

植物 LTR 反转录转座子的研究进展

蒋 爽, 滕元文, 宗 宇, 蔡丹英*

(浙江大学 园艺系, 农业部园艺植物生长发育与品质调控重点开放实验室, 杭州 310058)

摘 要: 反转录转座子是真核生物基因组中普遍存在的一类可移动的遗传因子, 它们以 RNA 为媒介, 在基因组中不断自我复制。在高等植物中, 反转录转座子是基因组的重要成分之一。反转录转座子可以分为 5 大类型, 其中以长末端重复(LTR)类型报道较多。LTR 类型由于其首尾具有长末端重复序列, 内部含有 PBS、PPT、GAG 和 POL 开放阅读框、TSD 等结构, 可以采用生物信息学软件进行预测。LTR 反转录转座子的活性受到自身甲基化和环境因素的影响, DNA 甲基化抑制反转录转座子转座, 而外界环境的刺激能够激活转座子, 从而影响插入位点周边基因的表达。同时由于 LTR 反转录转座子在植物中普遍存在, 丰富的拷贝数以及多态性为新型分子标记(RBIP、SSAP、IRAP、REMAP)的开发提供了良好的素材。该文对近年来国内外有关植物反转录转座子的类型、结构特征、LTR 反转录转座子的活性及其影响因素、LTR 反转录转座子的预测以及标记开发等方面的研究进展进行综述。

关键词: 反转录转座子; 长末端重复(LTR); 预测; 功能; 分子标记

中图分类号: Q78

文献标志码: A

Review of LTR Retrotransposons in Plants

JIANG Shuang, TENG Yuanwen, ZONG Yu, CAI Danying*

(Department of Horticulture, The State Agricultural Ministry Laboratory of Horticultural Plant Growth, Development & Quality Improvement, Zhejiang University, Hangzhou 310058, China)

Abstract: Retrotransposons in eucaryon are mobile genetic elements with ubiquitous distribution, these elements can amplify themselves via RNA intermediates and increase their copy numbers in genome. In high plants, retrotransposon is an important component of genome. Retrotransposons can be divided into five types, the LTR retrotransposons, an important subclass of transposable elements, are well suited for computational identification, as they contain two long terminal repeats (LTRs) and features such as PBS, PPT, GAG and POL open reading frames, TSD. Recent research indicated that methylation of LTR retrotransposon and environments seemed to be affecting the jump of retrotransposons. Retrotransposons could be inhibited by high degree of DNA methylation and activated by environments, which influence the function of surrounding genes. The characteristics of LTR retrotransposons, such as ubiquitous distribution, abundant copy number and insertional polymorphism in plant species, provide an excellent basis for the development of molecular markers, e. g., RBIP, SSAP, IRAP, REMAP. In this review, we summarize the recent progress about the types and structures of retrotransposons, inhabitation and activation of LTR retrotransposons and their computational identification, molecular marker development.

Key words: retrotransposons; long terminal repeat(LTR); prediction; function; molecular markers

收稿日期: 2013-07-15; 修改稿收到日期: 2013-10-23

基金项目: 国家自然科学基金青年科学基金(31201592)

作者简介: 蒋 爽(1988—), 男, 在读博士研究生, 主要从事梨属植物进化与系统发育研究。E-mail: 271339616@qq.com

* 通信作者: 蔡丹英, 博士, 主要从事梨属植物进化与系统发育研究。E-mail: dycail@zju.edu.cn

植物基因组中含有大量可移动的遗传因子,统称为转座子^[1],其中第一大类型是反转录转座子(retrotransposons),是一类广泛存在于植物基因组中的特殊 DNA 序列^[2-4],植物中最初是在研究玉米突变体时发现的^[5]。反转录转座子是植物基因组中重要的组成成分,它们在植物中含有丰富的拷贝数。反转录转座子的转座机制不同于其他转座元件,它首先转录形成 RNA,再以 RNA 为媒介在逆转录酶作用下逆转录成 DNA,最终插入到基因组的靶基因片段上。这种“复制-粘贴”的转座机制使得反转录转座子能够快速自我扩增,从而在漫长的进化历史中改变整个基因组的大小^[6-7]。植物中的反转录转座子通常处于沉默状态,一部分转座子受到甲基化的影响导致沉默,另一部分转座子在长期的进化过程中由于碱基突变已丧失反转座功能。当植物受到外界条件影响,尤其是逆境时,反转录转座子会被激活并插入到靶基因中,同时,周边基因的表达也会受到影响^[8]。

