

# 植物反转录转座子的研究进展\*

刘 柱<sup>1)</sup> 朱建清<sup>1,2)</sup> 赵 建<sup>1)</sup> 陈东辉<sup>1)</sup> 杨志荣<sup>1)</sup>\*\*

(<sup>1)</sup>四川大学生命科学学院, 成都 610064; <sup>2)</sup>四川农业大学水稻研究所, 温江 611130)

**摘要** 反转录转座子是生物界中存在的一类可移动的遗传因子, 其转座功能通过 RNA 介导反转录来实现. 它在植物界中普遍存在, 并在植物基因和基因组进化中扮演了一个极其重要的角色. 概述了植物反转录转座子的类型结构及作为遗传工具在生物学中的作用.

**关键词** 逆转录因子, 植物反转录转座子, 遗传工具

**学科分类号** Q344

植物反转录转座子 (plant retrotransposon) 作为一类可移动的遗传因子在植物界广泛存在. 自 1984 年 Shepherd 等<sup>[1]</sup>发现植物反转录转座子以来, 许多研究小组进行了相关的研究. 研究表明它们是寄生性 DNA, 整合到宿主基因组之后, 和宿主基因遵循相同的遗传规律, 能进行复制和表达. 植物反转录转座子是构成植物基因组的主要成分, 以多拷贝形式出现. 其转座过程是转座因子的 DNA 先被转录成 RNA, 再借助反转录酶/RNase H 反转录成 DNA, 插入到新的染色体位点. 一般的 DNA 转座因子 (如 Ac, *Tam1* 和 *En/Spm*) 是利用切除/修复机制来实现转座, 不增加基因组的大小, 而反转录转座子通过复制实现转座, 增加了转座元件的拷贝数, 从而极大地增加了植物基因组的大小. 和 DNA 转座子一样, 反转录转座子也能通过插入基因附近或内部而导致基因突变或重排, 并且因其特有的复制模式, 保留了插入位点的序列, 引起的突变相对稳定. 由于反转录转座子具有存在广泛、高拷贝数、插入位点专一性等特点, 在研究植物基因组的组成, 表达调控、基因组进化、系统发育及生物多样性评价中, 目前已经受到了广泛的关注<sup>[2~5]</sup>.

## 1 植物反转录转座子的类型结构及分配

### 1.1 植物反转录转座子的类型

反转录转座子是真核转座元件中最为丰富的一类, 包括长末端重复反转录转座子 (the long terminal repeat retrotransposon, LTR) 和非长末端重复反转录转座子 (the non-long terminal repeat retrotransposon, non-LTR). LTR 反转录转座子又可分为 *Ty1-copia* 组和 *Ty3-gypsy* 组, 它们在植物基因组中都是高拷

贝 (每单倍体细胞核中的拷贝数在  $10^6$  以上). *Ty1-copia* 组在单细胞藻类、苔藓植物、裸子植物和被子植物中广泛存在; *Ty3-gypsy* 组存在于裸子植物和被子植物中. non-LTR 反转录转座子在植物物种中也是高拷贝 (拷贝数在  $2.5 \times 10^5$  以上), 它又分为长散在元件 (long interspersed elements, LINE) 和短散在元件 (short interspersed elements, SINE). LINE 在植物界中大量存在, SINE 仅出现在被子植物中.

### 1.2 反转录转座子的结构

目前已报道的反转录转座子结构主要有 4 种 (图 1).

**1.2.1 LTR 反转录转座子中的 *Ty1-copia* 组:** LTR 不编码任何已知的蛋白质, 但含有与该类转座子转录密切相关的启动子和终止子, LTR 通常以反向重复序列 5'-TG-3' 和 5'-CA-3' 结束. LTR 反转录转座子由一条含 5'-R-U5-PBS-编码蛋白区-PPT-U3-R-3' 的 mRNA 分子编码. 编码蛋白区 *gag*, *pol*, *int* 基因编码的多聚蛋白前体合成后, 被 *pol* 编码的蛋白水解酶 (PR) 切割为功能性的多肽. *gag* 基因编码的蛋白质负责反转录转座子 RNA 的成熟和包装; *pol* 编码反转录酶和 RNase, 用来实行反转录转座子的复制/转座功能; *int* 基因编码的整合酶协助反转录转座子以 DNA 形式插入一个新的染色体位点. 一般而言, *gag*, *pol* 和 *int* 基因所编码的蛋白质在同一个翻译阅读框架里, 但有时也有两三个阅读框架存在.

