

·专题介绍·

水稻转座子研究进展

高东迎*, 何冰, 孙立华

江苏省农业科学院粮食作物研究所, 南京 210014

摘要 转座子是植物基因组的重要组成部分, 对于研究植物基因组进化等具有重要意义。随着水稻全基因组测序计划的开展和完成, 水稻转座子研究取得了极大进展, 目前已经在水稻基因组中发现了几乎所有类型的转座子, 约占水稻基因组的35%。在正常情况下, 大多数水稻转座子不具有转座活性, 但是在特定的条件下(如组织培养或辐射等), 水稻基因组中沉默的转座子可以被激活, 从而可能导致插入突变并影响基因的表达。在水稻中已鉴定出6个有活性的转座子, 其中*Tos17*已被应用到水稻功能基因组研究中。转座子序列的新的分子标记转座子展示(transposon display, TD)现已被开发, 并在水稻遗传作图和遗传分化研究中得到应用。

关键词 基因表达, 水稻, 转座子, 转座子展示

高东迎, 何冰, 孙立华 (2007). 水稻转座子研究进展. 植物学通报 24, 667–676.

转座子(transposable elements 或 transposons)是指基因组中那些能够移动或复制自己并整合到新位点的DNA片段(Curcio and Derbyshire, 2003), 其对于研究植物基因组的组成、进化和基因的表达调控等都具有重要意义(Feschotte et al., 2002)。水稻是世界重要粮食作物, 禾本科植物分子生物学的模式植物。近年来, 水稻转座子研究受到越来越多学者的重视, 并已取得较大进展。本文将对水稻转座子研究所取得的一些新进展进行归纳。

1 水稻基因组中转座子的种类

传统观念认为, 水稻基因组中不存在转座子, 但随着水稻分子生物学的发展, 特别是水稻全基因组测序的开展和完成, 科学家们意外发现, 在水稻基因组中不仅有转座子, 而且几乎包括所有类型转座子(Mao et al., 2000; Turcotte et al., 2001; Jiang et al., 2004b; International Rice Genome Sequencing Project, 2005)。转座子约占水稻基因组组成的35%, 其中第1类转座子(Class

I, 也称反转录转座子)和第2类转座子(Class II, 也称DNA转座子)分别占19.4%和14.0%, 但从数目上讲, 第2类转座子要远多于第1类转座子, 这是因为第2类转座子包括了大量微小转座子(表1)(International Rice Genome Sequencing Project, 2005)。现对水稻的主要类型转座子介绍如下。

1.1 MITEs

微小反向重复转座子(miniature inverted repeat transposable element, MITEs)是水稻基因组中数量最多的一类转座子, 大约有 90 000 个(Jiang et al., 2004b)。MITEs为非自主DNA类转座子, 但是其序列小(一般为100 - 500 bp)且拷贝高, 具有插入位点偏爱性, 使得其与一般非自主DNA类转座子又有明显不同。由于MITEs不编码转座酶(transposase), 其分类主要依据非编码区的相似性, 如MITEs的末端反向重复(terminal inverted repeats, TIRs)及其插入到基因组后所形成的2 - 3 bp的同向重复序列(target site duplications, TSDs)。根据这个标准, 大多数水稻MITE被分为*Tourist*(3 bp的TSD,

收稿日期: 2006-11-08; 接受日期: 2007-04-09
基金项目: 国家自然科学基金(No. 30471066)
* 通讯作者。E-mail: gao.dongying@yahoo.com.cn

通常为TTA或TAA)和Stowaway(2 bp的TSD, 通常为TA)2类(Zhang et al., 2004)。

1.2 MULEs

MULEs(*Mutator*-like transposable elements)是水稻基因组中的另一大类DNA类转座子, 包括自主和非自主2种类型。此类转座子一般特征是: (1)具有较长的末端反向重复(一般大于100 bp); (2)大小从几百bp到30 kb左右不等, 且内部序列变化较大; (3)有9-11 bp的靶位点重复。与其它家族转座子不同的是, 非自主 MULEs 并不都是由对应的自主转座子基因序列缺失导致其转座酶失活; 在有些情况下, 非自主MULEs可以俘获寄主基因片段, 这种带有寄主基因片段的 MULEs 称为 Pack-MULEs, 其对水稻基因组进化具有重要作用(Jiang et al., 2004a)。目前在水稻基因组中已鉴定出 8 274 个 MULEs, 其中有 1 337 个带有寄主基因片段(Juretic et al., 2005)。

1.3 CACTA

CACTA 家族(也叫 En/Spm 家族)为 DNA 类转座子, 该家族有自主和非自主转座子 2 类。CACTA 类转座子最典型特征是反向重复序列末端是保守的CACTA; 同时, 该类转座子具有长 10-28 bp 末端反向重复, 是转座酶的识别位点。其另一个特点是反向重复序列区保守性较低。在水稻基因组中 CACTA 转座子有将近 1.1 万个拷贝, 约占整个基因组的2.69%(International Rice Genome Sequencing Project, 2005)。Tnr3 是最早发现的水稻CACTA转座子, 是作为插入片段被鉴定出来的, 其大小为 1 539 bp, 并有 13 bp 的反向末端重复(末端以 5'-CACTA-3'开始)(Motohashi et al., 1996)。

