

秀丽线虫细胞核内小干扰 RNA 调控基因表达的机制研究

冯雪竹, 光寿红

(中国科学技术大学生命科学学院, 安徽合肥 230027)

摘要:生物体内存在大量的不编码蛋白质序列的非编码 RNA(noncoding RNA, ncRNA), 这些非编码 RNA 广泛参与生命活动的各个过程, 包括基因表达调控、基因组稳定性维持、抵抗外源核酸侵染、发育的时序调节以及肿瘤发生等。越来越多的证据表明一系列重大疾病的发生、发展与这些非编码 RNA 的产生和调控失衡相关。小调节性 RNA 正在成为潜在的疾病标志物、药物靶点和生物分子药物。我们的研究主要集中在高等多细胞生物中细胞核里小干扰 RNA 调控基因表达的分子机制和生物学功能。我们在模式生物秀丽线虫中通过遗传筛选的方法发现了一条小干扰 RNA 在细胞核内调控基因表达的通路, 以及参与这条通路的几个关键的细胞核 RNA 干扰缺陷型基因(nuclear RNAi defective, Nrde)。这一发现不仅解决了高等多细胞生物中细胞核内是否存在小 RNA 干扰现象的争论, 而且发现小 RNA 可能通过主动转运的方式进入细胞核并调控 RNA 聚合酶 II(RNAP II)介导的转录延伸。这一通路还可能参与了生物体的获得性遗传过程。本文重点阐述这一小干扰 RNA 调控基因表达的分子机制, 并提出未来亟待解决的科学问题和发展方向。

关键词:非编码 RNA; 小干扰 RNA; RNA 干扰; 转录; 基因表达调控; 细胞核 RNA 干扰缺陷型

中图分类号: Q75

文献标识码: A

doi:10.3969/j.issn.0253-2778.2013.11.00

引用格式: Feng Xuezhu, Guang Shouhong. Small interfering RNA mediates gene regulation in the nucleus of *Caenorhabditis elegans*[J]. Journal of University of Science and Technology of China, 2013, 43(11):100-105.

冯雪竹, 光寿红. 秀丽线虫细胞核内小干扰 RNA 调控基因表达的机制研究[J]. 中国科学技术大学学报, 2013, 43(11):100-105.

Small interfering RNA mediates gene regulation in the nucleus of *Caenorhabditis elegans*

FENG Xuezhu, GUANG Shouhong

(School of Life Sciences, University of Science and Technology of China, Hefei 230027, China)

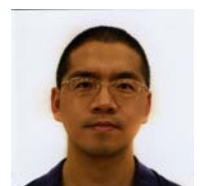
Abstract: Noncoding RNAs (ncRNAs) are widely present in higher eukaryotes and involved in a range of

收稿日期:2012--; **修回日期:**2012--

基金项目:国家重点基础研究发展(973)计划(2011CBA01100), 国家自然科学基金面上项目(31171254), 中国科学技术大学基础研究基金(WK2060190018)资助。

作者简介:冯雪竹, 女, 1975年生, 博士。研究方向: 细胞生物学。E-mail: wormrna@yahoo.com

通讯作者:光寿红, 中国科学技术大学生命科学学院教授, 博士生导师。1991~1996年在中国科学技术大学生物系学习并获得学士学位。1996~1999年师从中国科学技术大学施蕴渝院士进行多肽的溶液构象研究, 并获得硕士学位。1999~2004年, 在美国威斯康星大学癌症研究所师从 Janet Mertz 教授, 研究真核生物中病毒起源的无内含子 RNA 的表达和调控机制, 获得博士学位。2005~2010年, 作为博士后在威斯康星大学遗传系 Scott Kennedy 教授实验室, 从事模式生物中小干扰 RNA 的功能和调控机理的研究。2011年任中国科学技术大学生命科学院教授, 进一步从事非编码 RNA 与基因表达调控的研究, 同年入选中国科学院“百人计划”和中组部“青年千人计划”。主要研究领域包括真核细胞中 RNA 的表达与加工的调节、真核生物中转录调节机制、非编码 RNA 的表达与调节机制、模式生物的遗传与发育。研究成果发表于 Nature 和 Science 等国际期刊上。E-mail: sguang@ustc.edu.cn



biological processes to regulate gene expression, genome integrity, and development. Growing evidence has emerged that ncRNAs are also involved in the generation of diseases. Small regulatory RNAs have shown potential as disease marker, drug targets, and therapeutic drugs. Our research is focused on the molecular mechanism and biological roles of small interfering RNA (siRNA)-regulated gene expression in metazoans. By conducting genetic screenings in the model organism *C. elegans*, a novel nuclear RNAi defective (Nrde) pathway and the critical players have been identified through which siRNA performs gene silencing in the nucleus. This work not only demonstrated that nuclear RNAi machinery exists in metazoans, but also led to several discoveries. For example, we found that small RNAs are sorted between distinct subcellular compartments by associating with particular Argonaute proteins. siRNAs regulate transcription elongation by inhibiting RNA polymerase II and elicit premature termination. Furthermore, it was found that this pathway is necessary for the transgenerational maintenance of acquired RNAi.