\* 四川省杰出青年基金资助项目 (01ZQ052).

\*\* 通讯联系人.

Tel: 028-5412485, E-mail: liuzhuabel@yahoo.com.cn

收稿日期: 2001-12-18, 接受日期: 2002-01-29

LTR 反转录转座子的复制过程较为复杂. 首先要转录为 mRNA 分子, 该分子既作为复制所需的模板, 又要翻译为复制相关的蛋白质. mRNA 分子中的引物结合位点 (PBS) 与宿主细胞内对应的 tRNA 互补结合, 二者形成短的 RNA 双链, 反转录酶利用 tRNA 3' 端自由羟基为引物, 合成互补

的 DNA, 编码的 RNase H 专一地消化 RNA, 得到单链 DNA, 在 RNase H 和多聚嘌呤位点 (PPT) 的作用下合成双链线性 DNA 分子, 整合酶切割转座子 DNA 及靶基因产生 3~5 个碱基的切口, 插入到目标分子中, 结果在靶基因侧翼产生了 3~5 bp 的同向重复.

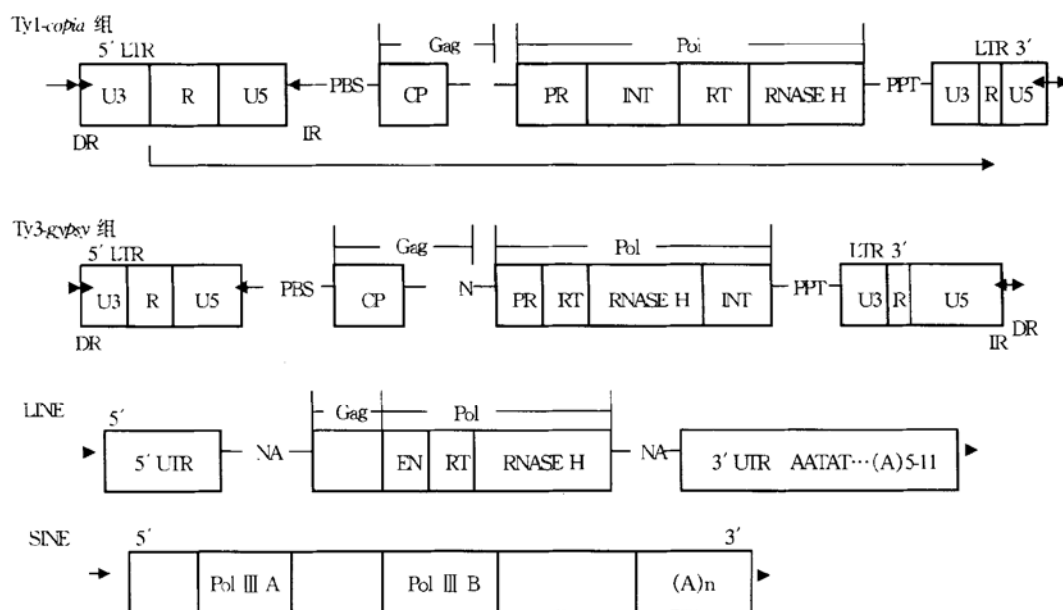


Fig. 1 Structures of the Ty1-*copia*, Ty3-*gypsy*, LINE, SINE retrotransposons

图 1 Ty1-*copia*, Ty3-*gypsy*, LINE, SINE 反转录转座子的结构

在植物中, 最早在拟南芥 (*Arabidopsis thaliana*) 中对来源于烟草的 *Tnt 1* 转座子进行关于反转录转座功能的研究<sup>[6]</sup>, 与反转录病毒和酵母反转录转座子 Ty3 一样, *Tnt 1* 转座元件的复制是以引物结合位点 (PPS) 和 PPT 作为正链和负链起始引物, 产生双链线性 DNA. 对大麦 *Hordeum* 中 *BARE-1* 反转录转座子的整合酶 (INT) 中心结构域的二级及三级结构研究表明, 它有反转录病毒 INT 的典型结构<sup>[7]</sup>. 至于 LTR 反转录转座子的复制是否和其他真核生物相同, 还需进一步研究.