1.4 hAT

hAT 是真核生物细胞中广泛存在的另一大类转座子。玉米*Ac*转座子是最早发现的该类转座子, 随后在金鱼草和黑腹果蝇中也发现了类似于 *Ac* 的转座子, 分别称为 *Tam3* 和 *hobo*。1994 年 Warren 等取最早发现的 3 个该类转座子(*hobo*、*Ac* 和 *Tam3*) 的第 1 个字母, 将该

类转座子命名为(*hAT*)(Warren et al., 1994)。*hAT* 转座子的特征是: (1)在转座过程中产生 8 bp 的靶位点重复; (2)有5-27 bp的短反向末端重复; (3)大多数*hAT*类转座子小于 4 kb (Kempken and Windhofer, 2001)。水稻中 *hAT* 类转座子有 1.1 万个, 数目与 CACTA 类转座子接近, 但其长度较小, 平均约为 1.27 kb, 因而在水稻基因组中的比例较低(International Rice Genome Sequencing Project, 2005)。

1.5 LTR retrotransposons

LTR类反转录转座子的典型特征是在转座子的两端具有正向长末端重复(long terminal repeat, LTR), 主要包括 *gypsy* 和 *copia* 两大类。在水稻基因组中 *copia* 类转座子长度一般为 4-6 kb, 而 *gypsy* 类转座子长度一般为 10-13 kb, 其大小约为 *copia* 类的 2 倍; 水稻 LTR 类反转录转座子的另一个特点是具有 5 bp 的靶位点重复(McCarthy et al., 2002)。LTR类反转录转座子不仅拷贝数高, 而且该类转座子的平均长度大于其它类型转座子(如MITes和CACTA等), 因而该类转座子在水稻基因组中所占的比例大约在 15% 或更高(McCarthy et al., 2002; Ma and Bennetzen, 2004; International Rice Genome Sequencing Project, 2005)。

1.6 其它TEs

*Helitron*是最近被鉴定出的另一类DNA类转座子, 在水稻基因组中大量存在, 与其它类型转座子最明显的区别是此类转座子不形成靶位点重复。水稻约有 10 000 多个 *Helitron*, 占水稻基因组的 2% 左右。另外, 在水稻基因组中也包括大量的LINEs 和SINEs, 但其数目不如上面提到的几大类转座子多(表 1)。

2 水稻转座子的发现鉴定

虽然转座子既古老又在水稻基因组中大量存在, 但是多年来一直没有被发现。直到 1992 年, 日本学者 Hirochika等(1992)首先发现水稻中存在反转录转座子。此后科学家们鉴定发现了大量水稻转座子, 使得人们对

表 1 水稻基因组中的转座子

Table 1 Transposons in the rice genome

转座子	拷贝数 ($\times 10^3$)	含量 (kb)	占水稻基因 组的比例 (%)
Class I			
LINES	9.6	4 161.3	1.12
SINES	1.8	209.9	0.06
<i>Ty1/copia</i>	11.6	14 266.7	3.85
<i>Ty3/gypsy</i>	23.5	40 363.3	10.90
其它 class I 转座子	15.4	12 733.3	3.43
总 class I 转座子	61.9	71 734.4	19.35
Class II			
<i>hAT</i>	1.1	1 405.9	0.38
CACTA	10.8	9 987.3	2.69
IS630/ <i>Tc1/mariner</i>	67	8 388.3	2.26
IS256/ <i>Mutator</i>	8.8	13 485.7	3.64
IS5/ <i>Tourist</i>	57.9	12 095.8	3.26
其它 class II 转座子	18.2	2 703.6	0.73
总 class II 转座子	163.8	48 066.6	12.96
其它转座子	23.6	6 797.7	1.80
总的转座子	249.3	129 019.3	34.79

其它转座子是指那些不能够被容易地划分为第 1 类或第 2 类转座子的 DNA 组分

Other TEs mean those DNA fractions that can not be easily classified into either class I or II

水稻转座子有了全新的认识,这在很大程度上得益于转座子鉴定思路和相关技术的创新以及水稻全基因组测序计划的开展和完成。鉴定水稻转座子的方法很多,但概括起来主要有以下几种。

2.1 传统方法

转座子可通过转座插入到基因组中某些位点(如基因或其调控序列),引起插入位点(靶位点)基因失活或影响其表达,从而导致表型变化形成突变型。有时转座子可以从原来插入位点跳出,使得靶位点基因恢复活性而表现正常性状。转座子从靶位点切除常发生在植株发育过程中,可以发现一些组织表型正常而另一些组织表型异常,即出现所谓的嵌合性状,这种不稳定的变异常常是由于活性 DNA 类转座子活动所致。由于转座子发生回复突变的频率一般较低,所以可通过对转座子插入导致的突变体和野生型的遗传分析,最终鉴定出转座子。如 Teraishi 等(1999)从 γ -射线辐照的粳稻 Gimbozu 种子后代中发现一个籽粒颖壳细长突变体 IM294,该突变体经

多次自交繁殖仍不能完全稳定,在同一单株不同穗之间或同一稻穗不同支梗之间籽粒形状存在嵌合情况。遗传分析表明,突变性状由一对隐性基因(*SLG*)控制,用三体和 RFLP 分析将基因定位到第 7 号染色体上(Teraishi et al., 1999)。图位克隆分析表明,突变体有 1 个 433 bp 插入片段,并发现插入片段序列为 1 个 MITEs 类转座子,命名为 *mPing* (Nakazaki et al., 2003)。

