2016年 第46卷 第1期:52~65

SCIENTIA SINICA Vitae

life.scichina.com



评 述

中国知名大学及研究院所专栏 中国科学院动物研究所专辑

果蝇卵巢中的 piRNA 发生

王海龙[†], 王晓娜[†], 陈大华^{*}

中国科学院动物研究所, 干细胞与生殖生物学国家重点实验室, 北京 100101 † 同等贡献

* 联系人, E-mail: chendh@ioz.ac.cn

收稿日期: 2015-07-23; 接受日期: 2015-08-20 国家自然科学基金(批准号: 2013CB945002, 31130036, XDA01010306)资助

摘要 piRNA(Piwi-interacting RNA)是一类长度约为 24~32 nt(nucleotides)的非编码 RNA. 在所有的非编码 RNA 中, piRNA 数量最多, 主要存在于生殖系统. piRNA 和 Argonaute 家族的 PIWI 蛋白特异结合并相互作用, piRNA 的生成及其功能依赖于 PIWI 蛋白. piRNA 序列来源于基因组的转座子、编码基因区和特异的基因间位点 (intergenic loci). piRNA 作为 PIWI 的向导调控靶基因的表达以及转录和转录后水平的修饰. 在果蝇卵巢生殖细胞 中, piRNA 簇的转录子转运到胞浆后经过初级加工途径形成初级 piRNA, 结合到 Piwi 和 Aub 上. 生殖细胞的初 级 piRNA 还会进入次级加工途径, 经 PIWI 家族蛋白的协同、加工, 使细胞中的 piRNA 大量扩增, 称为乒乓循环 ("Ping-Pong" cycle), 但是目前对乒乓循环的细节和具体机制还所知甚少. 本文以果蝇卵巢为例主要对 piRNA 的 生成、piRNA 的初级加工途径和"乒乓循环", 以及 piRNA 的转录沉默等方面的研究进展进行简要综述.

关键词 piRNA, PIWI, Argonaute, 乒乓循环, 转录沉默

小RNA参与的RNA沉默(RNA silencing)以时空 特异性的方式调节基因表达,在控制生物体的发育 和稳态中发挥着至关重要的作用.小RNA最初在 1993 年发现,在线虫(*Caenorhabditis elegans*)中小 RNA lin-4参与*lin14*基因的转录后调节,能够靶向其 mRNA^[1,2].后来发现更多种类的小RNA 广泛存在于 多个物种中,如植物、动物,甚至细菌.这些小RNA 结合到一类特定的蛋白——Argonaute 蛋白上,组装 成RNA沉默复合体(RNA inference silence complex, RISC),靶向到与小RNA互补的RNA或DNA上,调 节这些RNA或DNA的功能. 在动物体中,小 RNA 依其结构特征、来源、加 工因子和结合的 Argonaute 蛋白的不同分为 3 种,即 microRNA(miRNA), siRNA(small interference RNA) 和 piRNA(Piwi-interacting RNA). 这 3 种小 RNA 的 5' 端均含有一个单磷酸集团, siRNA 和 piRNA 的 3'端 2-OH 被甲基化,而 miRNA 没有甲基化. miRNA 的基 因组上特定的 miRNA 基因转录加工而成, siRNA 由 内源表达或外源导入的双链 RNA 加工而来. miRNA 和 siRNA 的加工都需要 RNaseIII蛋白 Dicer,而 piRNA 由基因组序列转录加工而来,且不需要 Dicer 的剪切^[3-7]. Argonaute 家族分为 Argonaute 亚家族和

引用格式: 王海龙, 王晓娜, 陈大华. 果蝇卵巢中的 piRNA 发生. 中国科学: 生命科学, 2016, 46: 52–65 Wang H L, Wang X N, Chen D H. piRNA biogenesis in *Drosophila* ovary. Sci Sin Vitae, 2016, 46: 52–65, doi: 10.1360/N052015-00224

www.scichina.com

PIWI 亚家族, 成熟的 miRNA 和 siRNA 结合到 Argonaute 亚家族蛋白上, 而 piRNA 结合到 PIWI 亚 家族蛋白上^[8]. 在果蝇(*Drosophila melanogaster*)中, 有 3 个 PIWI 亚家族成员, Piwi, Aub 和 AGO3, Piwi 蛋白在体细胞和生殖细胞中均表达, 主要定位在核 中, 其结合的 piRNA 最长, 平均 25.7 nt. Aub 与 AGO3 在生殖细胞中表达, 定位在胞浆, 它们结合的 piRNA 稍短, 平均为 24.7 和 24.1 nt^[3-5,9].

piRNA 序列来源于基因组的基因间重复序列区 域,且成簇分布,这些基因组位点称为 piRNA 簇 (piRNA cluster)^[9]. 在这些 piRNA 簇区域, 通常含有 大量的转座子及其残余序列. piRNA 簇有单向转录和 双向转录两类,前者从一端转录形成一条转录子,后 者从两端转录形成两条互补的转录子.这些转录产 物转运到胞浆后经过 piRNA 初级加工途径(primary processing pathway)形成 piRNA, 称为初级 piRNA (primary piRNA); 在生殖细胞中, 初级 piRNA 还要 进入次级加工途径(secondary processing pathway,也 称乒乓循环(ping-pong cycle)或扩增环(amplification loop), 使 piRNA 得到大量的扩增, 形成次级 piRNA(secondary piRNA). 由于这些 piRNA 主要来 源于转座子及重复序列, 其中对应于转座子序列反 义链的 piRNA 能够通过碱基互补识别和结合转座子 mRNA, 引起 RNA 干扰效应, 造成转座子在转录后 或转录水平上的沉默,因此 piRNA 在生殖系统中能 够控制转座子的表达,保护基因组免受转座造成的 DNA 损伤, 在维持基因组稳定性和保持物种代间遗 传稳定上发挥着极为重要的功能^[9~16]. 当参与 piRNA 发生的基因突变时,细胞中的转座子 mRNA 水平大 大升高, 基因组上转座事件的发生更加频繁, 引起 DNA 损伤, 激活 Chk2 DNA 损伤检验点(DNA damage checkpoint),导致生殖系统发育缺陷,如微管 组织和卵母细胞轴向发育缺陷,从而造成不育[10,17].

piRNA不同于 miRNA 和 siRNA 的序列特征和产 生方式,以及 piRNA 在动物发育中的重要作用,引 发研究者去试图揭示 piRNA 的生成和功能机制.现 在已经发现了很多蛋白参与 piRNA 簇的转录、加工 以及成熟 piRNA 对转座子表达的沉默过程,然而对 这些蛋白的作用机制知之甚少,亟需更深入的研究.

1 piRNA 簇及其转录

piRNA 簇广泛而不均一地分布在基因组上,长

度不等,大的可以达到数十万个碱基.除了少量的 piRNA 簇在常染色质区外,绝大部分在近着丝粒和 端粒区的异染色质区^[9,18].近着丝粒区的 piRNA 簇包 含有大量的转座子序列,而端粒区的 piRNA 簇主要 由大量的卫星序列组成,对应于亚端粒末端重复序 列(subtelomeric terminal associated sequence (TAS) repeats)^[9].根据 piRNA 对应于基因组链的分布情况, 可将 piRNA 簇分为两种:(i)转座子序列通常是同 一方向排列,产生的 piRNA 与这些转座子序列互补, 这种 piRNA 簇从其一端开始转录形成转录子后被加 工成 piRNA;(ii)转座子序列两种方向的均有分布, 产生的 piRNA 也对应于两条链,这种 piRNA 簇能够 从两端转录,形成互补的正义和反义转录子,加工后 形成 piRNA^[9].