1 植物反转录转座子类型和结构特征

1.1 植物反转录转座子类型

植物中反转录转座子根据其结构域可以分为 5 大类型^[9],分别为长末端重复(long terminal repeat,LTR)、DIRS 类型、PLE 类型(PLE)、长散布重复序列(long interspersed repetitive element,LINE)和短散布重复序列(short interspersed repetitive element,SINE)(表 1)。其中,LTR 是目前研究较多的一类反转录转座子,根据 POL 阅读框内部

酶基因的顺序,LTR 又可分为 Ty1-*copia* 和 Ty3-*gypsy* 2 个亚家族。这 2 个亚家族几乎存在于所有的高等植物中^[10]。

1.2 植物反转录转座子的结构特征

LTR 反转录转座子具有以下 5 个方面的结构特征(表 1,图 1)^[22-23]:(1)序列两端有 4~6 bp 的靶位点重复序列;(2)LTR 序列 5'-端以 TG 开始,3'-末端以 CA 结尾,两端含有相似的长末端重复序列;(3)LTR 之间存在 GAG 和 POL 开放阅读框,POL 中包含 4 种酶,分别是天冬氨酸蛋白酶(AP)、整合酶(INT)、逆转录酶(RT)和核糖核酸酶(RH);(4)在 5'LTR 末尾含有长度约为 18 bp,且与一些 tRNA 的 3'端互补的引物结合位点(primer binding site,PBS),该位点对于反转录转座子非常重要,因为它与 tRNA 相结合是反转录的第一步;(5)在 3'LTR 的起始位置含有多嘌呤序列(polypurine tract,PPT)。Ty1-*copia* 和 Ty3-*gypsy* 2 个亚家族的区别在于反转录转座子 POL 开放阅读框中 INT 的位置,Ty1-*copia* 亚家族中的 INT 位于编码区中间,而 Ty3-*gypsy* 亚家族中则位于编码区末端^[7](图 1)。虽然这两个亚家族含有相同的结构基因,但相同的结构基因之间存在较大差别^[24]。这类反转录转座子在反转录插入时会复制靶位点附近的区域形成长度为 4~6 bp 的靶位点重复(target site repeat,TSD)。

DIRS 类型的 5'和 3'端长末端重复序列为反向重复,在转座插入时并不形成 TSD 区域,它的内部与 LTR 家族一样均含有 GAG 和 POL 开放阅读框。

表 1 反转录转座子分类及结构
Table 1 The classification and structure of retrotransposons

分类 Classification		结构 Structure	重复序列 TSD	参考文献 Reference
类型 Order	亚家族 Subfamily			
长末端重复类型 LTR	<i>Copia</i>	→[GAG]-[AP]-[INT]-[RT]-[RH]→	4-6	[11-14]
	<i>Gypsy</i>	→[GAG]-[AP]-[RT]-[RH]-[INT]→	4-6	
DIRS	<i>DIRS</i>	→[GAG]-[AP]-[RT]-[RH]-[YR]←	0	[15-16]
PLE	<i>Penelope</i>	—[RT]-[EN]—	多变 Variable	[17]
长散布重复序列类型 LINE	<i>L1</i>	—[ORF1]-[APE]-[RT]—	多变 Variable	[18-19]
	<i>I</i>	—[ORF1]-[APE]-[RT]-[RH]—	多变 Variable	
短散布重复序列类型 SINE	<i>tRNA</i>	—□-□—	多变 Variable	[20]
	<i>7SL RNA</i>	—□-□—	多变 Variable	

注:参考[21]并修改。→,长末端重复序列;↔,末端反向重复序列;□,短散布重复序列;—,非编码区;AP,天冬氨酸蛋白酶;APE,无嘌呤核酸内切酶;EN,核酸内切酶;GAG,GAG 开放阅读框;INT,整合酶;RH,核糖核酸酶;RT,逆转录酶;YR,酪氨酸重组酶。
Note:Reference^[21] with modification。→,Long terminal repeat;↔,Terminal inverted repeats;□,Diagnostic feature in non-coding region;—,Non-coding region;AP,Aspartic proteinase;APE,Apurinic endonuclease;EN,Endonuclease;GAG,GAG open reading frame;INT,Integrase;RH,RNase H;RT,Reverse transcriptase;YR,Tyrosine recombinase.

