植物反转录转座子功能研究进展

林悦龙^{1,2,3}, 肖开转^{1,2,3}, 连玲^{1,2,3}, 何炜^{1,2,3}, 陈丽萍^{1,2,3}, 蔡秋华^{1,2,3}, 王颖姮^{1,2,3},
谢华安^{1,2,3*}, 张建福^{1,2,3*}

1. 福建省农业科学院水稻研究所, 福州 350019;
2. 农业部华南杂交水稻种质创新与分子育种重点实验室/福建省作物种质创新与分子育种省部共建国家重点实验室培育基地/水稻国家工程实验室/杂交水稻国家重点实验室华南研究基地/福州(国家)水稻改良分中心/福建省作物分子育种工程实验室/福建省水稻分子育种重点实验室, 福州 350003;
3. 闽台作物有害生物生态防控国家重点实验室, 福州 350003

* 联系人, E-mail: huaanxie@163.com; jianfzhang@163.com

2018-08-28 收稿, 2018-10-12 修回, 2018-10-12 接受, 2018-11-22 网络版发表

国家重点研发计划(2016YFD0101801)、福建省科技计划——省属公益类科研院所基本科研专项(2018R1101013-4)和福建省农业科学院创新团队项目(STIT17-1-1)资助

摘要 反转录转座子(retrotransposon), 是一类以RNA为中介, 在反转录酶的参与下, 进行自我复制并整合到基因组其他位点中的转座元件(transposable elements, TEs). 反转录转座子是许多真核生物基因组的主要组成部分, 在高等植物如玉米、小麦等作物中含量丰富. 这些反转座子的功能及生物学意义一直以来都存在争议, 但越来越多研究表明, 它们对于邻近基因的表达调控以及整个生物体基因组的进化有着深远的影响. 本文主要从反转座子与非编码RNA的联系及其转座过程所产生的基因结构变异、反转座基因等方面, 概述了植物中反转座子功能的相关研究进展.

关键词 反转座子, smRNA, lncRNA, 反转座基因, 嵌合基因, 胁迫, 反转录病毒

转座元件又称转座子(transposon), 是一类可以在基因组中某一位点转移或自我复制到另一个位点从而改变它们在基因组中原有位置或增加其拷贝数的遗传因子. 早在1950年, 美国著名遗传学家McClintock^[1]在研究玉米籽粒色斑变化现象时就已经提出了转座(transposition)的概念. 转座子按其转座的方式可以分为两大类: 一类直接以DNA→DNA的形式在转座酶的参与下进行“剪切-粘贴”式转座, 称为DNA转座子; 另一类则以RNA为中介, 在反转录酶参与下完成“复制-粘贴”式转座, 称为反转录转座子^[2,3]. 生物体基因组中大量的转座因子曾一度被视为冗余的“垃圾”基因与“寄生虫”, 但其实许多转座因子也参与到了宿主基因的表达调控、表观遗传及生物

适应性进化等方面内容, 并在植物的演化中扮演着重要角色.

1 反转座子的分类、结构与特征

反转录转座子属于Class I转座子, 简称反转座子, 是通过DNA→RNA→DNA的形式完成反转座过程的, 可以分为LTR(long terminal repeat, 长末端重复)和non-LTR(非长末端重复)反转座子^[4]. LTR反转座子可进一步分为Ty1/copia和Ty3/gypsy两个主类群^[5], 具有与反转录病毒基因组相似的长末端重复结构(图1), 长度几百到几千bp不等, 通常不编码蛋白, 但具有转录和终止等调控作用, 这两类结构完整的反转座子能编码转座所需要的酶类, 并自主完成转

引用格式: 林悦龙, 肖开转, 连玲, 等. 植物反转录转座子功能研究进展. 科学通报, 2019, 64: 35–48

Lin Y L, Xiao K Z, Lian L, et al. Functional progress of retrotransposons in plant (in Chinese). Chin Sci Bull, 2019, 64: 35–48, doi: 10.1360/N972018-00877

图1 (网络版彩色)反转座子主要类型及其基因结构^[5,8,9]. TSD: target site duplications, 插入位点重复序列; PBS: primer binding sites, 反转录第一条(-)DNA链引物结合位点; gag: group-specific antigen, 种群特异性抗原; pol: polymerase, 聚合酶类; PR: protease, 蛋白水解酶; IN: integrase, 整合酶; RT: reverse transcriptase, 反转录酶; RH: Ribonuclease H, RNA水解酶H; PPT: polypurine tract, 多嘌呤序列, 反转录第二条(+)DNA链引物结合位点; UTR: untranslated region, 非翻译区; ORF1: open reading frame 1, 开放阅读框1; EN: endonuclease, 核酸内切酶; L1 Element: 人类LINE1反转座子元件; A, B: BoxA, BoxB, RNA聚合酶III结合位点; L: Left monomer, 左臂序列; R: Right monomer, 右臂序列; Alu Element: 人类Alu反转座子元件

Figure 1 (Color online) The main types and gene structures of retrotransposons^[5,8,9]. TSD: Target site duplications, the (-) DNA strand binding site of reverse transcription; PBS: Primer binding sites; gag: Group-specific antigen; pol: Polymerase; PR: Protease; IN: Integrase; RT: Reverse transcriptase; RH: Ribonuclease H; PPT: Polypurine tract; the (+) DNA strand binding site of reverse transcription; UTR: Untranslated region; ORF1: Open reading frame 1; EN: Endonuclease; L1 element: the human LINE1 retroelements; A, B: BoxA, BoxB, the binding sites of RNA polymerase III; L: Left monomer; R: Right monomer; Alu Element: the human Alu retroelements

座过程。此外, LTR反转座子还包括非自主性转座的大型反转座子派生物(large retrotransposon derivative, LARD)和微型末端重复反转座子(terminal-repeat retrotransposons in miniature, TRIM), 它们都缺少中部负责编码gag, pol蛋白的基因, 其中, gag蛋白作为执行反转录过程时复合体的外壳结构, 并具有种群特异性, 而pol蛋白则能被进一步分解加工成整合酶、反转录酶等多个活性酶类。在LARD中, 中部取而代之的是具有保守RNA二级结构的一段长链非编码DNA序列^[6], 而TRIM元件则只剩下长度100~300 bp的中部结构^[7]; non-LTR反转座子则包括长散布重复序列(long interspersed nuclear elements, LINE)类和短散布重复序列(short interspersed nuclear elements, SINE)类(图1), SINE类不编码转座所需要的酶类, 是RNA聚合酶III的转录加工物, 常借助LINE反转座酶系统进行反转座^[10]。

反转座子是许多植物基因组的主要成分。在番茄中, 它们占据超过60%的基因组序列^[11], 而一些禾本科作物如小麦、玉米中则有近80%的序列是由反转座子序列所构成^[12,13], 它们的复制增殖是植物基因