1.1 单向转录的 piRNA 簇

在果蝇卵巢中, 最主要的单向转录的 piRNA 簇 是 X 染色体上的 flamenco 和 20A 两个位点(图 1)^[9]. flamenco 位点是沉默 gypsy 转座子所必需的, 在 flamenco 基因组位点缺失的果蝇中, gypsy 转座子在 卵巢体细胞中大量表达,并能够包装成病毒样颗粒, 感染临近的生殖细胞,从而引起果蝇雌性不育^[19~21]. 最近几年的研究将 flamenco 位点鉴定为一个 piRNA 簇, 位于 X 染色体真染色质区与近着丝粒异染色质 区之间, 跨度约 180 kb, 其上游靠近 DIP(Dipeptidase A)基因. 在 flamenco 区域, 含有大量的 gypsy, ZAM, Idefix 3 种转座子的序列, 而且这些序列在排列上都 是与该区域的转录方向相反的. 该区域产生了卵巢 中靶向 ZAM 的 piRNA 中的 79%, Idefix 的 30%, gypsy 的 33%^[9,22,23]. flamenco 只在卵巢体细胞中表达, 而 20A 簇在体细胞和生殖细胞中均有表达. 他们都由 RNA 聚合酶 II 转录, 转录产物也具有 5'端帽结构, 3' 端 AATAAA 剪切序列和 poly(A)(poly-adenylation (poly(A)))的保守序列^[18,24]. flamenco 启动子不含有典 型的 TATA 盒(TATA box), 其转录起始位点包含保 守的 Inr 序列, 下游含有 DPE(downstream promoter element)位点,上游还含有一个典型的 Ci(cubitus interruptus)结合基序, 通过与 Ci 结合调节其转录^[24]. 在 20A 簇的基因组位点还发现了典型的 RNA 聚合酶 II转录的组蛋白标签 H3K4me2 的富集^[18].

除了基因间区域的单向 piRNA 簇外,还有一类 定位于编码基因的非编码区.对果蝇卵巢体细胞系



图 1 piRNA 簇及转录

A:单向 piRNA 簇. piRNA 簇一端具有启动子,有 RNA 聚合酶 II 转录,该区域含有保守的转录起始序列(Inr),下游有 DPE 位点,上游有转录 因子(如 Ci)结合位点.转录产物 5′端有帽结构,3′端有 poly(A)尾,内含子被剪切掉;B:双向 piRNA 簇.该类 piRNA 簇两端有基因,由这些基 因转录时 RNA 聚合酶 II 未终止而延伸到 piRNA 簇区域.piRNA 簇区域基因组还有高的 H3K9me 水平,后者被 Rhi 结合,招募 Del 和 Cuff,保证 RNA 聚合酶 II 转录不会终止.基因区 mRNA 被剪切后,Cuff 结合 piRNA 簇转录子 5′端,保护其不被降解.新生的 piRNA 簇转录子 5′端没有帽结构,3′端没有 poly(A)尾,UAP56 结合其上保证内含子样序列仍然保留

OSC(ovary somatic cell)中的 piRNA 分析发现,来源 于转座子序列的 piRNA 最多,是来源于基因非编码 区的 piRNA 的 4 倍;在来源于基因非编码区的 piRNA 中,来源于 3'非翻译区(3' untranslated region, 3'UTR)区的 piRNA 超过 70%,小部分来源于 5'UTR, 只有少量来源于编码区^[25].这些 piRNA 结合在 Piwi 上,都对应于转录子的正义链,5'碱基一般是 U^[26]. 果蝇中,产生这类 piRNA 最丰富的基因为 *tj*(traffic jam)和 *brat*(brain tumor)^[25,26]. tj-piRNA 结合到 Piwi 后,通过序列互补靶向一些编码基因,如*fasIII*,调节 其表达^[26].这类 piRNA 的丰度与其来源基因 mRNA 的丰度并没有相关性,通过初级 piRNA 途径生成, 因此现在还不清楚这类 piRNA 的转录是不是也依赖 于 RNA 聚合酶 II 对该编码基因的转录.

1.2 双向转录的 piRNA 簇

双向转录的 piRNA 簇能够在其两端转录,形成 两条互补的转录子(图 1). 这种簇中的转座子的排列 比较随机,有些是同向的,有些是反向的,因此它们 转录出来的产物中既含有转座子序列的正义链,也 含有其反义链. 这种转录子将参与胞浆的乒乓扩增 过程,是生殖细胞中产生 piRNA 的主要方式.典型 的双向转录的 piRNA 簇是位于第二号染色体近着丝 粒区域的 42AB 区,该区域产生了卵巢中总 piRNA 的 30%^[9].

双向转录的 piRNA 簇只在生殖细胞中表达,一般没有明确的启动子,没有组蛋白标签 H3K4me2 的富集,没有转录终止信号,转录产物5′端没有帽结构. 其转录一般依赖于 RNA 聚合酶 II 转录临近基因时越过终止信号而继续转录合成 RNA,也有少数依赖于簇两端的非经典的转录起始,但具体机制不明^[18].而且这些 piRNA 簇的转录产物没有经过剪切步骤,形成了含有内含子样序列的长 RNA^[27].

目前已经鉴定出了一些参与调节 piRNA 簇转录的因子(图 1). Egg(eggless)能调节单向和双向 piRNA 簇的转录. Egg 是一种组蛋白甲基转移酶,能够甲基 化组蛋白 H3 第 9 位赖氨酸. H3K9 甲基化主要与转录 抑制相关,在一些区域的 H3K9 甲基化是基因组异染 色质化所必需的. egg 的突变体中,异染色质状态发 生改变,从而影响了 piRNA 簇的表达^[28]. Rhino(Rhi), Deadlock(Del)和 Cutoff(Cuff)形成一个复合物,控制 双向转录的 piRNA 簇的表达. Rhi 是果蝇 HP1 家族蛋

白,同 HP1a 一样能够通过其 chromo 结构域结合到 H3K9me3 位点上[11,18,27]. 尽管 piRNA 簇与其他异染 色质区都有 H3K9me3 标志, 但 Rhi 的结合似乎是区 分它们的一个关键因子, Rhi 的结合与 piRNA 簇的转 录和 piRNA 的产生高度相关,将其结合到非 piRNA 簇区域甚至引起该区域的 piRNA 簇样转录和 piRNA 产生^[18,27]. Cuff 是酵母 Rai1 或后生动物 Dom3Z 蛋白 的同源蛋白. Rail/Dom3Z 具有核酸外切酶和焦磷酸 水解酶活性,能够结合到新转录出的 mRNA 的 5'端, 降解具有不完整帽结构的 RNA, 从而参与 mRNA 帽 结构的质量控制. Cuff 蛋白的酶活位点缺失, 而且 piRNA簇5′端不含帽结构,很可能只是通过结合到新 生 piRNA 簇转录子的 5'端, 起保护作用^[18]. Del 蛋白 没有典型的已知结构域, 其 N 端结合到 Rhi 蛋白的 chromo 结构上, C 端结合 Cuff 蛋白, 桥接形成 Rhi-Del-Cuff(RDC)复合物, 通过 Rhi 的 Chromo 结构 域与染色质上H3K9me3的结合而结合到piRNA 簇区 域^[18]. 这 3 个蛋白在染色质 piRNA 簇区域有很强的 富集,而且是相互依赖的,任何一个蛋白的缺失都造 成其他两个蛋白不能聚集到 piRNA 簇位点, 从而造 成 piRNA 簇转录的降低和 piRNA 形成的急剧减 少^[11,18,27]. 由于 piRNA 簇缺少典型的启动子, 通常是 RNA 聚合酶 II 转录临近基因时,跨越和忽视该基因 的转录终止信号后接着转录 piRNA 簇的, RDC 复合 物很可能是在阻止 RNA 聚合酶 II 转录终止的过程中 起关键作用^[18]. 双向 piRNA 簇转录的产物的另一个 重要特征是没有剪切, 而仍然保留有内含子样序列, 这一过程很可能是通过 UAP56 蛋白结合到 Rhi 和 Cuff上起作用的^[27]. UAP56 是一个广谱表达的 DEAD 结构域的蛋白,参与 RNA 剪切和外运. 在果蝇生殖 细胞中, UAP56 表达模式与 Rhi 蛋白相同, 定位到染 色质 piRNA 簇位点,可能通过其 DEAD 结构域结合 到新生的 RNA 上,与同样结合在 5′端的 Cuff 协同作 用,抑制 piRNA 簇转录子的剪切^[27,29]. 但是到底是如 何抑制和为什么需要抑制剪切还不是很清楚. 另外 Piwi 蛋白对一些双向转录的 piRNA 簇的表达也是必 需的,很可能是通过影响Egg蛋白在这些位点的定位, 进而影响 H3K9me3 的形成和 Rhi 的结合而发挥作 用的.

