

DOI: 10.13376/j.cblls/2016082

文章编号: 1004-0374(2016)06-0640-05



陈月琴, 中山大学生命科学学院教授、博士生导师, 主要从事非编码 RNA 研究。1984 年毕业于西北大学生物系, 获理学学士学位; 1987 年毕业于北京大学生物系, 获理学硕士学位; 1996 年毕业于中山大学生命科学学院, 获博士学位。2001—2005 年, 先后在美国哈佛大学、美国 Walter Reed Army Institute of Research 做客座研究员。在 *Nature Biotechnology*、*Genome Biology*、*Autophagy* 等国际重要期刊上发表了一系列文章。获国家自然科学奖二等奖 1 项 (2007), 广东省自然科学奖一等奖 1 项 (2005); 广东省丁颖科技奖 (2013); 获国家发明专利 4 项; 主持国家及省部级科研项目近 20 项。现任中山大学有害生物控制与资源利用国家重点实验室副主任, 中山大学生物工程中心主任; 中国生化学会 RNA 专业委员会秘书长。

长链非编码 RNA 在植物生殖发育中的调控作用

张玉婵, 陈月琴*

(中山大学生命科学学院, 基因工程教育部重点实验室, 广州 510275)

摘要: 植物长链非编码 RNA (long noncoding RNA, lncRNA) 是调节植物生长发育的一类分布广泛和种类丰富, 且作用机制多样的非编码 RNA。虽然植物 lncRNA 还处于研究的起步阶段, 很多相关信息还不清楚, 但近期的研究显示, 它们在植物的发育, 尤其是生殖发育中具有重要功能。植物的生殖发育不仅是植物生长中最重要的一环之一, 也是作物育种中最关键的步骤, 因此, 发掘植物生殖发育相关的 lncRNA 具有重要的意义。现总结了已有的植物 lncRNA 的鉴定和特征分析, 着重介绍调节植物开花和有性生殖相关的 lncRNA 及功能, 以及 lncRNA 是通过何种作用方式参与到植物生殖调控中的。

关键词: 长非编码 RNA; 植物; 生殖发育

中图分类号: Q945.6; Q522 **文献标志码:** A

Long non-coding RNAs in plant reproductive growth

ZHANG Yu-Chan, CHEN Yue-Qin*

(Key Laboratory of Gene Engineering of the Ministry of Education, School of Life Science, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: Plant long noncoding RNAs (lncRNAs) play important roles in a wide range of biological processes through various mechanisms. Although the lncRNAs related researches are limited, recent progresses have shown that lncRNAs are crucial in regulating plant development especially in regulating reproductive growth. Plant reproductive growth is one of the most fundamental development processes, and is the critical step in crop breeding, thus it is important to explore reproductive development related lncRNAs in plants. In this review, we summarized the current knowledge of lncRNAs discoveries in plants including their feature analysis, with an emphasizing on the plant flowering and sexual reproduction related lncRNAs and the regulatory mechanisms by which lncRNAs

收稿日期: 2015-07-24

基金项目: 国家自然科学基金项目(91335104)

*通信作者: E-mail: lsscyq@mail.sysu.edu.cn

regulating plant reproduction.

Key words: long noncoding RNAs; plants; reproductive development

基因组是由蛋白质编码区域和非蛋白质编码区域组成。近年来的研究显示,在高等物种中基因组的非编码区域远远多于蛋白质编码区域,并且这些非编码区域转录出了大量的非编码RNA(ncRNA)。这些ncRNA按照长度可以分为短于200 nt的小非编码RNA(sRNA),以及长度超过200 nt的长非编码RNA(lncRNA)^[1]。sRNA已经被研究的较为系统,如microRNA等具有重要调控功能的sRNA。然而,作为发现较晚的lncRNA,人们对于它们的功能还所知甚少;但是,近期的研究显示,lncRNA同样具有重要的功能,并参与到调控动植物生长发育的各个方面^[2]。相较于动物lncRNA,植物lncRNA的研究较为初步,但越来越多的研究不断证明了植物lncRNA的重要作用。有趣的是,在目前鉴定到的植物lncRNA中,很大一部分都与植物的生殖发育有关。生殖发育是植物最重要的生命过程之一,包括从花序形成到成熟的种子形成的整个阶段,以植株从营养生长向生殖生长阶段的转变为开始。本综述将总结探讨植物lncRNA与植物生殖发育的关系及其在生殖发育中的重要调控作用。

