

植物长链非编码RNA的研究进展

张芳 董衡 刘丹丹 黄鹂* 曹家树

(浙江大学蔬菜研究所, 细胞与分子生物学实验室, 杭州 310058)

摘要 长链非编码RNA(long non-coding RNAs, lncRNAs)是真核生物中一类长度大于200个核苷酸、无蛋白编码能力或编码能力极低的RNA转录本。lncRNA种类和功能的多样性导致了对其进行研究的复杂性,特别是对植物中lncRNA的认识相当有限。该文就近年来植物中已发现的lncRNA的种类、相关转录酶、参与的生物学过程、发挥功能的分子机制以及其相关的研究策略等方面进行综述和展望,以期对深入认识植物lncRNA提供借鉴。

关键词 长链非编码RNA; 植物; 分子机制

The Research Progress of Long Non-coding RNA in Plants

Zhang Fang, Dong Heng, Liu Dandan, Huang Li*, Cao Jiashu

(Laboratory of Cell & Molecular Biology, Institute of Vegetable Science, Zhejiang University, Hangzhou 310058, China)

Abstract Long non-coding RNAs (lncRNAs), widely found in eukaryotes, are tentatively defined as a class of RNA transcripts, which are greater than 200 nt in length and have extremely low protein-coding ability or without protein-coding ability. The type and function diversity of lncRNAs make them difficult to be understood. Especially, the knowledge about lncRNAs in plants is still quite limited. In this review, the types of characterized lncRNAs, the RNA polymerases involved in lncRNAs transcription, the associated biological functions and the corresponding molecular mechanisms, as well as the methods and strategies for lncRNAs research were introduced and discussed.

Keywords long non-coding RNA; plant; molecular mechanism

长链非编码RNA(long non-coding RNA, lncRNA)是一类长度大于200个核苷酸、无蛋白编码能力或编码能力极低的RNA转录本^[1-3],具有诸如5'帽子、poly(A)尾巴及选择性剪接位点等与mRNA类似的结构特点^[4-5]。近年来,随着高通量测序技术的深入发展,越来越多的研究表明,lncRNA在真核生物中广泛存在,是构成转录组的重要组成部分^[1,6]。在以哺乳动物为主的研究中,现已明确lncRNA参与了包括干细胞命运决定、肿瘤发生等在内的众多生物学过程,并作为顺式作用元件或反式作用因子、染色质修饰蛋白复合物的骨架等,在不同水平大范围地影

响基因的表达^[2,7-9]。相比之下,我们对lncRNA在植物生长发育过程中所起作用的认识还非常有限。所幸,通过梳理近年植物中研究过的lncRNA,我们发现,植物lncRNA在结构、起源及作用的分子机制上,都与动物lncRNA具有一定的相似性,同时也显示出一些其特有的规律性^[10]。

1 植物lncRNA的种类

1.1 与其在基因组所处位置相关的分类

lncRNA可由基因组上的任意位点转录产生^[11],根据其在基因组中与蛋白质编码基因的相对位置,

收稿日期: 2014-10-09 接受日期: 2015-01-13

国家自然科学基金(批准号: 31372078)和浙江省重点科技创新团队项目(批准号: 2013TD05)资助的课题

*通讯作者。Tel: 0571-88982597, E-mail: lihuang@zju.edu.cn

Received: October 9, 2014 Accepted: January 13, 2015

This work was supported by the National Natural Science Foundation of China (Grant No.31372078) and the Key Technology Innovation Team of Zhejiang Province (Grant No.2013TD05)

*Corresponding author. Tel: +86-571-88982597, E-mail: lihuang@zju.edu.cn

网络出版时间: 2015-03-26 15:45 URL: <http://www.cnki.net/kcms/detail/31.2035.Q.20150326.1545.003.html>

一般将其分为正义lncRNA(sense long non-coding RNA)、反义lncRNA(antisense long non-coding RNA)、双向lncRNA(bidirectional long non-coding RNA)、基因内lncRNA(intronic long non-coding RNA)和基因间lncRNA(intergenic long non-coding RNA)五类,其中,基因间lncRNA也被称为大型介入性非编码RNA,即lincRNA(large intervening noncoding RNA)^[2,12]。在此基础上,Kung等^[13]将lncRNA的起源和功能引入作为分类需考虑的因素,对哺乳动物中的lncRNA进行了新的命名和划分,主要包括独立的lncRNA(stand-alone lncRNA)、天然反义转录本(natural antisense transcripts, NATs)、长内含子ncRNA、假基因、离散转录本、启动子相关转录本以及增强子RNA等。

独立的lncRNA也即上述所说的lincRNA,是一类产生于蛋白编码序列间隔区的独立转录单位,不与蛋白编码基因重叠^[14-16]。lincRNA在lncRNA中占了较大的比重,例如在拟南芥中发现的16 227条lncRNA中就有40%(6 480条)是lincRNA,其具体命名和编号可参见PlncDB(<http://chualab.rockefeller.edu/gbrowse2/homepage.html>)数据库中的信息;另外,Li等^[17]在对玉米的全基因组进行lncRNA鉴定时发现,在20 163条lncRNA中,绝大部分(93%)转录自蛋白编码基因间区域。

天然反义转录本(natural antisense transcripts, NATs)产生于其相应蛋白编码序列的同一或不同基因座,通过与正义转录本互补配对,使其不能正常翻译,在转录后沉默基因的表达^[18]。NATs一般起始于正义转录本编码区域(coding sequence, CDS)的起始密码子或终止子附近,与正义转录本完全或部分重叠。据报道,在哺乳动物(主要为老鼠和人)中,70%的正义转录本具有相应的NATs^[19-21],但大部分的NATs是否有生物学功能还有待研究。在植物中,目前比较明确的NATs有两个,一个是与春化相关的COOLAIR(COLD INDUCED LONG ANTISENSE INTRAGENIC RNA),它形成于植物开花抑制基因FLC(FLOWERING LOCUS C)的3'端,具有多种剪接方式,长度在400~750 nt不等^[1,22];另一个是与拟南芥生物和非生物胁迫响应相关的npc536(non-protein coding 536)(拟南芥基因ID为: AT1G67930),该基因全长597 bp,其3'端有71 bp与相邻反向基因AT1G67930的3'端重叠^[23]。