1.3 piRNA 的加工与形成

piRNA 簇的转录产物被转运到胞浆, 经过多步

的加工,形成 23~30 nt 的成熟 piRNA,结合到 PIWI 家族蛋白上.在果蝇卵巢体细胞中,单向转录的转录 子在胞浆通过初级加工途径形成 piRNA,结合到 Piwi蛋白上.而在生殖细胞中,单向和双向转录出的 转录子在胞浆也经过初级加工途径,形成的 piRNA 结合到 Piwi和 Aub蛋白上.双向转录子及其产生的 初级 piRNA 还要进一步加工,使 piRNA 得到大量扩 增,这个过程称为次级加工过程(图 2).

2 初级加工途径

piRNA 簇转录子转运到胞浆后,经过潜在的外 切酶切割形成 5'端,然后可能再次被剪切形成 3'端, 或者有 3'→5'核酸外切酶从 3'端逐个切除核苷酸,形 成一定长度的 piRNA 后,再通过甲基转移酶 Hen1 使 2-OH 甲基化,形成成熟的 piRNA,结合到 Piwi 或 Aub 蛋白上^[30-33].可能因为 Piwi 和 Aub 对 5'端为 U 的 RNA 有结合优先性,初级 piRNA 有强烈的 5'U 偏 向性^[33].

卵巢体细胞中的 piRNA 的初级加工途径主要在 一种无膜的颗粒状细胞器——Yb 小体(Yb body)中进 行(图 2)^[34,35]. Yb 小体位于胞浆靠近细胞核的位置, 以 Yb 蛋白特异性定位其中得名^[36]. 除了 Yb 蛋白外, 已发现其中还有多个组分,包括另一个 Tudor 结构域 蛋白 Vreteno(Vret), SDE 结构域蛋白 Armitage (Armi), FKBP 家族蛋白 Shutdown(Shu)^[34,35,37-41]. 除了 Yb 小 体外,初级加工途径组分 Zucchini(Zuc)与GASZ 位于 线粒体中^[35,39,42-45]. 这些蛋白的缺失都造成体细胞中 piRNA 产生的急剧减少.

Yb 蛋白含有一个 Tudor 结构域和一个 DEAD/ DEAH 盒 RNA 解旋酶结构域,只在体细胞中表达. Armi 是一个不含 DEA(H/D)盒的 ATP 依赖的 RNA 解 旋酶,与拟南芥(*Arabidopsis thaliana*)中参与转录后 基因沉默的蛋白 SDE3 同源^[46,47]. 生化方面的研究证 明了 Armi 蛋白参与 RISC 的组装^[46]. 对体外培养的 OSC 的研究发现, Armi 蛋白能够结合 Yb 和 Piwi,在 Yb 小体中形成 Yb-Armi-Piwi 的复合体^[35,48]. Yb 缺失 后, Armi 不能够聚集到 Yb 小体上,但是 Armi 仍然能 够与 Piwi 形成复合物,可能 Yb 能够影响 Armi-Piwi 复合物的定位和动态^[35,48]. 研究发现, piRNA 在胞浆 加工成熟后,在 Yb 小体上加载到 Piwi 蛋白上,随后



图 2 piRNA 形成

A: 初级加工途径.体细胞 piRNA 簇转录子转运到胞浆,在线粒体(mitochondria)中由 Zuc 剪切形成 5'端,在 Yb 小体由 Yb, Armi, Vret, Shu 等参与进一步加工,再由 Trimmer 蛋白去掉 3'端多余的碱基形成特定长度的 piRNA,由 Hen1 使 2-OH 甲基化而成熟,结合到 Piwi 上.在体细胞初级加工途径中,Yb 的功能由 SoYb 和 BoYb 代替,Yb 小体由 nuage 代替,加工过程可能类似;B: 次级加工途径.生殖细胞中,初级 piRNA 进入乒乓循环,经 6 个步骤形成新的 piRNA:(i) Aub/piRNA 识别和结合转座子 mRNA 并剪切;(ii) 3'端剪切产物在 Vasa 和 Qin/Kumo 的作用下转移到 AGO3,游离的 AGO3 可能由 Hsp90 和 Krimp 结合而保护不被降解;(iii) AGO3 结合的 piRNA 前体被进一步加工,在 3'2-OH 加上甲基成熟;(iv) AGO3/piRNA 识别和结合 piRNA 簇转录子并剪切,该步骤可能需要 Armi 与 Zuc 的作用;(v) 3'端剪切产物结合到 Aub 上;(vi) Aub 结合的 piRNA 前体被进一步加工,在 3'2-OH 加上甲基成熟.在乒乓循环过程中,还有很多蛋白(如 Vret, Shu, Tud, SpnE, SoYb 和 BoYb) 参与其中,但是分子功能需要进一步的鉴定

Piwi/piRNA 复合物进入到核内. Yb 和 Armi 缺失时, 结合到 Piwi 上的 piRNA 急剧减少, Piwi 蛋白被阻滞 在胞浆而不能进入核内. Armi 蛋白还能够直接结合 25~70 nt 长的 RNA(piR-ILs),这可能是 piRNA 加工过 程中的中间体,被进一步剪切后才能形成成熟的 piRNA. 与成熟的 piRNA 结构不同,这些 piR-ILs 在 5'和 3'端都没有游离磷酸基团, 3'端 2'和 3'羟基通过一 个 磷 酸 相 连,形成 2',3'环 磷 酸 (2',3'-cyclic phosphate)^[35].

Vret 含有两个 Tudor 结构域,还有一个 RNA 识别 结构 域 (RNA recognition motif, RRM) 和一个 MYND 结构域,在体细胞中定位于 Yb 小体,与 Yb, Armi 和 Piwi 有物理相互作用^[37,38].而且 Vret 在 Yb 中的定位依赖于 Yb 和 Armi,不依赖于 Piwi,而 Vret

的缺失不影响 Armi 和 Yb 在 Yb 小体中的定位^[38]. Shu 是 FK506 结合蛋白(FK506-bind protein, FKBP)家族 成员,是一种免疫亲和素(immunophilin)类伴侣蛋白, 含有一个肽酰脯氨酰顺反异构酶(peptidyl-prolylcis/trans-isomerase, PPIase)结构域和一个 TPR 结构域. 在体细胞中, Shu 也定位在 Yb 小体中,且其定位依赖 于 Yb 组分 Yb, Armi 和 Vret,也依赖于 Piwi,相反 Shu 的缺失不影响 Yb, Armi 和 Vret 在 Yb 小体中的定位, 因此 Shu 在这些 Yb 组分的下游发挥作用,可能未结 合 piRNA 的 Piwi 在 Yb 小体招募 Shu^[40].

