

植物长链非编码RNA的生物学功能研究进展

李宁^{1,2}, 王柏柯¹, 王娟¹, 黄少勇², 戴麒¹, 帕提古丽¹, 高杰², 余庆辉^{1,*}

¹新疆农业科学院园艺作物研究所, 乌鲁木齐830091

²新疆农业大学林学与园艺学院, 乌鲁木齐830052

摘要: 长链非编码RNA (long non-coding RNA, lncRNA) 是一类由RNA聚合酶II转录产生长度大于200 nt并且不具备编码蛋白功能的RNA, 广泛存在于动植物及病毒中, 具有表观遗传标记、发育阶段和组织特异性表达等特征。研究表明lncRNA不仅参与植物生长发育等生理过程, 而且还在植物抗病和响应逆境胁迫中起到重要作用。该文简要介绍近年来国内外有关lncRNA的形成机制、结构特征及其数据库, 主要阐述lncRNA在植物生长发育、生物和非生物胁迫等方面发挥的生物学功能, 对lncRNA存在的问题及其在作物生产育种中的利用前景进行了分析和展望。

关键词: lncRNA; 生长发育; 生物和非生物胁迫

随着高通量转录组测序和表达谱分析技术的快速发展, 动植物转录组数据的不断积累, 促进科研人员在探索生物生长发育、应激反应的转录调控等基因分子机制方面取得巨大进展(Cho 2018)。在植物RNA方面的研究表明, 非编码RNA (non-coding RNA, ncRNA) 特别是长链非编码RNA (long non-coding RNA, lncRNA), 已经成为真核生物转录的关键调节因子(Ariel等2015), lncRNA在植物生长发育、胁迫响应机制以及重要的生物代谢途径中扮演了不可忽视的关键角色(Sun等2018)。

如今lncRNA的鉴定和功能研究在科研领域越来越受到密切关注, 许多具有关键调控功能的lncRNAs被挖掘验证(Nejat和Mantri 2018)。目前, 功能性lncRNA的鉴定和效应机制研究虽处于起步阶段, 但前景不可限量。大规模转录组研究导致中心法则和传统的基因调控等经典理论受到前所未有的挑战。全基因组阵列分析和RNA-seq结果显示, 植物体内存在大量lncRNAs, 其通过DNA甲基化、组蛋白甲基化和转录干扰等机制介导基因表达, 在植物雄性不育、开花、转座子、生物和非生物胁迫等生物过程中起着重要的调节作用(Mishra和Bohra 2018)。本文主要概述了国内外对植物lncRNA的结构特征和分类, 以及生物学功能机制的研究报道, 以期为今后lncRNA的相关研究提供理论基础和思路。

1 lncRNAs概述

1.1 lncRNAs的来源和分类

根据分子生物学的中心法则, RNA经常被认为是连接DNA和蛋白质的桥梁。以“人类基因组计划”为代表的研究发现, 人类基因组中不超过2%~5%长度的基因可以编码蛋白质, 剩余的大部分是ncRNA (Snijders等2018)。ncRNA广泛存在于许多生物体中, 根据其功能不同可分为核糖体RNA (ribosomal RNA, rRNA)、转运RNA (transfer RNA, tRNA)、小核RNA (small nuclear RNA, snRNA) 和核仁小分子RNA (small nucleolar RNA, snoRNA) (Gloss和Dinger 2016)。又根据ncRNA分子链长短的不同(Guo等2016), 分为短链非编码RNA (small non-coding RNA, sncRNA) 和lncRNA。其中lncRNA是指由RNA聚合酶II转录产生的长度在200 nt以上的不具备编码蛋白功能的RNA, 开放阅读框 (open reading frame, ORF) 较短, 一般少于100个氨基酸, 在体内的表达水平很低。lncRNA与细胞核中转录合成蛋白质的mRNA在亚细胞定位上并不相同(Kornienko等2016), 大多数lncRNAs主要存在于细胞核和亚核室染色质中, 或者存在于2个亚细

收稿 2019-05-20 修定 2019-09-11

资助 国家重点研发计划(2017YFD0101906)和新疆农业科学院重点项目前期预研专项(xjzdy-003)。

* 通讯作者(yuqinghui98@sina.com)。

胞室中, lncRNA的二级结构较为复杂, 存在于细胞核和细胞质中, 具有时空表达特异性(Liao等2018)。

lncRNA的形成大致通过以下5种方式完成(Rai等2018): (1)蛋白质编码基因造成阅读框的插入, 插入的阅读框和之前的编码序列形成新的功能性lncRNA; (2)染色体重组后, 2个非转录并且相隔很远的序列区域合并到了一起, 产生1个含有多个外显子的lncRNA; (3)非编码基因通过反转录转座作用复制, 产生1个有功能的非编码逆转基因, 或产生1个无功能的反转录基因; (4) 2次连续重复事件在ncRNA内部形成相邻的重复序列产生lncRNA; (5)转位因子插入产生一个有功能的lncRNA。