从上个世纪90年代开始,我们就陆续知道,诸

如小核仁RNA(small nucleolar RNA, snoRNA)和微小RNA(microRNA, miRNA)等小RNA可以转录自内含子。目前,在动物中通过高通量测序和生物信息学分析发现,许多由基因的内含子转录而来的长转录本,如在人的78 147条以及老鼠的39 660条内含子转录本中,均有98%以上属于长内含子ncRNA^[24]。这些长内含子ncRNA的功能目前只有少数被研究过,但人们已经明确,许多长内含子转录本在不同的组织器官中具有不同的表达模式^[13]。在植物中的研究表明,FLC基因的第二个内含子中隐藏的一个启动子可以响应低温信号,转录形成一个名为COLDAIR(COLD ASSISTED INTRONIC NONCODING RNA)的长内含子非编码RNA,参与植物春化过程。

在植物中发现的一些还不能够明确归类的lncRNA,统一归为其他类(表1)。

1.2 按其作用的分子机制进行的分类

按lncRNA发挥作用的分子机制,Wang等^[25]将lncRNA分为了以下四类,即信号分子(signals)、诱饵分子(decoys)、引导分子(guides)、骨架分子(scaffolds)。关于该分类方法的具体划分规则将在下文介绍lncRNA作用的分子机制部分详细展开。值得注意的是,某些lncRNA可能同时扮演多个角色。从而在这种分类方式中归属一个以上的类别。

无论是按位置关系还是按分子作用机制进行的分类,都是出于研究的需要,侧重点不同而已,两者并不冲突。本文参照两种分类方法,将植物中已发现的lncRNA作了相应划分(表1)。

2 植物lncRNA相关转录酶

目前一般认为,真核生物中存在三种RNA聚合酶(RNA polymerase),其中聚合酶I(Pol I)存在于核仁中,负责rRNA的转录;Pol III存在于核质中,转录少数几种基因,如tRNA以及5s rRNA;而大多数蛋白编码基因以及lncRNA由核质中的Pol II转录产生^[2,23,43]。

植物中的lncRNA也主要由Pol II转录形成^[44]。此外,植物中两种特有的RNA聚合酶——Pol IV和Pol V同样能转录形成lncRNA^[45]。Pol V产生一组独特的lncRNA,其可三磷酸化或加帽,但是没有poly(A)尾巴,它们在RNA介导的DNA甲基化(RNA-directed DNA methylation, RdDM)中起作用^[46-49],而Pol IV则可以产生小干扰RNA(siRNA)的前体^[45,50]。

此外,在拟南芥中还发现了一组lncRNA可能是由Pol III催化转录,其中一个名为*At8*的lncRNA被证实可由烟草核酸提取物通过Pol III在体外有效地转录得到^[51],但这是一个特殊例子还是一个普遍机制,目前尚无定论,还需要更多的研究支撑。

3 植物lncRNA参与的生物学过程

3.1 lncRNA与成花转变

成花转变是植物生殖发育的第一个关键环节,由多种途径控制,其中,春化途径中研究得最为清楚的一个关键基因是开花抑制因子*FLC*,植物感受低温后通过抑制*FLC*的表达促进开花。*FLC*的表达受一系列转录因子及组蛋白共价修饰的调控^[11,31,52]。最近研究发现,由其基因座鉴定出的两个lncRNA(*COOLAIR*和*COLD AIR*)也参与了这一过程。如前所述,*COOLAIR*是*FLC*的一类天然反义转录本,而*COLD AIR*转录自*FLC*的第一个内含子,其中*COOLAIR*的表达量在春化初期急剧上升,在低

温处理10 d左右达到最高峰,而*COLD AIR*在未春化之前就有少量表达,春化开始后,表达量持续增加,在低温春化约20 d时达到峰值,表明*COOLAIR*和*COLD AIR*在春化过程的不同阶段发挥作用^[1,31]。

根据剪切方式的不同,*COOLAIR*又分为近端和远端两种类型,即剪切位点距离*FLC*的最后一个外显子近的称为近端,远的称为远端。Liu等^[53]通过筛选突变体抑制子,得到了2个与3'端加工相关的突变体*CstF*(*Cleavage stimulatory Factor*)77和*CstF64*。研究表明,*CstF77*和*CstF64*通过清除近端*COOLAIR*的poly(A)位点附近组蛋白激活型甲基标记,沉默*FLC*正向转录本的转录,继而改变植株的抽薹时间^[52-53]。这一研究说明了在依赖于一定的条件或发育阶段的作用方式中,通过起始于基因3'端的反义转录来调控其相应的正义转录,是lncRNA一种普遍存在的作用机制。在*COOLAIR*抑制*FLC*的表达促进开花的这个过程中,所需要的仅是低温诱导的低水平的天然反义转录本*COOLAIR*的持续产生^[22]。

表1 植物中已发现的长链非编码RNA

Table 1 The discovered lncRNA in plants

按位置关系的分类 Types related to genome position	lncRNA名称 lncRNA	生物学功能 Biological Function	分子作用机制 Molecular function	按作用机制的分类 Archetypes of molecular function	参考文献 References
Stand-alone lncRNAs (lincRNAs)	<i>IPS1</i>	Responses to phosphorus stress	Inhibits the activity of miRNA-399	Signals, decoys	[26]
	<i>At4</i>	Responses to phosphorus stress	Inhibits the activity of miRNA-399	Signals, decoys	[27-29]
	<i>Npc48</i>	Participates in the development of leaves and flowers in <i>Arabidopsis</i>	Regulates the accumulation of a series of miRNAs and AGO1 (a member of Argonaute protein family)	Scaffolds	[23,30]
Natural antisense transcripts (NATs)	<i>COOLAIR</i>	Regulation of vernalization process	Regulates the epigenetic silencing of <i>FLC</i> , a flowering inhibition factor	Signals	[22,31]
	<i>Npc536</i>	Responses to biotic and abiotic stress	Regulates the translation of sense transcript	Signals	[23]
Long intronic ncRNAs	<i>COLD AIR</i>	Regulation of vernalization process	Recruits polycomb repressive complex -PRC2 to <i>FLC</i> locus to maintain the expression inhibitory state of <i>FLC</i>	Signals, guides	[1,31]
Others	<i>TERRA</i>	Maintains telomeric chromatin by promoting telomere methylation	Take parts in RNA-mediated DNA methylation process	Decoys	[32]
	<i>LDMAR</i>	Causes the pollen abortion of photoperiod-sensitive genic sterility rice	As the precursor of a small RNA	Signals	[33-35]
	<i>Zm401</i>	Functions in maize pollen development	The exact mechanism is unclear	Signals	[36-37]
	<i>Enod40</i>	Take parts in RNA binding proteins (RBP) relocation process	Binds to RBP directly	Guides	[38-42]

