

# 转座子与转座子相关的非编码小RNA

李首道 周明兵\*

(浙江农林大学林业与生物技术学院, 亚热带森林培育国家重点实验室, 杭州 311300)

**摘要** 转座子是一种在真核生物基因组中大量存在的可移动DNA序列。非编码小RNA具有广泛的生物学功能,能够在不同层面影响基因的表达。在许多真核生物基因组中,转座子是非编码小RNA的重要来源。最近的一些研究发现,转座子及其衍生非编码小RNA在基因调控中发挥着重要作用,但是国内相关的综述却较少。所以该文从转座子与其衍生的非编码小RNA的关系,这些小RNA在基因组和生物进化中扮演的角色两方面,对现阶段关于转座子和非编码小RNA的相关研究成果进行综述。

**关键词** 转座子; 非编码小RNA; 微小RNA

## Transposable Element and Its Related Non-Coding Small RNA

LI Shoudao, ZHOU Mingbing\*

(State Key Laboratory of Subtropical Silviculture, College of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou 311300, China)

**Abstract** TEs (transposable elements) are a large number of mobile DNA sequences in eukaryotic genome. The non-coding small RNA has a wide range of biological functions and can affect gene expression at different levels. In many eukaryotic genomes, transposons are important sources of non-coding small RNAs. Recent studies have also found that transposons and their derived non-coding small RNAs play an important role in gene regulation. In this paper, the roles of transposons and their derived non-coding small RNAs in plant and animal genomes are discussed from the following aspects: the relationship between transposons and their derived non coding regulatory small RNAs, the roles of transposons and non-coding regulatory small RNAs in genome and biological evolution. The current research achievements of transposons and non-coding small RNAs are reviewed.

**Keywords** transposable elements; non-coding small RNA; mircoRNA

转座子(transposable elements, TEs)是一种可以自我拷贝,并在宿主基因组中移动,插入到染色体其他区域的DNA序列。转座子第一次是在1950年MCCLINTOCK<sup>[1]</sup>对玉米的籽粒颜色变化研究中被发现的。到1967年SHAPIRO<sup>[2]</sup>在对大肠杆菌的半乳糖操纵子的研究中发现转座子插入序列之后,转座子的概念才被广泛接受。随着测序技术的发展,研

究人员在越来越多的真核生物基因组中发现了转座子及类转座子序列。在植物金鱼草、矮牵牛、飞燕草、甜豌豆等中都发现了转座子的存在<sup>[3]</sup>。在动物中,比如大量哺乳动物、鱼类、鸟类的基因组中也发现了转座子序列。随着对转座子研究的不断深入,证据表明转座子对基因的进化、表达和功能等方面都有着重要的影响<sup>[4]</sup>。转座子按照其转座机理进

收稿日期: 2021-10-11 接受日期: 2021-11-08

浙江自然科学基金重点项目(批准号: LZ19C160001)和国家自然科学基金(批准号: 31870656、31470615)资助的课题

\*通讯作者。Tel: 0571-63732738, E-mail: zhomingbing@zafu.edu.cn

Received: October 11, 2021 Accepted: November 8, 2021

This work was supported by the Zhejiang Natural Science Foundation (Grant No.LZ19C160001), and the National Natural Science Foundation of China (Grant No.31870656, 31470615)

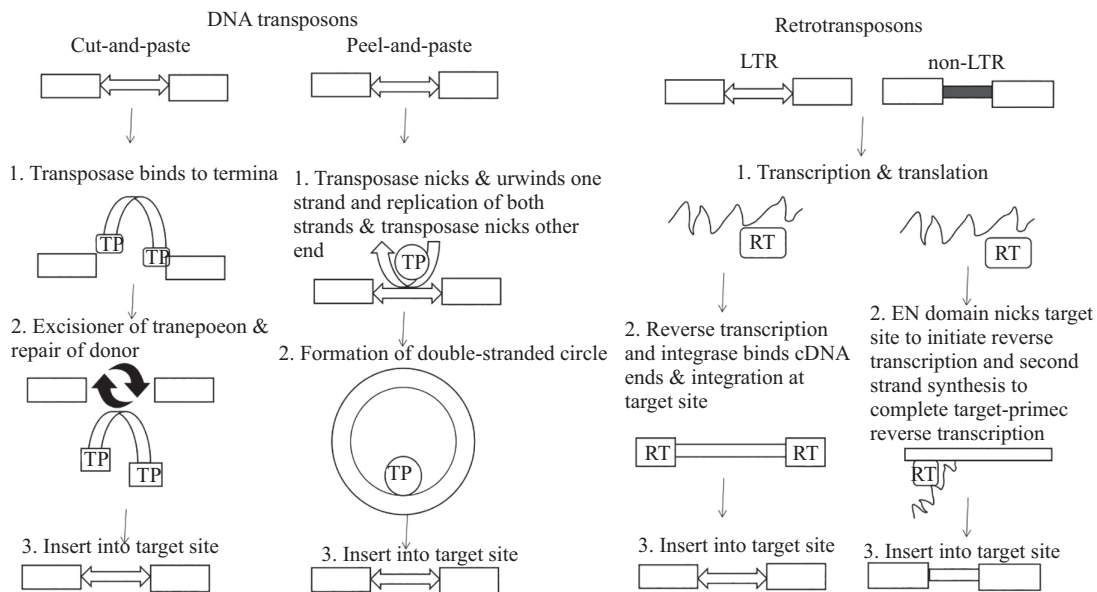
\*Corresponding author. Tel: +86-571-63732738, E-mail: zhomingbing@zafu.edu.cn

行划分,可被大致分为I类转座子和II类转座子。I类转座子,也被称为逆转录转座子,由反转录酶催化,合成cDNA,再插入到基因组中。转座过程为DNA-RNA-DNA的转座子。I类转座子根据是否具有长末端重复序列,可分为长末端重复序列(long terminal repeat, LTR)和非长末端重复序列(no-long terminal repeat, non-LTR)两类转座子。LTR逆转录转座子按照能否编码反转录酶,又可以分为自主性逆转录转座子和非自主性逆转录转座子。non-LTR类逆转录转座子按其结构的不同,可以分为长散布重复元件(long interspersed repetitive element, LINE)和短散布重复元件(short interspersed nuclear element, SINE)。II类转座子,也被称为DNA转座子,采用剪切-黏贴和脱落-黏贴两种机制,由转座酶催化,以DNA-DNA的方式进行转座的转座子。II类转座子的两端为反向末端重复序列(inverted terminal repeat, ITR),且通常含有能编码转座酶的基因序列,自身编码的转座酶能够识别转座子两端的ITR序列,从而介导DNA转座子的转座过程,将其插入到染色体其他位置。II类转座子按照是否能自主进行转座,可以分为自主性DNA转座子和非自主性DNA转座子。非自主性DNA转座子需要借助自主性DNA转座子才能进行转座<sup>[5]</sup>。两类转座子的大致转座过程如图1