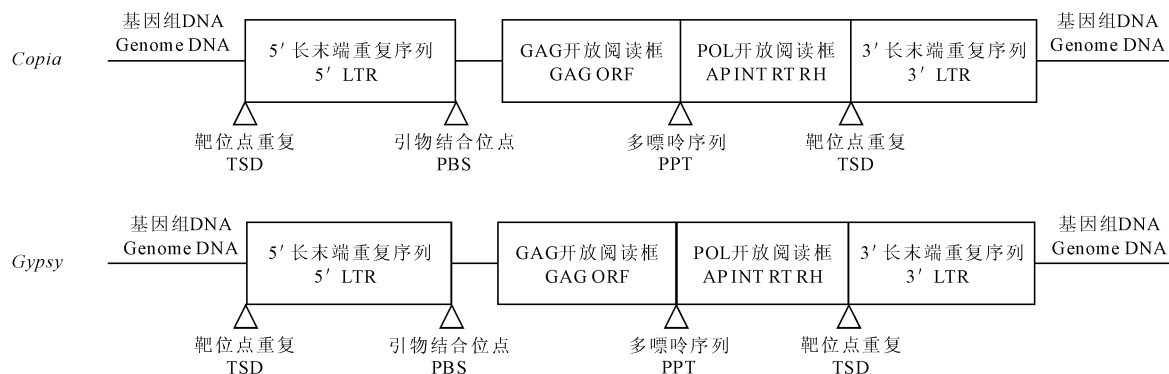


图 1 LTR 家族反转录转座子 *Copia* 和 *Gypsy* 的结构特征

AP. 天冬氨酸蛋白酶;INT. 整合酶;RT. 逆转录酶;RH. 核糖核酸酶;ORF. 开放阅读框

Fig. 1 Structure characteristics of LTR retrotransposons

AP. Asparticproteinase;INT. Integrase;RT. Reverse transcriptase;RH. RNase H;ORF. Open reading frame

PLE 类型也包含 LTR 区域,在编码区只含有 RT 和内切酶(Endonuclease, EN),这类反转录转座子在植物中鲜有报道。LINE 类型的 5' 和 3' 端均不含 LTR 区域,在其内部含有开放阅读框,并含有无嘌呤核酸内切酶(apurinic endonuclease, APE)和 RT,其中 I 类还包含 RH(表 1)。SINE 类型的长度一般在 100~600 bp,植物中该类型多以 tRNA, 7SL RNA 起始,在 3' 末端有一个长度不定的短重复尾巴。

2 LTR 反转录转座子活性及其影响因素

2.1 LTR 反转录转座子的活性

大部分的 LTRs 由于其结构基因存在终止子、移码突变和缺失等,形成无功能的蛋白质,所以不具有转座功能^[25-26]。但也仍存在部分有活性的 LTRs,它们有可能插入到功能基因内或周边,影响这些功能基因的表达。Kobayashi 等^[27]对不同色泽的葡萄果皮进行了研究,结果表明反转录转座子的转座导致了基因突变,从而引起葡萄果皮颜色的变化。红皮葡萄的色泽形成主要通过果皮中花青苷的积累,其花青苷的合成受到基因 *VvmybA1* 的调控;但是在白皮葡萄中,由于在 *VvmybA1* 基因的启动子区域插入了一个长度为 10 422 bp 的反转录转座子 Ty3-*gypsy*,影响了该基因的表达,从而使花青苷不能合成,形成白皮葡萄。反转录转座子的插入也可能导致其周边基因上调表达,Butelli 等^[28]在血橙中发现调控 MYB 启动子活性的 *RUBY* 基因上游插入了反转录转座子 *Copia-like*,该转座子在低温储藏过程中表达,同时调控 *RUBY* 基因的表达,促进果实合成花青苷,导致果实颜色变红。

2.2 LTR 反转录转座子活性的影响因素

LTR 反转录转座子的转座受到自身甲基化的影响,已有研究表明 DNA 甲基化酶可以抑制反转录转座子的迁移和插入。Mirouze 等^[29]利用拟南芥 *met1* 突变体(该突变体不能合成甲基化酶)与野生型杂交产生 F₁ 代,随后进行多代自交,在第 8 代中发现有些个体不存在雄性花器官,研究发现由于在控制成花的 *LEAFY* 基因中插入了一个反转录转座子而使该基因不能正常表达,导致无雄性花器官。另外,由于反转录转座子的插入,导致后代中出现了多种其他突变体类型,而野生型的拟南芥后代中并未发现这些突变体,从而证明 DNA 甲基化酶可以抑制反转录转座子的迁移。在野生型拟南芥中,DDM1 蛋白能够促进 DNA 甲基化,而在 *ddm1* 突变体多代自交形成的后代中,DDM1 蛋白缺失,引起基因组中甲基化程度下降,导致反转录转座子 *ATGP3* 出现大量复制^[30-31],从而推测甲基化程度降低能够激活反转录转座子的转座。