利用RNA干扰(RNAi)技术抑制COLDAIR的表达后同样发现植株开花延迟,但同时也发现,当温度回升后,FLC的mRNA含量又开始上升,这表明COLDAIR不仅在春化过程中抑制了FLC的表达,在春化过程后还起到了维持FLC沉默的作用^[31]。内含子lncRNA-COLDAIR调控春化过程中FLC的表达,表明内含子lncRNA对成花转变过程中的相关基因的转录和沉默起重要调控作用。

此外,在对拟南芥的全长cDNA文库进行生物信息学分析时得到的一条长983 nt的lncRNA——*Npc48*^[23,30]。*Npc48*在拟南芥的根、茎、叶、花等组织中均有表达^[30]。其过表达植株除表现出明显的叶锯齿(叶原基边缘细胞增殖的一种缺陷症状)以及莲座叶直径增大等营养生长的异常表型外,植株的开花时间也显著推迟。此种表型与*se(sucrose ester)*和*ago1*突变体的表型很相似,SE的作用是调控miRNA水平、影响miRNA的加工过程。而AGO1是大部分miRNA的目标剪切酶,但*npc48*的过表达对miRNA及AGO1的积累都无影响,推测其可能是通过miRNA以别的作用方式来参与转录后的调控^[23]。

由于成花转变在植物发育中占据着重要地位,对该过程的基因表达及调控网络的认识也相对清楚,因此,相较于其他生物学过程而言,对该过程lncRNA的研究较多,也相对容易。虽然目前的研究已经认识到,lncRNA在成花转变中起着重要的作用,但是,除了春化途径,lncRNA是否也在光周期途径、自主开花途径等其他成花转变机制中占据重要的位置还不得而知。

3.2 lncRNA与花粉发育

花粉发育不仅是植物生殖发育不可或缺的一环,也与农业生产上的雄性不育等现象密切相关。张启发等^[33]通过多年的研究,发现调控水稻光敏性雄性不育系‘农垦58S’育性的关键基因为一条水稻特异lncRNA——*LDMAR*(long-day-specific male-fertility-associated RNA)。该lncRNA长1 236 nt,其充分表达是长日照条件下水稻花粉正常发育所必需的。与野生型即可育系‘农垦58N’相比,不育系中的*LDMAR*产生了一个从C到G的单碱基突变(single nucleotide polymorphism, SNP),从而导致其在不育系中的表达量大幅降低,最终引起花药绒毡层的细胞程序性死亡,产生光敏雄性不育。这一结果对阐明光敏雄性不育的光调控途径的遗传和分子基础产

生了重要的影响,当然,该过程还需要进一步的研究,例如植物如何感知日长、光受体是什么等。

早年在玉米中发现了一条花粉特异表达的lncRNA——*Zm401*(*Zea mays* 401)。正向和反向遗传学研究表明,*Zm401*能够调控花粉发育关键基因(如*ZmMADS2*、*MZm3-3*和*ZmC5*等)的表达。在*Zm401*缺失突变体中,*MZm3-3*上调表达,而*ZmMADS2*和*ZmC5*都下调表达,导致小孢子和绒毡层发育异常,最终造成雄性不育。*Zm401*的过表达同样造成玉米花药的退化,进而严重影响花粉发育^[36-37]。在正义和RNAi构建的转基因植株中,*Zm401*的表达量都减少,且转录本的水平几乎一样,这也许是造成缺失和过表达具有相同表型的原因。*Zm401*在玉米、水稻、小麦以及小米中具有高度的序列保守性,且具有稳定的RNA二级结构,但*Zm401*到底是以RNA还是短肽在起作用仍存在争议^[37]。

本实验室在多年致力于白菜雄性不育调控机理的研究中,也发现了一条在‘矮脚黄’核不育两用系的可育系花粉中特异表达的lncRNA——*BcMF11*(*brassica campestris* male fertility gene 11)^[54]。*BcMF11*的异常表达造成绒毡层的提前降解和花粉败育,但与‘农垦58S’绒毡层的异常起始于花粉母细胞形成期及花粉能正常通过减少分裂的表型^[55]不同,*BcMF11*下调引起的绒毡层降解和花粉异常均起始于减数分裂时期,且未发现*BcMF11*引起的花粉败育依赖于光照条件。花粉作为雄配子体,包含了与雌配子体结合完成受精过程所需要的全部遗传信息,在植物的整个生殖周期中发挥着重要作用。但目前花粉发育过程中发现的lncRNA还很少,大规模鉴定lncRNA的工作将是研究者所期待的。

3.3 lncRNA与逆境胁迫

植物在生长发育过程中往往会经历诸如磷胁迫、干旱胁迫、盐胁迫等一系列非生物及生物胁迫。研究发现,在不同的逆境胁迫应答中,lncRNA也扮演了重要角色。其中,*TPSII*(tomato phosphate starvation induced 1)/*Mt4*基因家族是研究地较为清楚的胁迫响应lncRNA,该家族成员在磷胁迫下诱导表达。*IPS1*(*INDUCED BY PHOSPHATE STARVATION1*)最早在番茄中被发现,之后在拟南芥和水稻中也检测到其同源基因,过表达该基因能够导致植物根中磷含量的减少^[26]。

At4(*AtIPS1*),长747 nt,是*TPSII/Mt4*基因家族的

另一成员, 该基因最早在低磷肥条件下的莨菪苜蓿根中被检测到^[27], 后在拟南芥的根和茎中也发现其在磷胁迫下诱导表达^[29]。*At4*的缺失导致在磷胁迫条件下磷元素不能重新分配到根中, 而是在茎中大量积累, 最终致使茎:根的磷比率提高^[28]。而*At4*的过表达能够导致根中磷含量减少^[26], 同时, 分别过表达*At4*和*IPSI*与同时过表达两者产生相同的磷含量变化, 表明两者在磷胁迫下表现出一定的功能冗余, 而这一功能冗余现象有可能存在于*TPSII/Mt4*基因家族的各成员中^[26]。

除*TPSII/Mt4*基因家族以外, *npc536*, 即如上所述的拟南芥高尔基体转运复合物相关基因AT1G67930的天然反义转录本也对磷胁迫有响应, 同时, 其还在其他非生物胁迫如干旱、盐胁迫等条件下以及在不同激素处理和生物胁迫下, 表现出一系列的动态表达变化^[23]。在盐胁迫下, *npc536*的T-DNA插入突变体无明显表型; 而过表达植株则较野生型植株表现出明显的主根生长加快、侧根长度增加等现象。在干旱和冷处理下, *npc536*与其正义转录本AT1G67930在根中的表达呈一定的负相关性。然而, 在*npc536*过表达的植株中或是*npc536*突变体中, AT1G67930的表达并没有明显的变化。同时, 由于*npc536*包含一个在水稻中保守的短开放阅读框(short open reading frame, sORF), 所以它也可能是通过编码多肽来发挥作用。因此, 它到底是直接作为反义NAT(trans natural antisense transcript, trans-NAT)来调控AT1G67930的翻译还是其他方式, 还需要进一步研究^[23]。