所示。

非编码RNA是一种不编码蛋白质的RNA分子。它具有非常广泛的生物学功能。非编码调控RNA在基因转录水平、转录后水平和基因翻译水平都具有重要的调控作用<sup>[6]</sup>。按其长度是否超过200个核苷酸,可以分为长度大于200 nt的长链非编码RNA(long non-coding RNA, lncRNA)和长度小于200 nt的短链非编码RNA。后者包括微小RNA(mircoRNA, miRNA)、小干扰RNA(short interfering RNA, siRNA)、PIWI蛋白相互作用RNA(PIWI interacting RNA, piRNA)等。这些非编码RNA分子从不同水平和方面参与基因的表达和调控。

在最近的研究中,转座子与其相关非编码调控RNA的关系得到了进一步的揭示。真核生物中含量丰富的小RNA,比如siRNA、miRNA和piRNA大多数都来源于转座子。在植物中,这些转座子衍生的小RNA能通过调控DNA的甲基化过程来调控基因的表达。在动物中,转座子衍生的piRNA能够导致精子细胞中mRNA的降解从而影响精子成为成熟精子的过程,在小鼠中的TEs的表达也会因为piRNA介导的乒乓周期而被沉默。本文着重对转座子与miRNA、piRNA和siRNA的关系,以及他们对宿主因子的表达调控机制做一系列的归纳和总结。



TR为转录酶, RT为反转录酶, 双向箭头为转座子, 黑色长方形为无LTR的反转录转座子, 白色长方形为附近的基因。

TR is a transcriptase. RT is a reverse transcriptase. The two-way arrow is a TE. The black rectangle is a reverse TE without LTR, and the white rectangle is a nearby gene.

图1 两类转座子的转座过程(根据参考文献[5]修改)

Fig.1 Transposition process of two kinds of TE (modified from reference [5])

## 1 转座子与其来源的miRNA对基因表达的调控

miRNA是一种长度为20~24 nt的非编码内源性单链RNA,它在生物基因调控网络中发挥重要作用。首先,miRNA能够与靶mRNA反向互补配对,从而在转录水平和转录后水平抑制mRNA的表达<sup>[7-8]</sup>。在动物中,miRNA和mRNA碱基互补配对程度较低时,会抑制mRNA的翻译。在植物中,miRNA与mRNA碱基配对匹配程度较高,miRNA会介导靶mRNA的降解<sup>[7,9]</sup>。其次,越来越多的研究证明,miRNA可以靶向编码许多调控蛋白和转座子,同时也参与植物对热胁迫的响应,在植物的生长发育和抗逆胁迫中扮演重要的角色<sup>[10]</sup>。miRNA的成熟和功能行使过程需要两种蛋白家族的参与,即DICER蛋白家族和ARGONAUTE蛋白家族。DICER蛋白是一种核酸内切酶,它的同源蛋白称为DCL(dicer-like)蛋白。21 nt大小的miRNA就是由DCL1/4(dicer-like 1/4)蛋白加工miRNA前体产生的,这些被加工的蛋白前体具有茎环结构<sup>[11-12]</sup>。ARGONAUTE蛋白在各种生物体中广泛存在,是真核生物RNA干扰(RNA interference, RNAi)途径的核心蛋白,siRNA、miRNA、piRNA等调控小RNA在经过DCL核酸内切酶加工后,能够与ARGONAUTE蛋白结合形成RNA诱导的沉默复合体(RNA-induced silencing complex, RISC),参与转录水平、转录后水平的基因表达调控<sup>[13]</sup>。

许多miRNA来源于转座子序列<sup>[14]</sup>。在哺乳动物中,许多miRNA是TEs或TEs的衍生物,TEs参与了miRNA茎环结构的形成。人类基因组中的已知miRNA的20%都来源于MITE(miniature inverted-repeat transposable elements)转座子<sup>[15]</sup>。带有反向TEs序列的miRNA主要来自SINE和LINE转座子,与非倒置的TEs序列部分重叠的miRNA主要来自SINE、LINE和DNA转座子,而完全衍生自TEs的miRNA主要来自DNA转座子和SINE序列。在其他动物中,HARDING等<sup>[16]</sup>对热带非洲爪蟾三个发育时间点胚胎中的半个原肠胚中的所有小RNA进行了高通量测序,发现了一种新的miRNA,并把它命名为siteRNA,而这种新发现的miRNA来自于内含子中残留的转座子序列。在植物中,转座子是miRNA的重要来源。LI等<sup>[17]</sup>在对CDS区与TEs同源序列的相关的miRNA原始靶基因的研究中,证明了在一些植物中miRNA进化自TEs,TEs能够在转录本的非编码区形成折

叠,最后形成miRNA。OUYANG等<sup>[18]</sup>使用深度测序数据集研究了水稻中存在的两种典型miRNA途径,DCL1/4途径和AGO1(argonaute 1)介导途径。发现在AGO1免疫沉淀样品中miRNA发夹结构特异性富集,在DCL1/4基因敲除的基因型中miRNA发夹结构的特异性富集相对较少。这些miRNA发夹结构与水稻中一种丰富的DNA转座子——MITE转座子具有显著的同源性。这说明在一些植物中TEs也参与了miRNA发夹结构的形成。在竹子快速生长的根茎叶中确定了2 297个miRNA<sup>[19-20]</sup>。本课题组发现这些已鉴定的miRNA中29.1%来自MITE相关序列,其中hAT转座子超家族的MITE贡献最大。这些研究说明MITE转座子是miRNA的主要来源之一<sup>[21]</sup>。

miRNA能通过抑制TEs转座来维持基因组的稳定。miRNA对L1(LINE1)转座子的翻译抑制就是一个例子。L1是哺乳动物最丰富的自主转座元件,占哺乳动物基因组的20%。研究发现,miRNA会诱导L1转座子沉默并且对L1转座子在基因组稳定性和可塑性方面有重要影响。在一些体细胞中,miRNA已经进化出了L1功能调节机制补偿在非生殖细胞中piRNA的缺乏<sup>[22]</sup>。在这些细胞中,miRNA诱导的沉默复合物由miR-128引导并直接结合到L1的靶位点ORF2 RNA上,这种相互作用会导致L1转录本的不稳定,进而导致L1转座子的转录抑制,从而降低L1介导的诱变发生的概率。另一项研究也说明了L1转座子和miRNA表达可能存在的联系。OHMS等<sup>[23]</sup>对L1活跃和L1沉默的乳腺癌细胞进行了深度测序和全局分析,发现L1被沉默时miRNA的表达会增加。这说明L1有可能与miRNA的表达激活有关联。miR-845是一种在植物中高度保守的miRNA。在拟南芥的花粉中,miR-845能够与LTR反转录转座子的PBS(primer binding site)位点结合,通过RNA聚合酶V激发21~22 nt小RNA的积累,被激活的小RNA能够介导拟南芥基因组的剂量反应<sup>[24]</sup>。

