

DOI: 10.13376/j.cblls/2016090

文章编号: 1004-0374(2016)06-0703-09



朱鸿亮, 教授、博士生导师。2013年入选中国农业大学拔尖人才计划。朱鸿亮博士长期致力于植物非编码 RNA (non-coding RNA, ncRNA) 的调控和功能研究, 在植物生长发育和果实成熟的 ncRNA 调控研究领域取得了较好的学术成绩, 得到国内外同行一致认可。主要成果: (1) 解析了 Argonaute10 (AGO10) 蛋白拮抗 miRNA 调控茎尖分生组织发育的分子机理; (2) 阐明了 Dicer-like 1 (DCL1) 双向剪切加工 miRNA 前体的生化和分子机制; (3) 探明了长链非编码 RNA (long non-coding RNA, lncRNA) 参与番茄果实成熟的分子调控; (4) 揭示了番茄果实成熟关键转录因子 RIN 直接调控 miRNA 前体转录生成的分子基础。近 5 年来, 相关研究成果先后发表在 *Cell*、*Nature Molecular & Structural Biology*、*Journal of Experimental Botany* 和 *Plant Biotechnology Journal* 等杂志上。目前是 *Plan*、*Cell & Environment*、*Journal of Experimental Botany* 和 *Scientific Reports* 等学术刊物的审稿人。

## 长链非编码RNA的功能及其作用机制

李睿, 杨永芳, 李冉, 朱鸿亮\*, 罗云波

(中国农业大学食品科学与营养工程学院, 北京 100083)

**摘要:** 近年来, 随着高通量深度测序技术的飞速发展, 大量长链非编码 RNA (lncRNA) 在真核生物体内不断被发现。lncRNA 是一类长度大于 200 nt, 生物学功能丰富, 高级结构保守的非编码 RNA (ncRNA) 分子, 能直接与蛋白质、微小 RNA (microRNA) 等分子作用, 在表观遗传水平、转录水平、转录后水平上调控靶基因的表达。主要就 lncRNA 的分类标准、调控基因表达模式及其与蛋白质、miRNA 的相互作用进行综述, 旨在为更好地理解和研究 lncRNA 生物学功能提供参考。

**关键词:** 长链非编码 RNA; 蛋白质; 微小 RNA; 生物学功能

**中图分类号:** Q522

**文献标志码:** A

## Functions of long non-coding RNA and its interacting mechanisms

LI Rui, YANG Yong-Fang, LI Ran, ZHU Hong-Liang\*, LUO Yun-Bo

(College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China)

**Abstract:** As with the rapid development of high through-put deep sequencing technology, a large number of long non-coding RNAs (lncRNAs) have been continuously discovered in eukarya in recent years. LncRNA is a group of non-coding RNA (ncRNA) which is longer than 200 nt, rich in biological function and highly structured. In addition, lncRNA could directly interact with protein, microRNA (miRNA) molecules to regulate target gene expression at epigenetic, transcriptional and post-transcriptional levels. This review systematically summarizes lncRNA classification standard, its gene expression regulatory patterns and interaction with protein and miRNA, which is aimed to provide better understanding of lncRNA biological functions.

**Key words:** long non-coding RNA (lncRNA); protein; microRNA (miRNA); biological function

收稿日期: 2016-03-15

基金项目: 中央高校基本科研业务费专项资金(2016- QC037); 国家自然科学基金重大研究计划(91540118); 中国农业大学教育基金会“大北农教育基金”(1061-2415003)

\*通信作者: E-mail: hlzhu@cau.edu.cn; Tel: 010-62737571

随着高通量测序技术的飞速发展, 数以万计的非编码转录物不断在生物体内被发现, 以 mRNA 为核心的遗传中心法受到空前挑战。人类全基因组转录组分析表明, 人体内虽有大量转录物产生, 仅有 1%~2% 的基因组序列具备蛋白质编码功能, 而不具备蛋白质编码能力的非编码区含量竟高达 98% 以上, 暗示着生物体内有大量非编码 RNA (non-coding RNA, ncRNA) 产生<sup>[1]</sup>。ncRNA 根据其功能可分为持家 ncRNA (house keeping ncRNA) 和调控 ncRNA (regulatory ncRNA)<sup>[2]</sup>。持家 ncRNA 主要包含 rRNA、tRNA、核内小 RNA (small nuclear RNA, snRNA) 以及核仁小 RNA (small nucleolar RNA, snoRNA)<sup>[2-3]</sup>。调控 ncRNA 根据其分子链长短差异可进一步分为短链非编码 RNA (small non-coding RNA, sncRNA) 和长链非编码 RNA (long non-coding RNA, lncRNA)<sup>[2-3]</sup>。以 miRNA (microRNA)、siRNA (small interfering RNA)、piwiRNA 为代表的 sncRNA 在过去十年中作为分子生物学领域的研究热点, 其生物学功能被广泛揭示, 它们能在转录水平、转录后水平调控真核生物基因表达<sup>[4-5]</sup>。与 sncRNA 相比, 表达量更低、分子结构更为复杂的 lncRNA 研究才刚起步。即便如此, 作为分子生物学领域内的新星, lncRNA 在近年来取得了迅猛的发展和丰硕的成果。本文主要就 lncRNA 分类标准、调控基因表达模式及其与蛋白质、miRNA 的相互作用作一综述。