迄今为止, lncRNAs基本是根据自身特征分类, 包括转录长度、蛋白质编码基因的关联、其他已知功能的DNA元素和重复等(Kashi等2016), 这些特征影响作用于生物基因组中的DNA序列、生化途径或稳定性、序列或结构保护、生物学状态、功能和靶向机制(表1)。

1.2 lncRNAs的结构与特征

lncRNA的初级结构就是核苷酸的排列顺序, 目前仅有少量报道涉及高级结构。由于lncRNA分子与RNA分子存在多方面的相似性, 导致lncRNA三级结构的预测一般是以RNA为参考对象(Lennox和Behlke 2016)。lncRNA基本特征是长度大于200 bp并且无蛋白质编码能力。lncRNA有较少的外显子、内含子, 无ORF及起始密码子和终止密码子; 大部分lncRNAs均由RNA转录酶II转录而来, 具有5'帽子和3'多聚腺苷酸加尾; lncRNA转录物比

mRNA短, 外显子也少, 估计这也是lncRNA不能编码蛋白质的主要原因。

1.3 lncRNAs的数据库

迄今为止, 大量lncRNAs在不同的植物类群中被发现, 例如, 拟南芥(*Arabidopsis thaliana*)、小麦(*Triticum aestivum*)、水稻(*Oryza sativa*)、玉米(*Zea mays*)、苜蓿(*Medicago sativa*)、番茄(*Solanum lycopersicum*)等。为了协助科研人员分析lncRNA, 已有一些信息丰富且极具科研价值的数据库开发出可存储lncRNAs序列并提供较全面的数据信息平台。CANTATA数据库收录了大部分植物的lncRNAs信息(<http://yeti.amu.edu.pl/CANTATA/>), 主要通过软件预测的方式识别植物中的lncRNA, 目前包含来自39个物种共239 631个lncRNAs, 是最大的植物lncRNA数据库(Quek等2015)。

GreenC数据库(http://greenc.sciencedesigners.com/wiki/Main_Page)收录了植物和藻类的lncRNAs信息(Quek等2015)。该数据库预测lncRNA的策略较为严格, 主要是通过与Swiss-Prot蛋白数据库比对, 筛选没有比对上的转录本的方法, 或者采用CPC软件预测蛋白编码潜能, 最终取2种方法的并集作为候选的lncRNA数据, 其lncRNA数据分为high confidence和low confidence。目前该数据库收录了49个物种的lncRNAs信息。

NONCODE是一个系统化的数据库(<http://www.noncode.org/>), 致力于提供非编码RNA-ncRNA尤其是lncRNA最完整的注释(Yi等2015)。通过收集和整合, 该数据库中的lncRNAs数量从

表1 长链非编码RNA的类型
Table 1 Classification of lncRNA

分类依据	lncRNA类型	参考文献
lncRNAs功能	诱饵分子型(decoy archetype)	Laurent等2015
	骨架分子型(scaffold archetype)	Takenaka等2016
	导向分子型(guide archetype)	
	信号分子型(signal archetype)	
在基因组上相对于蛋白编码基因的位置	增强子lncRNA (enhancer lncRNA)	Quinn和Chang 2016
	反义链lncRNA (antisense lncRNA)	Kopp和Mendell 2018
	双向lncRNA (bidirectional lncRNA)	Platt等2018
	内含子lncRNA (intronic transcript lncRNA)	
	基因间lncRNA (large intergenic noncoding RNA)	

527 336个增加到548 640个, 且数据库物种总数达到17个。

PlncRNADB (<http://bis.zju.edu.cn/PlncRNADB/index.php>)是一个专注分析植物lncRNA的数据库(Xuan等2015)。该数据库共包含了拟南芥、白杨(*Populus trichocarpa*)、玉米的lncRNAs信息。从GEO数据库中采集物种对应的RNA-seq测序reads, 通过特殊算法策略组装得到转录本序列, 识别到lncRNA之后, 通过catRAPID在线服务预测lncRNA和蛋白之间的相互作用, 这个数据库提供了研究植物中lncRNA的思路, 其预测lncRNA序列和lncRNA-protein相互作用的方法值得借鉴。

2 lncRNAs在植物中表达及可能的生物学功能

2.1 lncRNAs参与植物的生殖发育

植物lncRNA可以在转录水平、转录后水平或表观遗传学水平上发挥作用, 对植物生殖发育的影响表现在调节植物开花, 影响植株雌雄分化以及花粉发育等过程。破译lncRNA在植物生殖发育中的作用, 有助于更好地理解细胞器和核基因组间串扰受损导致的花粉不育(Kiegle等2018)。长日照特异性雄性不育相关RNA (long-day-specific male-fertility-associated RNA, LDMAR)是长度为1 236 nt的lncRNA, 研究表明其参与调控水稻生殖发育。在长光照条件下, LDMAR的表达维持水稻花粉正常发育。由于单碱基突变造成LDMAR二级结构的变化, 引起LDMAR启动子区域的甲基化程度升高, LDMAR的表达量下降, 花粉程序化死亡, 最终导致光敏不育系(Ding等2012)。