**1.2.2 LTR 反转录转座子中的 Ty3-*gypsy* 组:** Ty3-*gypsy* 组和 Ty1-*copia* 组在基因组结构和编码蛋白的功能上类似, 仅 *int* 和 *pol* 基因的顺序不同. 反转录酶序列比较, 进一步暗示这二组反转录转座子由不同的血统形成, 通过序列及进化树的分析, 表明动物反转录病毒来源于 Ty3-*gypsy* 组. 当前, 人们普遍认为 Ty3-*gypsy* 组、Ty1-*copia* 组和反转录病毒在细胞内的转录、复制和整合的机制是一样的.

**1.2.3 non-LTR 反转录转座子中长散在元件 LINE:** LINE 元件较 LTR 反转录转座子简单, 也含 *gag* 和 *pol* 基因, 但缺少鉴定识别的整合酶. *gag* 位点产生具有核酸内切酶 (EN) 活性的蛋白质, 把 LINE 的 DNA 整合到宿主 DNA 中; 另外, LINE 也可能通过 DNA 切除/修复合活性整合进宿主染色体. 序列比较说明 LINE 元件是真核生物中最原始的, LTR 反转录转座子可能由 LINE 进化而获得.

在植物中, 对 LINE 转录和整合的机制研究较少. 植物中的 LINE 和人的 L1 因子一样, 其 5' 端被截短了. RNA 聚合酶 II 和 RNA 聚合酶 III 参与 L1 因子的转录, 进而双顺反子 L1 mRNA 被翻译为 *gag* 和 *pol* 蛋白; 核酸内切酶 (EN) 同时切开靶位点和 L1 因子反转录后获得的 DNA, 进行非同源重组.

**1.2.4 non-LTR 反转录转座子中短散在元件 SINE:** 被称为 SINE 的小反转录转座子不同于其他反转录转座子, 因为它不编码任何起顺式作用功

能的蛋白质。所知的 SINE 都源于 RNA 聚合酶 III 的产物 (如 tRNA), 它能利用 LINE 或 LTR 反转录转座子所编码的蛋白质来进行自身的复制、整合, 因此认为 SINE 非常象无内含子的假基因。与 RNA 聚合酶 II 产物所转录的 mRNA 相比, RNA 聚合酶 III 转录的基因通常包括 RNA 编码区在内的启动子序列。在进化过程中, SINE 序列向有利于反转录和整合的方向进化。尽管 SINE 如何复制整合还不清楚, 但侧翼靶 DNA 改变类型的相似性及 poly A 尾的共通性都暗示了它采用了一些 LINE 特有的功能<sup>[8]</sup>。

## 2 植物反转录转座子在生物学中的应用

### 2.1 作为分子标记用于生物多样性及遗传连锁分析

用反转录转座子序列来进行分子标记极为有利。它们作为高度异源种群以多拷贝的形式广泛分布在染色体上, 并在物种间或物种内表现出插入位点的多态性; 而且, 活跃的反转录转座子可以在基因组中制造新的插入位点, 用来在谱系中建立系统发生关系。它作为 DNA 标记广泛用于玉米、豆类及大麦中的生物多样性研究<sup>[9, 10]</sup>, 并用于大麦和豆类中作遗传连锁图<sup>[11]</sup>。因为许多反转录转座子广泛分布在染色体的常染色质区内, 它可能产生与农业重要基因相连锁的标记。利用反转录转座子 *Tnd-1* 在烟草中已找到了一个与黑根抗性相连锁的标记<sup>[12]</sup>。

序列特异性扩增多态性 (sequence-specific amplification polymorphism, SSAP) 是一种以反转录转座子为基础的多态性标记技术, 它与 AFLP 标记技术类似。如果单个反转录转座子在基因组的某个位点插入, 凝胶电泳时就表现为一条带。因为 SSAP 标记时其中一条引物是根据反转录转座子的碱基序列设计的, 所以估计植物系统发生关系时比传统的 AFLP 标记更为有利; 多反转录转座子方法也用来研究豆科和禾本科作物中种内种间的系统发育关系。

利用反转录转座元件的序列及其侧翼 DNA 序列, 设计引物进行 PCR 来检测单个反转录转座子插入位点, 这种方法已得到了广泛的应用。位点上等位基因不同的状态 (反转录转座子的存在和缺失) 都可被发现。这种以反转录转座插入位点多态性 (retrotransposon-based insertion polymorphism, RBIP) 的优点是快捷方便。对一个富含微卫星的大麦品种序列分析表明, 反转录转座子序列和微卫星密