根据关于miRNA起源的研究以及已发现的miRNA对基因及转座子的调控,似乎可以这样推测:某些生物中的TEs能够通过自身衍生的miRNA来调控自身以及其他基因的表达。一些研究为这个假说提供了证据,miR820是一种在小麦中的来源于TEs的miRNA,它的目标对象是DRM2(domains rearranged methyltransferase)基因转录本,它对DRM2基因的抑制会导致DNA甲基化的降低和对应TEs

转录产物表达量的增加<sup>[25]</sup>。这是TE衍生miRNA调节植物DNA甲基化例子。在人的大脑组织中,许多miRNA来源于L2(LINE2)转座子。这些TEs衍生的miRNA被定位在编码蛋白质基因的内含子上。人类基因组中的L2元件缺乏自己的启动子,在研究中发现lncRNA会利用它们的启动子来驱动插入其中的L2片段的pri-miRNA转录本的转录<sup>[26]</sup>。这是TEs衍生miRNA被lncRNA驱动从而转录的例子。在人类基因组中,一些TEs转录本衍生miRNA的靶基因被确定为嵌入3'UTR(3'untranslated region)的TEs序列,它们能够通过3'UTR中的TEs序列结合来调控基因的表达<sup>[27-28]</sup>,比如与AGO2蛋白相关的功能miRNA的靶位点来自3'UTR中的L2转座子序列<sup>[29]</sup>。

## 2 转座子与其来源的piRNA对基因表达的调控

piRNA是一种非编码小RNA,它的长度为24~31 nt。这种非编码小RNA大量存在于生殖细胞中,能够与生殖细胞中的调节蛋白PIWI蛋白相互作用,影响生物的生殖发育和表观调控<sup>[30-31]</sup>。piRNA的调控作用主要体现在对mRNA的降解和对TEs的转录抑制上,它与PIWI蛋白的相互作用能够引导PIWI蛋白降解目标mRNA,同时也能促进DNA甲基化来沉默TEs的表达。关于piRNA产生的主流观点是小的piRNA前体产生于长前体转录本的内切核裂解,之后内切酶ZUC(endonuclease zucchini)会对这些前体进行切割。切割后,piRNA的5'端会结合到PIWI蛋白上发挥作用。piRNA按照它在生物中产生方式的不同,可以分为初级piRNA和次级piRNA。初级piRNAs来源于单链piRNA前体分子的转录,由RNA聚合酶II催化。这种前体分子也被称为RNA簇。再由一种具有RNA结合活性且包含一个保守HMG-box(high mobility group box)结构域的蛋白MAEL(maelstrom, MAEL)运输到细胞质中<sup>[31-32]</sup>。在进入细胞质后,piRNA会通过MAEL蛋白和PLD6内切酶的相互作用产生初级piRNA的5'端<sup>[33-34]</sup>。piRNA前体会优先与PIWI蛋白结合从而达到稳定,然后PIWI-piRNA复合体会进行3'端的产生和修饰,产生初级piRNA<sup>[35-36]</sup>。初级piRNA可以引导PIWI蛋白介导的反义转录本的切割产生11~12 nt的次级piRNA,这个产生次级piRNA的过程就被称为乒乓周期。次级piRNA也可以反过来促进初级piRNA的产生。VANDEWEGE等<sup>[37]</sup>用统计学的方法

研究了狗、马和蝙蝠三种哺乳动物中TEs和piRNA在进化中的相互作用,他们发现被转录次数最多的TEs亚家族能够引发最剧烈的乒乓循环,这说明piRNA的产生与TE的转录有重要联系。在一些真核生物中,TEs就是piRNA的主要来源。GAN等<sup>[38]</sup>在对小鼠的A型精原细胞、粗胚精母细胞、圆形精子细胞进行深度测序和生物信息学分析后,发现TEs是前两个细胞时期piRNA的主要来源之一,小鼠中含量丰富的三种反转录转座子LTRs、LINEs和SINEs都产生了大量piRNA。研究人员在对人和果蝇的相关研究中也得到了相似的结论。

piRNA可以在转录水平和转录后水平沉默TEs。乒乓周期同时也是一种piRNA介导的转录后转座子抑制机制,piRNA识别转座子靶点,PIWI蛋白切割转座子结构域使其沉默,同时产生一个方向相反的新的piRNA。这些新产生的piRNA再与AGO3蛋白结合并且经过一系列ZUC内切酶和某些未知核酸参与的修饰后形成AUG-piRNA复合体。AUG-piRNA复合体能够识别和切割更多的聚类转录本,并且产生更多的反义piRNA<sup>[39-40]</sup>。许多物种中都发现了类似机制。比如在黑腹鼠和小鼠中,乒乓周期循环的功能被鉴定为转座子转录降解和小RNA扩增的有效手段<sup>[41-42]</sup>。HOUWING等<sup>[28]</sup>对三个成人睾丸样本的piRNA进行了高通量测序的数据也说明了这一点。他们发现LTR1/huers-P2家族的LTR元件可能在人睾丸中受到一个piRNA介导的乒乓周期的调控。piRNA也能在转录前沉默TEs。核PIWI蛋白与piRNA结合后会被诱导转入细胞核中,之后通过piRNA的碱基配对与新生的RNA结合,这一过程会导致DNA甲基化和组蛋白修饰从而抑制TEs的转录。许多研究都提供了证据,SASKIA等<sup>[43]</sup>发现在斑马鱼中甲基化L1转座子的分布和在相似染色体位置上的piRNA簇之间存在显著的正相关关系,从而揭示了piRNA-Piwi系统对L1(LINE-1)沉默的区域机制。缺乏小鼠PIWI家族蛋白MILI或MIWI2的小鼠突变体表现出L1转座子和细胞内LTR-反转录转座子的DNA甲基化缺失,之后会导致雄性生殖细胞中TEs转录激活<sup>[44]</sup>。缺乏MILI相互作用蛋白的小鼠也表现出DNA甲基化缺失和L1表达重新激活<sup>[45]</sup>。在剪切水平上piRNA也能导致TEs的沉默。在果蝇中的p转座子是一种能够诱发杂种不育的DNA转座子,piRNA能够通过剪切p转座子的最后一个内含子来抑制p转座子