Kim等(2017)揭示了多梳组蛋白(Polycomb Repressive Complex 2, PRC2)在拟南芥春化发育过程中的分子作用。*FLC*是开花途径中的核心因子, COOLAIR与COLDAIR是*FLC*附近的2个lncRNAs, 通过结合PRC2的方式帮助PRC2维持在*FLC*染色体上的覆盖, 在低温处理时COOLAIR的启动子接受到低温信号, COOLAIR被表达, 并靶定*FLC*的3'端序列, 使其发生降解。Zhao等(2018)利用链特异性RNA-seq数据系统鉴定分析拟南芥的lncRNAs, 从中发现与拟南芥成花抑制基因*MAF4* (mads af-

fecting flowering 4)相关的1个调节开花时间的天然反义转录长链非编码RNA (natural antisense transcript-lncRNAs, NAT-lncRNAs), 命名为MAS, 并阐明了其正向调控正义链基因转录的作用机制。Zhang等(2014)研究发现水稻lncRNA表现出极高的组织特异性和发育阶段特异性, 其中一些lncRNAs为内源性竞争RNA, 削弱细胞内miR160、miR164的功能; XLOC_057324作为众多生殖相关lncRNAs中的一种, 和圆锥花序的发育及生殖能力紧密相关; 该研究还建立了水稻lncRNA的插入突变库。

lncRNA可以作为内源模拟靶标(endogenous microRNA target mimics, eTMs)来调节miRNA的功能, 通过互补序列与miRNA结合, 从而抑制miRNA的相互作用。采用RACE和过表达实验, 对白菜花粉发育中2种lncRNAs作为miR160的eTMs进行验证。不仅为白菜基因组提供了独特的注释资源, 并解析了白菜雄性育性潜在的可能调控机制(Huang等2018)。

2.2 lncRNAs参与植物的生长发育

在生长素运输和发育信号输出方面, lncRNA也发挥调控作用。在拟南芥中, 生长素调控的启动子环(auxin regulated promoter loop, *APOLO*)基因位于*PID*基因上游5 148 bp位置, 被RNA聚合酶II和V转录以响应生长素, 转录产生的lncRNA可以动态调控*PID*启动子的环化过程, 从而影响*PID*基因的表达模式(Ariel等2014)。基因表达的昼夜节律是由一系列生物过程的组合实现的, 在拟南芥的lncRNAs中鉴定到*CDF5*的天然反义转录本*FLORE*, 它属于与生物钟调控相关的lncRNAs。通过实时定量RT-PCR验证了*FLORE*的昼夜节律调控, 最终提出昼夜节律调节的新模块即*CDF5/FLORE*的天然反义转录对, 其具有保守相互抑制和独特的反式作用特性, 并能够精确调控昼夜节律变化, 最终促使植物启动开花(Henriques等2017)。

为深入研究*lncRNA1459*在番茄果实成熟中的功能, 利用CRISPR/Cas9基因编辑技术获得了*lncRNA1459*功能缺失突变体(Li等2018)。与野生型相比, 番茄*lncRNA1459*功能缺失突变体在成熟过程中受到显著抑制, 而且乙烯合成与番茄红

素积累也被大量抑制。在*lncRNA1459*敲除突变体中,许多与番茄果实成熟相关基因及lncRNAs的表达也发生了显著变化。在突变体纯系中有81个差异表达的lncRNAs与*lncRNA1459*表达呈正相关,31个差异表达的lncRNAs呈负相关。这些数据表明*lncRNA1459*在番茄果实成熟过程中对基因和lncRNAs的转录都起到了调节作用。

拟南芥的核斑RNA结合蛋白(nuclear speckle RNA-binding proteins, NSRs)参与调控可变剪切过程,可以结合mRNA以及*ENOD40*或*lnc351*等lncRNAs。其中,这些结合的lncRNAs被统称为ASCO-RNA (alternative splicing competitor RNA)。ASCO-RNA能与mRNA竞争性结合NSRs,干扰NSR蛋白对下游生长素响应基因的可变剪切,从而影响侧根的生长(Bardou等2014)。Hidden treasure 1 (HID1)是长度为236 nt的lncRNA,其T-DNA突变体*hid1*在持续红光条件下下胚轴伸长(Wang等2014)。HID1具有4个茎环结构(stem-loop structure, SL),其中SL2和SL4是HID1在红光下发挥功能的关键结构,HID1主要通过反式作用方式直接参与光形态建成关键抑制因子光敏色素互作因子(phytochrome-interacting factor 3, PIF3)的转录。