组不断变大的重要因素之一。

植物基因组中的反转录转座子具有多种潜在的功能与作用, 一些反转录转座子元件本身在激活转录时会参与到非编码RNA的产生, 通过这些非编码RNA的靶向作用, 它们能进一步调控下游相关基因的表达。另外, 反转座子的复制转位常影响到插入位点及其附近基因的结构及表达, 例如, 它们可以作为一种顺式或反式作用元件作用于基因的上游或下游, 或者直接作为启动子及终止子等来影响靶基因的表达、转录与功能。此外, 自主性反转座子自身所产生的反转座酶促系统还可能作用于基因组中的其他基因, 形成嵌合基因或者是完整的基因副本, 而一些反转座子序列还可能在植物体长期的进化中渗入到了新基因的生成与编码, 反转座子的存在对于植物体本身及进化具有重要意义。

2 反转座子通过非编码RNA调控基因表达

非编码RNA(non-coding RNA, ncRNA), 包括siRNA(short interfering RNA), miRNA(micro RNA), piRNA(Piwi interacting RNA)等长度为18~30 nt小

RNA(small RNA, smRNA)和长链非编码RNA(long non-coding RNA, lncRNA), 它们不编码蛋白质, 但却参与了基因的转录调控、个体发育及表观遗传等多种生物学过程, 而转座元件的活动与非编码RNA的起源、进化关系密切。

2.1 反转座子参与small RNA的产生

一些研究表明, 不少非编码RNA直接或间接来源于转座元件序列. Li等人^[14]在挖掘植物转座子相关的miRNA(transposable element related MIRs, TE-MIRs)时发现, 一些已注释的miRNA基因位点本身就是转座子或其派生物. 在水稻中, TE-MIRs 80%来源于名为MITE(miniature inverted-repeat transposable elements)的转座子, 但也有约10%来源于反转座子, 例如1个水稻miRNA基因位点, osa-MIR1879, 是由2个非自主的同源反转座子以倒置的方向并列在一起形成的, 这种倒置并排的同源重复序列有利于发夹结构的产生, 是新miRNA起源的一种重要方式. 另外, 在对这些TE-MIRs的靶基因进行预测时发现, 只有少数为其他转座子序列, 而多数则是非转座子基因或表达的未知蛋白。

小RNA的另一位成员, siRNA, 有别于茎环发夹结构所产生的miRNA, 通常来源于长的双链RNA^[15]. 拟南芥的一个小RNA, siRNA854, 产生于Athila6家族的LTR反转座子, 当Athila6转录激活时, 其转录本会被加工产生大量siRNA854. 通过RNA干扰(RNA interference, RNAi)机制, 产生的siRNA854还会协同AGO1蛋白对靶标基因UBP1b进行转录和翻译水平的抑制^[16,17](图2), UBP1b是一个RNA结合蛋白, 为植物胁迫颗粒(stress granules, SGs)的一种组分, 在mRNA的剪切及成熟过程中起着重要调节作用, 当UBP1b的表达与翻译被抑制时, 较之野生型植株, 则对热胁迫反应更为敏感^[18].

一些反转座子派生的siRNA会在其解除表观沉默时大量产生, 这类由解除表观沉默的转座子所产生的siRNA也称为easiRNAs(epigenetically activated small interfering RNAs)^[19], Zhang^[20]同样也发现, 水杨酸处理拟南芥植株后, 会从去除甲基化而激活的转座元件中产生大量的长度为21 nt的小RNA, 并推测部分小RNA的靶基因可能和细胞的防御反应有关. 此外, Hsieh等人^[21]用低磷处理拟南芥时发现, 植株根中会大量产生名为smRPi1^{LTR}的小RNA, 它们来源

图2 (网络版彩色) Athila6 反转座子通过siRNA调控UBP1b基因的表达^[17]

Figure 2 (Color online) Athila6 retrotransposons regulate the expression of UB1b gene through siRNA^[17]

于快速重排的Copia95反转座子, 其靶标基因很可能和机体缺磷时的适应性生长有关。

2.2 反转座子参与lncRNA的产生

lncRNA是指长度大于200 nt的非编码RNA, 包括基因间、反义、内含子lncRNA三类, 它们具有多种生理功能, 例如参与了拟南芥磷酸盐调节代谢途径^[22]、春化作用^[23]、水稻光周期变化所产生的雄性不育^[24]等。

lncRNA多数起源于转座元件富集的基因间隔区, Wang等人^[25]发现, 它们不仅在不同物种间保守性差, 即使在同一物种的不同品种间也存在较大差异, 这与基因间隔区内活动的转座子密切相关, 转座子的复制转位促成了许多lncRNA基因结构的变异及新生. 例如, 在Heinz1706番茄品种特异的lncRNA中, 存在启动子区和转录区分别来源于2个反转座子LTR的lncRNA-263以及转录区来源于反转座子序列的lncRNA-314等. 有趣的是, lncRNA-314只在番茄果实的破色期阶段特异性地转录表达, 这个新生的lncRNA很可能参与到了果实成熟的调节过程中。

有一种假说认为, 转座子所衍生的lncRNA功能就来源于外显子化的转座子本身所固有的功能, 比如DNA, RNA及蛋白质结合结构域等, 也称之为lncRNA的重复序列获得结构域repeat insertion domains of lncRNAs (RIDLs)^[26]. 因为反转座子通常以高拷贝数存在于基因组中, 并存在于许多功能基因的邻近区域^[27], 当它们转录激活时可能会影响到邻近基因的转录与表达, Kashkush等人^[28]发现, 在双二

倍体小麦激活的 $Wis\ 2-1A$ 反转座子两端的LTR都会产生转录读出,并生成邻近基因的反义或正义链转录本,从而抑制或促进邻近基因的表达.这些转录本可能部分是未发现的lncRNA类群,它们还可能进一步形成双链RNA参与到siRNA介导的转录沉默机制而作用于其他靶基因^[29].此外,一些lncRNA还会与小RNA不完全配对,这种配对方式不会引起lncRNA的降解反而降低了游离小RNA的水平,削弱了它们的干扰作用,使之也参与到了靶基因的调控平衡中^[30].

3 反转座子改变基因的结构与功能

基因是严密的遗传序列,基因序列的突变可能造成功能缺陷、失活,或者形成有益突变.转座子的活动,是基因突变的重要来源,并能形成可遗传性的改变.

3.1 介导基因失活及沉默

转座子的移位存在一定的随机性,并能插入基因中破坏原有的结构,使之失活.例如,紫色葡萄表皮颜色的转变,是由一个名为 $Gret1$ 的Ty3-gypsy反转座子整合到了与花青素合成相关的 $VvmybA1$ 基因上游,直接导致了该基因的转录失活,由于花青素无法合成,所以表皮转变为白色^[31].有趣的是,当该基因位点发生重组而只剩下solo LTR作为启动子时, $VvmybA1$ 基因又得到一定表达,还会产生红色葡萄

的表型(图3(a)).类似地,一个Ty3-gypsy反转座子破坏了萝卜中类黄酮3'羟化酶基因的结构并使之转录沉默,导致在根中所积累的花青素成分的改变,并形成红色萝卜的性状^[33].