Key words: noncoding RNA; siRNA; RNAi; transcription; gene regulation; nuclear RNAi defective (Nrde)

0 引言

在真核生物中,长的双链 RNA (double-stranded RNA, dsRNA)会通过 RNA 干扰(RNA interference, RNAi)的机制来抑制基因表达^[1].双链 RNA 被摄取到细胞内后,首先被保守的 RNase III——类似的核酸酶 Dicer 剪切成 21~23 个核苷酸长度的小干扰 RNA (small interfering RNA, siRNA)^[2-3]. siRNA 结合到保守的 Argonaute 蛋白里,形成 RNA 诱导的干扰复合物(RNA induced silencing complex, RISC). siRNA 通过碱基互补配对的方式将 RISC 招募到具有同源序列的核酸序列上,通过一系列的方式调控基因表达,如抑制蛋白质翻译、降解 mRNA、修饰染色质,以及抑制转录延伸等^[3-8].

RNA 干扰过程在细胞质和细胞核内都可以发生^[3].迄今为止,小调节性 RNA 在细胞质中的作用机理研究得相对比较清楚,其缄默基因表达的手段主要集中在抑制翻译和降解 mRNA 上.如微 RNA (microRNA)在细胞质中结合 mRNA 来抑制蛋白质的翻译和诱导 mRNA 的降解,外源性的小干扰 RNA (siRNA)在进入细胞后也在细胞质内诱导形成 RISC 复合物,从而造成 mRNA 的降解.在细胞核内小调节性 RNA 则有多种作用方式,而且存在明显的物种特异性.例如在裂殖酵母 *S. pombe* 里, microRNA 诱导中性粒区域的异染色质 (heterochromatin)形成;在植物中, microRNA 诱导组蛋白 3 赖氨酸 9 的三甲基化(H3K9me3)和 DNA 的甲基化,从而抑制基因表达;在嗜热型四膜虫 *Tetrahymena thermophila* 内, scanRNA 可以诱导具有同源序列的基因组序列的丢失;在另一中纤毛

虫 *Oxytricha trifallax* 中, piRNA 可以保护具有同源序列的基因组,相反不被保护的基因组序列则被清除;在哺乳动物细胞内, piRNA 可以在生殖细胞的细胞核内诱导 H3K9 三甲基化和 DNA 的甲基化,也有报道认为 microRNA 可以调控启动子区域表达活性,还有报道认为小干扰 RNA 可以调控转录过程中基因的选择性剪切^[8-9].

我们在模式生物秀丽线虫中研究了细胞核内 RNA 干扰的分子机理,我们发现小干扰 RNA 可以通过结合一个 Argonaute 蛋白 NRDE-3 转运到细胞核内,在细胞核内,进一步结合正在被转录的前 mRNA (pre-mRNA),然后次序招募 NRDE-2, NRDE-1, 和 NRDE-4 到 pre-mRNA 上,通过尚未明确的机理,诱导 RNA 干扰位置的 H3K9 的三甲基化和 RNA 聚合酶 II 在转录延伸过程中的暂停,从而抑制转录的进一步延伸,并造成转录的提前终止.这一通路还可以介导 RNA 干扰现象从亲代遗传到子代中^[7,10-14].

1 秀丽线虫中 RNA 干扰的基本原理

双链 RNA 可以通过多种方式进入秀丽线虫体内.如纯化过的 dsRNA 可以被直接显微注射到线虫体内,线虫可以浸泡在含 dsRNA 的溶液中,或者使用转基因技术在线虫体内表达 dsRNA.目前线虫研究常用的方法是喂食法 (feeding RNAi).秀丽线虫通过喂食大肠杆菌 (*E. Coli*)来生长,表达 dsRNA 的质粒可以转化到大肠杆菌体内诱导表达,线虫在喂食细菌的过程中将 dsRNA 摄入体内^[15].

dsRNA 进入细胞后,首先结合一个双链 RNA 结合蛋白 RDE-4. RDE-4 可以结合 Dicer 蛋白,从而将 Dicer 招募到 dsRNA 上,并将 dsRNA 剪切成 21

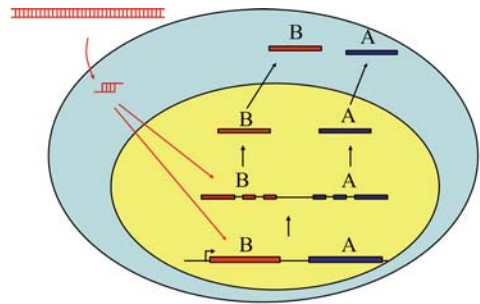
~23 核苷酸长度的双链的 siRNA, 称为初级 siRNA(primary siRNA). 初级 siRNA 接下来结合初级 Argonaute 蛋白 RDE-1, siRNA 中的一条链被丢失. 初级 siRNA 和具有同源序列的 mRNA 通过碱基互补配对. RDE-1 在 siRNA 的指导下, 在 siRNA 第 10 到 11 位碱基对应的 mRNA 的位置将 mRNA 切断^[16].

然而秀丽线虫里特别的地方在于起基因干扰主要作用的是次级 siRNA(secondary siRNA)和次级 Argonaute 蛋白, 这一过程依赖于 RNA 依赖的 RNA 聚合酶(RdRP)和 12 个次级 Argonaute 蛋白^[17-18]. *rrf-1*, *rrf-2*, *ego-1* 和 *rrf-3* 是秀丽线虫中的 4 个 RdRP. *rrf-1* 主要在体细胞中起作用, 对 RNA 干扰是必须的. *rrf-2* 和 *ego-1* 在生殖细胞中起作用, 目前机制并不明确. *rrf-3* 对内源性小 RNA 的产生起关键作用. RDE-1/siRNA 将 RRF-1 招募到 mRNA 上, 通过未知的机理, 诱导形成反义的次级 siRNA. 次级 Argonaute 蛋白会结合次级 siRNA, 然后进一步在细胞质内诱导 mRNA 的降解, 或者转运到细胞核内, 在细胞核内起基因干扰作用或染色体调控功能^[14, 19-20].