线粒体在 piRNA 初级加工途径中也起到重要的 作用(图 2). Zuc 定位于线粒体外膜,作为一个核酸外 切酶参与 piRNA 的加工^[35]. Zuc 是一个 PDL (phospholipase-D)家族的磷酸二酯酶,其酶活性位点

含有一个 HKD(H(X)K(X4)D)基序^[44,45]. 晶体结构和 体外活性分析表明, Zuc 蛋白没有磷酸二酯酶活性, 不能够从心磷脂中水解释放磷脂酸^[44,45]. Zuc 具有核 酸酶作用, 通过β折叠形成二聚体, 两个 HKD 基序组 成核酸酶活性中心,特异性地剪切单链RNA.其在N 端还有一个锌结合结构域,缺失时降低核酸酶活性, 它对 RNA 的剪切没有底物序列特异性, 剪切产物 5' 末端具有单磷酸基团^[45]. HKD 基序突变会使 Zuc 失 去核酸酶活性,在 OSC 和卵巢中, HKD 位点的突变 也会极大地影响转座子的沉默和 piRNA 的形成^[39,45]. 在 piRNA 初级加工途径中, Zuc 的缺失会使 Piwi 蛋白 不能进入核内而聚集在胞浆和 Yb 小体上, 而 Yb, Armi, Vret 和 Shu 仍然能够定位到 Yb 小体上, 但是 也发现这时的 Yb 小体会比正常的大很多, 可能 Yb 会通过某种方式使线粒体成簇地出现在 Yb 小体的位 置^[36,38,40,43]. Zuc 对 Armi, Yb 与 Piwi 的物理相互作用 也不是必需的,反而在 zuc 缺失时,有更多的 Piwi 与 Armi 相互作用. Zuc 缺失也会使 Armi 结合的 piR-ILs 稍有减少, 但是 Piwi 结合的 piRNA 大大减少, 而且 这些 piRNA 似乎比成熟的 piRNA 要长一个碱基, 推 测Zuc可能作为核酸酶对Armi结合的piR-ILs和Piwi 结合的成熟前 piRNA 进行加工^[35]. GASZ 是另一个参 与初级加工途径的线粒体蛋白,是哺乳动物 GASZ(germ cell protein with ankyrin repeats, sterile alpha motif, and leucine zipper)蛋白的同源物, 含有5 个 ankyrin 重复和 1 个 sterile alpha 基序. 其 C 端有一 个跨膜结构域,其定位不依赖Zuc,Zuc的线粒体定位 也不依赖 GASZ. 在 OSC 和卵巢滤泡细胞中, GASZ 缺失也不影响 Yb, Armi, Vret 在 Yb小体定位, 能够使 Piwi 聚集到 Yb 小体上而不能进入核内^[36].

此外,最近的研究证明了 nxt1, smt3, cls4 和 rrp6 等基因可能直接参与 piRNA 的初级加工途径 NXT1 蛋白与核孔复合物相互作用,可能参与 piRNA 簇转 录子从核内到胞浆的运输; SMT3 含有一个 SUMO 样 结构域,可能参与到某些 piRNA 途径蛋白的修饰; CLS4 和 RRP6 都是外切体(exosome)的组分,有可能 参与初级 piRNA 3'端的形成^[49].

3 乒乓循环

在果蝇卵巢生殖细胞中, piRNA 簇的转录子转运 到胞浆后经过与体细胞类似的初级加工途径形成初 级 piRNA,初级 piRNA 结合到 Piwi 和 Aub 上. 与体 细胞不同的是,生殖细胞中没有 Yb 蛋白和 Yb 小体,初级加工途径主要发生在 nuage 上. Nuage 是与 Yb 小体类似的细胞器,在电子显微镜下都呈现高电子密度,无膜,位于细胞核膜外周,其周围经常分布丰富的线粒体^[50]. Nuage 上没有 Yb 蛋白,但存在另外两个 Yb 家族的蛋白,BoYb(Brother of Yb)与 SoYb(Sister of Yb),参与 piRNA 的初级加工^[38]. 另外, Vret 和 Shu 也定位在 nuage 上,Zuc 和 GASZ 定位在线粒体上,而 Armi在 nuage 和线粒体中都有定位.目前对生殖细胞中 piRNA 初级加工所知甚少^[48].

生殖细胞的初级 piRNA 还会进入次级加工途径, 经过 PIWI 家族蛋白的协同作用,细胞中的 piRNA 大 量扩增(图 2). 首先, 结合初级 piRNA 的 Aub 蛋白通 过碱基互补识别结合到靶 RNA 上, 靶 RNA 被 Aub 的 Slicer 活性剪切, 其 3'端产物被加工而形成一个新 的 piRNA 与 AGO3 结合,这个 AGO3/piRNA 复合物 又识别、结合、剪切靶 RNA, 3'端产物加工成新的与 初始的初级 piRNA 完全相同的 piRNA, 再与 Aub 结 合,如此循环往复. Aub 结合的初级 piRNA 的 5'端第 一个碱基具有 U 的倾向性,由于 Slicer 活性剪切的特 点,剪切产物 5'端 10 个碱基会与该初级 piRNA 的 5' 端 10 个碱基互补, 其第 10 个碱基与初级 piRNA 的 第1个碱基互补,因此一般是A. 再者,Aub结合的初 级 piRNA 一般对应于转座子序列的反义链, 因此其 靶序列为转座子的 mRNA; 经剪切形成的次级 piRNA 对应于转座子的正义链,并结合到 AGO3 蛋 白,这个 AGO3/piRNA 的复合物靶向到转座子的反 义链,这个反义链来源于 piRNA 簇的转录.因此,在 乒乓过程中, 通过初级 piRNA 和 piRNA 簇转录子的 不断输入,能够形成大量的新 piRNA,在这个过程中 又将转座子的 mRNA 剪切降解而沉默, 成为一个 piRNA 形成与转座子沉默偶联的过程. 乒乓循环生 成的 piRNA 有明显的特点, 即 Aub 主要结合反义 piRNA, 第一个碱基一般是 U, 而 AGO3 结合正义 piRNA, 第 10 个碱基一般是 A, 且它们的 5'端有 10 个碱基的互补.

乒乓循环是一个很吸引人的模型,但是其中的 细节和具体机制,知道的还极少.有几个关键问题亟 待解决:(i)Aub/piRNA剪切转座子 mRNA 后,5'端 产物的命运如何,3'端产物是如何被加工和转移到 AGO3的,以及同样的AGO3/piRNA产物的命运和加

工的过程是如何,这些3'端产物是如何避免被降解的, 都很不清楚;(ii) 乒乓循环是如何选择在 Aub:AGO3 间进行, 而不是在通过 Aub:Aub 或者 AGO3:AGO3 的.现在,本研究组已经鉴定出了很多蛋白参与乒乓 过程,希望能够通过对这些蛋白的研究,了解它们与 PIWI家族蛋白的关系,它们怎样调节 PIWI 家族蛋白 的活性或动态过程,通过这些更多的深入理解乒乓 过程. 乒乓加工过程, 除了 PIWI 家族蛋白外, 现在 已知有其他很多蛋白参与其中,包括Vret和Shu这两 个在初级途径中必需的蛋白. 其中 Shu 参与 piRNA 的次级加工,可能还需要 Hsp90 的参与^[40]. Shu 可能 通过其 TPR 结构域与 Hsp90 蛋白 C 末端 MEEVD 的 肽端相互作用, 其 TRP 结构域中与 Hsp90 结合的关 键氨基酸突变时会造成果蝇不育, AGO3 和 Aub 蛋白 不能定位到 nuage. Hsp90 蛋白在滋养细胞中主要定 位在胞浆,小部分在核内,胞浆中在 nuage 上也有一 定程度的聚集. 在一些 piRNA 基因如 aub, vasa, vret, shu的突变体中,Hsp90会与AGO3/Krimp一同定位在 胞浆的颗粒中, Hsp90 与没有结合 piRNA 的 AGO3 蛋 白结合^[40].还有一些蛋白只参与次级加工过程,包括 Vasa, Tudor, Spindle-E(SpnE), Krimper (Krimp), Qin/Kumo, PAPI, 除了 Vasa 外, 都含有 Tudor 结构域, 也都定位在 nuage 上.