2.2 引物杂交法

由于反转录转座子引起的变异稳定遗传,通过传统遗传方法难以将由转座子插入引起的突变和其它变异区分开,反转录转座子必须通过其它方法才可以鉴定出来。引物杂交法鉴定反转录转座子的主要原理是:有些反转录转座子 5' 端 LTR 下游有一个包含十几个碱基组成的 PBS(primer binding site, 引物结合位点)与起始蛋氨酸的 tRNA(tRNA_{Met}) 3' 端序列互补,而且植物中 tRNA_{Met} 序列是高度保守的(Sprinzl et al., 1987);故可利用一段与起始蛋氨酸 tRNA 3' 端互补的寡核苷酸序列作为探针,筛选水稻基因组文库,获得水稻反转录转座子的片段,对这些片段进行测序或再作为探针继续筛选分析,最终可得到完整的水稻反转录转座子。由于 *Copia* 类和 *gypsy* 类反转录转座子均有 LTR,所以利用这一方法可以获得这 2 类转座子,如 Hirochika 等(1992)从水稻中鉴定出 3 个反转录转座子家族,分别命名为 *Tos1-Tos3*。但该方法要经过大量的分子杂交和测序,工作量大,而且假阳性较高,故在当前研究中已较少应用。

2.3 基于转座酶保守序列的 PCR 法

转座子可以分为不同的家族,不同家族的转座酶差异较大,但同一家族内的转座酶的特定区域具有保守性,因而可以根据其它生物中已克隆的转座酶序列设计简并引物,扩增水稻基因组 DNA。如反转录转座子的反转录酶编码序列具有较高的保守性,在水稻中对应 2 个保守单元(QMDVKT 和 YVDDM)的引物已经被成功应用到通过 PCR 方法扩增反转录转座子的反转录酶区(Hirochika et al., 1996)。由于 PCR 方法操作简单,省时省力,目前该方法已在鉴定和克隆水稻转座子研究中得到广泛应用。

2.4 计算机辅助数据库搜索

通过实验方法鉴定水稻转座子通常要花费大量的时间或财力,而且效率比较低。为此,一些新方法被发展并应用到水稻转座子鉴定中,其中最重要的是利用计算机辅助寻找转座子,而且水稻基因组测序计划的实施和完成也提供了大量基因组序列。利用计算机辅助寻找、鉴定水稻转座子的基本原理是:绝大多数情况下,水稻转座子在基因组中是以多拷贝形式存在的,同时转座子又具有一些特殊的特征,如具有转座酶、末端为反向或正向重复序列以及有靶位点重复,使得转座子序列和基因组中的其它重复序列(如着丝粒和端粒序列等)又有明显不同,因而可以据此开发一些生物学软件对这些序列进行鉴定。目前已经开发了许多软件,如 Repeat-Masker (<http://ftp.genome.washington.edu/RM/RepeatMasker.html>)和 RECON(Bao and Eddy, 2002)等。正是由于该方法的应用,极大地加快了水稻转座子的鉴定速度。

3 水稻转座子的失活与激活

3.1 水稻转座子的失活

3.1.1 转座子自身变异

有些转座子最初可能是有活性的,但在长期进化过程中,由于点突变和缺失等造成转座酶序列发生多种类型变异,使原来转座酶丧失活性,甚至直接缺失转座酶序列。由于突变是随机的,不仅可以发生在转座子内部(转座酶编码区),也可能发生在转座子末端重复区(如TIRs区),而TIR区包含转座酶的特异结合位点,其发生变化将可能导致转座子不能与转座酶特异结合而失去转座能力(Auge-Gouillou et al., 2001)。Asakura等(2002)利用与玉米MuDR转座酶同源序列筛选水稻基因组文库,获得2个Mutator类转座子,并利用这2个序列搜索水稻基因组序列数据库,获得6个序列,这些序列与前2个转座子的TIRs有很高的同源性。对这8个序列进行分析,发现最长的一个转座子OsMu4-2虽与玉米MuDR的转座酶具有很高的相似性,但在其编码区发现终止密码子,推测其转座酶为假基因,而在其余转座子中均发现内部有大片

段缺失(Asakura et al., 2002)。

3.1.2 甲基化

转座子发生转座至少造成两方面的后果:其一就是造成插入突变,假如这种插入发生在控制水稻生长发育的一些关键基因中,将可能导致水稻生长发育停止甚至死亡;另一个效应就是在基因组中增加自身的拷贝数,其无限转座将不断增加转座子成分在水稻基因组中所占比例,并使水稻基因组膨胀(genome expansion)。为防止或减少这种不利局面的出现,基因组进化产生出了某些应对措施,其中之一就是甲基化。目前已经比较清楚,在大多数或全部组织中,植物绝大部分高度重复DNA的甲基化都是5'-CG-3'或5'-CNG-3'序列胞嘧啶(C)的甲基化(Bennetzen et al., 1994);这些重复序列一般都高度浓缩和异染色质化,而且甲基化程度较高,推测这主要是抑制转座子的活性,以保持基因组的稳定性。虽然有少数学者对甲基化程度和转座子活动是否具有相关性存在质疑(Fedoroff, 2000),但是大多数人仍然相信二者之间存在相关性。在拟南芥中,DDM1 (decrease in DNA methylation)基因编码一个类似于染色质修饰因子SWI2/SNF2的蛋白质。如果该基因发生突变,则导致基因组甲基化程度降低,使得一些原来“沉默”的重复序列重新活化,同时发现ddm1突变体的表型不正常。Miura等(2001)发现有一个拟南芥内源的CACTA类转座子CAC1的插入与表型异常有关,同时发现在ddm1突变体低甲基化遗传背景下,CAC1的拷贝数大大增加,这说明基因组甲基化程度降低将可以增加内源转座子活性。有研究发现,转座子启动子及其毗邻TIRs区的转座酶结合位点甲基化对于转座子的转座活性是至关重要的,如玉米Ac转座子启动子或转座酶结合位点甲基化程度增加,则会降低其转座酶的转录及转座子的转座频率(Kunze and Starlinger, 1989; Schwartz, 1989)。