另外,在逆境、激素处理、机械伤等情况下,反转录转座子的转座功能也可以被激活。Felice 等^[32]发现在盐胁迫的环境下,柠檬组织培养再生植株中有 Tyl-*copia* 类转座子 *CLCoy1* 的表达。而在番茄中,机械损伤,胁迫以及激素等处理均可以诱导 Tyl-*copia* 类反转录转座子 *TLC1* 产生活性^[33]。

3 LTR 反转录转座子的预测方法

反转录转座子的预测方法分为 4 种^[34]:重复序列从头算法(de novo repeat discovery)、同源比对法(homology-based methods)、基于结构预测法(structure-based methods)和比较基因组法(comparative genomic methods)。

重复序列从头算法是根据序列的重复特性在基因组上识别全长 LTR 类反转录转座子,它的优点在于不需要序列特征及其分类信息,但是假阳性高。这类软件有 RepeatFinder^[35] 和 ReAS^[36]。

同源比对法是通过 Blast 同源比对现有的反转录转座子数据库,找出未知序列中的反转录转座子,这种预测方法还可以通过 HMMER 来寻找反转录转座子内部的结构域^[37],从而发现反转录转座子。

基于反转录转座子的种种特征(图 1),根据其结构预测包含 LTR 区域的反转录转座子的软件应运而生。Xu 等^[38] 开发了反转录转座子发掘工具 LTR_FINDER,它基于后缀树方法且效率高,可从基因组序列中快速鉴定出具有 LTR 反转录转座子一般结构特征的全长 LTR 类反转录转座子序列。Wang 等^[39] 使用 LTR_FINDER 从苜蓿基因组中预测出 85 个族含有 526 个反转录转座子序列,其中 64 个是首次报道的反转录转座子。Xu 等^[40] 从水稻基因组中预测出超过 6 000 个反转录转座子区域。Cossu 等^[41] 从杨树基因组预测出 1 479 个反转录转座子,这为木本植物中反转录转座子的研究提供了依据。随后,Ellinghaus 等^[42] 开发出更高效的反转录转座子预测工具 LTRharvest,它可以在短时间内分析大量的数据。与 LTR_FINDER 相比,LTRharvest 可以直接输出 fasta 格式的反转录转座子序列,方便后续分析。LTR_FINDER 在运行上较费时,虽然它能给出图形化的结果,在预测的反转录转座子上标注出 LTR 区域、PPT、PBS 等部件,但是其复杂的输出结果不利于后续分析;此外,可进行反转录转座子发掘软件还包括 LTR_STRUC^[43]、LTR_par^[44] 和 LTR_Rho^[45] 等。

比较法是通过比对基因组来发现反转录转座子的插入,相应的软件有 LTR_INSERT^[46],它适用于亲缘关系较近的 2 个基因组,但这种方法只能发现有插入的片段,而该片段是否是转座子还需要进一步鉴定。同时该方法不能发现两个比对基因组中特定区域共同拥有的反转录转座子。

4 LTR 反转录转座子标记开发

传统的分子标记包括 RFLP、RAPD、SSR、SNP 等,近年来,随着反转录转座子预测技术的发展,研究发现植物基因组中存在大量反转录转座子,通常情况下反转录转座子处于沉默状态,但在外界环境和内在生物因子的刺激下能够激发反转录转座子的转座功能^[6],引起基因突变,导致植物的遗传变异。

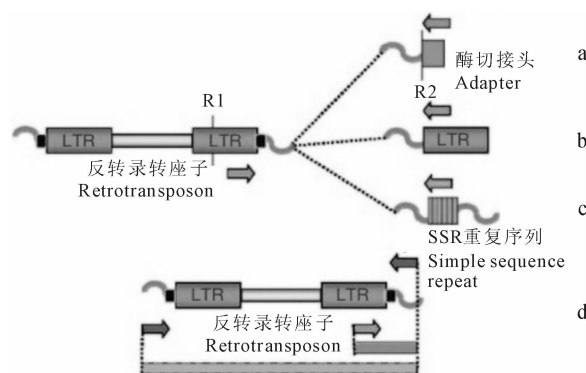


图 2 LTR 反转录转座子标记方法
(改编自 Kalendar 等^[50])

4 种标记上游引物均设计在 LTR 区域,(a)SSAP,基因组双酶切后加上 R1 和 R2 接头,下游引物使用 R2 接头引物。(b)IRAP,上下游引物均设计在 LTR 区域,(c)REMAP,下游引物设计在 SSR 重复序列上,(d)RBIP,下游引物设计在基因组区域