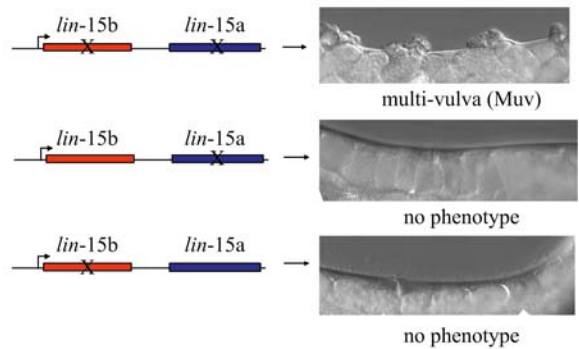
2 细胞核内 RNA 干扰通路(Nrde)的发现

秀丽线虫中细胞核内存在 RNA 干扰的现象最先由 Andy Fire 实验室发现, 他们在 RNA 干扰针对某个细胞核富集的 RNA 后, 通过原位杂交实验发现细胞核内的 RNA 水平降低^[21]. 后来 Michel Labouesse 实验室也发现 RNA 干扰针对 operon 中一个基因的时候同一 operon 中另一个基因的表达受到抑制^[22], 但其中的机理并不明确. 线虫中大概 12% 的基因位于 operon 结构中, 即两个或多个基因转录成同一个 pre-mRNA, 在细胞核内它们会分别剪接成成熟的 mRNA, 然后成熟的 mRNA 分别转运到细胞质里翻译蛋白质^[23-24]. 这样 operon 中不同的基因在细胞核内的时候是处于顺式位置(*cis*), 或者在染色体上相互连接, 或者在 pre-mRNA 上相互连接. 如果 RNA 干扰在细胞核内可以发生, 那么干扰 operon 中一个基因的时候会同干扰另一个基因(图 1(a)).

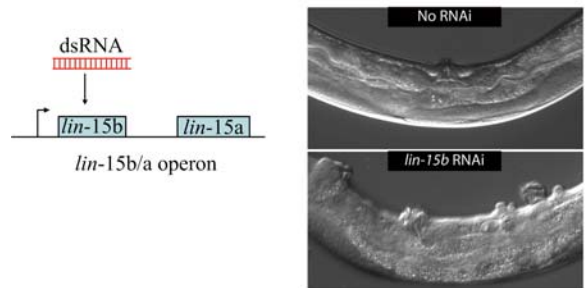
我们首先在 *lir-1/lin-26* 和 *lin-15b/lin-15a* 两个 operon 中, 通过喂食 RNA 干扰观察到了 RNA 干扰其中一个基因导致另一个基因受到抑制的现



(a) operon 结构中, 两个或多个基因转录成同一个 pre-mRNA, 在细胞核内它们会剪接成各自成熟的 mRNA, 然后成熟的 mRNA 分别转运到细胞质里. 如果 RNA 干扰在细胞核内可以发生, 那么干扰 operon 中其中一个基因可以同时造成干扰另一个基因的效果



(b) *lin-15b/lin-15a* 是一个 operon, 如果 *lin-15b* 和 *lin-15a* 同时突变, 线虫表现为多生殖孔表型(multivulva, Muv), 单独突变 *lin-15b* 或者 *lin-15a* 无此性状

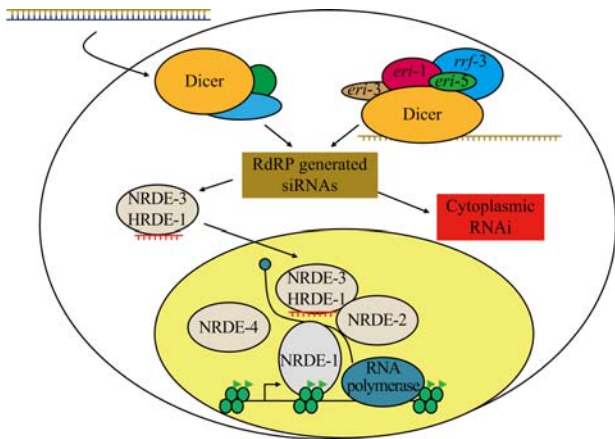


(c) RNAi 干扰 *lin-15b* 导致多生殖孔表型

图 1 细胞核 RNA 干扰通路的发现

象^[14](图 1(b), 1(c)). 从这一现象出发, 我们使用遗传筛选的方式找到了一批秀丽线虫的突变体. 在这些突变体中, 它们只在细胞核内的 RNA 干扰过程存在缺陷, 而细胞质中的 RNA 干扰过程是完全有功能的. 原位杂交实验证实了突变体中细胞核富集的 RNA 不会被 RNA 干扰诱导降解掉. 进一步的, 我们使用实时定量 PCR 证实了在突变体中 pre-mRNA 的表达量不受 RNA 干扰影响. 由此这些突变体被命名为细胞核 RNA 干扰缺陷型突变体(Nrde). 使用遗传图谱鉴定和 DNA 测序的方法, 我们发现了 4 个 *nrde*

基因,分别被命名为 *nrde-1/2/3/4*(图 2)^[10,13-14].