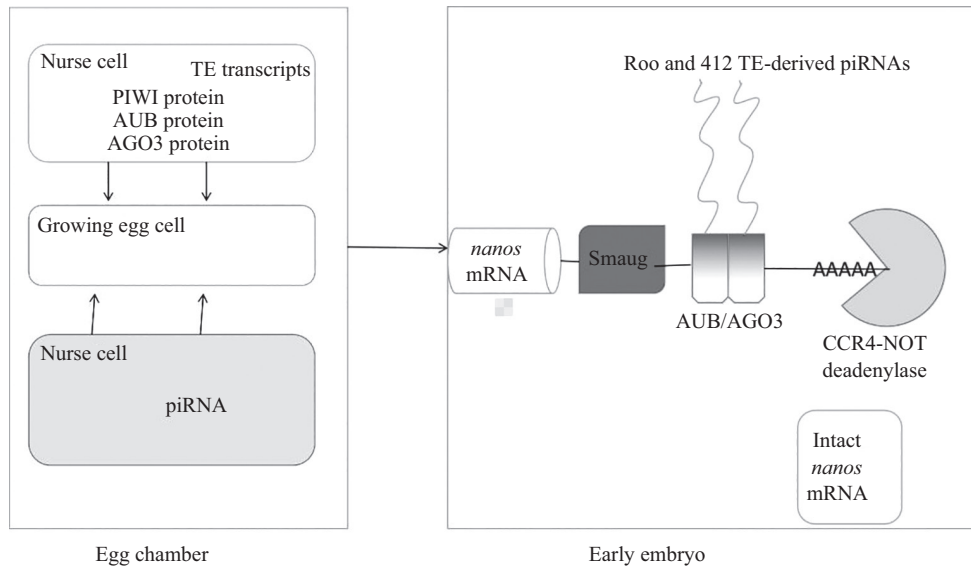


图2 果蝇TEs衍生piRNA的基因调控(根据参考文献[46]修改)

Fig. 2 Gene regulation of *Drosophila melanogaster* TE-derived piRNA (modified from reference [46])

的转座活性。piRNA介导的p转座子的转录抑制发生在剪切阶段,这可能是在染色体重塑期完成的<sup>[30]</sup>。

在果蝇中piRNA对mRNA的调控就是TEs和其衍生piRNA影响mRNA表达的一个典型例子。在果蝇的早期胚胎中TE衍生的piRNA能够结合到nanos mRNA的UTR(untranslated region)区,调节其表达(图1)。早期胚胎中,营养细胞将细胞质内容物沉积到卵细胞中,从而激活卵细胞和营养细胞中的TEs转录本。被激活的TEs转录本通过PIWI、AUB(aubergine)和dmAGO3(*drosophila melanogaster argonaute 3*)蛋白被加工成piRNA(图2)。在受精时,衍生自TEs的piRNA是在胚胎中母系遗传的。在早期胚胎中,roo系和412系转座子衍生的piRNA与AUB蛋白和dmAGO3蛋白结合,抑制母体遗传的nanos mRNA<sup>[46]</sup>。与这种mRNA结合的Smaug蛋白和piRNA有利于吸引CCR4-NOT(recruitment of the C-C chemokine receptor type 4-negative on TATA)灭活酶复合物,导致随后的翻译抑制(图2)<sup>[47-48]</sup>。小鼠精子产生的过程中也存在piRNA介导的TEs对mRNA的调控。ZHANG等<sup>[49]</sup>的研究发现,piRNA介导的mRNA衰变是圆形精子细胞变为成熟精子的必要过程。在未出生和新出生的雄性小鼠性腺的生殖细胞中,衍生自TEs的粗线前期piRNA(pre-pachytene piRNAs)大量存在。这种piRNA能被装载到小鼠PIWI蛋白家族的MIWI2蛋白和MILI蛋白上,这种MIWI2蛋白和MILI蛋白突变体能够分别阻止细线期、粗线期和偶线期的精子

发生,同时伴随着TEs的上调。有一些研究也发现piRNA对在精子形成阶段miRNA的翻译激活有潜在的间接作用<sup>[50]</sup>。在圆形精子细胞中,发现MIWI蛋白与内源性mRNA大量结合,并且在与MIWI蛋白结合的mRNA的切割位点中识别到了乒乓周期的信号位点。同时,此研究中也发现介导mRNA衰变的MIWI蛋白的剪切活性在这一过程中发挥重要作用。这说明mRNA的衰变过程中piRNA的产生机制可能是乒乓周期,也说明在圆形精子变为成熟精子的过程中TEs相关piRNA发挥着重要作用。在精子细胞伸长的过程中,大量的粗线期piRNA能够以精子形成时圆形和被拉长的精子细胞中的mRNA为靶点,形成CAF1(chromatin assembly factor-1)核酸酶的富集和MIWI蛋白的裂解,最后导致mRNA的降解<sup>[51]</sup>。总的来说,在动物的精子形成过程中,TEs衍生的piRNA能够影响相关mRNA的降解来调控精子的形成。

### 3 转座子与其来源的siRNA对基因表达的影响

siRNA,是一种长度在20~25 nt的双链RNA分子,它能够调控基因的表达和维持基因组的稳定性。它的主要功能有三个:一是通过RNA干扰(RNA interference, RNAi)切割mRNA转录本阻止其翻译,二是通过RdDM过程参与DNA的甲基化,三是参与了植物对热胁迫的响应。在植物和动物中,TEs是siRNA的主要来源。比如endo-siRNA(endogenous

short interfering RNA)是在非性腺细胞中发现的一种siRNA。在果蝇的非性腺细胞中,TEs已被证明是endo-siRNA的重要来源之一<sup>[52]</sup>。在本课题组关于毛竹LTR序列的研究中,发现15%的21 nt siRNA和18%的24 nt siRNA与LTR逆转录因子相关序列完全匹配,并且发现21 nt siRNA主要来源于LTR转座子的Tat和Oryco系,24 nt RNA主要来源于LTR转座子的Tat和Reina系<sup>[21]</sup>。在课题组对毛竹基因组的另一项研究中,研究人员在毛竹的节间组织中鉴定出23 154个siRNA,这些siRNA的67%来自于MITE转座子<sup>[53]</sup>。值得注意的是,许多TEs相关siRNA都在一定程度上影响着动物和植物中基因的表达。