著名孟德尔遗传实验中的皱粒豌豆,其成因则是由于一个长约800 bp的Ac/Ds家族类转座子插入到了淀粉分支酶 $SBE1$ 基因的外显子中,造成了该基因的转录水平大大降低,且翻译产物被截短而失活,淀粉合成受阻则增加了种子中蔗糖的积累,进一步影响到渗透压、水分等性状的塑造^[34].而由反转座子活动所造成相似的基因转录失活或功能缺陷同样在其他一些植物的研究中屡见不鲜.例如,水稻的低植酸突变体 $XS-lpa$,是由于一个LINE反转座子出现在 $OsMIK$ 基因的内含子区,它提高了该基因启动子区的甲基化水平,从而导致该基因转录水平降低,造成植株低植酸的表型^[35].另外,水稻单穗^[36]及松散株型等^[37]突变体的形成都和反转座子的插入所导致的基因失活有关.

3.2 介导基因结构及转录本变异

正常基因的读码区如果发生转座子的插入则可能提前终止,造成基因结构的不完整性.水稻白叶枯抗性基因 $Xa21$ 家族成员 $Xa21D$,其实质是一个被反转座子插入到外显子中所截短的 $Xa21$ 基因副本,与完整的 $Xa21$ 基因相比,由于翻译的蛋白缺失了行使完整功能所必需的跨膜结构域与激酶结构域,虽然

图3 LTR反转座子介导的葡萄(a)、血橙(b)变种形成分子机制^[32]

Figure 3 The molecular mechanism of the grape (a) and orange (b) variants formation mediated by LTR retrotransposons^[32]

和Xa21有着相同的抗谱, 却仅表现出Xa21完整基因抗性的一部分^[38]。

当转座子插入到基因的内含子时, 还能通过改变基因的可变剪切来影响基因功能。在玉米中, *waxy(Wx)*基因存在多种等位突变基因, 其中有近一半是由转座子活动所造成的^[39]。研究者从其等位突变基因中克隆出多个转座子, 包括activator(*Ac*), suppressor-mutator(*Spm*), enhancer-inhibitor(*En-I*)等转座家族元件^[40,41]。其中, 3个由反转座子插入其内含子中所形成的等位基因, 即*wx-Stonor*, *wx-G*, *wx-B5*, 造成了该基因可变剪切的形成。有些剪切本跳跃了一个或多个外显子, 有些剪切本则被截短或包含了部分反转座子序列而终止, 虽然它们都还能剪切出正常mRNA, 但由于正确剪切的效率低下, 直链淀粉合成受阻而表现出糯玉米的特征^[42]。

3.3 介导基因活化及异位表达

反转座子具有转录起始和调控功能的LTR区, 本身就是一种可移动的基因调控因子。例如, 一个名为*Renovator*反转座子的LTR区直接作为启动子提高了稻瘟病抗性基因*Pit*的转录水平, 并使其重新表现出相应的抗病活性^[43]。而在人类基因组的研究中, 许多基因受到了内源性反转录病毒LTR区的调控^[44], 内源性反转录病毒是人类祖先感染反转录病毒的产物, 因而也是现存体内的一种反转座元件。某些时候, 反转座子的存在还会改变基因表达的时空位置, 黄橙的变种——血橙, 是由于一个Copia-like的反转座子插入到了*Ruby*基因编码区的上游, *Ruby*能激活调控花青素的合成, 但不在果肉中表达。而当反转座子出现并受到一定寒冷诱导时, 却能够驱动*Ruby*基因在果肉中表达, 并最终表现出红色果肉的表型^[45], 类似于红色葡萄的表型, 在基因重组后仅存的LTR结构区的控制下, 则会产生颜色更深的橙肉(图3(b))。

4 反转座子参与基因组进化

4.1 介导反转座-产生嵌合基因与基因副本

当反转座子的反转座过程偶然地作用于其他基因的mRNA时, 便会产生一个反转座基因(retrogene), 由于缺失内含子及调控序列, 其结果常形成一个无法正常表达的假基因, 也称为加工假基因(processed pseudogenes), 它们带有明显的反转座痕迹, 例如3'

端有polyA尾类似序列以及两侧存在插入位点的重复序列等^[46]。

反转座基因作为基因的一种特殊副本, 如果整合到了一段具有启动子功能序列的下游, 则可能被重新激活转录, 或者和邻近基因的结构域进行外显子的重排而产生嵌合基因(chimeric gene), 这种新生的嵌合基因往往具有与父母本基因不一样的功能^[47]。1993年, Long和Langley^[48]在果蝇中发现的第一个年轻基因——精卫基因(*jgw gene*)。该基因是由*Adh*基因经过反转座后, 捕获到*ymp*基因的部分结构域而蜕变形成的。后来又陆续发现了编码一个lncRNA的*Sphinx*基因^[49]、猴王基因^[50]等20多个嵌合基因。Wang等人^[51]在研究水稻基因组中也发现, 这种基因的反转座然后又以嵌合基因形式被转录的现象是很普遍的, 同时也是新基因起源的一种重要模式。

基因一直被认为是静止的线性排列, 然而反转座子的存在却可能改变这一不变的状态。例如, 在番茄品种Sun1642中, 一个名为*Rider*的反转座子介导了*IQD12*基因副本的形成, 并最终造成了番茄长果实质性状的出现。在这次反转座过程中, 通过转录读出与模板交换(图4), *Rider*在10号染色体上产生了长度为24.7 kb的转录本, 并将它完整地反转座到7号染色体的*DEFL1*基因中。*IQD12*是一个控制番茄果实形状的数量性状基因, 被反转座到7号染色体后, 其基因位点被重排, 原来*DEFL1*基因的启动子此时便驱动*IQD12*在果实中大量表达, 导致番茄的果实变长^[52]。由此可见, 反转座子所介导的反转座过程对植物基因组的进化有着重要意义, 反转座过程不仅能创造出新基因, 也能创造基因副本到其他染色体上, 它们使得整个基因组都具备了跳动的潜能。

4.2 反转座子序列渗入基因编码与功能进化

在转座子与生物体漫长的共进化过程中, 许多已演化为重要的功能基因。人们所熟知的端粒酶, 具有类似反转录酶的功能, 通过修复染色体的端粒末端, 维持细胞的正常分裂增殖, 基于反转录酶的进化树分析推测, 端粒酶可能起源于一类penelope-like elements (PLEs)的反转座子^[53]。在哺乳动物中, 还发现了许多由转座子转变而来基因, 例如, 由反转座子*Gag*蛋白基因进化而来的*Mart*家族基因、*Ma*、*Fv1*基因, 它们有着和*Gag*蛋白相似的保守锌指结构域; 由整合酶起源的*Gin-1*基因、内源性反转录病毒*env*基因

图4 (网络版彩色)Sun1642的*IQD12*基因副本形成原理^[52]

Figure 4 (Color online) The formation mechanism of the other *IQD12* gene copy in Sun1642^[52]

起源的*Syncytin-1*, *Syncytin-2*基因等^[54]. Nekrutenko和Li^[55]则发现,许多转座子都是先插入到基因的内含子区,而后被部分招募为新的外显子而融入到新的蛋白编码区中.