双链 RNA 会被 Dicer 剪切成 siRNA,进一步激活 RNA 依赖的 RNA 聚合酶合成次级 siRNA,一部分次级 siRNA 可能通过细胞质里的 RNA 干扰通路调控基因表达,另一部分次级 siRNA 可以结合 NRDE-3(体细胞)或者 HRDE-1(生殖细胞)转运到细胞核,结合 pre-mRNA,招募 NRDE-2 和 NRDE-1. 在 NRDE-4 的作用下,NRDE-1 还可能结合靶基因的 DNA 序列,造成该位点组蛋白的甲基化和抑制 RNA 聚合酶 II 的转录延伸. 绿色三角表示 H3K9 甲基化

图 2 *Nrde* 通路的基本原理

Fig. 2 A working model of nuclear RNAi in *C. elegans*

3 NRDE-3 将 siRNA 从细胞质转运到细胞核

NRDE-3 是一个次级 Argonaute 蛋白,含有 3 个结构域,分别是核定位序列(nuclear localization signal, NLS),PAZ 结构域,和 PIWI 结构域^[14]. PAZ 结构域是 Argonaute 蛋白结合 siRNA 3'末端的地方,PIWI 结构域在一些 Argonaute 蛋白中起到剪切 mRNA 的作用.在具有剪切活性的 Argonaute 蛋白中,PIWI 结构域含有 3 个保守的氨基酸天冬氨酸-天冬氨酸-组氨酸(DDH).NRDE-3 不具有这 3 个氨基酸,因此推测 NRDE-3 可能不具有剪切活性.NRDE-3 在不结合 siRNA 的时候定位于细胞质,结合了 siRNA 后定位于细胞核.其亚细胞定位依赖于 PAZ 和 NLS 两个结构域.如果 PAZ 结构域突变,NRDE-3 不结合 siRNA,或者 NLS 结构域突变,NRDE-3 都将定位于细胞质.而结合 siRNA 和定位到细胞核对于 NRDE-3 在细胞核 RNA 干扰通路中的作用都是必须的,只有同时结合 siRNA 和定位到细胞核内,转基因表达 NRDE-3 才能恢复 *nrde-3* 突变体的功能.

NRDE-3 结合的 siRNA 是次级 siRNA.在喂食

RNA 干扰过程中,*rrf-1* 对产生次级 siRNA 非常关键.如果 *rrf-1* 突变,NRDE-3 不能结合 siRNA,同时不能转运到细胞核内,体细胞 RNA 干扰不起作用.RRF-1 合成次级 siRNA 需要 mRNA 作为模板,如果某个基因只在某些细胞中特异性地表达,对应的次级 siRNA 也只在这些细胞中合成,从而在喂食 RNA 干扰的过程中,NRDE-3 也只在这些细胞中定位于细胞核内.

NRDE-3 可以结合内源性的 siRNA (endo-siRNA),这些内源性 siRNA 的产生依赖于很多基因,如 *eri-1*,*eri-6*,*eri-7*,*rrf-3*,*mut-2*,*mut-7*,*rrf-1*,以及 *ergo-1* 等,当这些基因突变时,NRDE-3 将定位于细胞质.NRDE-3 结合的次级 siRNA 末端存在特别的化学修饰,如 5'端有三磷酸,3'端是两个羟基,进一步证明这些次级 siRNA 是来源于 RNA 依赖的 RNA 聚合酶合成的产物.

4 NRDE-3 招募 NRDE-2, NRDE-1, 和 NRDE-4,抑制转录延伸

siRNA/NRDE-3 复合物进入细胞核内以后,首先通过碱基配对的方式和 pre-mRNA 结合,然后 NRDE-3 和 NRDE-2 通过蛋白-蛋白相互作用的方式将 NRDE-2 招募到 pre-mRNA 上,接下来是 NRDE-1 和 NRDE-4^[10,13].NRDE-1 也可以结合到 siRNA 靶位的 DNA 序列上.NRDE-1 结合 pre-mRNA 不依赖于 NRDE-4,但是其结合 DNA 依赖于 NRDE-4.和 NRDE-3 可以在细胞质与细胞核间转运不同的是,NRDE-1 和 NRDE-2 一直定位于细胞核中.我们可以使用 RNA 免疫沉淀(RNA immunoprecipitation)结合实时定量 PCR(real time PCR)的技术检测遗传突变体中,单个 *nrde* 基因突变对其他 NRDE 蛋白与靶基因以及靶 DNA 位点相互结合的影响,从而解析整个遗传通路中各个基因的相互依赖关系和作用次序^[10,13].