3.4 植物lncRNA参与的其他生物学过程

TERRA(telomeric repeat-containing RNA)即端粒重复RNA, 是一类大小在100~9 000 nt之间的lncRNA, 从亚端粒区域转录而来, 由Pol II催化, 具有poly(A)尾^[56-57]。TERRA的表达有多种机制, 特别是端粒长度调控。在癌细胞中, 端粒越长, TERRA的RNA含量越高。它们是哺乳动物端粒染色质的特异结构成分^[57]。目前, 在拟南芥中发现也存在端粒重复RNA, 这些TERRA转录本中的一部分被加工为siRNA, 这些siRNA(主要是24 nt的siRNA)进一步通过促进端粒重复序列(CCCTAAA)(n)中不对称胞嘧啶的甲基化来维护端粒染色质, 这个过程依赖于RdDM途径。但我们目前还不明确具体哪些TERRA是被DCL3(Dicer-like 3)酶作用而降解成siRNA的, 以及普

遍在RdDM过程中起重要作用的AGO4复合物是否也参与了这个过程^[32]。

在黄瓜中发现一个位于6号染色体的端粒区域、名为*CsM10*(*C. sativus* male specific clone 10)的lncRNA, 在不同的组织中差异表达, 且与种子发育和光周期相关。同时, 它与生物胁迫信号相关的三个非编码RNA(GenBank Accession No: BX827695、D79216和BX819089)具有相同的179 bp长的保守区域^[58], 且这段保守区域都处在各基因的核心位置, 表明其也可能参与生物胁迫响应过程。后续的研究可能会采取RNAi技术来确定此段保守区域的具体功能以及过表达*CsM10*或抑制*CsM10*表达, 来观察其在各阶段植株的具体表型^[58]。

从以上的研究结果看, 如同在动物中的研究发现一样, 非编码端粒RNA在植物细胞端粒结构的形成过程中也起到了一定的调节作用, 同时这也为lncRNA调节染色体末端染色质重塑提供了有力的佐证。另外, 推测某些lncRNA在光周期调控、生物胁迫响应以及别的尚未发现的生物学过程中都发挥着不可估量的作用。

4 植物lncRNA发挥功能的分子机制

在动物中的研究表明, lncRNA可以作为蛋白复合体的骨架分子起作用^[59], 也可以通过招募转录因子或与转录因子发生特异性结合来激活/抑制基因的表达^[7,60-62], 还可以调控mRNA的选择性剪接^[63]及成熟mRNA的稳定性^[64-65]。根据目前的研究, lncRNA在植物中也基本以类似的机理发挥作用, 其分子功能也即分子作用模式可以归结为以下几个方面。

4.1 信号分子(signals)

lncRNA具有组织和发育阶段特异性表达以及响应不同刺激的特点, 说明lncRNA的转录受到了严格的调控。lncRNA的转录一般发生在生物体发育过程中特定的时间和特定的组织中, 其转录本有可能作为信号分子, 进一步调控其他基因的表达。有些lncRNA在这种作用模式中具有调控功能, 而其他的仅仅是转录的副产物。但在这两种情况下, 我们都可以仅仅根据相关lncRNA的表达来推断调控因子的染色质状态。动植物中的一些例子都说明, lncRNA作为信号标记基因调控的空间、时间以及表达情况。特别是, 在这种作用模式中lncRNA可以作为功能重大的生物学事件发生的标志^[25]。

总结来说,作为信号分子的lncRNA的功能就是作为转录活性的指示剂,独立于别的功能角色以及简单的一对一的关系。植物中的*COOLAIR*和*COLD AIR*是这类lncRNA的典型代表。

4.2 诱饵分子(decoys)

增强子和启动子的普遍转录^[66],暗示了lncRNA不管是在正向调控还是反向调控中都可能具有重要作用。这类非编码RNA调控转录的方式囊括了多种多样的机制,其中最主要的一种就是作为诱饵分子,这种作用模式的lncRNA可以通过招募其他RNA结合蛋白(RNA binding protein, RBP),共同实现对目标基因表达的调控。而这些RBP本身通常是转录因子、或是染色质修饰物或者是别的调控因子^[25]。前文所述植物中的*TERRA*可以归为此类。

此外,作为诱饵分子的lncRNA,亦可诱捕miRNA和剪切因子,即作为假基因或是miRNA的伪靶基因(target mimicry)而发挥生理作用,具备该作用的lncRNA也被称为竞争性内源的RNA(competing endogenous RNA, ceRNA)^[67]。植物中的*IPSI*、*At4*可以归为此类。

“伪靶基因”是指一类可以与互补的miRNA相互作用、抑制miRNA活性但同时又不被降解的内源非编码RNA。这个概念最早由Franco-Zorrilla等于2007年在研究*IPSI*对*miR-399*活性的抑制作用时提出^[26]。*IPSI*含有一个能够和*miR-399*互补配对、长为23 nt的基序,该基序在*miR-399*的剪切位点存在一个错配环,使得*IPSI*能够被*miR-399*错误地识别为它的靶基因而与之结合,但同时又不被*miR-399*所降解^[26-28]。*IPSI*通过捕获*miR-399*,而使其不能与它本身的靶基因*PHO2*(*PHOSPHATE 2*, 一个编码泛素结合酶相关蛋白的基因)的mRNA结合,导致*PHO2* mRNA积累, *PHO2*活性增加,而高活性的*PHO2*能够导致根中的两个特异的磷转运基因*Pht1;8*和*Pht1;9*的表达降低,从而减少根对磷的吸收^[26]。

*At4*与*IPSI*类似,也是通过伪靶基因的方式抑制*miR-399*、上调*PHO2*的表达量,导致根中磷含量的变化。此外,通过构建人工伪靶基因, Franco-Zorrilla等^[26]证明了这种作用方式是磷平衡调控过程中最常见的调控方式。