在植物中,TEs和其衍生siRNA能够通过植物的RdDM途径参与DNA的甲基化过程从而影响基因的表达。植物的RNAi是指双链RNA(double-stranded RNA, dsRNA)被双链家族成员切割成21~25 nt的siRNAs,这些siRNA会被导入不同的途径中从而调控植物或动物转录后基因表达。在植物中,siRNA通过指导互补DNA序列的胞嘧啶甲基化活动,在基因沉默中发挥作用。这种核调控系统被称为RNA导向的DNA甲基化(RNA-directed DNA methylation, RdDM)。关于开花位点c(flow ering locus c, *FLC*)基因的调控就是一个例子。*FLC*基因是开花的主要阻抑基因,它受到春化(vernalization)机制的负调控。这种表型机制与1 224 bp的非自主Mutator-like转座子插入*FLC*的第一内含子有关,这种转座子的插入会抑制*FLC*等位基因的表达。同时也会使*FLC*基因受到来自基因组其他位置的同源转座子产生的siRNA介导的抑制性染

色质修饰<sup>[54-55]</sup>。另一个例子是花期的印记基因(flow ering wageningen, *FWA*),它在拟南芥的胚乳中特异性表达,在拟南芥的营养组织中不表达。*FWA*的组织特异性表达依赖于*FWA*启动子的去甲基化,*FWA*启动子由两个直接重复序列组成,其中一个序列与SINE逆转录转座子有关。*FWA*启动子的甲基化会导致局部异染色质的形成,进而导致转录沉默和SINE转座子相关siRNA的产生。这些siRNA也能通过RdDM途径在未甲基化的*FWA*基因中产生新的沉默<sup>[56]</sup>。另一个例子是,当拟南芥中的Athila反转录转座子被表观基因激活时,它的转录本被加工成siRNA,直接与*UBP1b*(3'UTR of the genic oligouridylylate binding protein 1B) mRNA结合,影响*UBP1B* mRNA的表达(图2)<sup>[57-58]</sup>。当野生型的表观基因被DNA甲基化沉默时,LTR反转录转座子将不会生成Athila反转录转座子衍生的siRNA,这样Athila转座子就不会影响*UBP1B* mRNA的功能(图3)<sup>[59-60]</sup>。当来自Athila转座子的表观遗传沉默被移除时,这种转座子会通过RDR6(RNA-dependent RNA polymerase 6)酶和DCL1/4蛋白衍生出21~22 nt的endo-siRNAs,这些siRNA中的Athila siRNAs和siRNA854会与蛋白AtAGO1(*Arabidopsis thaliana* AGO1)结合,通过其3'UTR中的四个结合位点调节*UBP1b* mRNA,抑制*UBP1b*的翻译<sup>[57,60]</sup>(图3)。转座子和siRNA在植物的配子产生和胚胎发育中也存在着RNA干扰和RdDM对基因的调控。在配子中的转座子活性抑制机制是防止突变保证宿主基因组稳定性的重要防御措施。DME(demeter)是一种螺旋发夹DNA糖基化酶,它去除甲基化的胞嘧啶,导致胚

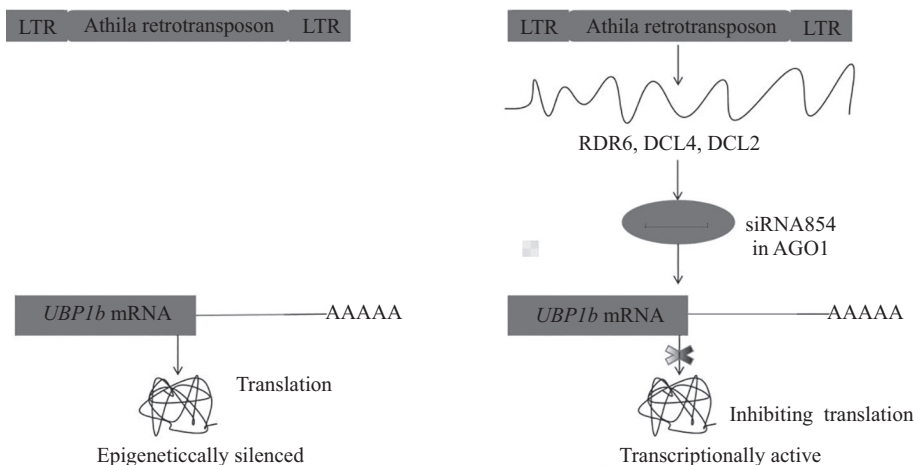


图3 拟南芥中TEs衍生siRNA的基因调控(根据参考文献[57]修改)

Fig. 3 Gene regulation of TE-derived siRNA in *Arabidopsis* (modified from reference [57])

乳的整体低甲基化<sup>[61]</sup>。DME的主动去甲基化激活转座子的表达,将转座子转录本引入RNAi通路,产生引导DNA甲基化的siRNAs<sup>[62]</sup>。胚乳中产生的siRNAs进入卵细胞,引导DNA甲基化,导致生殖细胞中转座子的沉默<sup>[63]</sup>。这一机制可能是抑制胚胎转座子活性的机制之一。此外,ANDREA等<sup>[64]</sup>发现,由RNAi降解mRNA和TEs后得到的产物siRNA会直接结合到AGO6蛋白中,之后AGO6蛋白又与TE染色质结合,使siRNA能在RdDM过程中发挥作用。在编码ATP依赖的染色质重塑基因突变的拟南芥突变体中,整体的DNA甲基化水平会降低,相关的TEs被激活。这些TEs的激活会伴随着内源性小RNA的产生,它们被称为easiRNA(epigenetically activated siRNAs)。这些easiRNA能够切割TEs转录本,导致TEs的沉默<sup>[65]</sup>。在ZHANG等<sup>[66]</sup>关于水稻TEs来源的siRNA的研究中,TEs衍生的siRNA815可通过RdDM途径诱导靶基因位点的DNA甲基化。综上所述,TEs相关siRNA在植物的DNA甲基化中发挥着重要的作用。