1 植物中长链非编码RNA的鉴定及特征分析

在基因组中大规模鉴定lncRNA是发现新的lncRNA最为有效的方法,并且为不同物种中的lncRNA特征和潜在功能分析提供重要的信息。lncRNA可以产生自蛋白质编码基因的反向序列,称为天然反义转录本(NATs),以及基因间序列,称为基因间长非编码RNA(lincRNAs)。这些lncRNA既可以是末端有多聚腺苷酸尾的转录本,也可以是末端没有该修饰的转录本,它们需要由不同的测序方法来鉴定。在一些作为模式植物的植物中,lncRNA已经被系统鉴定过,如拟南芥中已有6480个lincRNA和3.7238万个NAT通过tiling array技术及RNA-seq技术鉴定出^[3-4]。我们在水稻中用全转录组测序的方法鉴定了2224个lncRNA,包括了有多聚腺苷酸尾的lncRNA以及没有该修饰的lncRNA,其中1624个为lincRNA,600个为NATs^[5]。在小麦中有125个对逆境响应的lncRNA通过计算机预测及实验验证而被鉴定出来^[6]。在玉米中,从全长cDNA序列中鉴定了一批潜在的玉米lncRNA^[7]。Wang等^[8]也从玉米和水稻中鉴定了一批lincRNA

和NAT,并分析了玉米和水稻中保守的lncRNA。在杨树中也有2542个干旱响应的lincRNA被通过RNA-seq方法鉴定出来^[9]。在土豆种属*Solanum commersonii*中有1.8882万个lncRNA通过计算机预测和RNA-seq技术被鉴定出^[10]。

这些不同物种中鉴定出的lncRNA在特征上都有一些相似点和一定程度上的差异,如它们最明显的相似点就是都具有高度的组织特异性表达模式,以及高度的不保守性。高度的组织特异性可能暗示了lncRNA在植物器官发育过程中起着特异的调节作用,而高度的不保守性说明了lncRNA是一类进化上很活跃的非编码RNA。此外,这些lncRNA在不同物种中的平均长度、外显子个数等特征有着细微的不同^[3-5]。由于不同物种的lncRNA很不保守,所以,对它们的功能预测较为困难。有趣的是,在水稻lncRNA鉴定过程中发现,有相当大比例的lncRNA在雄蕊中特异表达^[5]。这一点也同人类lncRNA相似,人睾丸中有大量lncRNA特异表达^[11]。虽然目前还不知道这是否为所有植物所共有的特征,但这一发现暗示着lncRNA在植物生殖发育或雄配子体建成过程中所起的重要作用。

2 长链非编码RNA与植物开花

lncRNA引起植物研究人员的兴趣和重视很大程度上来自人们对lncRNA和植物开花关系的发现。这一过程是通过lncRNA调节春化作用实现的。春化作用是指植物需要通过冬季的低温阶段实现从营养生长到生殖生长的时期转化^[12]。植物的FLOWERING LOCUS C (FLC)蛋白是调控春化作用的关键蛋白。FLC表达量高时会延迟植物的开花时间。在春化过程中,FLC受到表观遗传学负调控,即低温促进FLC上H3K4me3修饰的去除和H3K27me3修饰的添加,从而抑制FLC的表达,进而解除其对下游开花基因的抑制作用来实现开花转变^[13]。

有趣的是,近年来的研究发现至少有两种lncRNA参与调节FLC基因,从而参与到调节植物开花过程中。在拟南芥中,第一种是FLC的反向转录本NAT lncRNA COOLAIR;另一种是来自于FLC的内含子区域的lincRNA COLDAIR。COOLAIR是5'末端具有帽子结构,3'末端具有POLY A尾的lncRNA,而COLDAIR是5'末端具有帽子结构,但3'末端没