近来,从拟南芥和水稻中均鉴定得到一些可能为miRNA候选内源伪靶基因的lncRNA,且部分lncRNA(如*ath-eTM160-1*、*ath-eTM166-1*、*osa-*

eTM160-3)已被实验证明能有效地抑制相应miRNA的活性^[68]。

4.3 引导分子(guides)

lncRNA的第三种作用模式是引导分子,即指导核糖核蛋白(RNA结合蛋白)复合体在特定目标位点的定位^[25]。研究表明, lncRNA可以以顺式作用(cis)或反式作用(trans)指导基因表达,顺式作用是以共转录(关联Pol II)或是作为小调控RNA的互补目标的方式起作用;反式作用是以lncRNA结合到目标DNA,形成RNA:DNA双螺旋、形成RNA:DNA:DNA三螺旋或RNA识别特定染色质特性的表面复合物的方式起作用^[69-70]。由lncRNA形成的基因调控成分包括激活复合物、抑制复合物以及普通的转录因子集合体。这种作用模式的调控通常是利用lncRNA特定的空间,而非特定序列来实现的。这类lncRNA在发挥基因调节作用时,靶位点染色质局部结构的变化不仅与该区域内的改变相关,也可能与染色质远端结构的改变有关^[25]。

在大豆和蒺藜苜蓿中鉴定出的根瘤素基因*Enod40*^[40-41],由存在于根中柱鞘中的根瘤菌和共生反应阶段的根瘤原基中正在分裂的皮层细胞中的根瘤菌迅速诱导表达^[71]。*Enod40*在豆科植物中非常保守,并且也可从水稻和烟草等非豆科植物中分离得到^[42,72]。

在蒺藜苜蓿中发现,*Enod40*的RNA能够直接与根瘤中的MtRBP1(medicago truncatula RNA binding protein 1)蛋白结合,在MtRBP1从细胞核的核小点(一种动态亚核结构)到细胞质颗粒的重定位中扮演重要角色^[39]。此外,*Enod40*也被证明参与了另外两个小根瘤素酸性RNA结合蛋白——MtSNARP1(medicago truncatula small nodulin acidic RNA-binding protein 1)和MtSNARP2的重定位^[73]。

一般而言, lncRNA可以引导染色质变化。染色质修饰对于发育过程中组织特异性基因表达和染色体重塑都非常重要^[74-75],而特定基因位点的染色质修饰起始于染色质修饰复合物在该位点上的募集^[31]。有证据表明, lncRNA是调控染色质状态的关键因子,通常通过与染色质重塑复合物作用,将其募集到特定的基因座上而起作用^[76-78]。在植物中,由lncRNA参与调节的染色质修饰,已在拟南芥*FLC*基因上得到证实。

PRC2(polycomb repressive complex 2)蛋白复合

体能够通过三甲基化组蛋白H3K27位点抑制基因表达。通过体外RNA结合等实验证明,在春化过程中,*COLD AIR*与植物PRC2复合体中的亚基CLF(CURLY LEAF)特异结合,在募集PRC2到*FLC*的过程中起了最为直接和关键的作用^[1,32,79]。这一过程直接改变了*FLC*基因座的染色质结构,使得在此之后,直到双受精前的整个生长发育阶段*FLC*的表达均处于被抑制的状态,受精后,在胚胎发育过程中*FLC*才被重新激活^[80]。

4.4 骨架分子(scaffolds)

lncRNA可以作为一个中央平台,多个相关的分子成员在这里进行装配;这对许多生物信号的传递、分子间互作以及对信号本身的特异性和动态性的精确调控非常重要^[81]。传统上,我们认为蛋白质在各种骨架复合物中扮演着主要的角色^[82]。然而,近来的证据表明,lncRNA也可能扮演着类似的角色。

因此,按作用模式分类方法中的第四类lncRNA是骨架分子。这类lncRNA可能是功能上最复杂的一类,因为这类lncRNA具有多个不同的结构域,从而能够结合不同蛋白质或其他效应分子,且能够同时结合多个效应元件,这样可以把激活转录或抑制转录的效应元件聚集在相同的时间和空间里^[25]。一旦我们能够弄清这些骨架复合物是如何实现组装和调控的,那么我们就可以选择性地利用特定的信号组分和信号产物重塑细胞行为。

在植物中,Pol V可以产生siRNA及相关蛋白的骨架转录本,通过RdDM途径来修饰染色质^[45,83]。Argonaute蛋白ARGONAUTE4(AGO4)是RdDM途径中的一个关键因子,这些lncRNA可以和AGO4相互作用并且指导siRNA-AGO4复合物转运到它们在染色质上的目标位点^[45-46,84-85]。

4.5 小RNA生物合成的前体

研究表明,许多lncRNA是siRNA、miRNA等小RNA生物合成的前体。一些研究表明,在植物中,Pol IV可以产生siRNA生物合成的前体,而这些前体就是一些lncRNA^[45,50]。

在植物中,lncRNA作为小RNA生物合成前体的另一个例子是,前面所述的水稻光敏雄性不育关键调控基因*LDMAR*最终被发现通过剪切加工形成了长21 nt的小RNA *osa-smR5846w/m*^[34]。两个实验室的独立研究均显示,1 236 nt长的*LDMAR*可能是初级转录本,该初级转录本会被加工为一条长136 nt的中

间转录本,最终才形成21 nt的小RNA^[33-35]。目前还不清楚,如果21 nt的小RNA是最终的功能形式,它的作用靶标是什么? 1 236 nt及136 nt长的这两条转录本是否也具有功能? 另一方面,研究发现,那些从长的前体产生而来的小RNA的生物合成在体内通常依赖于RNAi通路中核心蛋白Dicer/Dicer-like或其他核酸内切酶^[35,86-87]。*osa-smR5846w/m*的生物合成是否也是如此,还有待于进一步阐明。研究还发现,那些从长的前体产生而来的小RNA通常具有RNAi通路中另一个核心蛋白Argonaute复合物装载偏好,从而将小RNA本身和其功能模式联系起来^[35]。*osa-smR5846w/m*是否与Argonaute蛋白有关系,目前还不得而知。此外,有意思的是,可育系‘农垦58N’和不育系‘农垦58S’的DNA甲基化水平有差异,推测在不育系中,*LDMAR*启动子区域的甲基化水平升高是导致其在长日照条件下表达下降的原因^[33]。那么,是否是*LDMAR*上的SNP突变导致的DNA甲基化水平变化也是值得进一步研究的内容。

除此以外,在玉米的一组全长cDNA文库中鉴定得到的1 802条lncRNA中,发现其中有60%可能是miRNA的前体序列^[88]。

事实上,一些lncRNA可能同时具有多个作用模式的特点,结合起来才能发挥其最终的生物学功能。例如,*COOL AIR*和*COLD AIR*响应环境的低温而转录,因此它们的转录是作为一个重大的生物事件的信号分子,为持续寒冷之后的正常开花作准备。开花抑制物的表观抑制则是通过*COLD AIR*结合PRC2而实现的,而此过程中lncRNA则是作为引导分子沉默*FLC*的,进而导致春化。因此,作为信号分子的lncRNA,通常兼有作为诱饵分子、引导分子和骨架分子的作用。