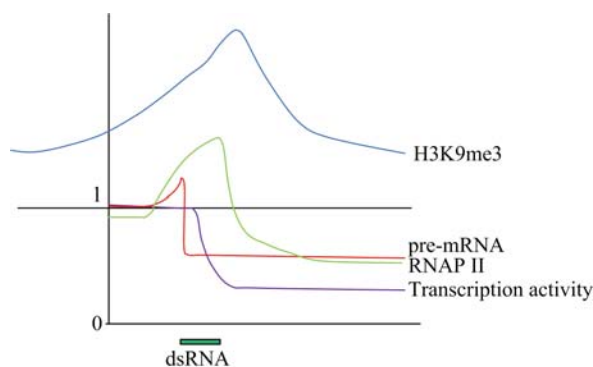
NRDE-2 是一个保守的 RNA 结合蛋白,其在人类细胞中也有单一的同源蛋白^[13].NRDE-2 具有的丝氨酸-精氨酸(SR)结构域表明 NRDE-2 可能参与 pre-mRNA 的加工.SR 结构域广泛存在于与 pre-mRNA 加工和代谢相关的蛋白中,最近的研究表明有些 SR 蛋白也可能直接调控转录过程本身^[25-26].遗传研究证明 NRDE-2 位于 NRDE-3 的下游,因为 NRDE-2 的突变并不影响 NRDE-3 和小干扰 RNA 结合,细胞核定位,以及 NRDE-3 和 pre-

mRNA 的结合, 而 NRDE-3 的突变则会抑制 NRDE-2 和 pre-mRNA 的结合。

NRDE-1 是一个线虫特有的非保守基因, NRDE-1 一直定位于细胞核^[10]。NRDE-1 位于 NRDE-2 和 NRDE-3 的下游。当 NRDE-1 突变时, NRDE-3 和小干扰 RNA 的结合, 亚细胞定位, 以及和 pre-mRNA 的结合不受影响。在 NRDE-2 和 NRDE-3 的作用下, NRDE-1 也结合到小干扰 RNA 针对的 pre-mRNA 序列上。NRDE-1 诱导靶点位置的 H3K9 的三甲基化, 并且其自身也结合到靶点位置的染色体上。当 NRDE-1 突变时, 内源性 siRNA 针对位点的 H3K9 的三甲基化会降低。

NRDE-4 也是一个线虫特有的非保守基因^[10]。NRDE-4 从序列上分析有一个细胞核定位序列。NRDE-4 对 H3K9 甲基化和细胞核 RNA 干扰都是必须的。从 Nrde 通路的次序出发, NRDE-4 位于 NRDE-1 结合 pre-mRNA 与 NRDE-1 结合染色质之间, NRDE-4 的突变不影响 NRDE-1 结合 pre-mRNA, 但影响 NRDE-1 与染色质的结合。

这样当 siRNA/NRDE 复合物结合 pre-mRNA 和 DNA 后会最终造成几个结果(图 3)^[13]: ①对应位置的组蛋白 3 赖氨酸 9 的三甲基化, 当 NRDE 基因突变的时候, 这部分 RNA 干扰造成的 H3K9 甲基化被阻止; ②RNA 聚合酶 II 暂停在 siRNA 针对的模板序列的位置, 而在 RNA 干扰区域的下游, RNA 聚合酶 II 在模板上的结合量会减少, 从而转录延伸无法继续进行下去; ③如果我们使用 nuclear



X 轴表示基因从 5' 到 3',

Y 轴表示 RNA 干扰后相对 RNA 干扰前的变化

图 3 细胞核 RNA 干扰会造成靶基因位点的 H3K9 甲基化, RNA 聚合酶 II 的暂停, 转录延伸活性的下降, 以及下游区域 pre-mRNA 的降低

Fig. 3 Nuclear RNAi induces multi-facet change at the targeted locus

run on (NRO) 实验来直接测量转录活性, 我们可以观察到基因中 RNA 干扰区域的上游的转录活性不受影响, 而在干扰区域的下游转录活性被抑制; ④对终产物 pre-mRNA 来说, RNA 干扰区域的上游的量不受影响, 而 RNA 干扰下游的量会下降。从这些结果出发, 我们推测 RNA 干扰可以抑制转录延伸过程的进行, 从而造成转录的提前终止。

5 NRDE 介导的 RNA 干扰的遗传

早先的研究表明, RNA 干扰现象可以从被干扰的亲代遗传到子代, 即获得性遗传^[1, 27]。线虫试验中发现 RNA 干扰的遗传过程甚至可以在线虫子代中传递到第 80 代。这一现象依赖于参与 RNA 干扰过程的各个因子。

Nrde 通路参与了这一 RNA 干扰的遗传过程^[11-12, 19]。在 *nrde* 突变体中, RNA 干扰现象只存在于亲代, 而不会传递到子代中去。亲代线虫被 RNA 干扰以后会积累小干扰 RNA, 同时也会形成 H3K9 甲基化。二者都被传递到下一代。在子代个体中, NRDE-3 结合 RNA 干扰诱导形成的小干扰 RNA。通过染色质免疫沉淀的方法, 我们发现 RNA 干扰造成的小干扰 RNA 和 H3K9 甲基化都可以至少遗传 3~4 代。然而我们并不清楚 H3K9 甲基化在 RNA 遗传中起着多大的作用, 也不清楚子代中的 H3K9 甲基化是来源于亲代中已被甲基化的模板 DNA 的指导, 还是在子代的小干扰 RNA 指导下重新形成的。有报道表明某些 H3K9 甲基转移酶, 如 *set-25/set-32* 等, 可能影响线虫中 H3K9 甲基化, 然而它们在 RNA 干扰的建立以及遗传中到底起着什么样的作用尚不明确^[28]。