3.1.3 自主转座子的丢失

对于大多数DNA类转座子而言,一个家族自主转座子在基因组中的拷贝数较少,通常只有1个或几个,而非自主转座子的拷贝数相对较高,甚至多达几千个(Jiang et al.,

2004b)。在漫长进化过程中, 由于自然或人为因素造成的选择和遗传重组, 使原来在一个材料中的自主和非自主转座子发生分离, 有些水稻材料基因组中仅存在非自主转座子, 因而不能发生转座。如果将其与原来自主转座子组合到一起, 其仍可以发生转座。如 Tsugane 等 (2006) 发现一个水稻自然突变体, 其苗期叶片为黄绿斑相间, 图位克隆发现突变型是由于有 1 个 607 bp 片段插入到 *OsClpP5* 基因中, 该基因编码叶绿体蛋白酶分析发现插入片段为一个非自主的 DNA 类转座子, 属于 *hAT* 家族, 并将该转座子命名为 *nDart1*。将突变体和具有 *nDart1* 自主转座子 *aDart1* 的材料杂交, 诱导 *nDart1* 发生转座, 但在无自主转座子存在情况下, *nDart1* 是稳定的 (Tsugane et al., 2006)。

3.2 水稻转座子的激活

转座子是基因组中能够移动的组分, 具有转座潜能是其区别于基因组其它遗传成分的本质所在。由于长期进化, 某些原来具有转座活性的转座子发生变异而丧失移动性, 但有些转座子仍具有这种潜能。基因组中“沉默”的转座子可以被一些外界条件激活, 主要有 3 种情况。

3.2.1 组织培养

组织培养激活水稻转座子是非常普遍的。Hirochika 等 (1996) 发现水稻反转录转座子 *Tos10*、*Tos17* 和 *Tos19* 在正常生长条件下是失活的, 但是在组织培养条件下恢复活性; 同时发现随着继代培养时间的延长, *Tos17* 的拷贝数不断增加。另一个最近鉴定的水稻转座子 *mPing* 同样可以被组织培养 (Jiang et al., 2003) 和花药培养所激活 (Kikuchi et al., 2003)。组织培养激活水稻转座子的原因之一可能是在组织培养过程中基因组 (包括转座子成分) 的甲基化程度降低, 使原来“束缚”转座子的外界因子消除或减少。最近, Cheng 等 (2006) 研究了水稻转座子 *Tos17* 和甲基化程度的关系, 发现在组织培养过程中位于 Nipponbare 第 7 染色体上的 1 个 *Tos17* 拷贝 (*Tos17^{chr.7}*) 甲基化程度降低, 其转座活性增加; 而在再生植株的生长过程中, 随着 *Tos17* 的甲基化程度逐渐增

加, 转座子的转座活性也降低或失活, 这从另一方面也证实 *Tos17* 的转座是受甲基化控制的。

3.2.2 理化处理

对水稻种子进行理化处理可以活化基因组中的转座子。其中最为常见的是 γ -射线辐照, 如 Tanisaka 和 Tsugane 分别于 2003 年和 2006 年从不同水稻材料 γ -射线辐照后代中获得突变体, 并在此基础上鉴定出 2 个活性水稻转座子 *mPing* 和 *nDart1* (Nakazaki et al., 2003; Tsugane et al., 2006), 说明通过 γ -射线等激活水稻转座子是可行的。由于转座子活性与甲基化程度具有相关性, 用去甲基化试剂 (如 5-azaC) 对水稻进行处理, 将可能降低基因组甲基化水平, 增加水稻转座子转座活性。Cheng 等 (2006) 发现大多数水稻材料基因组中 *Tos17* 可以被组织培养激活, 但粳稻 Moritawase 基因组中的 *Tos17* 却不能, 说明其是高度甲基化的, 对粳稻种子用去甲基化试剂 5-azaC 进行处理, 发现 Moritawase 中的 *Tos17* 可以被激活。

3.2.3 病菌侵染

Pouteau 等 (1994) 首先发现, 烟草反转录转座子 *Tnt1* 可以被多种病原菌或其提取物所激活。Hirochika 等 (1996) 接着发现, 接种病毒也可以活化 *Tnt1* 转座子。He 等 (2000) 对一个水稻 CACTA 类转座子 *Rim2* 的研究发现, 接种水稻稻瘟病菌后, *Rim2* 的转录活性增加, 并发现 *Rim2* 和水稻抗白叶枯病基因 *Xa21* 家族的 A1 和 C 成员的侧翼序列有 82% 的同源性, 且在 *Rim2* 的上游发现了与欧芹 *PR-1* 启动子激发子反应因子序列高度同源的 4 个单元, 说明 *Rim2* 的积累可能与水稻对病原菌的反应有关, 但病菌侵染引起转座子活化的机制尚不清楚。