同时,由于我们对植物中已发现的这些lncRNA的某些具体的机制还不是了解地特别清楚,本文只是列出了其已知的分子作用模式;另一方面,对lncRNA的认识还处于探索阶段,对其分子作用机制的划分可能还存在不完善之处。

5 lncRNA的研究策略和方法

目前在动物和医学领域,lncRNA研究的基本策略一般分为三步:(1)lncRNA的鉴定和表达的高通量分析;(2)高通量分析所得lncRNA表达结果的验证;(3)目标lncRNA的生物学功能和作用机制研

究^[3]。其中,微阵列芯片(microarray)和RNA-seq是高通量分析常用的两种技术,目前已经有商品化的人、大鼠、小鼠的lncRNA芯片。对高通量分析得到的批量数据进行生物信息学分析是非常重要的,其中包括lncRNA表达差异分析、新lncRNA预测、lncRNA-mRNA共表达分析以及lncRNA生物学功能预测。高通量分析所得结果的进一步验证可使用较多的成熟技术,包括Northern印记、实时荧光定量PCR(qRT-PCR)、RNA荧光原位杂交(fluorescence *in situ* hybridization, FISH)等。对lncRNA的生物学功能研究,包括了两方面,即功能获得性研究和功能缺失性研究。其中,功能获得性研究主要通过导入过表达载体实现;功能缺失性研究则通过构建siRNA、短片段发夹RNA(short hairpin RNA, shRNA)等载体进行RNA干扰而实现。另外,关于lncRNA与蛋白质的互作研究,常用的有RNA-pulldown实验、RNA结合蛋白免疫沉淀(RNA-binding protein immunoprecipitation, RIP)、RNA纯化的染色质分离(chromatin isolation by RNA purification, ChIRP)、目标RNA的捕捉杂交分析(capture hybridization analysis of RNA targets, CHART)和交联免疫沉淀(crosslinking-immunoprecipitation, CLIP)等方法。

在植物领域,目前已在拟南芥、蕨藜苜蓿、水稻和玉米中进行过全基因组lncRNA的检索及相关研究^[17]。利用约200个拟南芥tilling array数据库和RNA-seq技术,在拟南芥中鉴定出了成千上万的lncRNA^[89]。近来关于拟南芥NATs的相关研究中,使用不同的技术平台,包括表达微阵列、特异链RNA-seq、qRT-PCR等发现了拟南芥中6 000多条lincRNA、37 238对转录本正反义对。研究人员进一步进行了正反义对在各种环境或胁迫下以及各生长发育阶段的表达谱分析,从而确定特定lncRNA在特定阶段和特定环境的特定功能^[90]。这个研究思路为在别的植物种类中鉴定lncRNA提供了借鉴。关于lncRNA突变体的构建目前还没有报道,相信随着植物中越来越多的lncRNA被发现和鉴定,在不久的将来,突变体也将成为研究lncRNA的重要途径。

6 结论与展望

随着高分辨率tilling array技术、RNA-seq等技术的出现和发展,已使得研究者们可以大量地鉴别出转录本^[17]。转录组分析表明,真核基因组90%以

上都是转录的^[91],但仅有1%~2%的基因组可以编码蛋白质^[92]。虽然已经在动物中通过高通量的方法发现了大量的lncRNA,但大部分的表达特点和功能还不是很清楚;且研究表明,有些高丰度表达的lncRNA保守性很差,而有些低表达的反反而具有重要的功能,所以试图用lncRNA存在的稳定性、lncRNA在物种间的保守性和lncRNA的表达水平等为标准来找到其中的规律意义不大,可行性也不高^[13]。

与动物的研究相比,对植物lncRNA的研究则更加有限,尽管我们已经通过高通量测序等方法在水稻、拟南芥等物种中发现了数量可观的植物lncRNA,也知道lncRNA参与了植物成花、花粉发育及胁迫响应等过程,但仍然有相当多的lncRNA需要挖掘,它们所扮演的生物学功能及分子功能也需要进一步探索。总体来说,目前植物lncRNA的相关研究还处于探索阶段,主要存在的问题有:(1)lncRNA基因芯片缺乏,除拟南芥以外还没有其他物种的商品化芯片;(2)基因组、蛋白质组数据库不完善;(3)植物lncRNA相关数据库寥寥无几;(4)适合植物lncRNA结构和功能预测的软件较少;(5)功能研究的思路和技术不成熟。这些问题对新lncRNA的发现及预测、lncRNA的生成、结构、功能等的研究造成了极大的阻碍。

针对以上问题,近期及将来在植物中开展lncRNA的研究需重点关注以下几点:(1)健全和完善植物基因组、蛋白组数据库;(2)对新一代转录组测序技术的进一步利用和推广;(3)加强生物信息学理论和方法在植物lncRNA的发现、表达和功能预测上的运用;(4)根据植物生长发育和细胞组成特点等设计其lncRNA的研究方法。

可以预见的是,未来在不同物种中以及在特定生物学过程中发现大量的lncRNA将不再是困难的事情;通过生物信息学分析,也能较容易进行lncRNA表达量及在不同样品间的表达差异分析。因此,未来植物lncRNA研究中的难点是对其生物学功能和分子功能进行研究,即某个或某类lncRNA,到底在哪个生物学过程或事件中起作用以及以怎样的分子机制起作用。

参考文献 (References)

- 1 Yamaguchi A, Abe M. Regulation of reproductive development by non-coding RNA in *Arabidopsis*: To flower or not to flower. J