6 HRDE-1 在生殖细胞中介导 RNA 干扰的遗传

遗传筛选发现在生殖细胞中特异表达的另一个定位于细胞核的 Argonaute 蛋白, *heritable RNAi deficient-1* (HRDE-1), 参与了 RNA 干扰的多代遗传过程^[19, 28]。在野生型和突变体线虫中, RNA 干扰在同一代都能有效地抑制基因表达。在野生型线虫的子代中, 基因干扰还能继续进行, 而突变体的子代线虫里的基因干扰被取消, 被干扰的基因重新恢复表达。在子代中可以检测到 HRDE-1 和小干扰 RNA 相互结合。试验也表明 HRDE-1 在子代中起记忆 RNA 干扰的作用, 而对 RNA 干扰信号从亲代传递到

子代本身是不必要的. HRDE-1 位于 NRDE-2 的上游. 生殖细胞中 NRDE-2 和靶标 pre-mRNA 的结合依赖于 HRDE-1. 生殖细胞中 RNA 干扰诱导的 H3K9 甲基化的遗传也依赖于 HRDE-1. 在 *hrde-1* 突变体中, 亲代和子代的 H3K9 甲基化都会被抑制.

7 Nrde 通路的生物学功能

nrde-1, *nrde-2*, *nrde-4* 和 *hrde-1* 在线虫生殖系统发育中起着重要作用. 当这些基因突变时, 它们子代的数目比野生型线虫减少^[19]. 在正常生理条件下, HRDE-1 可以和内源性的小干扰 RNA (22G-RNA, 长度为 22 个核苷酸, 5' 末端为鸟嘌呤) 结合, 来调控生殖细胞中的基因表达. 这些 22G-RNA 在生殖细胞中调控大约 1 500 个基因、假基因以及奇异染色质位点. H3K9 甲基化的染色质免疫沉淀和深度测序也发现 Nrde 通路调控线虫基因组里大约 300 个基因左右, 这些基因主要在生殖细胞中表达. 如果 *glp-4* 基因突变, 线虫缺失生殖细胞, 对应基因的 H3K9 甲基化会降低.

有意思的是, 在 *nrde-1*, *nrde-2*, *nrde-4* 和 *hrde-1* 突变体线虫中, 内源性小 RNA 介导的 H3K9 甲基化会逐代消失, 靶点位置的基因表达逐代增加 (*Mort* 性状). 同时, 突变体线虫的配子形成会逐代恶化, 直至不育. 在后期子代中, 染色体的配对和分离过程出现异常, 雄性子代的比例也会增加.

8 piRNA 和基因组监控

piRNA 对生殖发育与基因组的完整性维持是非常关键的. 在秀丽线虫中, piRNA 的长度是 21 核苷酸, 5' 末端是尿嘧啶, 因此也被称为 21U-RNA. piRNA 主要从染色体 IV 上的一个特别区域产生, 其他染色体也会产生少量 piRNA. PRG-1 和 PRG-2 对 piRNA 的产生非常关键, 它们都特异地表达在生殖细胞中. piRNA 可以和靶基因非完美配对, 招募 RdRP 来合成 22G-RNA^[28-31]. piRNA 通常缄默转座子的起始和终止区域. 在 *prg-1* 突变体里, 转座子的活性增加. piRNA 对起始 RNAi 的多代遗传也是非常重要的, 其介导的生殖细胞中转基因的表达抑制可多达 20 代. 然而 piRNA 对 RNAi 状态的维持反而是不必要的. 一旦 RNA 干扰状态建立, 其维持将依赖于 Nrde 通路以及染色质修饰因子^[28], 包括 *hrde-1*, *nrde-1*, *nrde-2*, *nrde-4*, *hpl-2* (异染色质结合蛋白), 以及 *set-25* 和 *set-32*. 这部分的分子机制尚不明确.

piRNA 通路被认为主要是通过比较生殖细胞中外源的 RNA 序列和自身基因组序列的差别来识别自我和非我. 非我序列激活 piRNA 通路而被抑制, 从而保护宿主本身基因组的完整性.

9 CSR-1 和染色体分离

CSR-1 也是秀丽线虫中的一个次级 Argonaute 蛋白, 它在染色体的分离过程中起着关键的作用^[18,20]. CSR-1 和其辅助因子 EGO-1 (生殖细胞特异的 RdRP), DRH-3 (Dicer 相关的解旋酶) 和 EKL-1 (Tudor 结构域蛋白) 都位于生殖细胞的染色体上, 其在染色体上的结合位点依赖于 CSR-1 结合的内源性 siRNA. siRNA/CSR-1 复合物结合到染色体上后, 并不抑制靶基因的表达. 相反, 这些 CSR-1 结合的位点往往编码蛋白质, 在生殖细胞中表达. 这些位点也不结合 CENP-A. CENP-A 是组蛋白 H3 的变体, 是中心粒区域染色质的分子标记. 因而推测 siRNA/CSR-1 可能通过标记染色质的不同区域来监控染色体分离. 当 CSR-1 缺失的时候, 染色体不能正确排列在赤道板上, 动粒也不能正确定位到纺锤体的相反方向.

CDE-1 是一个生殖细胞特异表达的核苷转移酶, 对线虫中染色体的正确分离也是必需的^[32]. CDE-1 特异结合在有丝分裂的染色体上, 这一结合依赖于 EGO-1 和 CSR-1. 其中 CDE-1 和 EGO-1 存在蛋白-蛋白相互作用. CDE-1 主要作用可能在于尿嘧啶化 CSR-1 结合的 siRNA, 从而避免这一类 siRNA 的量无限增加.