4 水稻转座子与基因组进化和基因的调节

最初认为转座子在基因组中的唯一作用就是繁殖自己, 因而转座子被称为基因组中自私的 DNA (Doolittle and Sapienza, 1980)。实际上转座子对基因组的形成和进化以及基因活动等都有重要作用。

4.1 转座子与基因组进化

转座子可能参与了早期基因组的形成,因而从这个意义上讲,转座子从一开始就是基因组的重要组成部分(Brosius and Tiedge, 1995)。初看起来,高等植物中存在大量转座子似乎对基因组是不利的,因为它们过度繁殖和频繁插入将可能导致植物整个遗传结构体系的消失,但是所有目前存在的基因组实际上是转座子和寄主基因组共同成功进化的结果(Kazazian, 2004)。在真核生物中,转座子是基因组中主要非基因类型的DNA,它为寄主提供了组装或修饰遗传功能的原料。同时,许多转座子活动增加了基因组结构的变化,如形成染色体断裂、染色体重排、插入突变和特定的序列扩增,并常和基因活性变化相联系。所以有人猜测转座子在进化中主要作用之一就是作为超级诱变的来源,使得整个群体在遭受逆境胁迫而面临灭绝的情况下有可能产生存活个体(Bennetzen, 2000),因此转座子可能与基因组的可塑性有关。此外,转座子(特别是LTR类反转录转座子)的存在也与基因组的大小和分化有关。Ma等(2004)对具有共同祖先的非洲栽培稻和亚洲栽培稻的2个亚种粳稻和籼稻的基因组序列进行比较,发现粳稻和籼稻的基因组大小分别比非洲栽培稻增加了6%和2%,主要原因是亚洲栽培稻基因组中LTR类反转录转座子的扩增,这可能是这2种栽培稻在不同条件下基因组对周围生态环境适应的结果。转座子也参加了染色体基本结构的形成,如异染色质、核仁组织区、着丝粒和端粒等(Dimitri and Junakovic, 1999)。由于着丝粒和端粒等对维持染色体在细胞中的相对独立性和稳定性都有重要作用,所以该区域维持一定遗传惰性是必需的,在水稻中这些区域主要由LTR类反转录转座子组成(Cheng et al., 2002)。

4.2 转座子与基因调节

转座子与基因活动密切相关。没有基因,转座子将不可能在细胞中生存,但转座子对基因进化和活性调节等也有重要影响,如相邻转座子可以通过不等交换等方式,使某些转座子俘获寄主基因片段,有时甚至可以将两个不同位点基因片段融和到一起,形成新基因。在水稻中已

经发现有部分MULEs(Pack-MULEs)转座子带有寄主基因或基因片段(Jiang et al., 2004a),这可能是一条形成水稻新基因的方式,对水稻基因进化有积极作用。转座子能通过不同方式改变基因表达,其可插入到基因不同部位(如内含子、编码区)或其上游调节区,导致基因功能发生变化(Wessler, 1998)。一些人认为反转录转座子比DNA类转座子对基因表达影响更大,主要理由是反转录转座子的拷贝数高且比较大,如果其插入到基因中将会使基因表达下降甚至失活,但其实并不一定是这样。水稻中LTR类反转录转座子就占水稻基因组的15%,甚至更高,但大多集中在基因贫瘠区或异染色质区(如着丝粒或近着丝粒区)(International Rice Genome Sequencing Project, 2005),因而此类转座子插入到基因中并不常见,其对基因的表达影响可能不如DNA类转座子。有些反转录转座子倾向于插入到基因富集区,*Tos17*就是其中之一(Hirochika et al., 1996)。相比而言,DNA类转座子倾向于插入或存在于含有基因的基因组活性区。如*Mu1*和*Spm*都倾向于插入到玉米基因组少部分区域,这些区域的基因在成熟组织中是非甲基化的,而且是低拷贝的(Cresse et al., 1995)。最近研究发现,水稻中一些转座子超家族如IS5/Tourist和IS630/Tc1/mariner等也倾向插入到基因富集区(Kikuchi et al., 2003; International Rice Genome Sequencing Project, 2005)。总之,在正常条件下转座子被严格控制,但在一定条件下(如基因组胁迫)转座子可能会被激活。转座子与基因互作的结果并不仅仅造成基因突变或失活,也可能表现为增强基因的表达活性。

5 水稻转座子的应用

转座子不仅对于研究水稻起源和进化等理论问题具有重要意义,而且也具有较高的应用价值。有关水稻转座子应用研究已有一些报道,主要包括以下两方面。

5.1 利用转座子标签鉴定水稻功能基因

利用转座子标签法发现、鉴定功能基因是目前水稻遗传学研究的一项重要内容。由于最初在水稻中没有发

现活性转座子, 所以一些科学家设想将玉米转座子(如 *Ac-Ds*、*Spm*)导入到水稻中, 建立一个外源水稻转座子标签体系(Izawa et al., 1997)。目前已得到大量 *Ds* 插入突变株, 并正在验证表型突变是否由 *Ds* 插入引起(Szeverenyi et al., 2006), 但外源转座子鉴定基因尚存在一些问题, 如 *Ds* 因子的沉默、转基因(转座标记)不稳定表达以及在同世代单株中缺少自主转座(Kolesnik et al., 2004)。