切相关<sup>[13]</sup>。反转录转座-微卫星扩增多态性 (retrotransposon microsatellite amplified polymorphism, REMAP) 能区分大麦的不同品种及对同一属中不同物种间构建指纹图<sup>[14]</sup>。

Non-LTR 反转录转座子也可作为分子标记用来对物种及属间的遗传多样性分析。例如, 利用水稻 *p-SINE1* 因子论证了带有 AA 染色体基因组的 *Oryza rufipogon* 和 *O. sativa* 发生分离之后, *p-SINE-1-r2* 才转座到 *wx* 位点上, 而在不同的水稻品种分离之前, *p-SINE-1-r1* 已转座到 *wx* 位点的另一位置上<sup>[15]</sup>。*Brassica* SINE 元件用在 *Cruciferae* 物种间遗传多样性的研究, 并用来分析 *Brassica* 栽培种和野生种间渐渗现象的发生程度<sup>[16]</sup>。

### 2.2 基因标签及基因的功能分析

反转录转座子的特征使之作为理想的遗传工具用于植物基因标签研究。a. 反转录转座子的插入突变都是通过复制的模式发生; b. 产生大量因饱和和诱变而获得的随机插入转座靶位点; c. 通过生物和非生物的压力条件来实现转座; d. 导致基因的高度诱变; e. 低拷贝数对因特殊突变引起的反转录转座子插入位点的鉴别; f. 内生性反转录转座子在植物中广泛存在, 并十分活跃。

在水稻中, 内生性 *Tos17* 反转录转座子在上述条件下可用作基因标签。它在组织培养条件下非常活跃, 优先插入基因内或基因附近; 它还可通过饱和和突变获得大量随机插入位点, 并用于克隆水稻中的基因<sup>[17]</sup>。另外, *Tos17* 还通过反遗传方法鉴定 *OSH15* 基因 (*Oryza sativa homeobox*) 中的无效突变种系, 证明了 *OSH15* 在水稻节间伸长时充当了重要角色<sup>[18]</sup>。

### 2.3 体细胞变异的进化

植物的遗传操纵及营养繁殖的许多方法都依靠植物的离体再生系统。然而, 体外操作过程中都倾向于采用体细胞变异, 离体培养过程中, 常需选择突变频率高的时间。在细胞和组织培养条件下, 反转录转座子被诱导的活性直接影响突变的频率, 如突变频率超过 10% 的水稻再生植株都经过反转录转座子诱变的过程<sup>[19]</sup>。

## 3 结 语

植物反转录转座子的转座复制模式使它能在宿主内快速地增殖, 极大地增加了植物基因组的大小, 成为高等植物基因组的重要组成部分。另外, 在基因内或基因附近, 反转录转座子的插入突变引

起基因表达的失活或改变, 因而许多包括反转录转座子片段的启动子都用来研究基因的调控, 同时位点之间的异位重组, 增加或减少了基因组的类型和基因家族的数量. 尽管它最初是寄生性的, 但反转录转座子还是会被寄主持持续地利用. 然而, 是否反转录转座子诱导的改变在自然群体中积极地得到选择还需进一步证明.

当前, 对植物反转录转座子研究并不多, 关于植物反转录转座子家族的来源, 还不太清楚. 某些LTR反转录转座子是通过缺陷型昆虫反转录病毒经水平传播到达植物基因组, 还是细胞内因子通过广泛杂交或部分吸收病原体的序列而获得? 家族内的反转录元件通过什么机制发生改变的? 基因组如何消除反转录转座子? 哪些反转录元件利用另外反转录元件的顺式作用因子来实现复制和插入的? 这些都因反转录转座子的多样性、高拷贝、不可预测的扩增潜力及快速进化等特点而难以得到答案, 所有的问题还需要在今后的工作中解决.