Fig. 2 LTR retrotransposon-based molecular marker methods (adopted from Kalendar *et al.*^[50])

The LTR generally serves as the retrotransposon priming site. (a)SSAP. Amplification is carried out from genomic DNA cut with two restriction enzymes (R1 and R2), containing a retrotransposon and ligated to an adapter (R2). (b) IRAP. The second priming site is also a retrotransposon. (c) REMAP. Amplification takes place between a microsatellite domain (SSR) and a retrotransposon. (d)RBIP. Full sites are scored by amplification between a primer in the flanking genomic DNA and a retrotransposon primer

据此,一种基于反转录转座子插入的标记应运而生。该标记的特点主要是:普遍存在于植物基因组中,拷贝数丰富,高度异质性,插入多态性好^[47]。目前基于 LTR 反转录转座子开发的标记(图 2),包括 SSAP (sequence-specific amplification polymorphisms)^[48]、IRAP (inter-retrotransposon amplified polymorphism)^[49]、REMAP (REtrotransposon-microsatellite amplified polymorphism)^[49]、RBIP (retrotransposon-based insertion polymorphisms)^[50]。

4.1 SSAP 标记

SSAP 标记的开发过程与 AFLP (amplified fragment length polymorphism) 标记相似,首先对基因组 DNA 进行双酶切 (*Mse* I / *Pst* I 或 *Mse* I / *Eco*R I),酶切后加上对应接头引物用于预扩增,而选扩时引物一端使用接头引物,另一端引物设计在 LTR 反转座子内部^[51],而与之不同的是 AFLP 在选扩时采用两端的接头引物进行扩增。另外,SSAP 基于已知 LTR 反转座子序列,但它得到的多态性比 AFLP 更丰富。Ellis 等^[2] 根据豌豆 *Ty1-copia*

家族的 *PDR1* 中多嘌呤序列 (PPT) 设计了 SSAP 标记的一端引物, 另一端引物则为 DNA 酶切后连接的接头引物, 他们采用这些 SSAP 标记并结合 AFLP 标记成功区分了豌豆属的 3 个种源。Yu 等^[52] 通过 372 个 AFLP 标记并结合 78 个 SSAP 标记构建了二倍体燕麦属的遗传连锁图谱。Woodrow 等^[53] 根据香桃木的基因组数据和 *Ty1-copia* 家族中 *Tmc1* 的特征设计了 SSAP 标记引物, 并将它们用于香桃木属植物遗传多样性的分析, 结果表明 SSAP 引物能很好地区分开 4 个供试品种。Tam 等^[54] 使用 SSAP、AFLP 和 SSR 分别对番茄和甜椒的遗传多样性进行了分析, 得出 3 种标记适用于不同的类群, 并强调 SSAP 标记技术是一种高效且用途广泛的标记。

4.2 RBIP 标记

近年来, 随着基因组测序技术的不断成熟, 越来越多的植物全基因组序列测序完成^[10]。基于基因组特异序列的 RBIP 标记显得更有针对性, 它主要用于确定反转录转座子在特异位点上是否存在插入和缺失。RBIP 包含两种类型, 在第一种类型中, 两端的引物都设计在反转录转座子区域的侧翼 DNA 序列上, 扩增结果反映反转录转座子的有无, 第二种类型是一端引物设计在 LTR 区域, 另一端引物设计在反转录转座子的侧翼序列上, 这种引物特异性高, PCR 扩增后通过电泳检测目的片段的有无来确定该位置是否有反转录转座子插入^[55]。第一种类型受到反转录转座子长度的限制, 一般的反转录转座子在 5 000 bp 以上^[38], 不利于 PCR 扩增, 而第二种类型不存在该问题, 因此在植物中得到了广泛应用。Vitte 等^[56] 采用 9 对 RBIP 引物分析了印度水稻和日本水稻的进化关系, 结果显示 2 个地域的水稻分化时间至少在 20 万年前。Kim 等^[57] 利用在‘丰水’梨上开发的 22 对 RBIP 引物对亚洲梨和欧洲梨进行了亲缘关系的分析, 并将这 22 对标记定位到连锁图谱上, 结果表明这些引物可以很好地区分日本梨, 该研究中将一端引物修饰荧光基团, 更有利于目的条带的分离与鉴定。

4.3 IRAP 和 REMAP 标记

IRAP 标记的引物两端都设计在 LTR 区域, REMAP 标记在引物设计时一端引物设计在 LTR 区域, 另一端引物设计在 SSR 区域内。目前针对这 2 种标记的报道较少。杜晓云等^[58] 根据罗田甜柿基