## 1 lncRNA概述

一般认为 lncRNA 是一类生物来源广泛, 长度大于 200 nt, 主要由 RNA 聚合酶 II (PolII) 转录生成, 并具备甲基鸟苷帽子和多聚腺苷酸尾 (polyA) 结构的 ncRNA<sup>[6]</sup>。lncRNA 在生物体内含量巨大, 种类繁多, 至今没有统一的生物学分类标准。为便于研究, 研究者通常会根据 lncRNA 与蛋白质编码基因的相对位置将其分为 4 类: 重叠型、双向型、内含子型、顺式反义型<sup>[7]</sup>。起初, 这类“暗物质”并未引起科学家的广泛关注, 它们被认定为基因转录过程中的噪音, 不具备任何生物学意义<sup>[7-8]</sup>。近年来, 越来越多的研究已证实, lncRNA 广泛参与 DNA 甲基化、组蛋白修饰、染色质重塑等生物学过程, 能与 DNA、RNA、蛋白质分子作用, 顺式或反式调控靶基因表达<sup>[6,9-10]</sup>。介于 lncRNA 丰富的生物学功能, Chang 实验室在 2011 年首次提出将 lncRNA 行使生物学功能的方式分为 4 类 (信号分子、诱饵分子、引导分子、支架分子), 极大地推进了 lncRNA

研究领域的发展<sup>[11]</sup>。

lncRNA 不仅生物学功能复杂, 在表达方式上还具有极强的时空特异性<sup>[12]</sup>。虽然 lncRNA 初级结构保守性差, 但其二级、三级结构保守性却很强, 这些高度保守的结构与 lncRNA 生物学功能密切相关<sup>[8,13]</sup>。例如, 参与染色体剂量补偿效应的经典 lncRNA——X 染色体特异性失活转录物 (X-inactive specific transcript, Xist) 中重复单元 A (repeat A, RepA) 和重复单元 C (RepC) 分别与多梳抑制复合物 2 (polycomb repressive complex 2, PRC2) 和二价体蛋白 YY1 招募有关<sup>[14-15]</sup>。此外, Iliket 等<sup>[16]</sup>还在果蝇 (*Drosophila melanogaster*) lncRNA——X 染色体上的 RNA 1 (RNA on the X 1, roX1) 和 X 染色体上的 RNA 2 (roX2) 中发现了一系列串联茎环结构, 并证实该结构的改变将直接影响雄性特异致死复合物 (male-lethal specific, MSL) 招募, 进一步提示结构与功能的相关性。

## 2 lncRNA调控基因表达

### 2.1 lncRNA参与基因表观遗传调控

目前已知的表观遗传学现象有基因沉默、DNA 甲基化、组蛋白修饰、休眠转座子激活、RNA 编辑和基因甲基化<sup>[17]</sup>。lncRNA 作为一种新的表观遗传调控分子, 在表观遗传学调控中具有十分重要的作用。研究报道, 在哺乳动物染色体剂量补偿效应中, Xist 首先从 X 失活中心 (X inactivation center, Xic) 转录, 通过招募 PRC2 靶向作用于 X 染色体特定位点, 促进组蛋白 H3 第 27 位赖氨酸三甲基化 (histone H3 lysine K27 trimethylation, H3K27me3), 介导相关基因沉默<sup>[14-15,18-19]</sup>。此外, lncRNA 表观遗传调控对植物的生殖发育和胁迫应答也具有重要意义。由开花抑制基因 *FLC* (*Flowering locus C*) 的第一个内含子和 3' 末端转录出的 2 条经典 lncRNAs——冷协助内含子区非编码 RNA (cold assisted intronic noncoding RNA, COLDAIR) 和冷诱导反义 RNA (cold induced antisense intragenic RNA, COOLAIR) 已被证实能在春化过程中通过招募 PRC2 介导染色质重塑, 来抑制 *FLC* 的表达, 进而调控拟南芥 (*Arabidopsis thaliana*) 开花时间<sup>[20-21]</sup>。Ding 等<sup>[22]</sup>在水稻 (*Oryza sativa*) 光敏性雄性不育系“农垦 58S”体内发现了与其性状表达密切相关的 lncRNA——LDMAR (long-day-specific male-fertility-associated lincRNA)。研究表明 LDMAR 长 1236 nt, 在相关酶作用下, 能产生许多 21 nt 功能性 siRNAs, 它们进一步通过

RNA介导的DNA甲基化(RNA-directed DNA methylation, RdDM)途径使 *LDMAR* 启动子区甲基化水平升高,阻遏 *LDMAR* 转录,导致水稻雄性不育<sup>[23]</sup>。

## 2.2 lncRNA参与基因转录调控

基因转录是一个严密复杂的生物过程, lncRNA 能够通过模仿 DNA 元件,竞争性结合转录因子或进行可变剪切等来调控基因转录<sup>[24]</sup>。由 PolII 转录的长约 200 bp 的生长阻滞特异转录物 5 (growth arrest-specific transcript 5, Gas5) 是哺乳动物细胞凋亡和生长的关键调控因子。Gas5 通过模拟糖皮质激素应答元件来结合糖皮质激素受体 (glucocorticoid receptor, GR) 的 DNA 结合结构域,阻止糖皮质激素受体与糖皮质激素应答元件的相互作用,从而抑制下游基因的转录表达,促进细胞凋亡发生<sup>[25]</sup>。Feng 等<sup>[26]</sup>在远端缺失同源异型盒子 -5/6 (distal-less homeobox-5/6, *Dlx-5/6*) 基因序列间发现了 2 个超保守增强子,其中由增强子 ei 转录生成的长为 3.8 kb 的 lncRNA——胚胎前脑 -2 (embryonic ventral forebrain-2, *Evf-2*) 作为 *Evf-1* 的可变剪切体能直接与转录因子 *Dlx-2* 结合形成功能性复合物,靶向作用于 *Dlx-5/6* 基因转录增强子,促进靶基因的表达,调控动物神经元的发育。动物体内细胞周期常受到二氢叶酸还原酶相关基因 (dihydrofolatereductase, *DHFR*) 的调控。近年来研究人员在 *DHFR* 基因上游关键启动子可变区域产生的转录物中发现了一类 lncRNA 能与 PolII 竞争性结合转录因子 IIB, 阻断相应转录因子在转录过程中的招募,进而介导 *DHFR* 相关基因沉默,调控细胞周期<sup>[27]</sup>。