Vasa 是一个 DEAD 盒(DEAD box)家族的 ATP 依赖的 RNA 解旋酶, 早期的研究鉴定出其在胚胎极 粒(polar granules)中表达, 对早期胚胎发育和生殖细 胞形成是必需的. Vasa 在卵巢中定位在 nuage 上, 并 且在卵母细胞后端聚集. 其在卵巢生殖细胞中发挥 着多种重要的功能,除了参与 piRNA 通路外还调节 mRNA 翻译和生殖细胞分化、极浆组装(pole plasm assembly)和有丝分裂的进行^[51]. vasa 突变极大地影响 了乒乓循环的进行而减少了生殖细胞特异的 piRNA 的表达^[12]. 后来的研究发现, Vasa 蛋白能够直接与从 核内经核孔转运出的 piRNA 转录子前体结合^[29], 是 piRNA 转录子进入乒乓循环的入口. 最近通过对昆 虫 Bombyx mori 细胞系 BmN4 中 Vasa 的研究, 对理 解乒乓过程提供了极好的参考^[52]. Vasa在BmN4细胞 中也定位在 nuage 上, 但是没有检测到 Vasa 与其他 nuage 组分的相互作用, 然而当其 DEAD 盒中活性氨 基酸突变使其失去 ATP 水解酶活性时, Vasa 会高度 富集在一些胞浆大颗粒中,相互作用分析也发现突 变型的 Vasa 能与 AGO3, Qin/Kumo 有物理上的相互 作用,同时 AGO3,Qin/Kumo 也定位在那种大颗粒中, 而非 nuage 蛋白 Siwi(Aub 同源蛋白)也会富集其中. 因此在乒乓循环过程中,Vasa 与 AGO3,Qin/Kumo 会 形成瞬时的结构,称为扩增子(amplifier).对这种结 构体的进一步分析发现,其中含有成熟的反义 piRNA,正义的转座子 mRNA 序列及 3'端剪切产物. 因此推测在乒乓过程中,Vasa 通过 ATP 水解造成构 象改变,结合并保护 Aub/piRNA 的 3'端的剪切产物, 并将其转移到 AGO3 上,以便能够进行下一步的加 工,形成 AGO3 结合的成熟 piRNA.

Tudor 蛋白是一类含有 Tudor 结构域的蛋白, 多 种 Tudor 蛋白参与 piRNA 途径^[53,54]. Tudor 结构域能 够结合到甲基化的精氨酸残基上而与靶蛋白相互作 用[53,54]. 蛋白精氨酸的甲基化有3种,单甲基化、对 称型和非对称性二甲基化,甲基化的精氨酸残基可 以被特定的结构域(如 Tudor 结构域)结合或者从空间 上阻碍其他蛋白对该精氨酸临近氨基酸的修饰,从 而介导蛋白之间的相互作用或调节蛋白功能[54].蛋 白精氨酸甲基化由蛋白精氨酸甲基转移酶(protein arginine methyltransferase, PRMT)催化. 在果蝇中, PRMT5/DART5/Capsleen 参与 PIWI 家族蛋白的对称 性二甲基化修饰而参与 piRNA 通路^[55,56]. 果蝇的 PIWI 蛋白 Piwi, Aub 和 AGO3 的 N 端均含有富含精 氨酸-甘氨酸的基序(RG-rich motif),从而被 PRMT5 识别^[55]. 现已发现多种 Tudor 蛋白参与 piRNA 途径, Tudor可能通过结合 PIWI 家族蛋白,调节其活性,影 响其结构,或者介导 PIWI 家族蛋白与其他蛋白的相 互作用.

Tudor(Tud)是最早发现含有Tudor结构域的蛋白, 有11个Tudor结构域.它在nuage上和卵母细胞后端 聚集.研究发现,Tud通过其Tudor结构域与Aub和 AGO3的甲基化精氨酸残基相结合^[56-58].tudor 突变 不影响Aub与AGO3在nuage的定位,但是会增强它 们的相互作用以及结合 piRNA 的量,但是这些 piRNA 很可能是异常的,其种类和数量都不同于野 生型^[56].Tudor蛋白抗体的免疫沉淀中还发现了 piRNA 前体样的长 RNA,推测其可能直接参与 piRNA的加工,对piRNA的质量控制至关重要^[56].

SpnE 是另一个含有 ATP 依赖的 RNA 解旋酶活性的蛋白,它有解旋酶活性相关的 HELICc, HA2 结构域和一个 Tudor 结构域. SpnE 蛋白在生殖细胞中特异表达,对卵巢卵母细胞的极性发育起重要作用,对

睾丸的有丝分裂过程及沉默 Stellate 元件是必需的^[59,60].研究发现, SpnE 蛋白也定位在 nuage 上,对 乒乓循环也是必需的,其缺失完全阻断了乒乓循 环^[12]. SpnE 似乎并不是通过结合甲基化精氨酸来结 合 AGO3 或 Aub 来发挥作用^[56].

Krimp 蛋白在 C 端含有一个 Tudor 结构域^[50], 目 前对其生化功能还不清楚. krimp 基因突变使 AGO3 蛋白不能定位在 nuage, 而弥散在胞浆, 而且导致乒 乓循环的缺陷^[12]. Krimp 似乎是与 AGO3 蛋白特异结 合, 因为在一些 piNRA 途径基因的突变体, 如 aub, spnE 中, AGO3 与 Krimp 都不再位于 nuage, 而位于 胞浆的小颗粒中; 而当 krimp 也缺失时, AGO3 不能 聚集到这些颗粒中^[40].

Tejas 蛋白 C 端含有一个 Tudor 结构域, 其 N 端还有一个 Tejas 结构域^[53,61]. 其 Tudor 结构域对它在piRNA 途径的功能不是必需的, 反而其 Tejas 结构域 是至关重要的^[61]. 而且 Tejas 蛋白这两个结构域都能 够与 piRNA 途径蛋白相互作用, 而且其与 Aub 的相 互作用并不依赖于 Aub 的精氨酸甲基化^[61]. 它与 piRNA 途径蛋白的另一个不同在于, 其突变体并 不会造成在其他突变体中很常见的 Osk 的提前翻 译和 Grk 的表达减少, 因此不影响卵母细胞极性的 形成^[61].

Qin/Kumo的C端有5个Tudor结构域,N端含 有一个RING结构域和两个B-Box结构域,定位在 nuage上. qin/kumo的缺失会造成Aub与AGO3相互 作用减弱,使细胞中Aub:AGO3之间的乒乓转换成 Aub:Aub乒乓,导致piRNA的种类发生变化,但似乎 对piRNA的数量影响不大,但仍然会造成某些转座 子的去抑制^[62,63]. Qin/Kumo在卵巢原卵区(germarium) 定位在核内,且能够与HP1a有相互作用,推测其可 能也与piRNA簇的双向转录有关^[63]. PAPI蛋白C端 有一个Tudor结构域,N端有两个KH结构域,定位在 nuage上. PAPI通过Tudor结构域与AGO3的N端甲 基化精氨酸结合,同时PAPI/AGO3也与P小体(P body)组分TRAL和ME31B在nuage上形成复合体^[64].

最近,本研究组发现 Armi 与 Zuc 这两个线粒体 蛋白很可能也直接参与乒乓过程. 检测到 AGO3 能 够特异地与 Armi 和 Zuc 相互作用,而且一小部分 AGO3 定位在线粒体上,而生殖细胞中 Armi 也定位 在线粒体和 nuage 上,本研究组^[65]认为, AGO3/Armi 复合体能够在线粒体与 nuage 之间动态地穿梭,在 nuage上共同参与 AGO3 RISC 复合体的组装,而 Zuc 对 Armi 的穿梭以及与 AGO3 的相互作用都有调节作用.线粒体蛋白 GASZ 也参与次级 piRNA 途径.生殖 细胞中,与 Armi 或 Zuc 缺失类似, GASZ 的缺失也会 引起 Piwi 不能聚集到核内,一部分 AGO3 不能定位到 nuage,而 Aub 定位没有影响,也会造成 piRNA 的缺失^[42,43]. 另外 gasz 缺失时, Armi 蛋白不能够定位到 线粒体上,推测 GASZ 在生殖细胞中可能作为一个接 头蛋白将 Armi 招募到线粒体上^[42,43].