而在高等植物中, Jiang等人^[56]则发现,仅在水稻日本晴参考基因组中,就有超过3000多个的Pack-MULEs(mutator-like transposable elements)转座子捕获了来源于1000多个不同的细胞基因片段. 这些不同来源的基因片段在转座子区域内进行复制、重排,或者融合了转座子及附近的DNA序列而产生各式各样新的嵌合基因转录本,并被认为是植物基因进化的新机制. 类似的,一些反转座子同样具有捕获细胞其他基因片段并进行重组的功能,它们有时候可能会牺牲其*pol*基因而转导进细胞的一些其他基因. 例如玉米的*Bs1*反转座子,其内部*Gag*编码区被截短,后部也没有*pol*基因,取而代之的是3个细胞内源基因的部分片段(图5),分别为质子依赖的膜蛋白ATP酶(*c-pma*)、木聚糖内切水解酶(*c-xe*)、 β -1,3葡聚糖酶(*c-bg*)^[58,59]. 这种结构和哺乳动物反转录病毒捕获细胞内基因的现象十分相似,随后研究发现, *Bs1*是1个反转座子介导的新嵌合基因,4个基因在反转座子基因结构内进行外显子重排,同时被证实存在着转录与翻译,推测其可能在玉米生殖器官形成中发挥着一

图5 (网络版彩色)玉米*Bs1*转导基因结构示意图^[57]. *r-bg*, *r-xe*, *r-pma*: 对应转导的细胞基因部分序列

Figure 5 (Color online) The structures of gene transduction by the *Bs1* in maize^[57]. *r-bg*, *r-xe*, *r-pma*: Corresponding partial transduction sequence of cellular gene

定作用^[57]. 此外, Duan等人^[60]也发现了一个发源于拟南芥*Copia95*反转座子家族的新基因*AtCopegl*, 且该基因存在组织表达的特异性,同时还受到外源激素及非生物胁迫的诱导,并可能在其中发挥着一定作用.

5 反转座子在生理变化与胁迫中的潜在功能

反转座子通常受到了机体控制而处于沉默或低转录水平的状态,因为转座容易造成一些不利的突变. 生物体对这些转座子的沉默机制主要有转录水平的抑制(transcriptional gene silencing, TGS)-包括甲基化、组蛋白修饰和后转录水平的抑制(post-transcriptional gene silencing, PTGS)-RNA干扰(RNAi)等. 在甲基化修饰的有关基因如*DDM1*^[61,62]或siRNA合成缺陷^[63,64]的突变体中,转座子都会解除抑制而大量转录复制并转座.

值得注意的是,当植物体所处的外界环境条件改变,受到干旱、高温、病菌侵染等胁迫或处在特殊生理状态如离体组织培养时,许多反转座子会特异性地转录表达并可能进行转座^[65,66]。例如在进行原生质体培养时,烟草Tnt1家族反转座子成员的转录水平会显著提高,另外,包括机械割伤、病毒感染、水杨酸、茉莉酸等处理都会使其转录水平发生较大变化^[67~71]。而其他植物如番茄、拟南芥中的Tnt1成员则与烟草有所不同,Tnt1在烟草进行组培时表达很少,而在番茄和拟南芥中则表达较强烈^[72]。除Tnt1外,不少反转座子都有类似的反应,比如大麦的BARE反转座子也响应于干旱^[73]、水稻的Tos家族反转座子也受组织培养所诱导^[74]、而拟南芥的ONSEN反转座子则会被高温所激活^[75]等。徐玲等人^[76]以水稻Ty3-gypsy反转座子为研究对象,通过反转录酶的保守区域对5种胁迫处理(稻瘟菌、水杨酸、2,4-D、高盐及组织培养)所诱导的反转座子进行扩增分类后则发现,对不同胁迫响应的反转座子多数同源性较高,但有少数具有胁迫响应的特异性。

反转座子在不同条件下的激活,一方面可能是由于环境改变造成植物DNA的甲基化水平发生改变;另一方面,一些反转座子的LTR序列本身具有特殊的响应元件,比如短回文结构BI box、串联重复BII box、G、W-box等调控元件序列。这些结构在许多已知的诱导型防御基因中也被发现^[71,72,77],例如烟草Tnt1A反转座子中的G-box类元件是一个典型的ABA诱导元件,它与一些受干旱诱导基因的MYC转录因子识别序列很相似^[78]。在拟南芥ONSEN反转座子的启动子中也发现了类似HRE的元件。HRE为热响应因子HSFs(heat shock factors)三聚体结合位点,当HSFs与HRE结合时,会诱导下游目标基因的转录表达^[79,80]。Ohama等人^[81]认为,热胁迫下新生的反转座子拷贝可能作为一种适应性记忆而传递给子代,以长期对抗高温环境。

一些观点认为,胁迫条件下的反转座子活性可能促进了基因组的快速重塑,这有利于提供选择压力下的进化材料^[82]。Zhang等人^[20]也提到,激活的转座子往往随机插入基因组各个位置,这让子代能有大量的自然变异生态型来对抗病原菌;或者也可能诱导了细胞重编程以调节各种基因表达来对抗和适应不利条件^[83]。如果从反转座子“寄生性”分子本身来看,是否更有利于它们逃避宿主控制以进行垂直

或水平方向的复制传播,特别是病毒家族类的某些反转座子^[84,85]?徐玲等人^[86]则发现,一些反转座子相关基因的旁系同源物功能都与抗病有关,这意味着它们的转录可能会影响到植物的抗病性。沙爱华和张端品^[87]在研究水稻白叶枯病成株抗性机理时发现,虽然病原菌所诱导产生的反转座子数目与成株期相当,但植株在苗期生长发育所诱导激活的反转座子数却高于成株期,推测白叶枯病成株抗性的机制可能与不同生理状态下的反转座子活性有关。到目前为止,关于反转座子在植物胁迫反应过程中的功能与机制尚不十分清楚,但这种特异的转录活性也有部分可能在长期进化中而演变为植物体一种自我调节的信号,并借此调控相关基因的表达来改变机体的生理状态或应对外界环境的变化。

如果从smRNA, lncRNA的角度来看,这些激活的转录本还可能参与形成smRNA或lncRNA再执行对其他基因的表达调节。由于许多LTR结构具有特定条件的响应元件,可以想象,当成千上万个散布在基因组中的LTR结构同时激活所产生的转录本,必然对整个机体的转录组有着不可忽视的冲击与改变。Skibbe等人^[88]也发现,当玉米中Mutator家族的反转座子在早熟花药中激活时,有近25%的转录组发生了改变,而在蛋白丰度上改变最大的则是一些和新陈代谢相关的酶类。当然,若从转座子序列渗入到基因编码的角度来看,这种转录活性可能有部分来源于新的基因结构,在研究人和小鼠的基因时也发现,有超过25%的基因参考序列在3'UTR区包含有反转座子序列,并且这些基因位点的转录水平往往低于3'UTR区没反转座子的基因位点,当它们坐落于基因位点的邻近5'区时,则通常直接作为启动子或者表达非编码RNA^[89]。事实上,不少人类基因从转座子中直接获得了基因表达的调控元件,其中包括有大量的顺式作用元件^[90]。

此外,因为许多转座子本身具有蛋白结合结构域的性质,它们散布在基因组的各个位置,当转座子结合的DNA结合蛋白(DNA binding protein, DBP)为一个转录因子时,它便能调控所结合区域附近基因的转录表达。它们复制移动时,这种调控区也会随之发生移动,这也是转座子调控整个基因系统活动的一种潜在方式(图6(a))。当然,通过融合进基因的转录区或者和邻近基因共转录,将使这些基因同样会受到RNA结合蛋白(RNA-binding protein, RBP)的调

图6 转座元件通过DBP, RBP调控基因表达示意图^[91]. DBP:DNA结合蛋白; RBP:RNA结合蛋白; 彩色矩形: 细胞基因; 白色矩形: 转座元件重复序列; 黑色三角形: 末端倒置重复序列

Figure 6 TEs regulated gene expression by DBP and RBP^[91]. Annotation: DBP:DNA binding protein; RBP:RNA binding protein; coloured boxes: cellular gene; white boxes: repeat sequence of transposable elements; black triangles: terminal inverted repeats

控(图6(b))^[91]. Elbarbary等人^[92]也指出, LINEs和SINEs类反转座子可以通过影响染色质结构、基因转录物、pre-mRNA的加工和mRNA的新陈代谢来调控基因的表达.