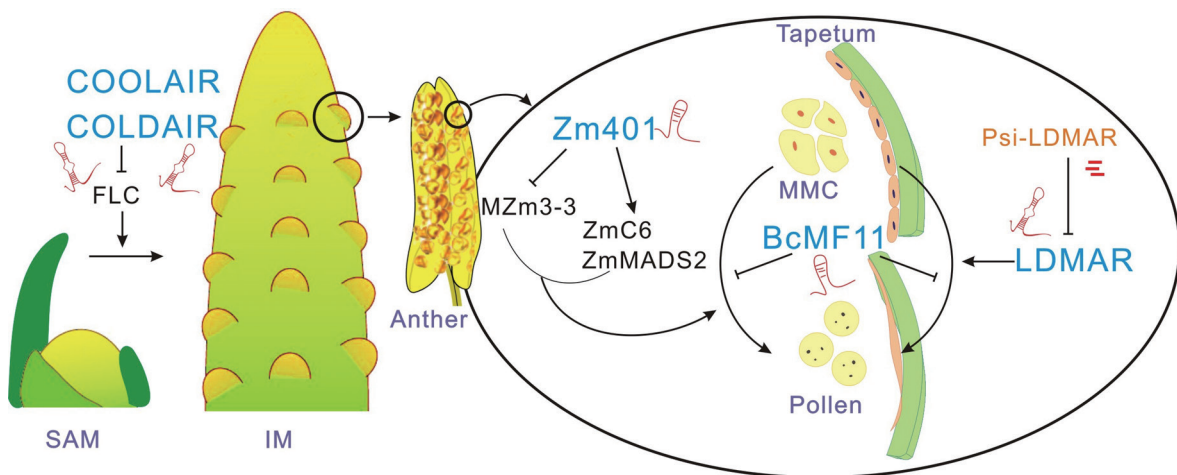
有 POLY A 修饰的 lncRNA^[14-15]。COOLAIR 的表达会通过吸引相关蛋白清除 FLC 上激活型组蛋白甲基标记, 从而引起 FLC 正向转录本的沉默。而 COLDAIR 则通过结合 PcG 蛋白复合体从而使 FLC 染色质组蛋白抑制型甲基化, 引起 FLC 的沉默。这两种 lncRNA 都通过沉默 FLC 而参与调节春化过程, 并影响植物营养生长向生殖生长的转变以及开花时间。生殖转变是植物生殖发育的第一步, lncRNA 在该过程中行使如此重要的调控作用, 暗示着 lncRNA 在植物生殖发育中的重要性。

3 长非编码RNA与植物有性生殖

植物有性生殖过程中, 首先花序将发育成熟, 期间雄配子体和雌配子体在花药和胚珠中分别发育成熟, 包括花粉母细胞通过有丝分裂和减数分裂形成花粉粒以及卵母细胞形成八核囊胚。随后将开花及授粉。精子通过花粉管进入囊胚, 受精后逐渐形成成熟的胚胎。有趣的是, 目前研究过的 lncRNA

中很多都是参与调控这个过程的。

LDMAR 是目前报道的调节水稻雄性不育的一个重要的 lincRNA^[16]。LDMAR 长 1 236 nt, 它参与长日照特异的雄性不育调控, 是保证长日照下正常花粉发育的重要因子。LDMAR 表达量降低会造成未成熟花药的过早程序性细胞死亡, 从而造成光敏型雄性不育^[16]。在其他物种中也发现有调节花粉发育的 lncRNA, 如在玉米中有一个雄蕊特异表达的 lncRNA Zm401。该 lncRNA 通过调控花药发育中的关键基因 Zm3-3、ZmMADS2 和 ZmC5, 而影响为花粉粒发育养分供应的绒毡层及花粉粒的发育。降低 Zm401 的表达会造成玉米的雄性不育^[17]。大白菜中也有一个名为 BcMF11 的 lncRNA 在调节花粉发育和雄蕊育性中行使功能。该 lncRNA 长 828 nt, 在整个花粉发育阶段都有表达。当降低其表达时, 绒毡层降解会延迟, 从而造成花粉粒无法成熟, 但对营养生长阶段没有任何影响。因此, BcMF11 是专门调节生殖发育过程的一类 lncRNA^[18-19] (图 1)。



蓝色字标识 lncRNA; 黑色字标识蛋白质编码基因; 红色字标识 sncRNA。SAM: 顶端分生组织; IM: 花序分生组织; MMC: 花粉母细胞。

图1 调控植物生殖发育的 lncRNA

虽然目前关于植物 lncRNA 的功能研究还较少, 作用机制也不太清楚, 但现已研究的植物 lncRNA 与植物有性生殖具有密切的联系, 相信进一步的研究将发现更多的调控植物生殖的 lncRNA。