2.3 lncRNAs参与植物非生物胁迫

2.3.1 响应干旱胁迫

干旱是世界面临的主要生态环境问题,中国是世界荒漠化面积最大的国家。水分缺失是限制植物生长的主要因素之一,植物生长在干旱环境中,植物体内消耗的水量多于吸收量,最终导致干旱胁迫(Zandalinas等2018)。干旱胁迫时,植物的消耗量要大于吸收量,植物体内细胞水势和膨压降低,导致叶片凋萎、死亡和落果,甚至整株死亡。

目前研究表明,lncRNA与植物干旱胁迫反应相关。然而,lncRNA在植物抗旱-盐中的作用鲜有报道。At5NC056820是Liu等(2012)在拟南芥受干旱胁迫和脱落酸处理后发现的一种lncRNA,其表达变化显著,可能对干旱胁迫有响应。毋若楠等(2017)在拟南芥中导入At5NC056820,构建得到转基因株系A-3、A-7和A-8,其抗干旱性检测均优于野生型,尤其是游离脯氨酸含量是野生型2.2~2.5倍。最终证明At5NC056820一定程度上可以提高

拟南芥的抗旱性。Qin等(2017)发现了一种新的干旱诱导的lncRNA (drought induced RNA, DRIR)可提高拟南芥对干旱和盐胁迫的耐受性,其作为植物对干旱和盐胁迫响应的新型正调节因子,在干旱、盐胁迫以及脱落酸(abscisic acid, ABA)处理后被激活,表达量显著提升。突变体转录组分析,DRIR调控参与ABA信号、水运输和其他胁迫减轻过程的相关基因。Li等(2019)运用转录组链特异性测序(whole-transcriptome strand-specific RNA sequencing)得到大量与水稻干旱“记忆”相关的候选差异表达基因。这一研究不但证明水稻在合适的反复干旱处理条件下可以形成干旱“记忆”,还进一步对其机制进行探究,发现lncRNAs和植物激素(特别是ABA)均参与到这一短期干旱“记忆”的形成过程中,并激活光合作用、脯氨酸合成等代谢途径中“记忆”转录本的表达,从而提高植物在应对后期干旱胁迫中的能力。

利用RNA-seq技术研究棉花(*Gossypium hirsutum*) lncRNA响应干旱胁迫,Lu等(2016)鉴定筛选到基因间lncRNA、内含子lncRNA和反义lncRNA。作为lncRNAs的XLOC_063105和XLOC_115463可能分别在调节2个相邻编码基因CotAD_37096和CotAD_12502中起作用。通过研究在干旱胁迫和复水处理下lncRNA的特性和表达模式,发现lncRNA可能参与调节植物激素对干旱胁迫的反应途径。这些研究表明,lncRNA在植物对干旱胁迫的反应中起着不可忽视的作用。

2.3.2 响应盐胁迫

土壤盐渍化已成为世界土地治理面临的主要问题之一。目前,联合国教科文组织和粮农组织统计全球盐碱地面积9.54亿公顷,约占陆地面积的7%,其中9 913万公顷在中国(Ruiz等2016)。盐胁迫影响植物光合作用、种子萌发,一定浓度的盐离子会引起植物中毒、渗透胁迫和次生氧化胁迫,从而导致植物养分失衡,抑制植物生长,最终植物死亡。

虽然耐盐基因的功能验证工作进行的如火如荼,但盐胁迫相关的lncRNA研究鲜见报道。Huanca-Mamani等(2018)利用转录组分析盐胁迫处理后的智利耐盐玉米‘Lluteño’,共鉴定出48 345个不同的lncRNAs,其中41.9%属于‘Lluteño’玉米独有的ln-

cRNA转录组。鉴定并验证了一组848个应激反应反义转录物的存在和表达模式, 这些反义转录物似乎对与转录调节、应激反应、对非生物刺激的反应和参与烟碱代谢过程相关的基因有调节作用。Wang等(2015)利用全转录组分析盐胁迫下的苜蓿lncRNA, 值得注意的是根中盐胁迫差异表达的lncRNAs数量高于叶, 而在对渗透胁迫的反应中, 特异性变化的lncRNAs数量在叶中远远大于根。Deng等(2018)以盐胁迫处理的三叶期棉花幼苗为研究对象进行转录组分析, 共鉴定了1 117个独特的lncRNAs, 发现lnc_388可能是gh-A09G1182 MS的调节器, 而lnc_883可通过调节GH-D03G0339 MS通道从而参与调节盐胁迫耐受性, 在盐胁迫下, lnc_973和lnc_253可以调控ghr-miR399和ghr-156e的表达。

