Sanchez-Elsner T, Gou D, Kremmer E, Sauer F. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* 2006; 311: 1118-23.
  10. Petruk S, Sedkov Y, Riley KM, Hodgson J, Schweisguth F, Hirose S, *et al.* Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. *Cell* 2006; 127: 1209-21.
  11. Lempradl A, Ringrose L. How does noncoding transcription regulate Hox genes? *Bioessays* 2008; 30: 110-21.
  12. Peters J, Robson JE. Imprinted noncoding RNAs. *Mamm Genome* 2008; 19: 493-502.
  13. Royo H, Cavaille J. Non-coding RNAs in imprinted gene clusters. *Biol Cell* 2008;100: 149-66.
  14. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 2002; 415: 810-3.
  15. Seidl CI, Stricker SH, Barlow DP. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J* 2006; 25: 3565-75.
  16. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, *et al.* The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 2008; 322: 1717-20.
  17. Dreyfuss G, Kim VN, Kataoka N. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* 2002; 3: 195-205.
  18. Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Dev* 2006; 20: 1470-84.
  19. Shamovsky I, Ivannikov M, Kandel ES, Gershon D, Nudler E. RNA-mediated response to heat shock in mammalian cells. *Nature* 2006; 440: 556-60.
  20. Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol* 2004; 11: 822-9.
  21. Mariner PD, Walters RD, Espinoza CA, Drullinger LF, Wagner SD, Kugel JF, *et al.* Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* 2008; 29: 499-509.
  22. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 2008; 454: 126-130.
  23. Hogan PG, Chen L, Nardone J, Rao A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes & Dev* 2003; 17: 2205-32.
  24. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, *et al.* A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 2005; 309: 1570-3.
  25. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 2007a; 316: 1484-8.
  26. Fejes-Toth K, Sotirova V, Sachidanandam R, Assaf G, Hannon GJ, Kapranov P, *et al.* Post-transcriptional processing generates a diversity of 59-modified long and short RNAs. *Nature* 2009; 457: 1028-32.
  27. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 2004; 10: 1957-66.
  28. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 2004; 23: 4051-60.
  29. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 2007; 318: 761-4.
  30. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 2008; 322: 1845-8.
  31. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. *Science* 2008; 322: 1855-7.
  32. Preker R, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, *et al.* RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 2008; 322: 1851-4.
  33. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, *et al.* Sharp PA. Divergent transcription from active promoters. *Science* 2008; 322: 1849-51.
  34. Neil H, Malabat C, d' Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 2009; 457: 1038-42.
  35. Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Cambong J, *et al.* Bidirectional promoters generate pervasive transcription in yeast. *Nature* 2009; 457: 1033-7.
  36. Lamond AI, Spector DL. Nuclear speckles: A model for nuclear organelles. *Nat Rev Mol Cell Biol* 2003; 4: 605-12.
  37. Wilusz JE, Freier SM, Spector DL. 3' End processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* 2008; 135: 919-32.
  38. Lee Y, Han J, Yeom KH, Jin H, Kim VN. Drosha in primary microRNA processing. *Cold Spring Harb Symp Quant Biol* 2006; 71: 51-7.
  39. Jaskiewicz L, Filipowicz W. Role of Dicer in posttranscriptional RNA silencing. *Curr Top Microbiol Immunol* 2008; 320: 77-97.
  40. Ganesan G, Rao SM. A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. *RNA* 2008; 14: 1399-410.
  41. Ogawa Y, Sun BK, Lee JT. Intersection of the RNA interference and X-inactivation pathways. *Science* 2008; 320: 1336-41.
  42. Kanellopoulou C, Muljo SA, Dimitrov SD, Chen X, Colin C, Plath K, *et al.* X chromosome inactivation in the absence of Dicer. *Proc Natl Acad Sci* 2009; 106: 1122-7.
  43. Hellwig S, Bass BL. A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. *Proc Natl Acad Sci* 2008; 105: 12897-902.

44. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* 2007; 4: 721-6.
45. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* 2007; 39: 1033-7.
46. Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 2008; 453: 534-8.
47. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 2008; 453: 539-43.
48. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, *et al.* An endogenous small interfering RNA pathway in *Drosophila*. *Nature* 2008; 453: 798-802.
49. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, *et al.* Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* 2008; 320: 1077-81.
50. Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC. The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* 2008; 453: 803-6.
51. Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, *et al.* A natural antisense transcript regulates *Zeb2/Sip1* gene expression during *Snail1*-induced epithelial-mesenchymal transition. *Genes & Dev* 2008; 22: 756-69.
52. Spector DL. Nuclear domains. *J Cell Sci* 2001; 114: 2891-3.
53. Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, *et al.* Regulating gene expression through RNA nuclear retention. *Cell* 2005; 123: 249-63.
54. Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, *et al.* An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* 2009; 33: 717-26.
55. Sasaki YTF, Ideue T, Sano M, Mituyama T, Hirose T. MEN  $\epsilon/\beta$  noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci* 2009; 106: 2525-30.
56. Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. MEN  $\epsilon/\beta$  nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* 2009; 19: 347-59.
57. Kloc M, Wilk K, Vargas D, Shirato Y, Bilinski S, Etkin LD. Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of *Xenopus* oocytes. *Development* 2005; 132: 3445-57.
58. Kloc M, Bilinski S, Dougherty MT. Organization of cyto keratin cytoskeleton and germ plasm in the vegetal cortex of *Xenopus laevis* oocytes depends on coding and non-coding RNAs: Three-dimensional and ultrastructural analysis. *Exp Cell Res* 2007; 313: 1639-51.
59. Blower MD, Nachury M, Heald R, Weis K. A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell* 2005; 121: 223-34.
60. Lecuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, *et al.* Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 2007; 131: 174-87.
61. Zhang LF, Huynh KD, Lee JT. Perinucleolar targeting of the inactive X during S phase: Evidence for a role in the maintenance of silencing. *Cell* 2007; 129: 693-706.
62. Mohammad F, Pandey RR, Nagano T, Chakalova L, Mondal T, Fraser P, *et al.* *Kcnq1ot1/Lit1* noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. *Mol Cell Biol* 2008; 28: 3713-28.
63. Costa FF. Non-coding RNAs: New players in eukaryotic biology. *Gene* 2005; 357: 83-94.
64. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeneij LA, Aalders TW, *et al.* DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res* 2002; 62: 2695-8.
65. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of  $\beta$ -secretase. *Nat Med* 2008; 14: 723-30.
66. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, *et al.* Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* 2008; 451, 202-6.
67. Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG, *et al.* Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nature Genet* 2003; 34: 157-65.
68. Young TL, Matsuda T, Cepko CL. The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. *Curr Biol* 2005; 15: 501-12.
69. Ginger MR, Shore AN, Contreras A, Rijnkels M, Miller J, Gonzalez-Rimbau MF, *et al.* A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc Natl Acad Sci* 2006; 103: 5781-6.

## The Progress on Long Non-coding RNA Function

Liang-He Yu, Nan Li, Shu-Qun Cheng\*

(Second Military Medical University, Shanghai 200438, China)

**Abstract** Most of the eukaryotic genome is transcribed, yielding a complex network of transcripts that includes tens of thousands of long noncoding RNAs with little or no protein-coding capacity. Although the vast majority of long noncoding RNAs have yet to be characterized thoroughly, there are probably many functions of long ncRNAs awaiting discovery. Multiple studies have shown that significant numbers of long ncRNAs are regulated during development, exhibit cell typespecific expression, localize to specific subcellular compartments, and are associated with human diseases. Here, we review the recent evolutions of long noncoding RNAs, their functions and medical significance.

**Key words** long non-coding RNA; transcription; LncRNAs;

---

This work was supported by the Natural Science Foundation of China(No.30873352), Ministry of Science and Technology Key Program (No.2008zx10002-025), Shanghai Municipal Outstanding Leader Project Foundation(No.10XD1405800)

\*Corresponding author. Tel: 86-21-81875251, Fax: 86-21-65562400, E-mail: chengshuqun@yahoo.com.cn