在热胁迫下的拟南芥幼苗中,一种名为ONSEN的Copia型反转录转座子不仅具有转录活性,而且能够合成染色体外DNA拷贝<sup>[67]</sup>。在siRNAs产生过程中,热诱导的ONSEN积累被激活。应激后,ONSEN转录本和体外染色体外DNA逐渐减少,并且在缺乏siRNA的受胁迫植物后代中观察到了新的ONSEN的高频率插入。还发现在无法产生siRNA的突变体植物中,如果在器官分化过程中诱导ONSEN转座,突变植物的发育过程中就会维持应激机制,这说明siRNA通路能够调节由环境引起的反转录转座子转座,从而调控植物的环境应激基因网络。

siRNA能够沉默TEs的表达。在果蝇中,endo-siRNA能够与AGO2蛋白结合,与这些小RNA结合的AGO2蛋白能够介导与endo-siRNA序列互补的TEs转录本的切割,从而导致TEs转录本的降解和表达的抑制<sup>[68-69]</sup>。在小鼠的生殖细胞中,endo-siRNA通路也被认为能够导致TEs的沉默,造成小鼠的不育<sup>[70]</sup>。在关于肌萎缩侧束硬化症(amyotrophic lateral sclerosis, ALS)的果蝇模型的一项研究中发现<sup>[71-72]</sup>,一旦endo-siRNA介导的沉默机制受损,endo-siRNAs及其相应的反转录转座子的表达水平会在TDB-43(TAR DNA binding protein 43)蛋白异位表达时发生改变,这种蛋白与ALS的引起有关。这些研究说明了动物细胞中的siRNA起到了TEs沉默的作用。在拟南芥的

DDMI(decrease in DNA methylation1)基因突变体中,DNA的甲基化水平降低,许多转座子被重新激活。这些转座子的激活伴随着21~22 nt的小RNA的产生,这些小RNA被称作easiRNAs(epigenetically activated siRNAs),这些easiRNAs能够切割TEs的转录本来沉默TEs的表达。值得注意的是,其中的一些easiRNA能够作用于mRNA,降低其表达水平。比如siRNA854能够作用于编码应激颗粒蛋白的UBP1(upstream binding protein 1)基因的转录本3'UTR区域。

## 4 结论与展望

TEs对基因组的影响大致可以分为三类:一是TEs的插入改变了编码区的阅读框和拼接模式,导致基因功能的丧失和表达模式的改变;二是TEs的插入诱导了新的调节元件的出现;三是TEs通过改变染色质状态,使基因表达沉默。与TEs相关非编码RNA对基因组的影响也遵循着这些规律。首先,从进化和来源的关系上讲,TEs是生物体内小RNA的主要来源。其次,从功能和表达的角度来看,TEs对基因组的一些调控是通过它的衍生小RNA完成的。比如在植物中TEs衍生的siRNA会通过RDdM途径调控DNA的甲基化来调控基因的表达。小鼠和果蝇中的TEs衍生piRNA能够影响精子形成时期mRNA的降解。最后,转座子和非编码RNA也会相互影响和调节。果蝇和小鼠中siRNA会导致TEs的沉默。piRNA能够在转录水平和转录后水平沉默TEs的表达。这些都说明了转座子和非编码RNA存在着重要且紧密的联系(表1)。

高速发展的测序技术与生物信息学手段是转座子和非编码RNA研究迅猛发展的关键。高通量测序技术和microRNA-Sep技术的结合,能够更好地解析miRNA家族成员之间的相互关系,鉴定miRNA在目标物种下的特定生物行为,识别新的miRNA分子,也成为了研究miRNA和转座子关系的有利工具。RNA-Sep技术的进一步应用,使对低丰度转录物的检测成为可能。通过RIP(RNA binding protein immunoprecipitation assay)技术和CHIP(chromatin immunoprecipitation assay)技术对与蛋白质结合的RNA分子的测序,可以得到蛋白质和RNA的特异性结合信息,RIP技术和CHIP技术能够更好地研究RNA诱导沉默复合物,揭示小RNA与其靶基因之间

表1 TE衍生小RNA对基因的调控和与TE的关系

Table 1 Regulation of genes by TE derived microRNAs and their relationship with TEs

小RNA	与TEs的关系	对TEs的沉默效应	对动物基因的调控	对植物基因的调控
Small RNA	Relationship with TEs	Silencing effect on TEs	Regulation of animal gene	Regulation of plant genes
miRNA	TEs are involved in the formation of miRNA hairpin structure. Many miRNAs are TE or derivatives of TE	Transcriptional inhibition of L1 transposon by miRNA	The target site of L2 transposon is TE sequence	Affect DNA methylation level
piRNA	TEs are the main source of piRNA in some special cell stages In some mammals, the TEs subfamily with the most transcription times can trigger the most intense ping-pong cycle	Regulation of LTR transposon by Ping-Pong cycle Inhibition of TE transcription by DNA methylation and histone modification Inhibition of p transposon transposable activity by piRNA silencing mechanism of Zebrafish L1 transposon	piRNA in <i>Drosophila</i> affects sperm formation	
siRNA	Endo-siRNA is derived from TEs sequence, and 24 nt siRNA is also derived from TEs siRNA in <i>Phyllostachys heterocycla</i> genome comes from LTR TEs	EasiRNA silences the expression of TEs Inhibition mechanism of transposable activity in gametes	Endo-siRNA induces infertility in mice	Inhibition of <i>FLC</i> gene and <i>FWA</i> gene expression Inhibition of transposons in plant gamete and embryonic development. It is involved in the response of plants to heat stress