## 2.3 lncRNA参与基因转录后调控

转录后水平基因调控是指基因在转录后的一系列加工、修饰和调节过程,主要包括 RNA 剪切、加工、拼接、成熟、代谢以及稳定性调节等,在基因表达中起十分重要的作用<sup>[28]</sup>。mRNA 前体的剪切作为 mRNA 加工代谢中的重要步骤常受到 lncRNA 调控。Tripathi 等<sup>[29]</sup>发现肺腺癌转移相关 lncRNA MALAT1 (metastasis associated lung adenocarcinoma transcript 1) 能通过与丝氨酸/精氨酸 (serine/arginine, SR) 富含性剪切蛋白质因子相互作用,影响 SR 蛋白的亚细胞定位,调节其在细胞中的浓度,参与 mRNA 前体可变剪切。当真核细胞受到外界热刺激时,不仅会有热休克基因 (heat shock genes, *HS*) 的转录表达参与细胞防护,还会涉及 lncRNA 参与的 mRNA 前体剪切。例如,果蝇热

休克 RNA $\omega$ -n (heat shock RNA  $\omega$ -n, hsr $\omega$ -n) 和人类卫星 III (satellite III, sat III) 在细胞受到热刺激时会通过隔绝剪切因子调控 mRNA 前体剪切<sup>[30]</sup>。

## 3 lncRNA与蛋白质、RNA的相互作用

从 lncRNA 调控靶基因表达水平来看,可分为表观遗传水平、转录水平、转录后水平 (表 1); 根据 lncRNA 调控过程中的互作对象又可分为蛋白质、RNA、DNA 三类,目前研究主要集中于前面两类。功能性蛋白质可作为染色质重塑修饰物<sup>[18-21]</sup>、转录因子<sup>[26]</sup>、剪切因子<sup>[29-30]</sup> 等与 lncRNA 互作,分别在不同水平对靶基因表达进行调控。RNA 分子,尤其是 miRNA 主要通过 lncRNA 分子链上的特殊识别位点与其形成完全或不完全配对,在不同水平上调控基因转录表达<sup>[7,9]</sup>。因此,研究 lncRNA-蛋白质、lncRNA-miRNA 之间的互作机制对于揭开 lncRNA 复杂的生物学功能具有重大意义。下面将从蛋白质、miRNA 出发,分别总结归纳出三种 lncRNA-蛋白质互作、lncRNA-miRNA 互作模型,以期从另一视角对 lncRNA 生物学功能进行阐释。

### 3.1 lncRNA与蛋白质相互作用

#### 3.1.1 lncRNA作为蛋白质的引导分子

首先, lncRNA 可作为引导分子将蛋白质招募到基因转录发生位置,调控基因的表达 (图 1A)。通过研究体外表达的 MLL1 复合物亚基与 *HOXA* 远端转录物 (*HOXA* transcript at the distal tip, HOTTIP) 的结合能力发现, WDR5 蛋白能够特异性地结合 HOTTIP<sup>[31]</sup>。HOTTIP 通过与 WDR5 结合募集 MLL1 复合物并对组蛋白进行 H3K4me3 修饰,从而维持染色质的开放状态和活跃转录<sup>[31]</sup>。拟南芥中 COLDAIR 是一个具有 5' 帽子结构但缺乏 3' 多聚腺苷化 polyA 的 lncRNA。研究表明,在春化过程中 COLDAIR 可以引导 PRC2 至 *FLC* 基因组位置引起 H3K27me3, 抑制 *FLC* 基因表达<sup>[32]</sup>。lncRNA 除了能引导 PRC2 至靶基因特定位点,介导染色质重塑,抑制靶基因表达,还可作为其他功能性蛋白的引导分子参与靶基因表达调控。在 p53 信号通路研究中, Huart 等<sup>[33]</sup>发现由 p53 诱导产生的长链基因间非编码 RNA p21 (long intergenic ncRNA p21, lincRNA-p21) 可作为核不均一核糖核蛋白-K (heterogeneous nuclear ribonucleoprotein-K, hnRNP-K) 的引导分子,将其定位到 *p21* 基因启动子区,激活 *p21* 转录表达。