Wang等(2018a)以甲基甲烷磺酸盐(methyl methane sulfonate, MMS)处理泡桐(*Paulownia tomentosa*)幼苗的顶芽进行转录组测序分析, 共获得2 531个假定的lncRNAs。鉴定了7个lncRNAs作为13个已知小RNAs的前体, 15个lncRNAs可能作为19个小RNAs的靶向模拟, 鉴别出220个对MMS有反应的lncRNAs, 包括7个与激素相关的lncRNAs和1个参与碱基切除修复的lncRNA。lncRNA已被证实参与了植物的许多生物学过程(Nejat等2017), 但缺乏应激反应中lncRNA转录后调控的系统研究。Yuan等(2018)将水稻进行4种非生物胁迫(冷、热、干旱和盐), 转录组测序结果得到超过7 000个lncRNAs, 其中近一半是首次被鉴定出来的。值得注意的是, 发现在应激下差异表达的约500个poly(A) lncRNAs大部分显著下调。此外, 数百个下调的多聚腺苷酸化(downregulated polyadenylation, DPA) lncRNAs趋向于高度保守, 并与在胁迫下发挥作用的蛋白编码基因共同表达。特别是观察到在干旱和盐渍的水稻中积累了大量的DPA lncRNAs, 因此水稻lncRNA的多聚腺苷酸化和亚细胞定位可能在转录后水平受到调控。

2.4 lncRNAs响应生物胁迫

植物不仅受到非生物胁迫的作用, 而且在生长发育阶段受到生物胁迫(Seo等2017)。植物经常遭受细菌、真菌、病毒和害虫侵扰其生长发育(张

子杰等2018)。lncRNA在植物对病原体攻击的生物应激反应中的作用是复杂的。植物已经进化出一套有效的疾病抑制防御机制, 以最大限度地减少它们所遇到的疾病的损害。其对病原体攻击的反应依赖于细胞水平上的病原体识别, lncRNA作为小RNA的靶点, 会抑制相应的小RNA的功能, 在分子水平上触发复杂的防御信号网络来协调转录重编程。作为植物防御机制的一部分, lncRNA对植物病原体的反应至关重要(Nejat等2017)。

lncRNA在植物发育过程中扮演重要作用, 但其在植物抗病响应中的作用还不清楚。Zhang等(2018)比较分析棉花中lncRNA的表达及其抗病反应机制。2个保守lncRNAs (GhlncNAT-ANX2、GhlncNAT-RLP7)沉默的植株中都会表现出对棉花黄萎病菌(*Verticillium dahliae*)和灰霉病菌(*Botrytis cinerea*)的抗性增强, 可能与LOX1和LOX2基因的增强表达相关。通过鉴定lncRNA在真菌抗性中的作用, 为阐明棉花抗真菌疾病应答机制提供了新的线索。越来越多的研究表明, 大量lncRNAs在植物抵御病原物侵染过程中发挥着重要的作用(Nejat等2017)。番茄lncRNA16397作为SIGRX22的反义转录物调节其表达, 并且在lncRNA16397过度表达时诱导SIGRX21的表达。过表达lncRNA16397和SpGRX的番茄在感染马铃薯晚疫病病菌(*Phytophthora infestans*)后的症状和活性氧(reactive oxygen species, ROS)积累量低于野生型番茄。结果表明, 番茄lncRNA16397诱导SIGRX表达, ROS积累量更低, 增强了对马铃薯晚疫病的抵抗力。证明lncRNA16397与谷氧还蛋白(glutaredoxin, GRX)是番茄抵御马铃薯晚疫病网络中的一个重要组成部分, 通过这个网络番茄可以抵御疫病侵染, 并为番茄抗生物胁迫育种提供了依据(Cui等2017)。用野生型和WRKY1过度表达番茄植株之间的比较转录组分析, lncRNA33732通过序列特异性与启动子中W-Box元件相互作用被WRKY1激活。通过诱导呼吸爆发氧化酶(respiratory burst oxidase homolog, RBOH)的表达和增加H₂O₂的积累, 增强了番茄对感染性疟原虫的抵抗力, 最终WRKY1-lncRNA3373-RBOH新的抗性模式为提高番茄对病原体的广谱抗性提供了理论依据(Cui等2018)。

此外, Kwenda等(2016)通过比较转录组分析, 发现马铃薯(*Solanum tuberosum*)中大多数lncRNAs具有单外显子, 并在其12条染色体上都有表达。其中559个lncRNAs对胡萝卜软腐果胶杆菌巴西亚种(*Pectobacterium carotovorum* subsp. *brasiliense*)表达特异, lncRNA在马铃薯防御反应中具有潜在的功能作用。