10 Dicer 和细胞凋亡

最近的研究表明, 小干扰 RNA 在 DNA 损伤反应中也起着重要作用. 如在拟南芥中双链 DNA 断裂 (double strand break, DSB) 可以诱导 21 核苷酸长度的小 RNA 的产生^[33]. 这些小 RNA 产生于 DSB 附近, 并结合 AGO2 来参与 DNA 的修复. 类似的现象在其他物种中也被发现^[34].

在秀丽线虫中, DCR-1 参与了细胞的凋亡过程^[35]. DCR-1 突变会抑制细胞的凋亡和凋亡过程中染色体的碎片化. 与凋亡相关的蛋白酶 CED-3 剪切 DCR-1 而产生一个具有脱氧核糖核酸酶活性的 C-末端片断, 这一片断进一步剪切 DNA 造成 DNA 断裂而导致凋亡. 有意思的是, 这一过程可能并不依赖于 Nrde 通路, 其他与 RNA 干扰相关的突变体中并没有观测到细胞凋亡的改变.

11 展望

尽管过去几年里我们对秀丽线虫细胞核内小干扰 RNA 如何调控基因表达有了些初步的了解,然而这里面还存在大量的问题有待研究. 例如我们并不清楚细胞核 RNA 干扰造成 H3K9 甲基化和 RNA 聚合酶 II 的转录延伸暂停具体的分子机制是什么,甚至我们并不清楚这些甲基化和转录延伸暂停是 RNA 干扰现象的直接原因还是干扰以后的副产品.

秀丽线虫中有 38 个组蛋白甲基转移酶的类型基因,它们都含有 SET 结构域,这些基因各自的功能并不清楚. 我们在早先 Nrde 通路的遗传筛选过程中发现的 *nrde* 基因都不含有 SET 结构域. 尽管最近的研究表明 *set-25* 和 *set-32* 可能在 RNA 干扰的遗传中起作用,其机制并不清楚,而且尚不清楚其与 Nrde 通路的关系^[28]. 另一个有意思的地方是细胞核 RNA 干扰中, RNA 聚合酶 II、转录活性、以及终产物 pre-mRNA 都存在着非对称分布,而 H3K9 甲基化在 RNA 干扰的上下游区域呈现对称分布(图 3). 我们推测 H3K9 甲基化可能使靶基因位置被标记,从而间接对 RNA 聚合酶 II 的转录起调控作用.

我们不清楚抑制转录过程本身是 RNA 干扰的原因还是结果. 在遗传筛选的过程中,我们没有获得与转录相关的基因的突变体. 我们也没有检测到 NRDE 蛋白和 RNA 聚合酶 II 的直接蛋白-蛋白相互作用,以及没有检测到 RNA 聚合酶 II 的磷酸化修饰的改变. NRDE/siRNA 复合物结合到 pre-mRNA 上后,是如何让延伸中的 RNA 聚合酶 II 暂停的,还存在着大量的问题需要阐明.

在 RNA 干扰的遗传过程中真正被传递的信号也不清楚. RNA 干扰被遗传的过程包括几个步骤: ①亲代中可遗传状态的建立; ②干扰信号通过配子从亲代传递到子代; ③在子代中维持或者重建基因沉默状态. 虽然有很多证据表明 RNAi 和小干扰 RNA 参与了基因沉默里最先的建立过程和在于代中最终的维持过程,我们还是不清楚从亲代到子代实际传递的信号是什么. 小干扰 RNA 显然是其中的一个候选者^[19].

秀丽线虫中有 12 个次级 Argonaute 蛋白,小干扰 RNA 是怎样在这些 Argonaute 蛋白之间以及它们介导的通路中进行分配的也不清楚. 小 RNA 的长度、末端修饰,以及组织特异性的表达,可能决定了它们和那些 Argonaute 蛋白相互结合来进一步调

控基因表达或者染色体的分离.

小 RNA 如何参与染色体的配对分离以及损伤反应的机制也不清楚,一个可能的机制是小 RNA 通过 Argonaute 蛋白将 DNA 修复因子招募到染色质的断裂位点来进行 DNA 修复. 然而, Dicer 蛋白存在独立于 RNAi 通路外的其他功能也意味着问题的复杂性^[35].

RNA 干扰现象的发现给生命科学的研究和临床医学带来了革命性的变化,加深了我们对遗传和进化过程的理解. RNAi 作为一个基因表达调控机制不仅可以保护物种基因组的完整性,而且能够精确地控制染色体的功能. 因而进一步研究 RNA 干扰的分子机制,特别是小干扰 RNA 与染色质修饰的相互关系,将给我们对真核基因组表观遗传调控的研究带来新的启示.