#### 3.1.2 lncRNA作为蛋白质支架分子

lncRNA 还可作为两个或者两个以上蛋白质

表1 lncRNA参与基因表达调控

lncRNA	基因调控水平	生物学功能	文献
Xist	表观遗传水平	招募PRC2靶向作用于X染色体特定位点, 促进H3K27me <sub>3</sub> , 介导相关基因沉默	[14-15,18-19]
COLDAIR	表观遗传水平	在春化过程中通过招募PRC2介导染色质重塑, 来抑制 <i>FLC</i> 的表达, 从而使拟	[20-21]
COOLAIR		南芥( <i>A. thaliana</i> )正常开花	
LDMAR	表观遗传水平	在相关酶作用下, LDMAR能够产生许多21 nt功能性siRNAs, 它们进一步通过 RdDM途径使LDMAR启动子区甲基化水平升高, 阻遏LDMAR转录, 导致水稻( <i>O. sativa</i> )雄性不育	[22-23]
Gas5	转录水平	Gas5 通过模拟糖皮质激素应答元件来结合糖皮质激素受体的 DNA结合结构域, 阻止糖皮质激素受体与糖皮质激素应答元件的相互作用, 从而抑制下游基因的转录表达, 促进细胞凋亡的发生	[25]
Evf-2	转录水平	Evf-2直接与转录因子Dlx-2结合形成功能性复合物, 靶向作用于Dlx-5/6基因转录增强子, 促进靶基因的表达;	[26]
DHFR	转录水平	与Pol II 竞争性结合转录因子 II B, 阻断相应转录因子在转录过程中的招募, 进而介导DHFR相关基因沉默, 调控细胞周期	[27]
MALAT1	转录后水平	通过调节SR富含性剪切蛋白因子在细胞核中的分布水平, 调控mRNA前体可变剪切	[29]
hsr $\alpha$ -n	转录后水平	当果蝇细胞受到热刺激时通过隔绝剪切因子调控mRNA前体剪切	[30]
sat III	转录后水平	当人体细胞受到热刺激时通过隔绝剪切因子调控mRNA前体剪切	[30]

的支架分子, 将这些蛋白质组合成蛋白质复合物 (图 1B)。同源异型框基因反义基因间 RNA (HOX antisense intergenic RNA, HOTAIR) 位于同源异型框基因 C (homeobox C, *HOXC*) 基因簇, 并且在肢体后远端的细胞中表达。HOTAIR 能够通过反式作用抑制同源异型框基因 D (homeobox D, *HOXD*) 基因簇的转录, 抑制 HOTAIR 能够激活 *HOXD* 基因簇的表达<sup>[34]</sup>。此外, HOTAIR 能够作为支架结构同时结合 PRC2 和组蛋白赖氨酸特异性去甲基化酶 1 (lysine-specific demethylase1, LSD1) 两种染色质修饰复合物, 这两种复合物分别介导 H3K27me<sub>3</sub> 和组蛋白 H3 第 4 位赖氨酸去甲基化 (histone H3 lysine K4 demethylation) 修饰, 协同发挥转录抑制作用<sup>[35]</sup>。另一个 lncRNA——INK4 位点反义非编码 RNA (antisense non-coding RNA in the INK4 locus, ANRIL) 转录于周期蛋白依赖性激酶抑制因子 2B (cyclin-dependent kinase inhibitor 2B, CDKN2B, *p16<sup>INK4a</sup>*) 基因转录起始位点 (transcription start site, TSS), 也通过招募 PRC1 和 PRC2 两种蛋白质复合物到编码基因特定位置, 顺式调控基因表达<sup>[35]</sup>。果蝇中 roX 转录于 X 染色体, 同样是作为 MSL 剂量补偿复合物的支架分子, 使组蛋白 H4 第 16 位赖氨酸乙酰化 (histone H4 lysine K16 acetylation), 激活下游基因转录<sup>[36]</sup>。

### 3.1.3 lncRNA作为蛋白质的诱饵分子

最后, lncRNA 还可作为蛋白质的诱饵分子,

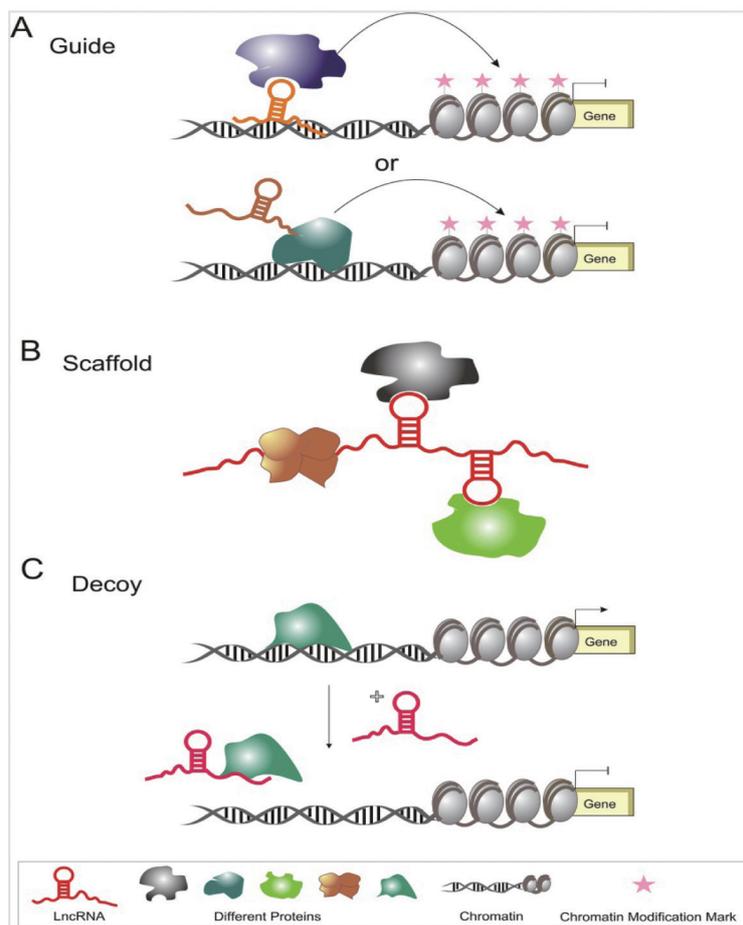
将结合在基因组的蛋白质调走, 从而抑制基因的转录和表达 (图 1C)。lncRNA 核副斑点组装转录物 1 (nuclear paraspeckle assembly transcript 1, NEAT1) 和 PTB 关联剪切因子 (PTB-associated splicing factor, PSF) 蛋白都是细胞核亚结构旁斑 (paraspeckle) 的必要组成成分, 病毒感染细胞能够导致 NEAT1 的表达量显著提高, 过表达的 NEAT1 结合 PSF, 将其从下游基因 IL8 等的启动子区解离并转移到细胞核亚结构旁斑中, 激活 IL8 的表达, 提高了细胞的抗病毒能力<sup>[37]</sup>, 这是 lncRNA 作为诱饵分子的一个典型例子。