因组开发出 5 对 IRAP 标记并对 7 个柿属植物进行扩增, 结果显示多态性良好, 可成功区分芽变品种。Kalendar 等^[59] 根据 *BARE-1* 反转录转座子中 LTR 区域设计出 IRAP 的上下游引物, 结合 SSAP、REMAP 标记对大麦不同品种进行遗传多样性分析, 并指出 SSAP 依赖酶切位点, 而 IRAP 和 REMAP 不存在酶切位点问题, 但它们依赖 2 个 LTR 之间的距离和 LTR 与 SSR 位点之间的距离。

5 问题与展望

植物基因组中含有大量反转录转座子, 研究发现, 在苹果基因组中, 反转录转座子占整个基因组的 30.7%, 水稻中占 25.8%, 黄瓜中为 6.9%, 拟南芥中占 6.6%^[10]。对水稻反转录转座子的研究发现, 大部分 LTR 反转录转座子的插入在近 800 万年内, 并且它可能与水稻基因组大小的进化有关^[60]。如此众多的反转录转座子在植物进化过程并没有丢失, 其在基因组中的功能以及在植物进化中的作用还有待于进一步研究。而植物中反转录转座子家族纷繁复杂, 已经报道的 LTR 家族有小麦中的 *Angela* 家族^[61], 大麦中的 *BARE1* 家族^[62], 水稻中 *Opie* 家族^[63], 高粱中 *Retrosor6* 家族^[64] 等。这些反转录转座子家族无论在 LTR 部分还是在内部结构上序列均表现出较大的差异, 关于这些反转录转座子家族的起源及进化史还有待进一步研究。

目前, 针对 LTR 反转录转座子的预测软件众多, 但是同一物种内的不同 LTR 反转录转座子差异大^[65], 在其分类过程中还存在诸多问题, 目前尚缺少系统的 LTR 反转录转座子分类软件以及内部原件预测软件, 随着生物信息学的发展, 预测软件能不断地得到开发和完善。

由于 LTR 反转录转座子的丰富性以及其突变性, 在植物基因组进化中起着重要作用, 已逐步应用于分子标记的开发, 基因标签和基因功能分析^[66] 等研究中。随着众多植物全基因组测序的完成, 必将极大地推动 LTR 反转录转座子的研究, 有利于获得更多种类的反转录转座子的家族以及家族成员的全长序列, 有助于进一步了解它们在植物基因组结构和进化中的作用, 不断地完善和发展基于反转录转座子的分子标记技术体系, 进一步拓展了分子标记在植物的系统发育和遗传多样性^[67]、图谱构建^[52] 以及品种鉴定^[68] 等方面的应用。

参考文献:

- [1] WICKER T, SABOT F, *et al.* A unified classification system for eukaryotic transposable elements[J]. *Nat. Rev. Genet.*, 2007, **8**(12): 973—982.
- [2] ELLIS T H, 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[J]. *Mol. Gen. Genet.*, 1998, **260**(1): 9—19.
- [3] KALENDAR R, TANSKANEN J, IMMONEN S, *et al.* Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence[J]. *Proc. Natl. Acad. Sci. USA*, 2000, **97**(12): 6 603—6 607.
- [4] PETIT M, LIM K Y, JULIO E, *et al.* Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (*Nicotiana tabacum*) [J]. *Mol Genet Genomics*, 2007, **278**(1): 1—15.
- [5] SHEPHERD N S, SCHWARZ-SOMMER Z, BLUMBERG V S J, *et al.* Similarity of the Cin1 repetitive family of *Zea mays* to eukaryotic transposable elements[J]. *Nature*, 1984, **307**(5 947): 185—187.
- [6] KUMAR A, BENNETZEN J L. Plant retrotransposons[J]. *Annual Review of Genetics*, 1999, **33**(1): 479—532.
- [7] HAVECKER E R, GAO X, VOYTAS D F, *et al.* The diversity of LTR retrotransposons[J]. *Genome Biology*, 2004, **5**(6): 225.
- [8] DOMINGUES D S, CRUZ G M, METCALFE C J, *et al.* Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns[J]. *BMC Genomics*, 2012, **13**: 137.
- [9] BEAUREGARD A, CURCIO M J, BELFORT M. The take and give between retrotransposable elements and their hosts[J]. *Annu. Rev. Genet.*, 2008, **42**: 587—617.
- [10] VELASCO R, ZHARKIKH A, *et al.* The genome of the domesticated apple (*Malus × domestica* Borkh.) [J]. *Nat. Genet.*, 2010, **42**(10): 833—839.
- [11] VICIENT C M, KALENDAR R, ANAMTHAWAT-JONSSON K, *et al.* Structure, functionality, and evolution of the BARE-1 retrotransposon of barley[J]. *Genetica*, 1999, **107**(1—3): 53—63.
- [12] SANMIGUEL P, GAUT B S, TIKHONOV A, *et al.* The paleontology of intergene retrotransposons of maize[J]. *Nat Genet.*, 1998, **20**(1): 43—45.
- [13] WICKER T, STEIN N, ALBAR L, *et al.* Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution[J]. *Plant J.*, 2001, **26**(3): 307—316.
- [14] PETERSON D G, SCHULZE S R, SCIARA E B, *et al.* Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery[J]. *Genome Res.*, 2002, **12**(5): 795—807.
- [15] GOODWIN T J, POULTER R T. A new group of tyrosine recombinase-encoding retrotransposons[J]. *Mol. Biol. Evol.*, 2004, **21**(4): 746—759.
- [16] POULTER R T, GOODWIN T J, DIRS-1 and the other tyrosine recombinase retrotransposons[J]. *Cytogenet Genome Res.*, 2005, **110**(1—4): 575—588.
- [17] EVGEN'EV M B, ARKHIPOVA I R. *Penelope*-like elements—a new class of retroelements; distribution, function and possible evolutionary significance[J]. *Cytogenet Genome Res.*, 2005, **110**(1—4): 510—521.
- [18] HIGASHIYAMA T, NOUTOSHI Y, FUJIE M, *et al.* Zepp, a LINE-like retrotransposon accumulated in the chlorella telomeric region [J]. *EMBO J.*, 1997, **16**(12): 3 715—3 723.
- [19] ZUPUNSKI V, GUBENSEK F, KORDIS D. Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons[J]. *Mol. Biol. Evol.*, 2001, **18**(10): 1 849—1 863.
- [20] KRAMEROV D A, VASSETZKY N S. Origin and evolution of SINEs in eukaryotic genomes[J]. *Heredity*(Edinb), 2011, **107**(6): 487—495.
- [21] WICKER T, SABOT F, *et al.* A unified classification system for eukaryotic transposable elements[J]. *Nat. Rev. Genet.*, 2007, **8**(12): 973—982.
- [22] MCCARTHY E M, MCDONALD J F. LTR_STRUC: a novel search and identification program for LTR retrotransposons[J]. *Bioinformatics*, 2003, **19**(3): 362—367.
- [23] KALYANARAMAN A, ALURU S. Efficient algorithms and software for detection of full-length LTR retrotransposons[J]. *J. Bioinform Comput Biol.*, 2006, **4**(2): 197—216.
- [24] 虞洪杰. 植物 LTR 反转录转座子的预测和注释及邻聚法构建系统进化树[D]. 杭州: 浙江大学, 2011.
- [25] DIAO W P(刁卫平), WANG SH B(王述彬), LIU J B(刘金兵), *et al.* Cloning and sequence analysis of reverse transcriptase of Ty1-*copia*-like retrotransposons in *Capsicum annuum* [J]. *Molecular Plant Breeding* (分子植物育种), 2012, **10**(1): 55—61 (in Chinese).
- [26] FAN F H(范付华), QIAO G(乔 光), ZHENG S CH(郑思成), *et al.* Cloning and analysis of reverse transcriptase of Ty1-*copia* retrotransposons in *Hylocereus undatus* [J]. *Acta Horticulturae Sinica* (园艺学报), 2012, **39**(2): 265—272 (in Chinese).
- [27] KOBAYASHI S, GOTO-YAMAMOTO N, *et al.* Retrotransposon-induced mutations in grape skin color[J]. *Science*, 2004, **304**(5 673): 982.
- [28] BUTELLI E, LICCIARDELLO C, ZHANG Y, *et al.* Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges[J]. *Plant Cell*, 2012, **24**(3): 1 242—1 255.
- [29] MIROUZE M, REINDERS J, *et al.* Selective epigenetic control of retrotransposition in *Arabidopsis* [J]. *Nature*, 2009, **461**(7 262): 427—430.
- [30] TSUKAHARA S, KOBAYASHI A, *et al.* Bursts of retrotransposition reproduced in *Arabidopsis* [J]. *Nature*, 2009, **461**(7 262): 423—426.
- [31] HIROCHIKA H, OKAMOTO H, KAKUTANI T. Silencing of retrotransposons in arabidopsis and reactivation by the *ddm1* mutation [J]. *Plant Cell*, 2000, **12**(3): 357—369.
- [32] De FELICE B, WILSON R R, ARGENZIANO C, *et al.* A transcriptionally active copia-like retroelement in *Citrus limon* [J]. *Cell Mol Biol Lett*, 2009, **14**(2): 289—304.
- [33] TAPIA G, VERDUGO I, YANEZ M, *et al.* Involvement of ethylene in stress-induced expression of the *TLC1.1* retrotransposon from *Lycopersicon chilense* Dun [J]. *Plant Physiol.*, 2005, **138**(4): 2 075—2 086.
- [34] BERGMAN C M, QUESNEVILLE H. Discovering and detecting transposable elements in genome sequences[J]. *Brief Bioinform.*, 2007, **8**(6): 382—392.
- [35] VOLFOVSKY N, HAAS B J, *et al.* A clustering method for repeat analysis in DNA sequences[J]. *Genome Biol.*, 2001, **2**(8): 21—27.
- [36] LI R, YE J, LI S, *et al.* ReAS: Recovery of ancestral sequences for transposable elements from the unassembled reads of a whole genome shotgun[J]. *PLoS Comput Biol*, 2005, **1**(4): e43.
- [37] EDDY S R. Profile hidden Markov models[J]. *Bioinformatics*, 1998, **14**(9): 755—763.
- [38] XU Z, WANG H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons[J]. *Nucleic Acids Res.*, 2007, **35**(Web Server issue): 265—268.
- [39] WANG H, LIU J S. LTR retrotransposon landscape in *Medicago truncatula*: more rapid removal than in rice[J]. *BMC Genomics*, 2008, **9**: 382.
- [40] XU L, ZHANG Y, SU Y, *et al.* Structure and evolution of full-length LTR retrotransposons in rice genome[J]. *Plant Systematics and E-*