中国番茄黄化曲叶病毒(*tomato yellow leaf curl virus*, TYLCV)是一种在全球范围内通过烟粉虱快速传播的毁灭性番茄病害。Yang等(2019)提出了TYLCV诱导的症状和宿主抗病毒免疫的新模型。番茄中lncRNA-*SILNR1*的下调与发育不全和卷曲的叶片表型有关, *SILNR1*有助于番茄的正常发育和TYLCV抗性。这为TYLCV诱导疾病和宿主抗病毒免疫提供了一个可信的非编码RNA介导的模型, 有助于制定行之有效的策略来控制病毒病原体。lncRNA和环状RNA (circular RNA, circRNA)在植物中相互作用抵御TYLCV病毒。Wang等(2018b)对感染TYLCV或未经处理的易感番茄进行链特异性测序, 共获得1 767个lncRNAs和289个NAT-lncRNAs。通过病毒诱导的基因沉默(virus-induced gene silencing, VIGS)发现lncRNA S-syl-lnc0957的表达降低, 导致易感番茄植株对TYLCV的抗性增强。其中83个circRNAs在TYLCV样本中特异表达, 通过PCR和Sanger测序验证了circRNA的存在, 并且其亲本基因的沉默将会导致TYLCV病毒的积累减少。

3 展望

总体来说, 植物lncRNA的研究尚处于相对早期的阶段。目前定义为lncRNA的核苷酸长短范围模糊, 甚至lncRNA还没有一个规范的命名方法, 通常研究者只能根据其功能、结构特点、作用方式等进行命名; 由于lncRNA种类和功能的多样性, 如何区分功能性和非功能性非编码转录物依然存在困难, 导致lncRNA生物学功能难以阐明。相对于其他数据库和生物信息学工具, lncRNA植物相关数据库的内容和注释还不够全, 难以对lncRNA二级结构和靶标等进行有效地预测。

迄今为止, 相关的植物研究都表明lncRNA在

植物对生物和非生物应激的免疫反应中起着关键作用。lncRNA不仅在应激反应中有功能, 而且在细胞和发育过程中起作用, 其参与激素信号、生理和发病机制等过程的生物学功能也开始被探索并深入研究。由于体内lncRNA的多样性和复杂性, lncRNA的研究仍是冰山一角, 除了识别应激相关的lncRNA及其转录谱外, 少数lncRNAs的功能特征和定位研究处于探索的初步阶段。此外, CRISPR-Cas9系统最近被开发用于靶向基因组编辑, 通过敲除或插入替换阻断转录起始来抑制lncRNA功能, 使其生物学功能和作用机制会进一步被阐明(Liu等2017)。目前, lncRNA参与植物的抗逆、生长发育等报道主要集中在机制研究中, 而在育种中鲜有报道, 随着大量lncRNAs的挖掘发现, 及探索其影响植物的生长、发育和应激反应等分子机制的成熟, lncRNA有望在农业育种、种质创新等方面发挥更重要的作用。

参考文献(References)

- Ariel F, Jegu T, Latrasse D, et al (2014). Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. *Mol Cell*, 55 (3): 383–396
- Ariel F, Romero-Barrios N, Jégu T, et al (2015). Battles and hijacks: noncoding transcription in plants. *Trends Plant Sci*, 20 (6): 362–371
- Bardou F, Ariel F, Simpson CG, et al (2014). Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*. *Dev Cell*, 30 (2): 166–176
- Cho J (2018). Transposon-derived non-coding RNAs and their function in plants. *Front Plant Sci*, 9: 600
- Cui J, Jiang N, Meng J, et al (2018). LncRNA33732-respiratory burst oxidase module associated with WRKY1 in tomato-*Phytophthora infestans* interactions. *Plant J*, 97 (5): 933–946
- Cui J, Luan YS, Jiang N, et al (2017). Comparative transcriptome analysis between resistant and susceptible tomato allows the identification of lncRNA16397 conferring resistance to *Phytophthora infestans* by co-expressing glutaredoxin. *Plant J*, 89 (3): 577–589
- Deng FN, Zhang XP, Wang W, et al (2018). Identification of *Gossypium hirsutum* long non-coding RNAs (lncRNAs) under salt stress. *BMC Plant Biol*, 18: 23
- Ding JH, Lu Q, Ouyang YD, et al (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci*