### 参 考 文 献

- 1 Shepherd N S, Schwarz-Sommer Z, Blumberg vel Spalve J, *et al.* Similarity of the *Cin1* repetitive family of *Zea mays* to eukaryotic transposable elements. *Nature*, 1984, **307** (5947): 185~ 187
- 2 Grandbastien M A. Activation of plant retrotransposons under stress conditions. *Trends Plant Science*, 1998, **3** (5): 181~ 187
- 3 Pearce S R, Harrison G, Flavell A J, *et al.* The *Ty1-copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Res*, 1996, **4** (9): 357~ 364
- 4 Hirochika H. Contribution of the *Tos17* retrotransposon to rice functional genomics. *Curr Opin Plant Biol*, 2001, **4** (2): 118~ 122
- 5 Kumar A, Pearce S R, McLean K, *et al.* The *Ty1-copia* group of retrotransposons in plants: genomic organization, evolution, and use as molecular markers. *Genetica*, 1997, **100** (3): 205~ 217
- 6 Lucas H, Feuerbach F, Kunert K, *et al.* RNA-mediated transposition of the tobacco retrotransposon *Tnt1* in *Arabidopsis thaliana*. *EMBO J*, 1995, **14** (10): 2364~ 2373
- 7 Suoneimi A, Tanskanen J, Pentikainen O, *et al.* The core domain of retrotransposon integrase in *Hordeum*. *Mol Biol Evol*, 1998, **15** (9): 1135~ 1144
- 8 Okada N, Hamada M, Ogiwara I, *et al.* SINEs and LINEs share common 3' sequences: a review. *Gene*, 1997, **205** (2): 229~ 243
- 9 Ellis T H N, Poyser S J, Knox M R, *et al.* Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. *Mol Gen Genet*, 1998, **260** (1): 9~ 19
- 10 Purugganan M D, Wessler S R. Transposon signatures-species-specific molecular markers that utilize a class of multiple-copy nuclear DNA. *Mol Ecol*, 1995, **4** (2): 265~ 269
- 11 Wessler S R, Bureau T E. LTR- retrotransposons and MITEs, important players in the evolution of plant genomes. *Curr Opin Genet Dev*, 1995, **5** (9): 814~ 821
- 12 Kenward K D, Bai D, Ban M R, *et al.* Isolation and characterization of *Tnd1*, a retrotransposon marker linked to black root resistance in tobacco. *Theor Appl Genet*, 1999, **98** (4): 387 ~ 395
- 13 Ramsay L, Macaulay M, Cardle L, *et al.* Intimate association of microsatellite repeats with retotransposons and other dispersed repetitive elements in barley. *Plant J*, 1999, **17** (8): 415~ 423
- 14 Kalendar R, Grob T, Regina M, *et al.* IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. *Theor Appl Genet*, 1999, **98** (6): 704~ 711
- 15 Hirano H Y, Mochizuki K, Umeda M, *et al.* Retrotransposon of a plant SINE into the *wx* locus during evolution of rice. *J Mol Evol*, 1994, **38** (2): 132~ 137
- 16 Lenoir A, Cournoyer B, Warwick S, *et al.* Evolution of SINE S1 retroposons in Cruciferae plant species. *Mol Biol Evol*, 1994, **14** (6), 934~ 941
- 17 Hirochika H. Activation of tobacco retrotransposons during tissue culture. *EMBO J*, 1993, **12** (6): 2521~ 2528
- 18 Sato Y, Sentoku N, Miur Y, *et al.* Loss-of-function mutations in the rice homeobox gene *OSH15* affect the architecture of internodes resulting in dwarf plants. *EMBO J*, 1999, **18** (4): 992~ 1002
- 19 Hillick J B, Dorweler J E, Chandler V L, *et al.* Paramutation and related allelic interaction. *Trends Gent*, 1997, **13** (10): 302~ 308

## Progress on Plant Retrotransposons\*

LIU Zhu<sup>1)</sup>, ZHU Jian-Qing<sup>1,2)</sup>, ZHAO Jian<sup>1)</sup>, CHEN Dong-Hui<sup>1)</sup>, YANG Zhi-Rong<sup>1)\*\*</sup>

<sup>1)</sup> Life Science College, Sichuan University, Chengdu 610064, China;

<sup>2)</sup> Rice Research Institute, Sichuan Agricultural University, Wenjiang 611130, China)

**Abstract** Retrotransposons are mobile genetic elements that transpose through reverse transcription of RNA intermediate. Retrotransposons are ubiquitous in plants and play a major role in plant gene and genome evolution. The types and structures of plant retrotransposons were summarized. The retrotransposons were used as genetic tools in plant biology.

**Key words** retroelement, plant retrotransposons, genetic tools

\* This work was supported by a grant from The Youth Foundation of Sichuan Province (01ZQ052).

\*\* Corresponding author. Tel: 86-28-5412485, E-mail: liuzhuabel@yahoo.com.cn

Received: December 18, 2001 Accepted: January 29, 2002