的结合关系。

### 参考文献 (References)

- [1] MCCLINTOCK B. Chromosome organization and genic expression; proceedings of the Cold Spring Harbor symposia on quantitative biology, F, 1951 [C]. CSHL, doi:10.1101/sqb.1951.016.01.004.
- [2] SHEPHERD N S, SCHWARZ-SOMMER Z, VEL SPALVE J B, et al. Similarity of the *Cin1* repetitive family of *Zea mays* to eukaryotic transposable elements [J]. *Nature*, 1984, 307(5947): 185-7.
- [3] AMUTAN M, NYSSÖNEN E, STUBBS J, et al. Identification and cloning of a mobile transposon from *Aspergillus niger* var *awamori* [J]. *Curr Genet*, 1996, 29(5): 468-73.
- [4] 田海霞. 转座子在基因组和基因进化方面的研究进展[J]. 安徽农业科学(TIAN H X. Research progress of transposable elements in genome and gene evolution [J]. *J Anhui Agric Sci*, 2011, 39(20): 12018-20.
- [5] MAXWELL P H. Diverse transposable element landscapes in pathogenic and nonpathogenic yeast models: the value of a comparative perspective [J]. *Mobile DNA*, 2020, 11(1): 1-26.
- [6] 杨福兰, 饶周舟, 陈汉春. 非编码RNA与基因表达调控[J]. 生命的化学(YANG F L, RAO Z Z, CHEN H C. RNA regulation for gene expression [J]. *Chemistry of Life*), 2014, 34(1): 119-25.
- [7] AXTELL M J. Classification and comparison of small RNAs from plants [J]. *Annu Rev Plant Biol*, 2013, 64: 137-59.
- [8] CHEN X. Small RNAs and their roles in plant development [J]. *Annu Rev Phytopathol*, 2009, 25: 21-44.
- [9] KIM V N. MicroRNA biogenesis: coordinated cropping and dicing [J]. *Nat Rev Mol Cell Biol*, 2005, 6(5): 376-85.
- [10] ZUO Z F, HE W, LI J, et al. Small RNAs: the essential regulators in plant thermotolerance [J]. *Front Recent Dev Plant Sci*, 2021, doi: 10.3389/fpls.2021.726762.
- [11] 李蕊, 王恬, 文莉薇, 等. 植物Dicer-like功能的研究进展[J]. 中国农学通报(LI R, WANG T, WEN L W, et al. Research progress of Dicer-like in plants [J]. *Chin Agric Sci Bull*), 2015, 31(30): 210-4.
- [12] 于爱萍. AGO1和DCL1转录调控的研究[D]. 上海: 上海师范大学, 2006.
- [13] 谷少伟. Argonaute蛋白结构及其在植物中的研究进展[J]. 湖北农业科学(GU S W. The structure of Argonaute protein and research progress in plants [J]. *Hubei Agric Sci*), 2020, 59(8): 11-6.
- [14] SMALHEISER N R, TORVIK V I. Mammalian microRNAs derived from genomic repeats [J]. *Trends Genet*, 2005, 21(6): 322-6.
- [15] QIN S, JIN P, ZHOU X, et al. The role of transposable elements in the origin and evolution of microRNAs in human [J]. *PLoS One*, 2015, 10(6): e0131365.
- [16] HARDING J L, HORSWELL S, HELIOT C, et al. Small RNA profiling of *Xenopus* embryos reveals novel miRNAs and a new class of small RNAs derived from intronic transposable elements [J]. *Genome Res*, 2014, 24(1): 96-106.
- [17] LI Y, LI C, XIA J, et al. Domestication of transposable ele-



- ments into microRNA genes in plants [J]. *PLoS One*, 2011, 6(5): e19212.
- [18] OUYANG F, LUO Q J, ZHANG Y, et al. Transposable element-associated microRNA hairpins produce 21-nt sRNAs integrated into typical microRNA pathways in rice [J]. *Funct Integr Genomics*, 2013, 13(2): 207-16.
- [19] XU P, MOHORIANY I, YANG L, et al. Small RNA profile in moso bamboo root and leaf obtained by high definition adapters [J]. *PLoS One*, 2014, 9(7): e103590.
- [20] ZHAO H, WANG L, DONG L, et al. Discovery and comparative profiling of microRNAs in representative monopodial bamboo (*Phyllostachys edulis*) and sympodial bamboo (*Dendrocalamus latiflorus*) [J]. *PLoS One*, 2014, 9(7): e102375.
- [21] ZHOU M, ZHU Y, BAI Y, et al. Transcriptionally active LTR retroelement-related sequences and their relationship with small RNA in moso bamboo (*Phyllostachys edulis*) [J]. *Molecular Breeding*, 2017, 37(10): 1-11.
- [22] PEDERSEN I M, ZISOULIS D G. Transposable elements and miRNA: regulation of genomic stability and plasticity [J]. *Mob Genet Elements*, 2016, 6(3): 824-31.
- [23] OHMS S, RANGASAMY D. Silencing of LINE-1 retrotransposons contributes to variation in small noncoding RNA expression in human cancer cells [J]. *Oncotarget*, 2014, 5(12): 4103.
- [24] BORGES F, PARENT J S, VAN EX F, et al. Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis [J]. *Nat Genet*, 2018, 50(2): 186-92.
- [25] MCCUE A D, NUTHIKATTU S, SLOTKIN R K. Genome-wide identification of genes regulated in trans by transposable element small interfering RNAs [J]. *RNA Biology*, 2013, 10(8): 1379-95.
- [26] SHIN C, NAM J W, FARH K K H, et al. Expanding the microRNA targeting code: functional sites with centered pairing [J]. *Mol Cell*, 2010, 38(6): 789-802.
- [27] SPENGLER R M, OAKLEY C K, DAVIDSON B L. Functional microRNAs and target sites are created by lineage-specific transposition [J]. *Hum Mol Genet*, 2014, 23(7): 1783-93.
- [28] HOUWING S, KAMMINGA L M, BEREZIKOV E, et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish [J]. *Cell*, 2007, 129(1): 69-82.
- [29] PETRI R, BRATTÁS P L, SHARMA Y, et al. LINE-2 transposable elements are a source of functional human microRNAs and target sites [J]. *PLoS Genet*, 2019, 15(3): e1008036.
- [30] SIGURDSSON M I, SMITH A V, BJORNSSON H T, et al. The distribution of a germline methylation marker suggests a regional mechanism of LINE-1 silencing by the piRNA-PIWI system [J]. *BMC Genet*, 2012, 13(1): 1-7.
- [31] CASTAÑEDA J, GENZOR P, VAN DER HEIJDEN G W, et al. Reduced pachytene piRNA s and translation underlie spermiogenic arrest in Maelstrom mutant mice [J]. *EMBO J*, 2014, 33(18): 1999-2019.
- [32] VOUREKAS A, ZHENG K, FU Q, et al. The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing [J]. *Genes Dev*, 2015, 29(6): 617-29.
- [33] ZHENG K, XIOL J, REUTER M, et al. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway [J]. *Proc Natl Acad Sci USA*, 2010, 107(26): 11841-6.
- [34] IPSARO J J, HAASE A D, KNOTT S R, et al. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis [J]. *Nature*, 2012, 491(7423): 279-83.
- [35] NISHIMASU H, ISHIZU H, SAITO K, et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis [J]. *Nature*, 2012, 491(7423): 284-7.
- [36] SAXE J P, CHEN M, ZHAO H, et al. Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline [J]. *EMBO J*, 2013, 32(13): 1869-85.
- [37] VANDEWEGE M W, PLATT R N, RAY D A, et al. Transposable element targeting by piRNAs in *Laurasiatherians* with distinct transposable element histories [J]. *Genome Biol Evol*, 2016, 8(5): 1327-37.
- [38] GAN H, LIN X, ZHANG Z, et al. piRNA profiling during specific stages of mouse spermatogenesis [J]. *RNA*, 2011, 17(7): 1191-203.
- [39] HAN B W, WANG W, LI C, et al. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production [J]. *Science*, 2015, 348(6236): 817-21.
- [40] MALONE C D, HANNON G J. Small RNAs as guardians of the genome [J]. *Cell*, 2009, 136(4): 656-68.
- [41] BRENNECKE J, ARAVIN A A, STARK A, et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila* [J]. *Cell*, 2007, 128(6): 1089-103.
- [42] BRENNECKE J, MALONE C D, ARAVIN A A, et al. An epigenetic role for maternally inherited piRNAs in transposon silencing [J]. *Science*, 2008, 322(5906): 1387-92.
- [43] KURAMOCHI-MIYAGAWA S, WATANABE T, GOTOH K, et al. MVH in piRNA processing and gene silencing of retrotransposons [J]. *Genes Dev*, 2010, 24(9): 887-92.
- [44] REUTER M, CHUMA S, TANAKA T, et al. Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile [J]. *Nat Struct Mol Biol*, 2009, 16(6): 639-46.
- [45] UPADHYAY M, MARTINO CORTEZ Y, WONG-DEYRUP S, et al. Transposon dysregulation modulates dWnt4 signaling to control germline stem cell differentiation in *Drosophila* [J]. *PLoS Genet*, 2016, 12(3): e1005918.
- [46] MCCUE A D, SLOTKIN R K. Transposable element small RNAs as regulators of gene expression [J]. *Trends Genet*, 2012, 28(12): 616-23.
- [47] ZHANG P, KANG J Y, GOU L T, et al. MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes [J]. *Cell Res*, 2015, 25(2): 193-207.
- [48] WATANABE T, CHENG E C, ZHONG M, et al. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline [J]. *Genome Res*, 2015, 25(3): 368-80.
- [49] VOUREKAS A, ZHENG Q, ALEXIOU P, et al. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis [J]. *Nat Struct Mol Biol*, 2012, 19(8): 773-81.
- [50] GOU L T, DAI P, YANG J H, et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis [J]. *Cell Res*, 2014, 24(6): 680-700.
- [51] CZECH B, MALONE C D, ZHOU R, et al. An endogenous small interfering RNA pathway in *Drosophila* [J]. *Nature*, 2008, 453(7196): 798-802.