参考文献 (References)

- [1] Fire A, Xu S Q, Montgomery M K, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*[J]. *Nature*, 1998, 391(6669): 806-811.
- [2] Hannon G J, Rivas F V, Murchison E P, et al. The expanding universe of noncoding RNAs [C]// Cold Spring Harbor symposia on quantitative biology. Cold Spring Harbor Laboratory Press, 2006, 71:551-564.
- [3] Ketting R F. The many faces of RNAi [J]. *Developmental cell*, 2011, 20(2):148-161.
- [4] Bernstein E, Allis C D. RNA meets chromatin[J]. *Genes & development*, 2005, 19(14):1 635-1 655.
- [5] Carthew R W, Sontheimer E J. Origins and mechanisms of miRNAs and siRNAs[J]. *Cell*, 2009, 136(4):642-655.
- [6] Czech B, Hannon G J. Small RNA sorting: Matchmaking for Argonautes[J]. *Nat Rev Genet*, 2011, 12(1):19-31.
- [7] Feng X Z, Guang S H. Small RNAs, RNAi and the inheritance of gene silencing in *Caenorhabditis elegans* [J]. *Journal of genetics and genomics*, 2013, 40(4): 153-160.
- [8] Castel S E, Martienssen R A. RNA interference in the nucleus: Roles for small RNAs in transcription, epigenetics and beyond[J]. *Nature reviews Genetics*, 2013, 14(2):100-112.
- [9] Matzke M A, Birchler J A. RNAi-mediated pathways in the nucleus[J]. *Nature reviews Genetics*, 2005, 6(1):24-35.
- [10] Burkhart K B, Guang S H, Buckley B A, et al. A pre-mRNA-associating factor links endogenous siRNAs to

- chromatin regulation [J]. *PLoS Genetics* 2011, 7(8):e1002249.
- [11] Burton N O, Burkhart K B, Kennedy S. Nuclear RNAi maintains heritable gene silencing in *Caenorhabditis elegans* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(49):19 683-19 688.
- [12] Gu S G, Pak J, Guang S H, et al. Amplification of siRNA in *Caenorhabditis elegans* generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint [J]. *Nature genetics*, 2012, 44(2):157-164.
- [13] Guang S H, Bochner A F, Burkhart K B, et al. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription[J]. *Nature*, 2010, 465(7301):1 097-1 101.
- [14] Guang S H, Bochner A F, Pavelec D M, et al. An Argonaute transports siRNAs from the cytoplasm to the nucleus[J]. *Science*, 2008, 321(5888):537-541.
- [15] Timmons L, Court D L, Fire A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans* [J]. *Gene*, 2001, 263(1/2):103-112.
- [16] Mello C C, Conte D, Jr. Revealing the world of RNA interference[J]. *Nature*, 2004, 431(7006):338-342.
- [17] Pak J, Maniar J M, Mello C C, et al. Protection from feed-forward amplification in an amplified RNAi mechanism[J]. *Cell*, 2012, 151(4):885-899.
- [18] Yigit E, Batista P J, Bei Y, et al. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi [J]. *Cell*, 2006, 127(4):747-757.
- [19] Buckley B A, Burkhart K B, Gu S G, et al. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality [J]. *Nature*, 2012, 489: 447-451.
- [20] Claycomb J M, Batista P J, Pang K M, et al. The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation[J]. *Cell*, 2009, 139(1):123-134.
- [21] Montgomery M K, Xu S, Fire A. RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95(26):15 502-15 507.
- [22] Boshier J M, Dufourcq P, Sookhareea S, et al. RNA interference can target pre-mRNA: Consequences for gene expression in a *Caenorhabditis elegans* operon [J]. *Genetics*, 1999, 153(3):1 245-1 256.
- [23] Reinke V, Cutter A D. Germline expression influences operon organization in the *Caenorhabditis elegans* genome[J]. *Genetics*, 2009, 181(4):1 219-1 228.
- [24] Chen N, Stein L D. Conservation and functional significance of gene topology in the genome of *Caenorhabditis elegans*[J]. *Genome research*, 2006, 16(5):606-617.
- [25] Pandit S, Zhou Y, Shiue L, et al. Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing[J]. *Molecular cell*, 2013, 50(2): 223-235.
- [26] Ji X, Zhou Y, Pandit S, et al. SR Proteins Collaborate with 7SK and Promoter-Associated Nascent RNA to Release Paused Polymerase[J]. *Cell*, 2013, 153(4): 855-868.
- [27] Vastenhouw N L, Brunschwig K, Okihara K L, et al. Gene expression: Long-term gene silencing by RNAi [J]. *Nature*, 2006, 442(7105):882.
- [28] Ashe A, Sapetschnig A, Weick E M, et al. piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*[J]. *Cell*, 2012, 150(1):88-99.
- [29] Shirayama M, Seth M, Lee H C, et al. piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline[J]. *Cell*, 2012, 150(1):65-77.
- [30] Lee H C, Gu W, Shirayama M, et al. *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts[J]. *Cell*, 2012, 150(1):78-87.
- [31] Bagijn M P, Goldstein L D, Sapetschnig A, et al. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs [J]. *Science*, 2012, 337 (6094): 574-578.
- [32] van Wolfswinkel J C, Claycomb J M, Batista P J, et al. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs[J]. *Cell*, 2009, 139(1):135-148.
- [33] Wei W, Ba Z Q, Gao M, et al. A role for small RNAs in DNA double-strand break repair[J]. *Cell*, 2012, 149(1):101-112.
- [34] Francia S, Michelini F, Saxena A, et al. Site-specific DICER and DROSHA RNA products control the DNA-damage response [J]. *Nature*, 2012, 488 (7410): 231-235.
- [35] Nakagawa A, Shi Y, Kage-Nakadai E, et al. Caspase-dependent conversion of Dicer ribonuclease into a death-promoting deoxyribonuclease[J]. *Science*, 2010, 328(5976):327-334.