6 反转座子的起源与进化

有趣的是,生物体内的LTR反转座子与哺乳动物反转录病毒在最初侵入宿主时所形成的“前病毒”状态有着许多惊人相似之处(图7),二者都有类似的细胞内增殖模式、具有LTR结构,也有同源的gag, pol基因. 在Ty1/copia和Ty3/gypsy的某些家族成员中,它们还拥有第3个开放阅读框,虽然编码蛋白与功能未知,但显然与反转录病毒包膜蛋白env具有一定的结构相似性,从氨基酸序列预测,一般具有明显的信号肽、C端跨膜结构域等特征^[93,94]. 例如在植物中,有拟南芥Tat1、Athila4,大豆SIRE-1、水稻Rigy-2等反转座子成员^[95]. 并且,LTR反转座子在激活转座进

行转座的时候,同样会形成一种类似于病毒的颗粒结构, gag蛋白作为壳结构,内部则有反转录酶、RNA水解酶等负责反转录过程,称之为病毒样颗粒(virus-like particles, VLPs)^[96,97]. 有些时候反转录病毒与LTR反转座子之间的界限是模糊不清的^[98],比如果蝇的gypsy(mgd4)反转座子,它被证实在某些情况下也具备感染细胞的能力,因而也称为无脊椎动物反转录病毒或内源性反转录病毒^[99]. 而在植物中,“env”基因同样存在,虽然至今没有明确发现植物中的反转录病毒,但这些拥有第三开放阅读框的反转座子,都是潜在的植物反转录病毒. 在最近的病毒分类学中, Ty1/copia家族被归类到假病毒科(Pseudoviridae),而Ty3/gypsy则归属于转座病毒科(Metaviridae). 根据“env”基因的有无, Ty3/gypsy科内还可进一步分为转座病毒属(Metavirus)和漂移病毒属(Errantivirus)^[100].

实际上,在研究人类基因组时早有发现,人类的基因组中至少有约8%的基因序列来自于古老的反

图7 病毒样反转座子与反转录病毒基因结构图^[86]

Figure 7 The gene structures of retrovirus and retrovirus-like retrotransposons^[86]

转录病毒^[101]，它们多数退化或只剩下一些残余序列，但少数至今仍具有完整的病毒序列，在漫长的共生中，它们最初的性质多数已经发生改变，有些甚至已经演化为寄主生命活动所不可缺少的元件，例如人类的HERV-W内源性反转录病毒家族被证实为在胎盘形成和保护胚胎发育过程中发挥着至关重要的作用^[102,103]。著名的克隆羊“多利”，由于缺少内源性绵羊肺腺瘤反转录病毒(enJSRV)的表达，而染上由外源性绵羊肺腺瘤(JSRV)反转录病毒所引起的肺腺瘤而死亡^[104,105]。显然，由反转录病毒→内源性反转录病毒→反转座子，是可能的一种退化起源模型，即反转座子可能是早先侵入生物体内古老病毒的遗迹，它们是被限制在细胞内活动的退化病毒元件。相反地，由于病毒难以离开细胞独立生存复制，更多观点

则认为转座子是细胞内原先固有的遗传结构，它们是病毒进化的原材料，有些LTR反转座子意外获得了*env*基因而具备感染细胞的能力，形成了反转录病毒。总而言之，生物体内的LTR反转座子与反转录病毒一定有着密切的关系，它们或许来源于同一祖先，或者它们存在进化上的母子关系^[106]，从植物反转座子起源与进化的研究中可能会获得一些有关哺乳动物中反转录病毒的启示。

7 展望

反转座子在许多植物基因组中的重要成分，但它们存在的意义及所发挥的功能一直存在争议。通过这些例子，可以总结并推测反转座子在植物基因组中的潜在功能与作用方式(图8)。反转座子的复制

图8 反转座子等TEs在生物系统中潜在的功能途径示意图

Figure 8 The potential functional pathway of retrotransposons and other TEs in the biological system

移动可以作为沉默子或终止子让插入区域的正常基因发生表达沉默或者翻译提前终止,进而使基因功能失活;也可以直接作为邻近基因的启动子、增强子或者通过改变基因的剪接模式以及产生甲基化等来调控相应基因的表达与功能行使;而它们通过“复制-粘贴”散布在基因组中各个位置的拷贝,则还能成为转录调控蛋白的结合位置而将许多基因调入了同一调控网络系统中,从而让机体实现对这些基因的协同调控。另外,由反转座子所介导产生的新基因副本、嵌合基因及其所发源的lncRNA, smRNA还会进一步参与到机体各种生理活动及逆境胁迫时整个基因系统的调控变化过程中,并持续推动着生物基因组的快速重塑与进化。

其实,早在麦克林托克提出转座子的概念之

后,就将其命名为“调控元件(controlling elements)”,她说:“不要认为这是理所当然的,真正的重点是调控,所有这些真实的秘密是调控,而不是转座^[107]”。越来越多研究也渐渐印证,基因组中大量存在的转座元件不仅仅是自私的垃圾DNA,它们能通过特定的方式对生物的基因表达系统进行调控,同时还是生物体创造新基因及进化的种子。目前,国内对于植物反转录转座子的研究并不多,而有关它们在植物中所发挥的具体功能尤其是在各种逆境胁迫中转录激活所担当的角色以及自身的起源和进化等方面的研究更显得不足,而这些都需要更多的研究者在今后的工作中不断探索,为进一步揭示反转录转座子在基因组中的功能与作用机制奠定良好的理论基础。