- Plant Res 2012; 125(6): 693-704.
- 2 Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* 2012; 81: 145-66.
 - 3 Zhu J, Fu HJ, Wu YG, Zheng XF. Function of LncRNAs and approaches to LncRNA-protein interactions. *Sci China Life Sci* 2013; 56(10): 876-85.
 - 4 Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, *et al.* Novel RNAs identified from an indepth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res* 2004; 14(3): 331-42.
 - 5 Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, *et al.* The transcriptional landscape of the mammalian genome. *Science* 2005; 309(5740): 1559-63.
 - 6 Bardou F, Merchan F, Ariel F, Crespi M. Dual RNAs in plants. *Biochimie* 2011; 93(11): 1950-4.
 - 7 Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 2010; 142(3): 409-19.
 - 8 Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, *et al.* LincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 2011; 477(7364): 295-300.
 - 9 Guttman M, Rinn J. Modular regulatory principles of large non-coding RNAs. *Nature* 2012; 482(7385): 339-46.
 - 10 Au PC, Zhu QH, Dennis ES, Wang MB. Long non-coding RNA-mediated mechanisms independent of the RNAi pathway in animals and plants. *RNA Biol* 2011; 8(3): 404-14.
 - 11 Kim ED, Sung S. Long noncoding RNA: Unveiling hidden layer of gene regulatory networks. *Trends Plant Sci* 2012; 17(1): 16.
 - 12 Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell* 2009; 136(4): 629-41.
 - 13 Kung JT, Colognori D, Lee JT. Long noncoding RNAs: Past, present, and future. *Genetics* 2013; 193(3): 651-69.
 - 14 Guttman MI, Amit M, Garber C, French MF, Lin MF, Feldser D, *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009; 458(7235): 223-7.
 - 15 Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, *et al.* Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011; 25(18): 1915-27.
 - 16 Ulitsky IA, Shkumatava CH, Jan H, Sive, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 2011; 147(7): 1537-50.
 - 17 Li Lin, Eichten SR, Shimizu R, Petsch K, Yeh CT, Wu W, *et al.* Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol* 2014; 15(2): R40.
 - 18 Erdmann VA, Barciszewski J. Natural antisense transcripts mediate regulation of gene expression. In: *Nucleic Acids Sequences to Molecular Medicine RNA Technologies* 2012; 247-74.
 - 19 Katayama SY, Tomaru T, Kasukawa K, Waki K, Nakanishi M. Antisense transcription in the mammalian transcriptome. *Science* 2005; 309(5740): 1564-6.
 - 20 He YB, Vogelstein VE, Velculescu N, Papadopoulos, Kinzler KW. The antisense transcriptomes of human cells. *Science* 2008; 322(5909): 1855-7.
 - 21 Faghihi MA, Wahlestedt C. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol* 2009; 10(9): 637-43.
 - 22 Swiezewski S, Liu F, Magusin A, Dean C. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* 2009; 462(7274): 799-802.
 - 23 Ben Amor B, Wirth S, Merchan F, Laporte P, d'Aubenton-Carafa Y, Hirsch J, *et al.* Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res* 2009; 19(1): 57-69.
 - 24 Louro R, El-Jundi T, Nakaya HI, Reis EM, Verjovski-Almeida S. Conserved tissue expression signatures of intronic noncoding RNAs transcribed from human and mouse loci. *Genomics* 2008; 92(1): 18-25.
 - 25 Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011; 43(6): 904-14.
 - 26 Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* 2007; 39(8): 1033-7.
 - 27 Burleigh SH, Harrison MJ. The down-regulation of *Mt4*-like genes by phosphate fertilization occurs systemically and involves phosphate translocation to the shoots. *Plant Physiol* 1999; 119(1): 241-8.
 - 28 Shin H, Shin HS, Chen R, Harrison MJ. Loss of *At4* function impacts phosphate distribution between the roots and the shoots during phosphate starvation. *Plant J* 2006; 45(5): 712-26.
 - 29 Martín AC, del Pozo JC, Iglesias J, Rubio V, Solano R, de La Peña A, *et al.* Influence of cytokinins on the expression of phosphate starvation responsive genes in *Arabidopsis*. *Plant J* 2000; 24(5): 559-67.
 - 30 Hirsch J, Lefort V, Vankersschaver M, Boualem A, Lucas A, Thermes C, *et al.* Characterization of 43 non-protein-coding mRNA genes in *Arabidopsis*, including the *MIR162a*-derived transcripts. *Plant Physiol* 2006; 140(4): 1192-204.
 - 31 Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 2011; 331(6013): 76-9.
 - 32 Vrbsky J, Akimcheva S, Watson JM, Turner TL, Daxinger L, Vyskot B, *et al.* siRNA-mediated methylation of *Arabidopsis* telomeres. *PLoS Genet* 2010; 6(6): e1000986.
 - 33 Ding J, Lu Q, Ouyang Y, Mao H, Zhang P, Yao J, *et al.* A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci USA* 2012; 109(7): 2654-9.
 - 34 Zhou H, Liu Q, Li J, Jiang D, Zhou L, Wu P, *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(4): 649-60.
 - 35 Zhu D, Deng X. A non-coding RNA locus mediates environment-conditioned male sterility in rice. *Cell Res* 2012; 22(5): 791-2.
 - 36 Dai XY, Yu JJ, Ma JX, Ao GM, Zhao Q. Overexpression of *Zm401*, an mRNA-like RNA, has distinct effects on pollen development in maize. *Plant Growth Regulation* 2007; 52(3): 229-39.
 - 37 Ma J, Yan B, Qu Y, Qin F, Yang Y, Hao X, *et al.* *Zm401*, a short-open reading-frame mRNA or non-coding RNA, is essential for tapetum and microspore development and can regulate the floret