内源活性转座子的发现为水稻功能基因组研究提供了一条新途径。迄今已发现6个活性水稻转座子, 分别是LTR类逆转座子 *Tos12*、*Tos17*和 *Tos19*(Hirochika et al., 1996)、LINE类逆转座子 *Karma* (Komatsu et al., 2003)、微小反向重复转座子 *mPing* (Jiang et al., 2003; Kikuchi et al., 2003; Nakazaki et al., 2003)和 *nDart* (Tsugane et al., 2006)。*Tos17* 是水稻中第1个被鉴定的活性转座子, 它在组织培养条件下可以被激活, 且随培养时间的延长其拷贝数会提高, 该转座子已经被用于鉴定水稻功能基因研究中。目前已经获得了约5万个 *Tos17* 插入系, 估计有50万个插入; 将其中11487个插入位点定位到水稻12条染色体上, 并鉴定了3243个基因(International Rice Genome Sequencing Project, 2005), 其中 *OSH15* (Sato et al., 1999)、*MSP1* (Nonomura et al., 2003)、*GAMYB* (Kaneko et al., 2004)、*PAIR2* (Nonomura et al., 2004)和淀粉合成酶基因(Fujita et al., 2006)等已经被分离。尽管 *Tos17*基因标签系统是十分有用的, 但该系统的效率仍然很低, 仅为5% - 10%, 这是因为大部分变异并不是由于 *Tos17*插入引起的, 而是由于组织培养诱导的其它变异引起。为了提高该系统的效率, 需要研究新条件, 在这种新条件下, 反转录转座子可以被激活但又可以减少其它变异, 或者利用其它新的内源转座子。

无论外源转座子标签, 还是内源水稻转座子标签, 都可能存在插入偏爱性, 这样可能会造成不均匀插入; 而且已经发现许多转座子有就近转座特性, 从原位点到插入点一般小于100 kb (Kazazian, 2000)。为此各国科学家已经开始合作, 正在建立一个公开的水稻插入数据库(<http://signal.salk.edu/cgi-bin/RiceGE>), 包括内、外

源转座子插入数据库及T-DNA插入数据库, 希望能将不同研究者的插入库数据整合起来, 以加快水稻饱和插入库构建步伐。但总的来说, 目前各国科学家获得的插入突变还远没有达到饱和插入的程度; 而且 *Ds* 和 *Tos17* 的片段大, 其插入到水稻某些关键基因可能引起致死突变, 所以开发鉴定新的插入体系仍很重要。最近发现的微小反向重复转座子 *mPing*有可能作为一个新的转座子标签体系, 由于其长度较小且具有较高的拷贝数, 因而其插入可能引起致死突变的几率较小。此外, 如何快速鉴定这些突变体并分离和利用水稻功能基因也是今后科学家需要研究的课题。

5.2 基于转座子的新的分子标记开发及应用

转座子大量散布在基因组中, 同时又具有重复末端和转座酶等, 因而可以根据这些特点开发一些新的分子标记。最早开发的基于转座子的分子标记为序列特异扩增多态性(sequence-specific amplification polymorphism, S-SAP), 其原理和操作与AFLP基本相同, 但S-SAP比AFLP具有更丰富的多态性且在染色体上分布更广(Waugh et al., 1997)。水稻全基因组测序计划的实施和完成, 大量转座子被发现鉴定, 给基于转座子序列的分子标记开发提供了极大便利。在水稻中被称作转座子展示(transposon display, TD)的新的分子标记已经被开发出来。目前TD主要应用于以下两个方面。(1)用于水稻遗传图谱构建。虽然目前水稻已经开发了大量分子标记, 并利用这些标记构建了饱和遗传图谱, 在大多数人看来, 已能够适应水稻遗传育种的需要。但由于转座子序列大量散布在水稻基因组中, 且多数是以多拷贝形式存在, 转座子序列富集区(如着丝粒)就成了目前水稻分子标记开发的“盲点”; 因而基于转座子序列的开发将可以弥补水稻其它分子标记的一些不足, 如Jiang等(2002)鉴定出1个LTR类反转录转座子 *Dasheng*, 其在水稻基因组中大约有1000个拷贝, 应用TD技术将其中215个定位到水稻12条染色体上, 同时发现大约有一半 *Dasheng* 分布在近着丝粒的异染色质区。Kwon等(2005)发现CACTA类转座子 *Rim2/Hipa*-TD分子标记也多位于近着丝粒区。(2)鉴定不同种

质的转座子拷贝数。有些转座子在水稻基因组中有很多拷贝,如水稻微小反向重复转座子 *mPing* 在粳稻 Nippobare 中有 70 个拷贝,在 Gimbozu 中可能有近 1 000个拷贝,这样多的拷贝数难以应用常规的Southern杂交方法准确鉴定,但可以通过TD分析鉴定出其拷贝数。

总之,转座子是水稻基因组中重要的组成部分,其对于水稻基因组组成进化和基因表达调控等都具有重要意义。如同著名遗传学家 Bennetzen 所言:“如果我们不了解转座子的功能,就不可能对植物基因组的结构和进化有一个明确的概念”(Bennetzen, 2000)。所以对于转座子的深入了解将为我们提供一把了解水稻基因组组成与进化秘密的钥匙,也必将促进水稻遗传育种相关理论和技术的发展。

致谢 美国密歇根州立大学园艺系蒋宁教授提供部分数据,并对本文提出修改意见,作者表示诚挚感谢!