4 piRNA 与表观沉默

在果蝇卵巢生殖细胞中, 通过乒乓循环, 在 piRNA 的产生过程中, Aub 利用 Slicer 活性不断剪切 转座子 mRNA, 使其降解, 从而使转座子表达受到抑 制. 然而在体细胞中, piRNA 只结合到 Piwi 上, 生殖 细胞中也有一部分 piRNA, 尤其是初级 piRNA 结合 到 Piwi 上, 而在这两种细胞中, Piwi 蛋白均进入核内 发挥功能. 尽管 Piwi 也具有 Slicer 活性[66], 但是似乎 不像 Aub 一样直接剪切转座子 mRNA, 因为缺失 Slicer活性的 Piwi 足够造成转座子的沉默^[35,67].相反, Piwi 的入核对其沉默转座子表达的功能是必需的, piwi^{NT}突变体表达一个缺失 N 端 26 个氨基酸而不能 入核形式的 Piwi 蛋白, 该突变体不影响生殖干细胞 的维持和分化,然而会造成多个转座子的去抑制[67]; 另一项研究也证明,同样缺失核定位信号的 Piwi 蛋 白也不能够拯救 piwi 突变造成的体细胞或生殖细胞 特异的转座子的表达[68].

结合 piRNA 的 Piwi 蛋白在核内以 piRNA 依赖 的方式结合到染色体上^[69]. Piwi 的 ChIP-seq (chromatin immunoprecipitation-squence)的数据表明, Piwi 主要结合在染色体的转座子和重复序列区域. Piwi 的 ChIP-seq 信号在富含基因的染色体臂上较低, 而在富含转座子和重复序列的 U, Uextra 等区域很强. 而且 Piwi 在染色质上的结合区域与 piRNA 在染色体 上的对应序列有高度的一致性. 在真染色体区, Piwi 可能通过 piRNA 结合到新合成的互补 RNA 上, 而 在异染色质区, Piwi 可能通过 piRNA 结合到互补 DNA 上^[69].

Piwi 结合到染色体上引起其表观状态的改变, 抑制转座子转录. 早期的研究发现 piwi 突变时, 染色

体上某些报告基因的表达会发生改变(图 3)^[6]. piRNA 通路基因 *spnE* 突变引起生殖细胞中 Het-A, I-element 与 copia 转座子 mRNA 的表达升高,这 3 个转座子的 启动子和编码区的 H3K9me2, H3K9me3 和 HP1a 水平 降低和 H3K4me2 水平升高^[70].在 *spnE* 突变体和生殖 细胞敲低 *piwi* 时,HeT-A 和 TART 区域活性的 RNA 聚合酶明显增多,转录加快^[15].近年通过对 piRNA 通路和 Piwi 对染色质结构的全局性的研究更进一步 验证了这些结果^[68,69,71,72].在卵巢体细胞和生殖细胞 中,piRNA 通路基因和 *piwi* 的缺失都会广泛引起转座 子序列区域 RNA 合成速率的增加、RNA 聚合酶水平 的升高、H3K9me3 水平降低.因此 Piwi/piRNA 会在 基因组上转座子区域建立异染色质结构,同时这种 异染色质结构能够从转座子插入位点向两边扩散, 可能会引起附近基因的表达受到抑制^[68].

在 Piwi 及 piRNA 途径参与染色质状态调节的过程中, HP1a发挥着重要的功能. HP1a蛋白C端还有一个 chromo 结构域 (chromodomain, CD), 介导 H3K9me2/3 的结合, N 端含有 chromoshadow 结构域 (chromoshadow domain, CSD), 介导其二聚化和与其 他蛋白的相互作用, N 端与 C 端通过一个铰链区 (hinge region)相连. Piwi 蛋白通过 N 端的 PXVXL 基 序与二聚化 HP1a的 CSD 区直接相互作用^[73]. 对果蝇 三龄幼虫的唾腺染色体的分析也发现, Piwi 与 HP1a 在染色体的很多位置有共定位^[73,74]. Piwi 通过 piRNA 的碱基互补作用结合到染色质后, 招募 HP1a, 然后 进一步招募组蛋白甲基转移酶 Su(Var)3-9, 使附近抑 制性的组蛋白标记 H3K9me2 和 H3K9me3 水平升高, 相 反 激 活性 组 蛋 白 标 记 H3K4me2, H3K4me3, H3K9ac 和 H3K36me3 水平都不会升高(图 3)^[69].

在生殖细胞和体细胞中, Maelstrom(Mael)也都参 与 Piwi 介导的转录沉默^[68]. Mael 含有一个 HMG 盒 (high-mobility group box)结构域和一个 Mael 结构域, 后者与 DnaQ-H 3'→5'核酸酶有低的同源性^[75].在生 殖细胞中, Mael 在胞浆、nuage 和核内均有定位; 在 体细胞中也在胞浆和核内均有表达,而且在核内有 更强的聚集. Mael 对 piRNA 的生成似乎不是必需的, 但是对生殖细胞和体细胞转座子的沉默都是至关重 要的. Mael 结构域保守氨基酸的突变会使 Mael 蛋白 不能聚集到核内, 也失去抑制转座子表达的功能, 而 HMG 结构域对其功能只有中度的影响. Mael 的缺失 不会影响 Piwi 蛋白的表达和定位, 但会影响 Piwi 对 转座子的抑制. 在 OSC 细胞中, mael 的缺失与 piwi 或 armi 缺失类似, 会造成一些特定的转座子 RNA 水 平升高和 RNA 聚合酶聚集增强^[68].另一个参与这一 过程的蛋白为GTSF1,GTSF1的N端含有两个CHHC 锌指,可能具有 RNA 结合活性,通过其 C 端与 Piwi 直接地相互作用. GTSF1 定位在核内, 且该定位依赖 于 piwi, 相反, GTSF1 的缺失不影响 Piwi 的表达和定 位.同 piwi 缺失的表型类似, gtsfl 缺失同源会造成转 座子表达的升高, RNA 聚合酶Ⅱ在转座子位点的聚 集和 H3K9me3 的降低^[76,77].

5 展望

尽管 piRNA 的研究已经取得了很大的进展,对 piRNA 的生成和了解已经有了一个总体上的认识, 但是在一些关键问题和具体机制方面还需要更多的 研究. 首先, piRNA 簇及其转录产物是如何同其他基 因组位点和转录产物区分的. 单向转录的转录子依



图 3 Piwi 参与的转录沉默(网络版彩图)

Piwi 通过 piRNA 的碱基互补识别和结合到基因组的转座子序列区域,招募 HP1a 蛋白,后者进一步招募组蛋白甲基转移酶 Su(Var)3-9,甲基 化该区域的组蛋白 H3K9,抑制 RNA 聚合酶 II 的结合和聚集,沉默转座子的转录.这种沉默还有可能向转座子区域两端延伸,使附近的基因 组 H3K9me 增强,转录减弱

赖于其自身的启动子,可以独立地启动转录,但是其 转录子似乎进入了不同与 mRNA 剪切的加工途径, 具体机制还不清楚.双向转录的 piRNA 簇几乎没有 独立的启动子,其似乎依赖于两端的基因和 Rhi 的结 合,而 Rhi靶向到某个非 piRNA 簇区域甚至能将其转 变为 piRNA 簇. Rhi 可能通过结合 H3K9me2/3 识别结 合染色质,但是 H3K9me2/3 广泛存在于异染色质区, H3K9me2/3 和两端基因的存在似乎不足以定义一个 piRNA 簇和 Rhi 的神奇功能,可能还有其他的染色质 结合蛋白与 Rhi 共同作用,有待进一步的鉴定.从这 种 piRNA 簇转录出的 RNA 进入了不同的剪切加工途 径,可能还有赖于 Del, Cuff 和 UAP56 的作用,但具 体的机制也不清楚.