参考文献

- McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA*, 1950, 36: 344–345
- Bainton R, Gamas P, Craig N L. Tn7 transposition *in vitro* proceeds through an excised transposon intermediate generated by staggered breaks in DNA. *Cell*, 1991, 65: 805–816
- Jr K H, Scott A F. ‘Copy and paste’ transposable elements in the human genome. *J Clin Invest*, 1993, 91: 1859–1860
- Wessler S R. Plant transposable elements: Where genetics meets genomics. *Nat Rev Genet*, 2002, 3: 329–341
- Schulman A H. Retrotransposon replication in plants. *Curr Opin Virol*, 2013, 3: 604–614
- Kalendar R, Vicient C M, Peleg O, et al. Large retrotransposon derivatives: Abundant, conserved but nonautonomous retroelements of barley and related genomes. *Genetics*, 2004, 166: 1437–1450
- Witte C P, Le Q H, Bureau T, et al. Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. *Proc Natl Acad Sci USA*, 2001, 98: 13778–13783
- Chen Z W, Wu W R. The plant retrotransposon and application (in Chinese). *Hereditas (Beijing)*, 2004, 26: 122–126 [陈志伟, 吴为人. 植物中的反转录转座子及其应用. *遗传*, 2004, 26: 122–126]
- Goodier J L. Restricting retrotransposons: A review. *Mobile DNA*, 2016, 7: 16
- Vassetzky N S, Kramerov D A. SINEBase: A database and tool for SINE analysis. *Nucleic Acids Res*, 2013, 41: D83–D89
- Consortium T G. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 2012, 485: 635–641
- Brenchley R, Spannagl M, Pfeifer M, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, 2012, 491: 705–710
- Schnable P S, Ware D, Fulton R S, et al. The B73 maize genome: Complexity, diversity, and dynamics. *Science*, 2009, 326: 1112–1115
- Li Y, Li C, Xia J, et al. Domestication of transposable elements into microRNA genes in plants. *PLoS One*, 2011, 6: e19212
- Carthew R W, Sontheimer E J. Origins and mechanisms of miRNAs and siRNAs. *Cell*, 2009, 136: 642–655
- 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. *PLoS Genet*, 2013, 8: e1002474
- Mccue A D, Slotkin R K. Transposable element small RNAs as regulators of gene expression. *Trends Genet*, 2012, 28: 616–623
- Chau N C, Kentaro N, Akihiro M, et al. Oligouridylation binding protein 1b plays an integral role in plant heat stress tolerance. *Front Plant Sci*, 2016, 7: 853–861
- Creasey K M, Zhai J, Borges F, et al. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature*, 2014, 508: 411–415
- Zhang X. Dynamic differential methylation facilitates pathogen stress response in *Arabidopsis*. *Proc Natl Acad Sci USA*, 2012, 109: 12842–12843
- Hsieh L C, Lin S, Shih C C, et al. Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing.

- Plant Physiol, 2010, 151: 2120–2132
- 22 Yuan J, Ye Z, Dong J, et al. Systematic characterization of novel lncRNAs responding to phosphate starvation in *Arabidopsis thaliana*. BMC Genomics, 2016, 17: 655
 - 23 Csorba T, Questa J I, Sun Q, et al. Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. Proc Natl Acad Sci USA, 2014, 111: 16160–16165
 - 24 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: 2654–2659
 - 25 Wang X, Ai G, Zhang C, et al. Expression and diversification analysis reveals transposable elements play important roles in the origin of Lycopersicon-specific lncRNAs in tomato. New Phytol, 2016, 209: 1442–1455
 - 26 Rory Johnson R G. The RIDL hypothesis: Transposable elements as functional domains of long noncoding RNAs. RNA, 2014, 20: 959–976
 - 27 White S E, Habera L F, Wessler S R. Retrotransposons in the flanking regions of normal plant genes: A role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA, 1994, 91: 11792–11796
 - 28 Kashkush K, Feldman M, Levy A A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nature Genet, 2003, 33: 102–106
 - 29 Hammond S M, Caudy A A, Hannon G J. Post-transcriptional gene silencing by double-stranded RNA. Nat Rev Genet, 2001, 2: 110–119
 - 30 Francozorrilla J M, Valli A, Todesco M, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet, 2007, 39: 1033–1037
 - 31 Kobayashi S, Gotoyamamoto N, Hirochika H. Retrotransposon-induced mutations in grape skin color. Science, 2004, 304: 982–982
 - 32 Lisch D. How important are transposons for plant evolution? Nat Rev Genet, 2013, 14: 49–61
 - 33 Masukawa T, Cheon K S, Mizuta D, et al. Insertion of a retrotransposon into a *flavonoid 3'-hydroxylase* homolog confers the red root character in the Radish (*Raphanus sativus* L. var. *longipinnatus* LH Bailey). Hort J, 2018, 87: 89–96
 - 34 Bhattacharyya M K, Smith A M, Ellis T H N, et al. The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. Cell, 1990, 60: 115–122
 - 35 Zhao H J, Cui H R, Xu X H, et al. Characterization of *OsMIK* in a rice mutant with reduced phytate content reveals an insertion of a rearranged retrotransposon. Theor Appl Genet, 2013, 126: 3009–3020
 - 36 Li X, Qian Q, Fu Z, et al. Control of tillering in rice. Nature, 2003, 422: 618–621
 - 37 Wu X, Tang D, Li M, et al. Loose Plant Architecture 1, an INDETERMINATE domain protein involved in shoot gravitropism, regulates plant architecture in rice. Plant Physiol, 2012, 161: 317–329
 - 38 Wang G L, Ruan D L, Song W Y, et al. *Xa21D* encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell, 1998, 10: 765–779
 - 39 Wessler S R, Varagona M J. Molecular basis of mutations at the waxy locus of maize: Correlation with the fine structure genetic map. Proc Natl Acad Sci USA, 1985, 82: 4177–4181
 - 40 Pereira A, Schwarz-Sommer Z, Gierl A, et al. Genetic and molecular analysis of the Enhancer (En) transposable element system of *Zea mays*. EMBO J, 1985, 4: 17–23
 - 41 Shure M, Wessler S, Fedoroff N. Molecular identification and isolation of the *Waxy* locus in maize. Cell, 1983, 35: 225–233
 - 42 Varagona M J, Purugganan M, Wessler S R. Alternative splicing induced by insertion of retrotransposons into the maize *waxy* gene. Plant Cell, 1992, 4: 811–820
 - 43 Hayashi K, Yoshida H. Refunctionalization of the ancient rice blast disease resistance gene *Pit* by the recruitment of a retrotransposon as a promoter. Plant J, 2009, 57: 413–425
 - 44 Cohen C J, Lock W M, Mager D L. Endogenous retroviral LTRs as promoters for human genes: A critical assessment. Gene, 2009, 448: 105–114
 - 45 Butelli E, Licciardello C, Zhang Y, et al. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell, 2012, 24: 1242–1255
 - 46 BRoslus J. Retroposons—Seeds of evolution. Science, 1991, 251: 753
 - 47 Betrán E, Thornton K, Long M. Retroposed new genes out of the X in *Drosophila*. Genome Res, 2002, 12: 1854–1859
 - 48 Long M, Langley C H. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. Science, 1993, 260: 91–95
 - 49 Wang W, Brunet F G, Nevo E, et al. Origin of *sphinx*, a young chimeric RNA gene in *Drosophila melanogaster*. Proc Natl Acad Sci USA, 2002, 99: 4448–4453
 - 50 Wang W, Yu H, Long M. Duplication-degeneration as a mechanism of gene fission and the origin of new genes in *Drosophila* species.