参考文献

Asakura N, Nakamura C, Ishii T, Kasai Y, Yoshida S (2002). A transcriptionally active maize MuDR-like transposable element in rice and its relatives. *Mol Genet Genomics* **268**, 321-330.

Auge-Gouillou C, Hamelin MH, Demattei MV, Periquet G, Bigot Y (2001). The TIR binding domain of the Mariner Mos-1 transposase. *Mol Genet Genomics* **265**, 58-65.

Bao Z, Eddy SR (2002). Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Res* **12**, 1269-1276.

Bennetzen JL (2000). Transposable element contributions to plant gene and genome evolution. *Plant Mol Biol* **42**, 251-269.

Bennetzen JL, Schrick K, Springer PS, Brown WE, SanMiguel P (1994). Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. *Genome* **37**, 565-576.

Brosius J, Tiedge H (1995). Reverse transcriptase: mediator of genomic plasticity. *Virus Genes* **11**, 163-179.

Cheng C, Daigen M, Hirochika H (2006). Epigenetic regulation of the rice retrotransposon *Tos17*. *Mol Genet Genomics* **276**,

378-390.

Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J (2002). Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell* **14**, 1691-1704.

Cresse AD, Hulbert SH, Brown WE, Lucas JR, Bennetzen JL (1995). Mu1-related transposable elements of maize preferentially insert into low copy number DNA. *Genetics* **140**, 315-324.

Curcio MJ, Derbyshire KM (2003). The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol* **4**, 865-877.

Dimitri P, Junakovic N (1999). Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. *Trends Genet* **15**, 123-124.

Doolittle WF, Sapienza C (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**, 601-603.

Fedoroff N (2000). Transposons and genome evolution in plants. *Proc Natl Acad Sci USA* **97**, 7002-7007.

Feschotte C, Jiang N, Wessler SR (2002). Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* **3**, 329-341.

Fujita N, Yoshida M, Asakura N, Ohdan T, Miyao A, Hirochika H, Nakamura Y (2006). Function and characterization of starch synthase I using mutants in rice. *Plant Physiol* **140**, 1070-1084.

He ZH, Dong HT, Dong JX, Li DB, Ronald PC (2000). The rice Rim2 transcript accumulates in response to *Magnaporthe grisea* and its predicted protein product shares similarity with TNP2-like proteins encoded by CACTA transposons. *Mol Gen Genet* **264**, 2-10.

Hirochika H, Fukuchi A, Kikuchi F (1992). Retrotransposon families in rice. *Mol Gen Genet* **233**, 209-216.

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996). Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci USA* **93**, 7783-7788.

International Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. *Nature* **436**, 793-800.

Izawa T, Ohnishi T, Nakano T, Ishida N, Enoki H, Hashimoto H, Itoh K, Terada R, Wu C, Miyazaki C, Endo T, Iida S, Shimamoto K (1997). Transposon tagging in rice. *Plant Mol Biol* **35**, 219-229.

Jiang N, Bao Z, Temnykh S, Cheng Z, Jiang J, Wing RA, McCouch SR, Wessler SR (2002). Dasheng: a recently am-

plified non-autonomous LTR element that is a major component of pericentromeric regions in rice. *Genetics* **161**, 1293-1305.

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003). An active DNA transposon family in rice. *Nature* **421**, 163-167.

Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004a). Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **431**, 569-573.

Jiang N, Feschotte C, Zhang X, Wessler SR (2004b). Using rice to understand the origin and amplification of MITEs. *Curr Opin Plant Biol* **7**, 115-119.

Juretic N, Hoen DR, Huynh ML, Harrison PM, Bureau TE (2005). The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res* **15**, 1292-1297.

Kaneko M, Inukai Y, Ueguchi-Tanaka M, Itoh H, Izawa T, Kobayashi Y, Hattori T, Miyao A, Kazazian HH Jr (2004). Mobile elements: drivers of genome evolution. *Science* **303**, 1626-1632.

Kazazian HH Jr (2004). Mobile elements: drivers of genome evolution. *Science* **303**, 1626-1632.

Kempken F, Windhofer F (2001). The *hAT* family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* **110**, 1-9.

Kikuchi K, Terauchi K, Wada M, Hirano HY (2003). The plant MITE mPing is mobilized in anther culture. *Nature* **421**, 167-170.

Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004). Establishing an efficient Ac/Ds tagging system in rice: large-scale analysis of Ds flanking sequences. *Plant J* **37**, 301-314.

Komatsu M, Shimamoto K, Kyojuka J (2003). Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon Karma. *Plant Cell* **15**, 1934-1944.

Kunze R, Starlinger P (1989). The putative transposase of transposable element Ac from *Zea mays* L. interacts with subterminal sequences of Ac. *EMBO J* **8**, 3177-3185.