- USA, 109 (7): 2654–2659
- Gloss BS, Dinger ME (2016). The specificity of long non-coding RNA expression. *Biochim Biophys Acta*, 1859 (1): 16–22
- Guo XL, Gao L, Wang Y, et al (2016). Advances in long non-coding RNAs: identification, structure prediction and function annotation. *Brief Funct Genomics*, 15 (1): 38–46
- Henriques R, Wang H, Liu J, et al (2017). The anti-phasic regulatory module comprising *CDF5* and its antisense RNA *FLORE* links the circadian clock to photoperiodic flowering. *New Phytol*, 216 (3): 854–867
- Huanca-Mamani W, Arias-Carrasco R, Cárdenas-Ninasivinchá S, et al (2018). Long non-coding RNAs responsive to salt and boron stress in the hyper-arid Lluetño maize from Atacama Desert. *Genes*, 9 (3): 170
- Huang L, Dong H, Zhou D, et al (2018). Systematic identification of long non-coding RNAs during pollen development and fertilization in *Brassica rapa*. *Plant J*, 96: 203–222
- Kashi K, Henderson L, Bonetti A, et al (2016). Discovery and functional analysis of lncRNAs: methodologies to investigate an uncharacterized transcriptome. *Biochim Biophys Acta*, 1859 (1): 3–15
- Kiegle EA, Garden A, Lacchini E, et al (2018). A genomic view of alternative splicing of long non-coding RNAs during rice seed development reveals extensive splicing and lncRNA gene families. *Front Plant Sci*, 9: 115
- Kim DH, Xi Y, Sung S (2017). Modular function of long non-coding RNA, COLDAIR, in the vernalization response. *PLoS Genetics*, 13 (7): e1006939
- Kopp F, Mendell JT (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell*, 172 (3): 393–407
- Kornienko AE, Dotter CP, Guenzl PM, et al (2016). Long non-coding RNAs display higher natural expression variation than protein-coding genes in healthy humans. *Genome Biol*, 17: 14
- Kwenda S, Birch PJ, Moleleki LN (2016). Genome-wide identification of potato long intergenic noncoding RNAs responsive to *Pectobacterium carotovorum* subspecies *brasiliense* infection. *BMC Genomics*, 17: 614
- Laurent GS, Wahlestedt C, Kapranov P (2015). The landscape of long noncoding RNA classification. *Trends Genet*, 31 (5): 239–251
- Lennox KA, Behlke MA (2016). Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Res*, 44 (2): 863–877
- Li P, Yang H, Wang L, et al (2019). Physiological and transcriptome analyses reveal short-term responses and formation of memory under drought stress in rice. *Front Genet*, 10: 55
- Li R, Fu DQ, Zhu BZ, et al (2018). CRISPR/Cas9-mediated mutagenesis of *lncRNA1459* alters tomato fruit ripening. *Plant J*, 94: 513–524
- Liao PR, Li SP, Cui XM, et al (2018). A comprehensive review of web-based resources of non-coding RNAs for plant science research. *Int J Biol Sci*, 14 (8): 819–832
- Liu J, Jung CY, Xu J, et al (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*. 24 (11): 4333–4345
- Liu X, Wu SR, Xu J, et al (2017). Application of CRISPR/Cas9 in plant biology. *Acta Pharm Sin B*, 7 (3): 292–302
- Lu XK, Chen XG, Mu M, et al (2016). Genome-wide analysis of long noncoding RNAs and their responses to drought stress in cotton (*Gossypium hirsutum* L.). *PLoS One*, 11 (6): e0156723
- Mishra A, Bohra A (2018). Non-coding RNAs and plant male sterility: current knowledge and future prospects. *Plant Cell Rep*, 37 (2): 177–191
- Nejat N, Mantri N (2018). Emerging roles of long non-coding RNAs in plant response to biotic and abiotic stresses. *Crit Rev Biotechnol*, 38 (1): 93–105
- Nejat N, Rookes J, Mantri NL, et al (2017). Plant-pathogen interactions: toward development of next-generation disease-resistant plants. *Crit Rev Biotechnol*, 37 (2): 229–237
- Platt EJ, Smith L, Thayer MJ (2018). L1 retrotransposon antisense RNA within ASAR lncRNAs controls chromosome-wide replication timing. *J Cell Biol*, 217 (2): 541
- Qin T, Zhao HY, Cui P, et al (2017). A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. *Plant Physiol*, 175 (3): 1321–1336
- Quek XC, Thomson DW, Maag JL, et al (2015). lncRNADB V2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res*, 43: D168–D173
- Quinn JJ, Chang HY (2016). Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet*, 17 (1): 47
- Rai MI, Alam M, Lightfoot DA, et al (2018). Classification and experimental identification of plant long non-coding RNAs. *Genomics*, 111 (5): 997–1005
- Ruiz KB, Biondi S, Martínez EA, et al (2016). Quinoa—a model crop for understanding salt-tolerance mechanisms in halophytes. *Plant Biosyst*, 150 (2): 357–371
- Seo JS, Sun HX, Park BS, et al (2017). ELF18-INDUCED LONG NONCODING RNA associates with Mediator to enhance expression of innate immune response genes in *Arabidopsis*. *Plant Cell*, 29 (5): 1024–1038
- Snijders C, Bassil KC, de Nijs LD (2018). Methodologies of