再者, piRNA 簇转录子进入胞浆中, 需要复杂的 加工, 尤其是生殖细胞中有初级和次级两条途径, 加 工的细节还知之甚少. 这些 piRNA 簇转录子很可能 是通过特定的核外运因子转运到胞浆的, 且马上进 入了 Yb 小体或 nuage 中同普通的 RNA 区分开来. 初 级途径中, piRNA 簇转录子很可能是被核酸酶 Zuc 剪 切而形成 piRNA 的 5'端的, 而在次级途径中, 是通过 Aub 和 AGO3 的 Slicer 活性剪切形成的, 但是现在还 没有鉴定出参与这些 piRNA 3'端形成的核酸酶. 在 乒乓循环中, Aub与AGO3都定位在nuage上, 而Piwi 也在 nuage 上结合 piRNA,因此需要保证乒乓在 Aub 与AGO3之间进行,从而产生有效的piRNA.现在已 知 Tud 和 Qin/Kumo 以及 Vasa 可能参与这个调节,但 具体的分子机制还不清楚. 这些 nuage 蛋白细微的定 位差异和动态变化可能是一个关键因素.特定的 PIWI 蛋白可能会与特定的蛋白定位在不同的 nuage 亚结构中,这在果蝇卵巢和小鼠(Mus musculus)精子 中均有发现^[78,79].参与 piRNA 途径的很多蛋白似乎 都处于动态变化中,有很多相互作用只是瞬时的,而 且很多蛋白在细胞中不同位置穿梭,包括 nuage、线 粒体、胞浆^[52,65,79],可能通过这些机制能保证 piRNA 的加工能够有序地进行.

最后, Piwi结合 piRNA 后进入核内, 参与表观调 节,有证据表明可能是转录沉默,也有可能是转录激 活, 似乎是依染色质类型而不同, 具体机制还不清楚. piwi 的缺失导致插入到真染色质位置的转座子的 H3K9me3 水平降低, RNA 聚合酶 II 聚集水平升高, RNA 合成增强,造成了该插入转座子的表观激活, 同时这种激活状态还有可能向转座子插入位点两端 延伸,因此 Piwi 蛋白通过表观抑制来沉默插入到真 染色质区的转座子. 然而也有研究发现, Piwi 的表观 激活功能, Piwi/piRNA 结合到亚端粒异染色质 3R-TAS 区,维持该区真染色质标记 H3K4me2/3 和 H3K9ac的水平,抑制H3K9me2/3和HP1a的水平,使 该区域的 piRNA 表达^[6,74]. 后来的研究验证了 Piwi 对一些插入到真染色质的转座子形成的 piRNA 簇的 转录是必需的, 在这些位点有较高的 H3K9me3 水平 和 Rhi 的结合^[27]. 直接的 Piwi 的 ChIP-seq 也提示, Piwi/piRNA 在真染色质和异染色质上有不同的结合 模式^[69]. 这种差异可能跟 HP1a 和 Rhi 相关, 但需要 更多的证据来解释.

果蝇是最早被发现 piRNA 存在的物种之一,对 其卵巢中 piRNA 的研究相对更多和更深入,但是现 在对其他物种和组织,包括果蝇睾丸、线虫、斑马鱼 (*Barchydanio rerio* var)和小鼠睾丸中 piRNA 的研究 有了很大的进展,这些 piRNA 通路与果蝇卵巢 piRNA 通路有很多相似的机制,很多同源蛋白参与 其中,但是也发现有一些不同点,这些相同和不同点 都能够促进对 piRNA 更加深入的研究.另外,除了 生殖系统,在很多类型的体细胞中也发现了 PIWI 家 族蛋白及 piRNA 类小 RNA,尤其是在癌细胞中^[80,81], 对这些细胞 piRNA 的研究也许会有一些医学上的价 值,同时这种不同组织的 piRNA 研究也可以相互借 鉴,对深入理解 piRNA 很有帮助.

参考文献。

¹ 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: 843–854

² Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell, 1993, 75: 855–862

³ Megosh H B, Cox D N, Campbell C, et al. The role of PIWI and the miRNA machinery in *Drosophila* germline determination. Curr Biol, 2006, 16: 1884–1894

- 4 Saito K, Nishida K M, Mori T, et al. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes Dev, 2006, 20: 2214–2222
- 5 Vagin V V, Sigova A, Li C, et al. A distinct small RNA pathway silences selfish genetic elements in the germline. Science, 2006, 313: 320–324
- 6 Yin H, Lin H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. Nature, 2007, 450: 304–308
- 7 Kim VN, Han J, Siomi M C. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol, 2009, 10: 126–139
- 8 Hutvagner G, Simard M J. Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol, 2008, 9: 22–32
- 9 Brennecke J, Aravin A A, Stark A, et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell, 2007, 128: 1089–1103
- 10 Khurana J S, Wang J, Xu J, et al. Adaptation to P element transposon invasion in Drosophila melanogaster. Cell, 2011, 147: 1551–1563
- 11 Klattenhoff C, Xi H, Li C, et al. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell, 2009, 138: 1137–1149
- 12 Malone C D, Brennecke J, Dus M, et al. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. Cell, 2009, 137: 522–535
- 13 Jensen P A, Stuart J R, Goodpaster M P, et al. Cytotype regulation of P transposable elements in *Drosophila melanogaster*: repressor polypeptides or piRNAs? Genetics, 2008, 179: 1785–1793
- 14 Shpiz S, Kalmykova A. Role of piRNAs in the Drosophila telomere homeostasis. Mob Genet Elements, 2011, 1: 274–278
- 15 Shpiz S, Olovnikov I, Sergeeva A, et al. Mechanism of the piRNA-mediated silencing of *Drosophila* telomeric retrotransposons. Nucleic Acids Res, 2011, 39: 8703–8711
- 16 Khurana J S, Xu J, Weng Z, et al. Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. PLoS Genet, 2010, 6: e1001246
- 17 Klattenhoff C, Bratu D P, McGinnis-Schultz N, et al. *Drosophila* rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. Dev Cell, 2007, 12: 45–55
- 18 Mohn F, Sienski G, Handler D, et al. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand pirna clusters in *Drosophila*. Cell, 2014, 157: 1364–1379
- 19 Robert V, Prud'homme N, Kim A, et al. Characterization of the flamenco region of the *Drosophila melanogaster* genome. Genetics, 2001, 158: 701–713
- 20 Prud'homme N, Gans M, Masson M, et al. Flamenco, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. Genetics, 1995, 139: 697–711
- 21 Pelisson A, Song S U, Prud'homme N, et al. Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the *Drosophila* flamenco gene. EMBO J, 1994, 13: 4401–4411
- 22 Mevel-Ninio M, Pelisson A, Kinder J, et al. The *flamenco* locus controls the *gypsy* and *ZAM* retroviruses and is required for *Drosophila* oogenesis. Genetics, 2007, 175: 1615–1624
- 23 Sarot E, Payen-Groschene G, Bucheton A, et al. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. Genetics , 2004, 166: 1313–1321
- 24 Goriaux C, Desset S, Renaud Y, et al. Transcriptional properties and splicing of the flamenco piRNA cluster. EMBO Rep, 2014, 15: 411-418
- 25 Robine N, Lau N C, Balla S, J et al. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr Biol, 2009, 19: 2066–2076
- 26 Saito K, Inagaki S, Mituyama T, et al. A regulatory circuit for piwi by the large *Maf* gene traffic jam in *Drosophila*. Nature, 2009, 461: 1296–1299
- 27 Zhang Z, Wang J, Schultz N, et al. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell, 2014, 157: 1353–1363
- 28 Rangan P, Malone C D, Navarro C, et al. piRNA production requires heterochromatin formation in *Drosophila*. Curr Biol, 2011, 21: 1373–1379
- 29 Zhang F, Wang J, Xu J, Zhang Z, et al. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell, 2012, 151: 871–884
- 30 Kawaoka S, Izumi N, Katsuma S, et al. 3' end formation of PIWI-interacting RNAs in vitro. Mol Cell, 2011, 43: 1015–1022