- Nature Genet, 2004, 36: 523–527
- 51 Wang W, Zheng H, Fan C, et al. High rate of chimeric gene origination by retroposition in plant genomes. *Plant Cell*, 2006, 18: 1791–1802
- 52 Xiao H, Jiang N, Schaffner E, et al. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science*, 2008, 319: 1527–1530
- 53 Arkhipova I R, Pyatkov K I, Meselson M, et al. Retroelements containing introns in diverse invertebrate taxa. *Nature Genet*, 2003, 33: 123–124
- 54 Volff J N. Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays*, 2006, 28: 913–922
- 55 Nekrutenko A, Li W H. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet*, 2001, 17: 619–621
- 56 Jiang N, Bao Z, Zhang X, et al. Pack-MULE transposable elements mediate gene evolution in plants. *Nature*, 2004, 431: 569–573
- 57 Elrouby N, Bureau T E. *BsI*, a new chimeric gene formed by retrotransposon-mediated exon shuffling in maize. *Plant Physiol*, 2010, 153: 1413–1424
- 58 Bureau T E, White S E, Wessler S R. Transduction of a cellular gene by a plant retroelement. *Cell*, 1994, 77: 479–480
- 59 Elrouby N, Bureau T E. A novel hybrid open reading frame formed by multiple cellular gene transductions by a plant long terminal repeat retroelement. *J Biol Chem*, 2001, 276: 41963–41968
- 60 Duan K, Ding X, Zhang Q, et al. *AtCopeg1*, the unique gene originated from *AtCopia95* retrotransposon family, is sensitive to external hormones and abiotic stresses. *Plant Cell Rep*, 2008, 27: 1065–1073
- 61 Hirochika H, Okamoto H, Kakutani T. Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. *Plant Cell*, 2000, 12: 357–368
- 62 Tsukahara S, Kobayashi A, Kawabe A, et al. Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature*, 2009, 461: 423–426
- 63 Ito H, Gaubert H, Bucher E, et al. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, 2011, 472: 115–119
- 64 Lippman Z, May B, Yordan C, et al. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol*, 2003, 1: 67
- 65 Grandbastien M A. Activation of plant retrotransposons under stress conditions. *Trends Plant Sci*, 1998, 3: 181–189
- 66 Grandbastien M A. LTR retrotransposons, handy hitchhikers of plant regulation and stress response. *Biochim Biophys Acta*, 2015, 1849: 403–416
- 67 Pouteau S, Huttner E, Grandbastien M A, et al. Specific expression of the tobacco Tnt1 retrotransposon in protoplasts. *EMBO J*, 1991, 10: 1911–1918
- 68 Pouteau S, Grandbastien M A, Boccara M. Microbial elicitors of plant defence responses activate transcription of a retrotransposon. *Plant J*, 2010, 5: 535–542
- 69 Mhiri C, Morel J B, Vernhettes S, et al. The promoter of the tobacco Tnt1 retrotransposon is induced by wounding and by abiotic stress. *Plant Mol Biol*, 1997, 33: 257–266
- 70 Grandbastien M A, Lucas H, Morel J B, et al. The expression of the tobacco Tnt1 retrotransposon is linked to plant defense responses. *Genetica*, 100: 241–252
- 71 Melayah D, Bonnard E, Chalhoub B, et al. The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. *Plant J*, 2010, 28: 159–168
- 72 Grandbastien M, Audeon C, Bonnard E, et al. Stress activation and genomic impact of Tnt1 retrotransposons in *Solanaceae*. *Cytogenet Genome Res*, 2005, 110: 229–241
- 73 Jääskeläinen M, Chang W, Moisy C, et al. Retrotransposon *BARE* displays strong tissue-specific differences in expression. *New Phytol*, 2013, 200: 1000–1008
- 74 Hirochika H, Sugimoto K, Otsuki Y, et al. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci USA*, 1996, 93: 7783–7788
- 75 Matsunaga W, Kobayashi A, Kato A, et al. The effects of heat induction and the siRNA biogenesis pathway on the transgenerational transposition of *ONSEN*, a *copia*-like retrotransposon in *Arabidopsis thaliana*. *Plant Cell Physiol*, 2012, 53: 824–833
- 76 Xu L, Yang J, Liu L, et al. Response of rice gypsy retrotransposons to different stress condition (in Chinese). *J Hunan Agr Univ (Nat Sci)*, 2012, 38: 591–596 [徐玲, 杨静, 刘林, 等. 水稻 gypsy 类逆转座子对不同胁迫条件的响应. 湖南农业大学学报(自然科学版), 2012, 38: 591–596]
- 77 Takeda S, Sugimoto K, Otsuki H, et al. A 13-bp *cis*-regulatory element in the LTR promoter of the tobacco retrotransposon *Tio1* is involved in responsiveness to tissue culture, wounding, methyl jasmonate and fungal elicitors. *Plant J*, 2010, 18: 383–393