4 长链非编码RNA调控植物生殖发育的作用方式

lncRNA 作为研究较晚的一类非编码 RNA, 不仅相关的功能研究较少, 人们对它们的作用机制了

解的则更少。lncRNA 并不像 miRNA 等其他研究较多的非编码 RNA 一样具有较为单一的作用方式, 目前发现的 lncRNA 有着十分多样的作用机制。因此, 对它们的作用机制研究是 lncRNA 功能研究的最大难点, 也是重点之一。在动物中 lncRNA 研究相对植物更加深入, 目前已经发现许多种完全不同的作用机制, 如顺式调控相邻基因表达、影响基因表观遗传修饰、作为招募蛋白质的“脚手架”、影响染色质结构、与小分子非编码 RNA 相互作用等

等^[20]。植物 lncRNA 作用机制研究还比较少, 但目前发现的作用方式与动物 lncRNA 作用方式比较相近。

目前发现的 lncRNA 调节植物生殖发育的作用机制大概有以下几种。一种, 是顺式作用调控临近基因表达。COOLAIR 调节 FLC 基因表达就是属于该类型的调控方式^[15,21]。这种作用方式是较早发现的, 并且这种类型的 lncRNA 较易进行功能预测。另一种, 是普遍的作用方式, 它是通过招募蛋白质改变靶向基因的甲基化修饰, 从而调节基因表达, 如 COLDAIR, 它可以结合 PRC2 复合体蛋白 CURLY LEAF (CLF), 招募该蛋白到 FLC 基因位点, 并引起 H3K27me3 染色质标记在该位点的沉积^[15]。这种结合的紧密程度在拟南芥植株暴露在冷气温中 20 d 后达到最大, 从而引起 FLC 基因的表达沉默。除了这两种常见的方式外, 另一种有趣的作用方式是通过产生小分子非编码 RNA 或者与小分子非编码 RNA 相互作用来调节基因表达, 如有一些 lncRNA 会像海绵一样吸附 microRNA, 使 microRNA 无法正常抑制其靶基因的表达而起到调控作用^[22]。最早发现的这类 lncRNA 为吸附 miR399 的 IPS1, 它在植物磷代谢平衡中起着调控作用^[23-25]。随后, Wu 等^[22] 在拟南芥和水稻中预测并验证了一批类似的 lncRNA, 分别对 miR160、miR166、miR156、miR159 和 miR172 有结合作用。我们在水稻中也大批量预测了潜在的该类 lncRNA, 并通过实验验证了其中两个 lncRNA 分别对 miR160 和 miR164 有吸附作用, 从而在调节水稻生殖发育中有潜在的功能^[5]。玉米中也鉴定出了一批该类 lncRNA^[26]。除了与 microRNA 互作外, 有些 lncRNA 还会与 siRNA 相互作用, 如 LDMAR 的启动子区会产生一个称为 Psi-LDMAR 的 siRNA, 该 siRNA 会使 LDMAR 的启动子区产生 RNA 引导的 DNA 甲基化, 从而沉默 LDMAR 的表达^[27]。另一个水稻光敏温敏雄性不育系中鉴定的 lncRNA P/TMS12-1 会产成一个 21 nt 的小非编码 RNA smR5864, 这可能是其行使功能的方式^[28]。还有科研人员从水稻生殖细胞发育和减数分裂相关的 AGO 蛋白 MEIOSIS ARRESTED AT LEPTOTENE1 (MEL1) 中分离出一批 21 nt 的 siRNA (phasiRNAs), 这些 siRNA 来自于 771 个花发育阶段特异表达的 lincRNA。这些 lincRNA 都具有 miR2118 的互补位点, miR2118 对它们的切割导致这些 phasiRNA 的产生, 从而在生殖细胞发育中起作用^[29]。

综上所述, lncRNA 通过多种多样的方式调节植物生殖发育。目前发现的调控机制还比较有限, 但是随着对植物 lncRNA 功能研究的深入, 人们将会更加系统地理解 lncRNA 的分类及调控基因表达的分子机制。