- 31 Horwich M D, Li C, Matranga C, et al. The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol, 2007, 17: 1265–1272
- 32 Saito K, Sakaguchi Y, Suzuki T, et al. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. Genes Dev, 2007, 21: 1603–1608
- 33 Ishizu H, Siomi H, Siomi M C. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev, 2012, 26: 2361–2373
- 34 Qi H, Watanabe T, Ku H Y, et al. The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. J Biol Chem, 2011, 286: 3789–3797
- 35 Saito K, Ishizu H, Komai M, et al. Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila*. Genes Dev, 2010, 24: 2493–2498
- 36 Szakmary A, Reedy M, Qi H, et al. The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster. J Cell Biol, 2009, 185: 613–627
- 37 Zamparini A L, Davis M Y, Malone C D, et al. Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in *Drosophila*. Development, 2011, 138: 4039–4050
- 38 Handler D, Olivieri D, Novatchkova M, et al. A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. EMBO J, 2011, 30: 3977–3993
- 39 Haase A D, Fenoglio, S, Muerdter F, et al. Probing the initiation and effector phases of the somatic piRNA pathway in *Drosophila*. Genes Dev, 2010, 24: 2499–2504
- 40 Olivieri D, Senti K A, Subramanian S, et al. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in *Drosophila*. Mol Cell, 2012, 47: 954–969
- 41 Preall J B, Czech B, Guzzardo P M, et al. shutdown is a component of the *Drosophila* piRNA biogenesis machinery. RNA, 2012, 18: 1446–1457
- 42 Czech B, Preall J B, McGinn J, et al. A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals factors of the germline piRNA pathway. Mol Cell, 2013, 50: 749–761
- 43 Handler D, Meixner K, Pizka M, et al. The genetic makeup of the Drosophila piRNA pathway. Mol Cell, 2013, 50: 762–777
- 44 Ipsaro J J, Haase A D, Knott S R, et al. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature, 2012, 491: 279–283
- 45 Nishimasu H, Ishizu H, Saito K, et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature, 2012, 491: 284–287
- 46 Tomari Y, Du T, Haley B, et al. RISC assembly defects in the Drosophila RNAi mutant armitage. Cell, 2004, 116: 831-841
- 47 Cook H A, Koppetsch B S, Wu J, et al. The *Drosophila* SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification. Cell, 2004, 116: 817–829
- 48 Olivieri D, Sykora M M, Sachidanandam R, et al. An *in vivo* RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. EMBO J, 2010, 29: 3301–3317
- 49 Muerdter F, Guzzardo P M, Gillis J, et al. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. Mol Cell, 2013, 50: 736–748
- 50 Lim A K, Kai T. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila* melanogaster. Proc Natl Acad Sci USA, 2007, 104: 6714–6719
- 51 Pe, J W, Kai T. A role for vasa in regulating mitotic chromosome condensation in Drosophila. Curr Biol, 2011, 21: 39-44
- 52 Xiol J, Spinelli P, Laussmann M A, et al. RNA clamping by vasa assembles a piRNA amplifier complex on transposon transcripts. Cell, 2014, 157: 1698–1711
- 53 Siomi M C, Mannen T, Siomi H. How does the royal family of Tudor rule the PIWI-interacting RNA pathway? Genes Dev, 2010, 24: 636–646
- 54 Chen C, Nott T J, Jin J, et al. Deciphering arginine methylation: Tudor tells the tale. Nat Rev Mol Cell Biol, 2011, 12: 629–642
- 55 Kirino Y, Kim N, de Planell-Saguer M, et al. Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. Nat Cell Biol, 2009, 11: 652–658
- 56 Nishida K M, Okada T N, Kawamura T, et al. Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. EMBO J, 2009, 28: 3820–3831
- 57 Kirino Y, Vourekas A, Sayed N, et al. Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. RNA,

2010, 16: 70–78

- 58 Liu H, Wang J Y, Huang Y, et al. Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. Genes Dev, 2010, 24: 1876–1881
- 59 Aravin A A, Naumova N M, Tulin A V, et al. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. Curr Biol, 2001, 11: 1017–1027
- 60 Stapleton W, Das S, McKee B D. A role of the *Drosophila* homeless gene in repression of Stellate in male meiosis. Chromosoma, 2001, 110: 228–240
- 61 Patil V S, Kai T. Repression of retroelements in *Drosophila* germline via piRNA pathway by the Tudor domain protein Tejas. Curr Biol, 2010, 20: 714–730
- 62 Zhang Z, Xu J, Koppetsch B S, et al. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. Mol Cell, 2011, 44: 572–584
- 63 Anand A, Kai T. The tudor domain protein Kumo is required to assemble the nuage and to generate germline piRNAs in *Drosophila*. EMBO J, 2011, 31: 870–882
- 64 Liu L, Qi H, Wang J, et al. PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition. Development, 2011, 138: 1863–1873
- 65 Huang H, Li Y, Szulwach K E, et al. AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification. J Cell Biol, 2014, 206: 217–230
- 66 Miyoshi K, Tsukumo H, Nagami T, et al. Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. Genes Dev, 2005, 19: 2837–2848
- 67 Klenov M S, Sokolova O A, Yakushev E Y, et al. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. Proc Natl Acad Sci USA, 2011, 108: 18760–18765
- 68 Sienski G, Dönertas D, Brennecke J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell, 2012, 151: 964–980
- 69 Huang X A, Yin H, Sweeney S, et al. A major epigenetic programming mechanism guided by piRNAs. Dev Cell, 2013, 24: 502–516
- 70 Klenov M S, Lavrov S A, Stolyarenko A D, et al. Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. Nucleic Acids Res, 2007, 35: 5430–5438
- 71 Le Thomas A, Rogers A K, Webster A, et al. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev, 2013, 27: 390–399
- 72 Rozhkov N V, Hammell M, Hannon G J. Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev, 2013, 27: 400-412
- 73 Brower-Toland B, Findley S D, Jiang L, et al. *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. Genes Dev, 2007, 21: 2300–2311
- 74 Lin H, Yin H. A novel epigenetic mechanism in *Drosophila* somatic cells mediated by Piwi and piRNAs. Cold Spring Harb Symp Quant Biol, 2008, 73: 273–281
- 75 Zhang D P, Xiong H L, Shan J F, et al. Functional insight into Maelstrom in the germline piRNA pathway: a unique domain homologous to the DnaQ-H 3'-5' exonuclease, its lineage-specific expansion/loss and evolutionarily active site switch. Biol Direct, 2008, 3: 48
- 76 Donertas D, Sienski G, Brennecke J. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. Genes Dev, 2013, 27: 1693–1705
- 77 Ohtani H, Iwasaki Y W, Shibuya A, et al. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. Genes Dev, 2013, 27: 1656–1661
- 78 Unhavaithaya Y, Hao Y, Beyret E, et al. MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. J Biol Chem, 2009, 284: 6507–6519
- 79 Aravin A A, van der Heijden G W, Castaneda J, et al. Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS Genet, 2009, 5: 1000764
- 80 Ross R J, Weiner M M, Lin H. PIWI proteins and PIWI-interacting RNAs in the soma. Nature, 2014, 505: 353-359
- 81 Mei Y, Clark D, Mao L. Novel dimensions of piRNAs in cancer. Cancer Lett, 2013, 336: 46-52

piRNA Biogenesis in Drosophila Ovary

WANG HaiLong, WANG XiaoNa & CHEN DaHua

State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

PIWI-interacting RNAs (piRNAs), a subset of small ncRNAs (24–30 nt in length), which are bound to and associated with Piwi proteins in animal gonads. piRNA are derived mainly from transposon-rich clusters, coding regions and intergenic loci. Loaded to Piwi proteins, piRNA-guided PIWI accesses promoters via base-pairing with the target loci, which is required for both initiating transcriptional silencing and post-translational modifications. In the model of germline of the Drosophila ovary, primary piRNAs are produced from piRNA cluster transcripts, which are then processed and loaded by Piwi and Aub. Additionally, the "Ping-Pong" model, has been proposed as an important mechanism to amplify piRNA biogenesis in the germline, but little is yet known about the detail and mechanisms of "Ping-Pong" cycle. In this review, we briefly summarize the piRNA biogenesis pathway, primary piRNA biogenesis and the Secondary piRNA biogenesis, and our current understanding of the piRNA pathway in transcriptional silence in Drosophila ovary.

piRNA, PIWI, Argonaute, "Ping-Pong" cycle, transcriptional silence

doi: 10.1360/N052015-00224