另一个经典的作为蛋白诱饵分子的 lncRNA——DNA 损伤活化 P21 相关非编码 RNA (P21 associated ncRNA DNA damage activated, PANDA) 常作为转录因子诱饵分子参与基因表达调控<sup>[38]</sup>。PANDA 位于周期蛋白依赖性激酶抑制因子 1A (cyclin-dependent kinase inhibitor 1A, CDKN1A, 即 p21) 基因上游 5 kb 处, 具有 5' 帽子和未经剪切修饰的 3' 多聚腺苷酸尾结构<sup>[38]</sup>。当 DNA 损伤发生时, PANDA 即被转录, 通过提前结合核转录因子 Y $\alpha$  亚基 (nuclear transcription factor Y subunit  $\alpha$ , NF-YA) 阻止其与靶基因互作, 进而调控细胞生存时间<sup>[38]</sup>。

## 3.2 lncRNA与miRNA相互作用

### 3.2.1 miRNA介导的lncRNA降解

lncRNA 在生物体中数量巨大, 且与细胞功能的发挥, 如增殖、分化、衰老、程序性死亡等密切



(A) lncRNA作为引导分子, 指导蛋白质复合物定位到特定位点, 调控靶基因表达。(B) lncRNA作为支架分子, 同时招募不同蛋白质复合物调控靶基因表达。(C) lncRNA作为蛋白质诱饵分子, 吸附目标蛋白, 阻遏其对靶基因的作用。

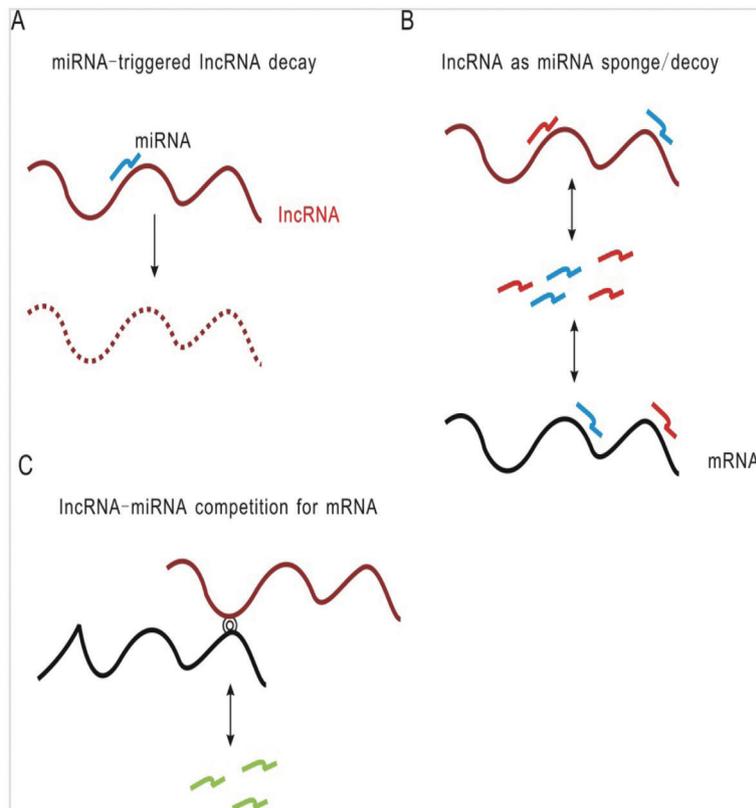
图1 lncRNA-蛋白质相互作用模型<sup>[11]</sup>

相关<sup>[39-40]</sup>。所以, lncRNA在生物体内含量的变化往往与细胞功能异常直接相关, 导致相应的生理病理症状出现<sup>[40]</sup>。此外, 由于lncRNA分子链长度的特殊性, 使得链上存在许多miRNA识别作用位点, 这些保守位点能特异地与miRNA结合, 影响lncRNA体内稳定性, 介导lncRNA降解, 调控细胞生物学功能<sup>[41]</sup>(图2A)。在人体宫颈癌细胞中, lincRNA-P21的稳定性不仅受到RNA结合蛋白(RNA binding protein, RBP), 如HUR与Ago2(Argonaute 2)的调控, 还直接受到miRNA let-7b的调节, 过表达miRNA let-7b将促进lincRNA-P21降解<sup>[42]</sup>。HOTAIR作为肿瘤发生相关lncRNA, 其在体内的稳定性也与miRNA let7的表达水平呈负相关<sup>[43-44]</sup>。此外, 在PC3和DU145人体前列腺癌细胞系中, Chiyomaru等<sup>[45]</sup>还发现miR-34a的表达量与HOTAIR低稳定性密切相关。在人类原发性胶质母细胞瘤细胞系U87MG中, MALAT1同样也作为miRNA-Ago2-RISC