Kwon SJ, Park KC, Kim JH, Lee JK, Kim NS (2005). Rim 2/Hipa CACTA transposon display: a new genetic marker technique in *Oryza* species. *BMC Genet* **6**, 1-13.

Ma J, Bennetzen JL (2004). Rapid recent growth and divergence of rice nuclear genomes. *Proc Natl Acad Sci USA* **101**, 12404-12410.

Ma J, Devos KM, Bennetzen JL (2004). Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res* **14**, 860-869.

Mao L, Wood TC, Yu Y, Budiman MA, Tomkins J, Woo S, Sasinowski M, Presting G, Frisch D, Goff S, Dean RA, Wing RA (2000). Rice transposable elements: a survey of 73 000 sequence-tagged-connectors. *Genome Res* **10**, 982-990.

McCarthy EM, Liu J, Lizhi G, McDonald JF (2002). Long terminal repeat retrotransposons of *Oryza sativa*. *Genome Biol* **3**, 1-11.

Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T (2001). Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* **411**, 211-214.

Motohashi R, Ohtsubo E, Ohtsubo H (1996). Identification of Tnr3, a suppressor-mutator/enhancer-like transposable element from rice. *Mol Gen Genet* **250**, 148-152.

Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, Inoue H, Tanisaka T (2003). Mobilization of a transposon in the rice genome. *Nature* **421**, 170-172.

Nonomura KI, Miyoshi K, Eiguchi M, Suzuki T, Miyao A, Hirochika H, Kurata N (2003). The *MSP1* gene is necessary to restrict the number of cells entering into male and female sporogenesis and to initiate anther wall formation in rice. *Plant Cell* **15**, 1728-1739.

Nonomura KI, Nakano M, Murata K, Miyoshi K, Eiguchi M, Miyao A, Hirochika H, Kurata N (2004). An insertional mutation in the rice *PAIR2* gene, the ortholog of Arabidopsis ASY1, results in a defect in homologous chromosome pairing during meiosis. *Mol Genet Genomics* **271**, 121-129.

Pouteau S, Grandbastien AM, Boccara B (1994). Microbial elicitors of plant defence responses activate transcription of a retrotransposon. *Plant J* **5**, 535-542.

Sato Y, Sentoku N, Miura Y, Hirochika H, Kitano H, Matsuoka M (1999). Loss-of-function mutations in the rice homeobox gene *OSH15* affect the architecture of internodes resulting in dwarf plants. *EMBO J* **18**, 992-1002.

Schwartz D (1989). Gene-controlled cytosine demethylation in the promoter region of the Ac transposable element in maize. *Proc Natl Acad Sci USA* **86**, 2789-2793.

Sprinzl M, Hartmann T, Meissner F, Moll J, Vorderwulbecke T (1987). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **15** (Suppl), 153-188.

Szeverenyi I, Ramamoorthy R, Teo ZW, Luan HF, Ma ZG, Ramachandran S (2006). Large-scale systematic study on stability of the Ds element and timing of transposition in rice.

Plant Cell Physiol **47**, 84-95.

Teraishi M, Okumoto Y, Hirochika H, Horibata A, Yamagata H, Tanisaka T (1999). Identification of mutable slender glume gene in rice (*Oryza sativa* L.). *Mol Gen Genet* **261**, 487-494.

Tsugane K, Maekawa M, Takagi K, Takahara H, Qian Q, Eun CH, Iida S (2006). An active DNA transposon nDart causing leaf variegation and mutable dwarfism and its related elements in rice. *Plant J* **45**, 46-57.

Turcotte K, Srinivasan S, Bureau T (2001). Survey of transposable elements from rice genomic sequences. *Plant J* **25**, 169-179.

Warren WD, Atkinson PW, O'Brochta DA (1994). The Hermes transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the hobo, Ac, and Tam3 (hAT) element family. *Genet Res* **64**, 87-97.

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997). Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol Gen Genet* **253**, 687-694.

Wessler SR (1998). Transposable elements and the evolution of gene expression. *Symp Soc Exp Biol* **51**, 115-122.

Zhang X, Jiang N, Feschotte C, Wessler SR (2004). PIF and Pong-like elements: distribution, evolution and relationship with the Tourist-like miniature inverted-repeat transposable elements. *Genetics* **166**, 971-986.

Research Advances on Transposable Elements in Rice

Dongying Gao*, Bing He, Lihua Sun

Institute of Food Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

Abstract Transposable elements are important components of plant genome and play critical roles in genome evolution. With the availability of complete rice genomic sequences, great progress has been made in identifying and characterizing rice transposons. To date, almost all types of transposons have been identified in rice, and they make up about 35% of the genome. Most transposons are silent during normal development, but they can be activated by factors such as tissue culture or radiation. The activation of transposons may induce the loss of gene function or affect rice gene expression. Six active transposons have been discovered, with *Tos17* the first active transposon widely used in rice functional genomics studies. Transposons can also serve as a new type of molecular marker with use of the technique transposon display (TD), a useful tool in rice genetic mapping and the studies of genetic diversity.

Key words gene expression, rice, transposable elements, transposon display

Gao DY, He B, Sun LH (2007). Research advances on transposable elements in rice. *Chin Bull Bot* **24**, 667-676.

* Author for correspondence. E-mail: gao.dongying@yahoo.com.cn

(责任编辑: 孙冬花)