- 78 Shinozaki K, Yamaguchi-Shinozaki K, Seki M. Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol*, 2003, 6: 410–417
- 79 Wu C. Heat shock transcription factors: Structure and regulation. *Annu Rev Cell Dev Biol*, 1995, 11: 441–469
- 80 Cavrak V V, Lettner N, Jamge S, et al. How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet*, 2014, 10: e1004115
- 81 Ohama N, Sato H, Shinozaki K, et al. Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci*, 2017, 22: 53–65
- 82 Jorgensen R A. Restructuring the genome in response to adaptive challenge: McClintock's bold conjecture revisited. *Cold Spring Harb Symp Quant Biol*, 2004, 69: 349–354
- 83 Arnholdt-Schmitt B. Stress-induced cell reprogramming. A role for global genome regulation? *Plant Physiol*, 2004, 136: 2579–2586
- 84 Flavell A J, Dunbar E, Anderson R, et al. *Ty1-copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Res*, 1992, 20: 3639–3644
- 85 Walsh A M, Kortschak R D, Gardner M G, et al. Widespread horizontal transfer of retrotransposons. *Proc Natl Acad Sci USA*, 2013, 110: 1012–1016
- 86 Xu L, Yang J, Liu L. Stress responses of three retrotransposon-related genes in rice (in Chinese). *Acta Phytopathol Sin*, 2014, 44: 147–155 [徐玲, 杨静, 刘林, 等. 三个水稻逆转座子相关基因的逆境响应特征. *植物病理学报*, 2014, 44: 147–155]
- 87 Sha A H, Zhang D P. Relations hip of activation of *Tos17* and rice adult plant resistance to bacterial blight (in Chinese). *Hereditas (Beijing)*, 2005, 27: 181–184 [沙爱华, 张端品. 逆转座子 *Tos17* 的激活与水稻白叶枯病成株抗性的关系. *遗传*, 2005, 27: 181–184]
- 88 Skibbe D S, Fernandes J F, Medzihradszky K F, et al. Mutator transposon activity reprograms the transcriptomes and proteomes of developing maize anthers. *Plant J*, 2009, 59: 622–633
- 89 Faulkner G J, Kimura Y, Daub C O, et al. The regulated retrotransposon transcriptome of mammalian cells. *Nature Genet*, 2009, 41: 563–571
- 90 Jordan I K, Rogozin I B, Glazko G V, et al. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet*, 2003, 19: 68–72
- 91 Feschotte C. Transposable elements and the evolution of regulatory networks. *Nature Rev Genet*, 2008, 9: 397–405
- 92 Elbarbary R A, Lucas B A, Maquat L E. Retrotransposons as regulators of gene expression. *Science*, 2016, 351: aac7247
- 93 Laten H M, Majumdar A, Gaucher E A. *SIRE-1*, a *copia/Ty1*-like retroelement from soybean, encodes a retroviral envelope-like protein. *Proc Natl Acad Sci USA*, 1998, 95: 6897–690
- 94 Wright D A, Voytas D F. Potential retroviruses in plants: Tat1 is related to a group of *Arabidopsis thaliana* *Ty3/gypsy* retrotransposons that encode envelope-like proteins. *Genetics*, 1998, 149: 703–715
- 95 Zaki E A. Plant retroviruses: Structure, evolution and future applications. *Afr J Biotechnol*, 2003, 2: 136–139
- 96 Garfinkel D J, Tucker J M, Saha A, et al. A self-encoded capsid derivative restricts Ty1 retrotransposition in *Saccharomyces*. *Curr Genet*, 2016, 62: 321–329
- 97 Laten H M, Gaston G D. Plant endogenous retroviruses? A case of mysterious ORFs. *Plant Transposable Elements*. Berlin, Heidelberg: Springer 2012. 89–112
- 98 Marco A, Marín I. Retrovirus-like elements in plants. *Recent Res Devel Plant Sci*, 2005, 3: 15–24
- 99 Kim A, Terzian C, Santamaria P, et al. Retroviruses in invertebrates: The gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc Natl Acad Sci USA*, 1994, 91: 1285–1289
- 100 Havecker E R, Gao X, Voytas D F. The diversity of LTR retrotransposons. *Genome Biol*, 2004, 5: 225
- 101 Subramanian R P, Wildschutte J H, Russo C, et al. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology*, 2011, 8: 90
- 102 Frendo J L, Olivier D, Cheynet V, et al. Direct involvement of HERV-W *Env* glycoprotein in human trophoblast cell fusion and differentiation. *Mol Cell Biol*, 2003, 23: 3566–3574
- 103 Ponferrada V G, Mauck B S, Wooley D P. The envelope glycoprotein of human endogenous retrovirus HERV-W induces cellular resistance to spleen necrosis virus. *Arch Virol*, 2003, 148: 659–675
- 104 Leroux C, Girard N, Cottin V, et al. Jaagsiekte sheep retrovirus (JSRV): From virus to lung cancer in sheep. *Vet Res*, 2007, 38: 211–228
- 105 Spencer T E, Mura M, Gray C A, et al. Receptor usage and fetal expression of ovine endogenous betaretroviruses: Implications for co-evolution of endogenous and exogenous retroviruses. *J Virol*, 2003, 77: 749–753
- 106 Eickbush T H, Malik H S. Origins and Evolution of Retrotransposons. *Mobile DNA II*. Washington: ASM Press, 2002. 1111–1144
- 107 Comfort N. When your sources talk back: Toward a multimodal approach to scientific biography. *J Hist Biol*, 2011, 44: 651–669

Summary for “植物反转录转座子功能研究进展”

Functional progress of retrotransposons in plants

Yuelong Lin^{1,2,3}, Kaizhuan Xiao^{1,2,3}, Ling Lian^{1,2,3}, Wei He^{1,2,3}, Liping Chen^{1,2,3}, Qiuhua Cai^{1,2,3}, Yingheng Wang^{1,2,3}, Hua'an Xie^{1,2,3*} & Jianfu Zhang^{1,2,3*}

¹ Rice Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou 350019, China;

² Key Laboratory of Germplasm Innovation and Molecular Breeding of Hybrid Rice for South China, Ministry of Agriculture, China/Incubator of National Key Laboratory of Fujian Germplasm Innovation and Molecular Breeding Between Fujian and Ministry of Sciences & Technology/ National Rice Engineering Laboratory of China/Base of South-China, National Key Laboratory of Hybrid Rice/Fuzhou Branch, National Rice Improvement Center of China/Fujian Engineering Laboratory of Crop Molecular Breeding/ Fujian Key Laboratory of Rice Molecular Breeding, Fuzhou 350003, China;

³ State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Fujian Academy of Agricultural Sciences, Fuzhou 350003, China

* Corresponding authors, E-mail: huaanxie@163.com; jianfzhang@163.com

Retrotransposon is a type of transposable elements (TEs), which can replicate itself and move to other loci through the form of DNA→RNA→DNA intermediated with RNA in the genome, and belongs to the class I type transposon.

Basically, retrotransposons can be divided into two mainly categories including LTR (long terminal repeat) and non-LTR (non-long terminal repeat). The LTR retrotransposon can be furtherly divided into two main groups including Ty1/copia and Ty3/gypsy, which have a long terminal repeat structure similar to the retrovirus genome. The LTR usually do not encode proteins, but it can serve as regulator for transcription and termination and so on. Non-LTR retrotransposon includes the LINE (Long interspersed nuclear elements) and SINE (Short interspersed nuclear elements). The SINE retrotransposons do not encode the enzymes required for transposition and cannot complete the retrotransposition autonomously.

Retrotransposons are the main components in the genome of most eukaryotic, which are also abundant in higher plants. For example, they account for more than 60 percent of the genome sequence in tomatoes, while nearly 80 percent of the genome sequence is made up of retrotransposons in corn and wheat.

As early as 1950s, the famous American geneticist, Barbara McClintock, had discovered the existence of transposon when studying the variation of color spots in corn kernels. However, the function and biological significance of those transposons including retrotransposons have always been controversial. But more and more researches reveal that the retrotransposons play an important role in the regulation of gene expression and evolution of plant genome in recent years. Some activated transcript of retrotransposons can participate the production of non-coding RNAs, such as small RNAs or lncRNAs, and rely on the targeting effect of these non-coding RNAs, and are further involved in the expression regulation of downstream-related genes. The transposition of retrotransposons always influences gene structure and expression of insertion loci and neighborhood. They can act as cis- or trans-acting elements in the upstream or downstream regions of genes, or directly as promoter and terminator and so on to regulate the expression, transcription and function of target genes. Besides, the enzyme system produced by autonomous retrotransposon may also act on the other genes of genome, which may result in the creation of chimeric genes or production of new gene copy. Furthermore, some retrotransposon sequences may have involved in the generation of new functional genes during the long-term evolution of plants. Obviously, they are of great significance to the plants.

Moreover, retrotransposon may play a role in physiological changes and stress responses of plants, for instance, the plant growth and development regulation, disease resistance and stress-related response and so on. They can even reprogramming plant cell gene expression or creating a considerable number of natural ecotype variations.

Intriguingly, the subtle relationship in the origin and evolution of retrotransposons and retroviruses deserves further exploration. They have too much in common and similarities, which indicate that they must have some evolutionary connections. From the study of retrotransposon, we may obtain some unknown secrets about mammalian retrovirus.

In this paper, the function progress of the retrotransposon in plants are briefly summarized from the above aspects, which may lay a good theoretical foundation for further revealing the function and mechanism of retrotransposon in the genome and incite the enthusiasm of researchers to study these elements.

retrotransposon, smRNA, lncRNA, retrogene, stress, retrovirus

doi: 10.1360/N972018-00877