复合物的靶标, 受到这些复合物的共调控。沉默Ago2或抑制miR-9表达将显著提升MALAT1在人体内的稳定水平<sup>[46]</sup>。

### 3.2.2 lncRNA作为miRNA的诱饵分子调控基因表达

由于lncRNA分子中存在许多miRNA识别位点, 所以lncRNA不仅可作为蛋白质的诱饵分子调控靶基因表达, 还可充当miRNA的诱饵分子, 通过靶目标模仿(target mimicry)方式吸附miRNA, 抑制其对靶标的进一步作用(图2B)。这种机制最早在植物中提出, Franco-Zorrilla等<sup>[47]</sup>在磷酸盐缺乏的拟南芥植株体内发现有大量磷酸盐缺乏诱导转录物1(induced by phosphate starvation 1, IPS1)产生, IPS1可作为miR-399的诱饵分子率先与其结合, 并抑制miR-399对PHO2 mRNA的降解, 从而促使PHO2转录表达, 提高植物体磷含量。随后, 动物中有关lncRNA作为miRNA诱饵分子调控靶基因表达模式也陆续被报道。Cesana等<sup>[48]</sup>在肌细胞分



(A) miRNA与lncRNA保守位点结合,影响lncRNA体内稳定性,介导lncRNA降解。(B) lncRNA作为miRNA的诱饵分子,通过靶目标模仿方式吸附miRNA,抑制其对靶标的进一步作用,调控基因表达。(C) lncRNA作为miRNA竞争性内源分子,直接与mRNA结合,减少miRNA结合位点,对靶基因实现调控。

图2 lncRNA-miRNA相互作用模型<sup>[40]</sup>

化的研究中发现一类肌分化特异性 lncRNA——linc-MD1 (long intergenic noncoding RNA muscle differentiation 1) 作为 miRNA-133 和 miRNA-135 诱饵分子,分别阻断 miRNA-133 与 miRNA-135 对决定因子样蛋白-1 (mastermind-like 1, *MAML1*) 和肌细胞增强因子 2C (myocyte enhancer factor 2C, *MEF2C*) mRNA 表达的抑制作用,调控肌细胞分化。在小鼠 (*Mus musculus*) 体内过表达 *linc-MD1* 将促进肌细胞分化,反之,*linc-MD1* 表达缺失则对肌细胞的分化起抑制作用<sup>[48]</sup>。最新研究表明,*linc-MD1* 分子诱饵效应在肌细胞分化过程中还参与 RBP HuR 生物学功能调控。HuR 在生物体内的表达量与 miR-133 的表达水平负相关,当 miR-133 被分子诱饵 *linc-MD1* 带走时,miR-133 对 HuR 的抑制作用即被解除,随即 HuR 在细胞内大量积累,介导肌源性基因表达,调控肌细胞分化<sup>[49]</sup>。起初,H19 常在胚胎时期高量表达,所以 H19 一直作为基因组印记相关 lncRNA 来研究<sup>[50]</sup>。最近,Kallen 等<sup>[51]</sup>发现 H19 在肌细胞分化过程中还可作为 let-7 诱饵分子调控肌细胞分

化。H19 分子链上存在 let-7 家族 miRNA 识别位点,可与 let-7e、let-7g 和 let-7i 形成不完全配对,阻遏 let-7 对靶标的作用。

### 3.2.3 lncRNA 作为 miRNA 竞争性内源分子与 mRNA 互作

lncRNA 可作为分子诱饵,通过阻遏 miRNA 对靶标的作用,间接调控靶基因的表达。此外,lncRNA 还可竞争性与 miRNA 结合 mRNA,直接对靶基因进行调控(图 2C)。 $\beta$  分泌酶 1 ( $\beta$ -site amyloid precursor protein cleaving enzyme 1, *BACE1*) mRNA 上存在其基因反义转录物 lncRNA *BACE1-AS* (*BACE1-antisense*) 识别位点,该识别位点包含 miR-484-5p 作用位点<sup>[52]</sup>。当 miR-484-5p 与 *BACE1* 结合时,*BACE1* 表达量会显著降低。在人体胚胎肾细胞 HEK-293 中,*BACE1-AS* 会与 miR-484-5p 竞争性结合 *BACE1*,减少 miR-484-5p 结合位点,进而提高 *BACE1* 稳定性<sup>[52]</sup>。促癌 lncRNA ncNRF (non-coding Nras functional RNA) 包含一段 22 nt 保守序列,该序列与 let-7a 极其相似,且同其他 miRNA,如 let-7b、let-7c、let-

7d 等表现出 1~4 nt 序列差异性<sup>[53]</sup>。研究表明,在结肠上皮细胞系 YAMC 中过表达 ncNRFR 会减弱 let-7 的生物学功能。尽管 ncNRFR 具体的作用机制并未被揭示,研究者在过表达 ncNRFR 的 YAMC 细胞系中发现了更高量的 let-7 靶标 mRNA,暗示 ncNRFR 可能通过其保守的 22 nt 序列与 let-7 靶标 mRNA 相互作用,从而阻遏 let-7 与其结合<sup>[53]</sup>。