- [52] KAWAMURA Y, SAITO K, KIN T, et al. *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells [J]. *Nature*, 2008, 453(7196): 793-7.
- [53] ZHOU M, TAO G, PI P, et al. Genome-wide characterization and evolution analysis of miniature inverted-repeat transposable elements (MITEs) in moso bamboo (*Phyllostachys heterocycla*) [J]. *Planta*, 2016, 244(4): 775-87.
- [54] MICHAELS S D, HE Y, SCORTECCI K C, et al. Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis* [J]. *Proc Natl Acad Sci USA*, 2003, 100(17): 10102-7.
- [55] GAZZANI S, GENDALL A R, LISTER C, et al. Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions [J]. *Plant physiology*, 2003, 132(2): 1107-14.
- [56] BOSS P K, BASTOW R M, MYLNE J S, et al. Multiple pathways in the decision to flower: enabling, promoting, and resetting [J]. *Plant Cell*, 2004, 16(s1): S18-S31.
- [57] CALARCO J P, MARTIENSSEN R A. Genome reprogramming and small interfering RNA in the *Arabidopsis* germline [J]. *Curr Opin Genet Dev*, 2011, 21(2): 134-9.
- [58] HIROCHIKA H, OKAMOTO H, KAKUTANI T. Silencing of retrotransposons in *Arabidopsis* and reactivation by the ddm1 mutation [J]. *Plant Cell*, 2000, 12(3): 357-68.
- [59] SLOTKIN R K, VAUGHN M, BORGES F, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen [J]. *Cell*, 2009, 136(3): 461-72.
- [60] SLOTKIN R K. The epigenetic control of the Athila family of retrotransposons in *Arabidopsis* [J]. *Epigenetics*, 2010, 5(6): 483-90.
- [61] GEHRING M, BUBB K L, HENIKOFF S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting [J]. *Science*, 2009, 324(5933): 1447-51.
- [62] LU J, ZHANG C, BAULCOMBE D C, et al. Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of *Arabidopsis* seeds [J]. *PNAS*, 2012, 109(14): 5529-34.
- [63] HSIEH T F, IBARRA C A, SILVA P, et al. Genome-wide demethylation of *Arabidopsis* endosperm [J]. *Science*, 2009, 324(5933): 1451-4.
- [64] MCCUE A D, NUTHIKATTU S, REEDER S H, et al. Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA [J]. *PLoS Genet*, 2012, 8(2): e1002474.
- [65] CREASEY K M, ZHAI J, BORGES F, et al. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis* [J]. *Nature*, 2014, 508(7496): 411-5.
- [66] ZHANG H, TAO Z, HONG H, et al. Transposon-derived small RNA is responsible for modified function of WRKY45 locus [J]. *Nat Plants*, 2016, 2(3): 1-8.
- [67] ITO H, GAUBERT H, BUCHER E, et al. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress [J]. *Nature*, 2011, 472(7341): 115-9.
- [68] TALIAFERRO J M, ASPDEN J L, BRADLEY T, et al. Two new and distinct roles for *Drosophila* Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression [J]. *Genes Dev*, 2013, 27(4): 378-89.
- [69] MATSUMOTO N, NISHIMASU H, SAKAKIBARA K, et al. Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA [J]. *Cell*, 2016, 167(2): 484-97.e9.
- [70] WATANABE T, TOTOKI Y, TOYODA A, et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes [J]. *Nature*, 2008, 453(7194): 539-43.
- [71] MURCHISON E P, STEIN P, XUAN Z, et al. Critical roles for Dicer in the female germline [J]. *Genes Dev*, 2007, 21(6): 682-93.
- [72] KRUG L, CHATTERJEE N, BORGES-MONROY R, et al. Retrotransposon activation contributes to neurodegeneration in a *Drosophila* TDP-43 model of ALS [J]. *PLoS Genet*, 2017, 13(3): e1006635.