5 结语与展望

lncRNA 作为一种数量庞大、种类繁多的非编码 RNA, 它们在动植物生长发育中的功能已经逐渐被人们发掘, 并引起了越来越多的关注。植物生殖发育不仅是植物发育中最重要的步骤之一, 也是与人们生活最为相关的过程, 在植物育种、作物生产中都是最关键的步骤。因此, 植物生殖相关的 lncRNA 发掘和研究是十分重要和紧迫的。现有对 lncRNA 的研究已经逐步揭示了植物 lncRNA 在植物生殖发育中的重要调控作用。虽然由于 lncRNA 发现较晚、序列不保守、作用方式多样等原因造成了 lncRNA 的研究进展相较蛋白质编码基因和小非编码 RNA 更为缓慢, 但人们对 lncRNA 的研究热度却不断增长, 在未来数年将会有更多调控植物生殖发育的 lncRNA 被发掘。大规模测序使 lncRNA 的鉴定更为便捷, 但 lncRNA 的功能研究仍是亟待攻克的难题之一。研究技术和手段的快速发展将会给 lncRNA 研究带来新的转机 and 突破, lncRNA 的未知领域将会逐渐被人们探知。

[参考文献]

- [1] Kapranov P, Cheng J, Dike S, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*, 2007, 316: 1484-8
- [2] Kim ED, Sung S. Long noncoding RNA: unveiling hidden layer of gene regulatory networks. *Trends Plant Sci*, 2012, 17: 16-21
- [3] Liu J, Jung C, Xu J, et al. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*. *Plant Cell*, 2012, 24: 4333-45
- [4] Wang HP, Chung J, Liu J, et al. Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*. *Genome Res*, 2014, 24: 444-53
- [5] Zhang YC, Liao JY, Li ZY, et al. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. *Genome Biol*, 2014, 15: 512
- [6] Xin M, Wang Y, Yao Y, et al. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol*, 2011, 11: 61

- [7] Zhang M, Zhao H, Xie S, et al. Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proc Natl Acad Sci USA*, 2011, 108: 20042-7
- [8] Wang H, Niu QW, Wu HW, et al. Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lincRNAs associated with agriculture traits. *Plant J*, 2015, 84: 404-16
- [9] Shuai P, Liang D, Tang S, et al. Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. *J Exp Bot*, 2014, 65: 4975-83
- [10] Aversano R, Contaldi F, Ercolano MR, et al. The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. *Plant Cell*, 2015, 27: 954-68
- [11] Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*, 2011, 25: 1915-27
- [12] Amasino RM. Vernalization and flowering time. *Curr Opin Biotechnol*, 2005, 16: 154-8
- [13] Michaels S, Amasino RM. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, 1999, 11: 949-56
- [14] Helliwell CA, Robertson M, Finnegan EJ, et al. Vernalization-repression of *Arabidopsis* FLC requires promoter sequences but not antisense transcripts. *PLoS One*, 2011, 6: e21513
- [15] Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 2011, 331: 76-9
- [16] 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
- [17] Ma J, Yan B, Qu Y, et al. *Zm401*, a short-open reading-frame mRNA or noncoding RNA, is essential for tapetum and microspore development and can regulate the floret formation in maize. *J Cell Biochem*, 2008, 105: 136-46
- [18] Song JH, Wang CG. *BcMF11*, a novel non-coding RNA gene from *Brassica campestris*, is required for pollen development and male fertility. *Plant Cell Rep*, 2013, 32: 21-30
- [19] Song JH, Cao JS, Yu XL, et al. *BcMF11*, a putative pollen-specific non-coding RNA from *Brassica campestris* ssp. *chinensis*. *J Plant Physiol*, 2007, 164: 1097-100
- [20] Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*, 2011, 43: 904-14
- [21] Swiezewski S, Liu F, Magusin A, et al. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* polycomb target. *Nature*, 2009, 462: 799-802
- [22] Wu HJ, Wang ZM, Wang M, et al. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Plant Physiol*, 2013, 161: 1875-84
- [23] Liu C, Muchhal US, Raghothama KG. Differential expression of TPS11, a phosphate starvation-induced gene in tomato. *Plant Mol Biol*, 1997, 33: 867-74
- [24] 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
- [25] Martin AC, del Pozo JC, Iglesias J, et al. Influence of cytokinins on the expression of phosphate starvation responsive genes in *Arabidopsis*. *Plant J*, 2000, 24: 559-67
- [26] Fan C, Hao Z, Yan J, et al. Genome-wide identification and functional analysis of lincRNAs acting as miRNA targets or decoys in maize. *BMC Genomics*, 2015, 16: 793
- [27] 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
- [28] Zhou H, Liu Q, Li J, et al. Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. *Cell Res*, 2012, 22: 649-60
- [29] Komiya R, Ohyanagi H, Niihama H, et al. Rice germline-specific Argonaute MEL1 protein binds to phasiRNAs generated from more than 700 lincRNAs. *Plant J*, 2014, 78: 385-97