#### 4 总结与展望

随着高通量测序技术的飞速发展和大规模应用,大量 ncRNA 被发现和预测出来,其中以 lncRNA 为代表的 ncRNA 已逐渐在分子生物学领域崭露头角<sup>[54]</sup>。近年来,虽有大量 lncRNA 不断在生物体内挖掘出来,但研究仍停留于表层阶段,距离完全揭示这类基因中“暗物质”的生物学功能还有很长的路要走。蛋白质作为 lncRNA 重要结合物,在 lncRNA 生物学功能的发挥中具有极其重要的作用,研究蛋白质分子如何与 lncRNA 互作将更便于研究者对 lncRNA 精细的分子调控机制作深层次理解。本文虽总结了 lncRNA 与蛋白质分子之间的互作模型,但这种划分方式还不够全面,并未完全考虑到 lncRNA 生物学功能的多样性。例如, HOTAIR 在招募 PRC2 和 LSD1 复合物时充当支架分子的角色,而将 PRC2 靶向定位到特定位点的过程中又起引导分子的作用<sup>[11]</sup>。所以,随着更多 lncRNA 生物学作用机制被揭示, lncRNA-蛋白质之间的相互作用方式将不断被更新。此外,在 lncRNA 与另一生物大分子 RNA 相互作用的研究中还存在诸多不足。就 RNA 分子种类而言,现有的研究大多集中于 miRNA 上,相对缺乏对其他 ncRNA (tRNA、rRNA、siRNA、piwiRNA 和 snoRNA 等)的报道<sup>[40]</sup>,因此,研究 lncRNA 如何与这些 RNA 分子作用将是其未来的发展方向,不仅能弥补 lncRNA 生物学调控网络的空白,还能为 lncRNA-RNA 分子间的互作模式带来新的理解。

最近研究者还发现,真核生物体内部分 lncRNA 可以自身环化的形式发挥生物学功能,该类 ncRNA 又称作环状 RNA (circular RNA, circRNA)<sup>[55]</sup>。circRNA 分子链上存在着 miRNA 识别元件,这些保守元件使得 circRNA 可作为许多 miRNA 的海绵分子,通过吸附 miRNA 来调控其对靶标的作用<sup>[56]</sup>。此外,部分 circRNA 还具备 RNA 结合蛋白 (RNA binding protein, RBP) 识别结构域,能与特异的 RBP 结合,顺式或反式调控靶基因表达<sup>[56]</sup>。circRNA 以其结构

的特殊性和功能的复杂性迅速受到了广泛关注。目前,我国在 circRNA 领域内也具有不错的研究成果。陈玲玲和杨力研究小组在人体细胞中发现了一类完全由内含子区环化而成的 circRNA (circular intronic RNA, ciRNA), ciRNA 富含于细胞核内,环形结构的产生依赖于剪切位点附近的特殊序列,敲除 ciRNA 将会减弱其亲本基因的表达<sup>[57]</sup>。随后,单革研究团队在人体细胞核内首次发现了一种由内含子和外显子共同组成的 circRNA (exon-intron circular RNA, EIciRNA)。研究表明, EIciRNA 可与 U1snRNP 作用,顺式增强亲本基因的转录表达,进一步提示 circRNA 可在转录水平上对靶基因进行调控<sup>[58]</sup>。随着这类基因中“暗物质”分子生物学机制不断被揭示,人们将有望从另一角度去阐释复杂的生命现象。

#### [参 考 文 献]

- [1] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 2004, 431: 931-45
- [2] Kim ED, Sung S. Long noncoding RNA: unveiling hidden layer of gene regulatory networks. *Trends Plant Sci*, 2012, 17: 16-21
- [3] Zhu QH, Wang MB. Molecular functions of long non-coding RNAs in plants. *Genes*, 2012, 3: 176-90
- [4] Bonnet E, Van de Peer Y, Rouzé P. The small RNA world of plants. *New Phytol*, 2006, 171: 451-68
- [5] Simon SA, Meyers BC. Small RNA-mediated epigenetic modifications in plants. *Curr Opin Plant Biol*, 2011, 14: 148-55
- [6] Chen LL, Carmichael GG. Decoding the function of nuclear long non-coding RNAs. *Curr Opin Cell Biol*, 2010, 22: 357-64
- [7] Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet*, 2009, 10: 155-9
- [8] Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol*, 2013, 20: 300-7
- [9] Bai Y, Dai X, Harrison AP, et al. RNA regulatory networks in animals and plants: a long noncoding RNA perspective. *Brief Funct Genomics*, 2015, 14: 91-101
- [10] Wierzbicki AT. The role of long non-coding RNA in transcriptional gene silencing. *Curr Opin Plant Biol*, 2012, 15: 517-22
- [11] Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*, 2011, 43: 904-14
- [12] Zhu J, Fu H, Wu Y, et al. Function of lncRNAs and approaches to lncRNA-protein interactions. *Sci China Life Sci*, 2013, 56: 876-85
- [13] Novikova IV, Hennelly SP, Sanbonmatsu KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res*, 2012, 40:

- 5034-51
- [14] Maenner S, Bland M, Fouillen L, et al. 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol*, 2010, 8: e1000276
- [15] Jeon Y, Lee JT. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*, 2011, 146: 119-33
- [16] Ilik IA, Quinn JJ, Georgiev P, et al. Tandem stem-loops in roX RNAs act together to mediate X chromosome dosage compensation in *Drosophila*. *Mol Cell*, 2013, 51: 156-73
- [17] Heo JB, Lee YS, Sung S. Epigenetic regulation by long noncoding RNAs in plants. *Chromosome Res*, 2013, 21: 685-93
- [18] Arthold S, Kurowski A, Wutz A. Mechanistic insights into chromosome-wide silencing in X inactivation. *Hum Genet*, 2011, 130: 295-305
- [19] Escamilla-Del-Arenal M, da Rocha ST, Heard E. Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum Genet*, 2011, 130: 307-27
- [20] Csorba T, Questa JI, Sun Q, et al. Antisense COOLAIR mediates the coordinated switching of chromatin states at *FLC* during vernalization. *Proc Natl Acad Sci USA*, 2014, 111: 16160-5
- [21] Yamaguchi A, Abe M. Regulation of reproductive development by non-coding RNA in *Arabidopsis*: to flower or not to flower. *J Plant Res*, 2012, 125: 693-704
- [22] Ding J, Lu Q, Ouyang Y, et al. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci USA*, 2012, 109: 2654-9
- [23] Ding J, Shen J, Mao H, et al. RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. *Mol Plant*, 2012, 5: 1210-6
- [24] Wang X, Arai S, Song X, et al. Induced ncRNA allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature*, 2008, 454: 126-30
- [25] Kino T, Hurt DE, Ichijo T, et al. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*, 2010, 3: ra8
- [26] Feng J, Bi C, Clark BS, et al. The Evi-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
- [27] Hainer SJ, Martens JA. Transcription of ncDNA: many roads lead to local gene regulation. *Transcription*, 2011, 2: 120-3
- [28] Filipowicz W, Jaskiewicz L, Kolb FA, et al. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr Opin Struct Biol*, 2005, 15: 331-41
- [29] Tripathi V, Ellis JD, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*, 2010, 39: 925-38
- [30] Jolly C, Lakhotia SC. Human sat III and *Drosophila* hsr $\omega$  transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res*, 2006, 34: 5508-14
- [31] Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, 2011, 472: 120-4
- [32] Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 2011, 331: 76-9
- [33] Huarte M, Guttman M, Feldser D, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, 2010, 142: 409-19
- [34] Rinn JL, Kertesz M, Wang JK, et al. Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell*, 2007, 129: 1311-23.
- [35] Tsai MC, Manor O, Wan Y, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, 2010, 329: 689-93
- [36] Maenner S, Müller M, Fröhlich J, et al. ATP-dependent roX RNA remodeling by the helicase maleless enables specific association of MSL proteins. *Mol Cell*, 2013, 51: 174-84
- [37] West JA, Davis CP, Sunwoo H, et al. The long noncoding RNAs *NEATI* and *MALATI* bind active chromatin sites. *Mol Cell*, 2014, 55: 791-802
- [38] Hung T, Wang Y, Lin MF, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet*, 2011, 43: 621-9
- [39] Guil S, Esteller M. RNA-RNA interactions in gene regulation: the coding and noncoding players. *Trends Biochem Sci*, 2015, 40: 248-56
- [40] Yoon J, Abdelmohsen K, Gorospe M. Functional interactions among microRNAs and long noncoding RNAs. *Semin Cell Dev Biol*, 2014, 34: 9-14
- [41] Johnsson P, Lipovich L, Grandér D, et al. Evolutionary conservation of long non-coding RNAs: sequence, structure, function. *Biochim Biophys Acta*, 2014, 1840: 1063-71
- [42] Yoon J, Abdelmohsen K, Srikantan S, et al. LincRNA-p21 suppresses target mRNA translation. *Mol Cell*, 2012, 47: 648-55
- [43] Yoon J, Abdelmohsen K, Kim J, et al. Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat Commun*, 2013, 4: 2939
- [44] Mukherjee N, Corcoran DL, Nusbaum JD, et al. Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol Cell*, 2011, 43: 327-39
- [45] Chiyomaru T, Yamamura S, Fukuhara S, et al. Genistein inhibits prostate cancer cell growth by targeting miR-34a and oncogenic HOTAIR. *PLoS One*, 2013, 8: e70372
- [46] Leucci E, Patella F, Waage J, et al. microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Sci Rep*, 2013, 3: 2535
- [47] Franco-Zorrilla JM, Valli A, Todesco M, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*, 2007, 39: 1033-7
- [48] Cesana M, Cacchiarelli D, Legnini I, et al. A long noncoding RNA controls muscle differentiation by

- functioning as a competing endogenous RNA. *Cell*, 2011, 147: 358-69
- [49] Legnini I, Morlando M, Mangiavacchi A, et al. A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis. *Mol Cell*, 2014, 53: 506-14
- [50] Reese KJ, Lin S, Verona RI, et al. Maintenance of paternal methylation and repression of the imprinted *H19* gene requires MBD3. *PLoS Genet*, 2007, 3: e137
- [51] Kallen AN, Zhou X, Xu J, et al. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell*, 2013, 52: 101-12
- [52] Faghihi MA, Zhang M, Huang J, et al. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol*, 2010, 11: R56
- [53] Franklin JL, Rankin CR, Levy S, et al. Malignant transformation of colonic epithelial cells by acolon-derived long noncoding RNA. *Biochem Biophys Res Commun*, 2013, 440: 99-104
- [54] Guo X, Gao L, Wang Y, et al. Advances in long noncoding RNAs: identification, structure prediction and function annotation. *Brief Funct Genomics*, 2016, 15: 38-46
- [55] Yang L, Chen LL. Competition of RNA splicing: line in or circle up. *Sci China Life Sci*, 2014, 57: 1232-3
- [56] Chen LL. The biogenesis and emerging roles of circular RNAs. *Nat Rev Mol Cell Biol*, 2016, 17: 205-11
- [57] Zhang Y, Zhang XO, Chen T, et al. Circular intronic long noncoding RNAs. *Mol Cell*, 2013, 51: 792-806
- [58] Li Z, Huang C, Bao C, et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol*, 2015, 22: 256-64