- formation in maize. *J Cell Biochem* 2008; 105(1): 136-46.
- 38 Girard G, Roussis A, Gulyaev AP, Pleij CWA, Spink HP. Structural motifs in the RNA encoded by the early nodulation gene *enod40* of soybean. *Nucleic Acids Res* 2003; 31(17): 5003-15.
- 39 Campalans A, Kondorosi A, Crespi M. *Enod40*, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. *Plant Cell* 2004; 16(4): 1047-59.
- 40 Yang WC, Katinakis P, Hendriks P, Smolders A, de Vries F, Spee J, *et al.* Characterization of *GmENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development. *Plant J* 1993; 3(4): 573-85.
- 41 Crespi MD, Jurkevitch E, Poiret M, d'Aubenton-Carafa Y, Petrovics G, Kondorosi E, *et al.* *enod40*, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. *EMBO J* 1994; 13(21): 5099-112.
- 42 Kouchi H, Takane K, So RB, Ladha JK, Reddy PM. Rice *ENOD40*: Isolation and expression analysis in rice and transgenic soybean root nodules. *Plant J* 1999; 18(2): 121-9.
- 43 Ganten D, Ruckpaul K. Encyclopedic reference of genomics and proteomics in molecular medicine. Springer Verlag Press, 2006, 1701-2.
- 44 Xin M, Wang Y, Yao Y, Song N, Hu Z, Qin D, *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.
- 45 Wierzbicki AT. The role of long non-coding RNA in transcriptional gene silencing. *Curr Opin Plant Biol* 2012; 15(5): 517-22.
- 46 Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 2008; 135(4): 635-48.
- 47 Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet* 2009; 41(5): 630-4.
- 48 Tran RK, Zilberman D, de Bustos C, Ditt RF, Henikoff JG, Lindroth AM, *et al.* Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. *Genome Biol* 2005; 6(11): R90.
- 49 Daxinger L, Kanno T, Bucher E, van der Winden J, Naumann U, Matzke AJM, *et al.* A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation. *EMBO J* 2009; 28(1): 48-57.
- 50 Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 2005; 120(5): 613-22.
- 51 Wu J, Okada T, Fukushima T, Tsudzuki T, Sugiura M, Yukawa Y. A novel hypoxic stress-responsive long non-coding RNA transcribed by RNA polymerase III in *Arabidopsis*. *RNA Biol* 2012; 9(3): 302-13.
- 52 Ietswaart R, Wu Z, Dean C. Flowering time control: Another window to the connection between antisense RNA and chromatin. *Trends Genet* 2012; 29(9): 445-53.
- 53 Liu FQ, Marquardt S, Lister C, Swiezewski S, Dean C. Targeted 3' processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* 2010; 327(5961): 94-7.
- 54 Song JH, Cao JS, 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(1): 21-30.
- 55 Shi Y, Zhao S, Yao J. Premature tapetum degeneration: A major cause of abortive pollen development in photoperiod sensitive genic male sterility in rice. *J Integr Plant Biol* 2009; 51(8): 774-81.
- 56 Azzalin CM, Reichenbach P, Khoraiuli L, Giulotto E, Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 2007; 318(5851): 798-801.
- 57 Schoeftner S, Blasco MA. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* 2008; 10(2): 228-36.
- 58 Cho JK, Koo DH, Nam YW, Han CT, Lim HT, Bang JW, *et al.* Isolation and characterization of cDNA clones expressed under male sex expression conditions in a monoecious cucumber paint (*Cucumis sativus* L. cv. Winter Long). *Euphytica* 2005; 146(3): 271-81.
- 59 Tsai MC, Manor O, Wan Y, Mosammamparast N, Wang JK, Lan F, *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010; 329(5992): 689-93.
- 60 Hung T, Wang YL, Lin MF, Koegel AK, Kotake Y, Grant GD, *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet* 2011; 43(7): 621-9.
- 61 Yap KL, Li SD, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, *et al.* Molecular interplay of the noncoding RNA *ANRIL* and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of *INK4a*. *Mol Cell* 2010; 38(5): 662-74.
- 62 Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA *gas5* is a growth arrest-and starvation- associated repressor of the glucocorticoid receptor. *Sci Signal* 2010; 3(107): ra8.
- 63 Jolly C, Lakhota SC. Human sat III and *Drosophila* hsr ω transcripts: A common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res* 2006; 34(19): 5508-14.
- 64 Gong C, Maquat LE. lncRNAs transactivate Staufen1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 2011; 470(7333): 284-8.
- 65 Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* 2008; 14(7): 723-30.
- 66 Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 2007; 130(1): 77-88.
- 67 Salmela L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A *ceRNA* hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011; 146(3): 353-8.
- 68 Wu HJ, Wang ZM, Wang M, Wang XJ. Widespread long noncoding RNAs as endogenous target mimics for MicroRNAs in plants. *Plant Physiol* 2013; 161(4): 1875-84.
- 69 Hung T, Chang HY. Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biol* 2010; 7(5): 582-5.
- 70 Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. *Science* 2010; 330(6004): 612-6.

- 71 Compaan B, Yang WC, Bisseling T, Franssen H. *Enod40* expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a pep tide. *Plant Soil* 2001; 230(1): 1-8.
- 72 Gulyaev AP, Roussis A. Identification of conserved secondary structures and expansion segments in *enod40* RNAs reveals new *enod40* homologues in plants. *Nucleic Acids Res* 2007; 35(9): 3144-52.
- 73 Laporte P, Satiat-Jeunemaitre B, Velasco I, Csorba T, van de Velde W, Campalans A, *et al.* A novel RNA-binding peptide regulates the establishment of the *Medicago truncatula*-*Sinorhizobium meliloti* nitrogen-fixing symbiosis. *Plant J* 2010; 62(1): 24-38.
- 74 Ho L, Crabtree GR. Chromatin remodelling during development. *Nature* 2010; 463(7280): 474-84.
- 75 Pfluger J, Wagner D. Histone modifications and dynamic regulation of genome accessibility in plants. *Curr Opin Plant Biol* 2007; 10(6): 645-52.
- 76 Koziol MJ, Rinn JL. RNA traffic control of chromatin complexes. *Curr Opin Genet Dev* 2010; 20(2): 142-8.
- 77 Nagano T, Fraser P. Emerging similarities in epigenetic gene silencing by long noncoding RNAs. *Mamm Genome* 2009; 20(9/10): 557-62.
- 78 Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007; 129(7): 1311-23.
- 79 Zhu Q, Wang M. Molecular functions of long non-coding RNAs in plants. *Genes (Basel)* 2012; 3(1): 176-90.
- 80 Yun H, Hyun Y, Kang MJ, Noh YS, Noh B, Choi Y, *et al.* Identification of regulators required for the reactivation of *FLOWERING LOCUS C* during *Arabidopsis* reproduction. *Planta* 2011; 234(6): 1237-50.
- 81 Spitale RC, Tsai MC, Chang HY. RNA templating the epigenome: Long noncoding RNAs as molecular scaffolds. *Epigenetics* 2011; 6(5): 539-43.
- 82 Good MC, Zalatan JG, Lim WA. Scaffold proteins: Hubs for controlling the flow of cellular information. *Science* 2011; 332(6030): 680-6.
- 83 Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X. The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. *EMBO J* 2011; 30(5): 814-22.
- 84 Lee TF, Gurazada SG, Zhai J, Li S, Simon SA, Matzke MA, *et al.* RNA polymerase V-dependent small RNAs in *Arabidopsis* originate from small, intergenic loci including most SINE repeats. *Epigenetics* 2012; 7(7): 781-95.
- 85 Zheng Q, Rowley MJ, Böhmendorfer G, Sandhu D, Gregory BD, Wierzbicki AT. RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. *Plant J* 2013; 73(2): 179-89.
- 86 Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, *et al.* A human snoRNA with microRNA like functions. *Mol Cell* 2008; 32(4): 519-28.
- 87 Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS. Small RNAs derived from snoRNAs. *RNA* 2009; 15(7): 1233-40.
- 88 Boerner S, McGinnis KM. Computational identification and functional predictions of long noncoding RNA in *Zea mays*. *PLoS One* 2012; 7(8): e43047.
- 89 Jin J, Liu J, Wang H, Wong L, Chua NH. PLncDB: Plant long non-coding RNA database. *Bioinformatics* 2013; 29(8): 1068-71.
- 90 Wang H, Chung PJ, Liu J, Jang IC, Kean MJ, Xu J, *et al.* Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*. *Genome Res* 2014; 24(3): 444-53.
- 91 Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, *et al.* Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 2008; 453(7199): 1239-43.
- 92 Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007; 447(7146): 799-816.