- neuroepigenetic research: background, challenges and future perspectives. *Prog Mol Biol Transl Sci*, 158: 15–27
- Sun X, Zheng HX, Sui N (2018). Regulation mechanism of long non-coding RNA in plant response to stress. *Biochem Biophys Res Commun*, 503 (2): 402–407
- Takenaka K, Chen BJ, Modesitt SC, et al (2016). The emerging role of long non-coding RNAs in endometrial cancer. *Cancer Genetics*, 209 (10): 445–455
- Wang JY, Yang YW, Jin LM, et al (2018b). Re-analysis of long non-coding RNAs and prediction of circRNAs reveal their novel roles in susceptible tomato following TYLCV infection. *BMC Plant Biol*, 18 (1): 104
- Wang TZ, Liu M, Zhao MG, et al (2015). Identification and characterization of long non-coding RNAs involved in osmotic and salt stress in *Medicago truncatula* using genome-wide high-throughput sequencing. *BMC Plant Biol*, 15: 131
- Wang YQ, Fan XD, Lin F, et al (2014). *Arabidopsis* noncoding RNA mediates control of photomorphogenesis by red light. *Proc Natl Acad Sci USA*, 111 (28): 10359–10364
- Wang Z, Li BB, Li YS, et al (2018a). Identification and characterization of long non-coding RNA in *Paulownia tomentosa* treated with methyl methane sulfonate. *Physiol Mol Biol Plants*, 24 (2): 325–334
- Wu RN, Wang H, Yang CC, et al (2017). Construction of lncRNA-At5NC056820 overexpression vector in *Arabidopsis thaliana* and study on drought resistance of transgenic plants. *Acta Bot Boreal-Occident Sin*, 37 (10): 1904–1909 (in Chinese with English abstract) [毋若楠, 王红, 杨成成等(2017). 拟南芥lncRNA-At5NC056820过表达载体构建及其转基因植株的抗旱性研究. 西北植物学报, 37 (10): 1904–1909]
- Xuan HD, Zhang LZ, Liu XS, et al (2015). PLNlncRbase: a resource for experimentally identified lncRNAs in plants. *Gene*, 573 (2): 328–332
- Yang YW, Liu TL, Shen DY, et al (2019). *Tomato yellow leaf curl virus* intergenic siRNAs target a host long noncoding RNA to modulate disease symptoms. *PLoS Pathog*, 15 (1): e1007534
- Yi X, Zhang ZH, Ling Y, et al (2015). PNRD: a plant non-coding RNA database. *Nucleic Acids Res*, 43: D982–D989
- Yuan JP, Li JR, Yang Y, et al (2018). Stress-responsive regulation of long non-coding RNA polyadenylation in *Oryza sativa*. *Plant J*, 93 (5): 814–827
- Zandalinas SI, Mittler R, Balfagón D, et al (2018). Plant adaptations to the combination of drought and high temperatures. *Physiol Plant*, 162 (1): 2–12
- Zhang L, Wang MJ, Li NN, et al (2018). Long non-coding RNAs involve in resistance to *Verticillium dahliae*, a fungal disease in cotton. *Plant Biotechnol J*, 6 (6): 1172–1185
- Zhang YC, Liao JY, Li ZY, et al (2014). Genome-wide screening and functional analysis identify a large number of long non-coding RNAs involved in the sexual reproduction of rice. *Genome Biol*, 15 (12): 512
- Zhang ZJ, Xiao WF, Qiu JR, et al (2018). Advances in small RNA regulating plant immune response. *Plant Physiol J*, 54 (4): 539–548 (in Chinese with English abstract) [张子杰, 肖文斐, 袁劫人等(2018). 小RNA调节植物免疫应答反应研究进展. 植物生理学报, 54 (4): 539–548]
- Zhao XY, Li JR, Lian B, et al (2018). Global identification of *Arabidopsis* lncRNAs reveals the regulation of *MAF4* by a natural antisense RNA. *Nat Commun*, 9: 5056

Advances in functional research of long non-coding RNAs in plants

LI Ning^{1,2}, WANG Bai-Ke¹, WANG Juan¹, HUANG Shao-Yong², DAI Qi¹, PATI Gu-Li¹, GAO Jie², YU Qing-Hui^{1,*}

¹*Institute of Horticultural Crops, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China*

²*College of Forestry and Horticulture, Xinjiang Agricultural University, Urumqi 830052, China*

Abstract: Long non-coding RNA (lncRNA) is a kind of RNA without the function of encoding protein and usually produced by transcription of RNA polymerase II. lncRNA is more than 200 nt and widely located in animals, plants and viruses with characteristics of epigenetic markers, developmental stage-specific and tissue-specific expression. Previous studies have shown that lncRNAs play an important role in resistance of plant disease and response to stress in addition to participating in growth and development. This article elaborates the formation mechanism, structural characteristics and database of lncRNAs in recent years briefly, and mainly focuses on their functions on the growth and biotic and abiotic stress. The utilization prospects of lncRNAs in production breeding of crops are also investigated and forecasted.

Key words: lncRNA; growth and development; biotic and abiotic stresses

Received 2019-05-20 Accepted 2019-09-11

This work was supported by the National Key R&D Program of China (2017YFD0101906) and Xinjiang Academy of Agricultural Science (xjzdy-003).

*Corresponding author (yuqinghui98@sina.com).