- volution, 2010, **287**(1–2):19–28.
- [41] COSSU R M, BUTI M, GIORDANI T, *et al.* A computational study of the dynamics of LTR retrotransposons in the *Populus trichocarpa* genome[J]. *Tree Genetics & Genomes*, 2012, **8**(1):61–75.
- [42] ELLINGHAUS D, KURTZ S, WILLHOEFT U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons[J]. *BMC Bioinformatics*, 2008, **9**:18.
- [43] MCCARTHY E M, McDONALD J F. LTR_STRUC: a novel search and identification program for LTR retrotransposons[J]. *Bioinformatics*, 2003, **19**(3):362–367.
- [44] KALYANARAMAN A, ALURU S. Efficient algorithms and software for detection of full-length LTR retrotransposons[J]. *J. Bioinform. Comput Biol.*, 2006, **4**(2):197–216.
- [45] RHO M, CHOI J H, KIM S, *et al.* De novo identification of LTR retrotransposons in eukaryotic genomes[J]. *BMC Genomics*, 2007, **8**:90.
- [46] WANG H, XU Z, YU H. LTR retrotransposons reveal recent extensive inter-subspecies nonreciprocal recombination in Asian cultivated rice[J]. *BMC Genomics*, 2008, **9**:565.
- [47] FLAVELL A J, DUNBAR E, ANDERSON R, *et al.* Ty1-copia group retrotransposons are ubiquitous and heterogeneous in higher plants[J]. *Nucleic Acids Res.*, 1992, **20**(14):3 639–3 644.
- [48] WAUGH R, MCLEAN K, FLAVELL A J, *et al.* Genetic distribution of *Bare-1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP)[J]. *Mol. Gen. Genet.*, 1997, **253**(6):687–694.
- [49] KALENDAR R, SCHULMAN A H. IRAP and REMAP for retrotransposon-based genotyping and fingerprinting[J]. *Nat Protoc*, 2006, **1**(5):2 478–2 484.
- [50] KALENDAR R, FLAVELL A J, ELLIS T H, *et al.* Analysis of plant diversity with retrotransposon-based molecular markers[J]. *Heredity*(Edinb), 2011, **106**(4):520–530.
- [51] SYED N H, FLAVELL A J. Sequence-specific amplification polymorphisms (SSAPs): a multi-locus approach for analyzing transposon insertions[J]. *Nature Protocols*, 2007, **1**(6):2 746–2 752.
- [52] YU G X, WISE R P. An anchored AFLP- and retrotransposon-based map of diploid Avena[J]. *Genome*, 2000, **43**(5):736–749.
- [53] WOODROW P, PONTECORVO G, CIARMIELLO L F. Isolation of Ty1-copia retrotransposon in myrtle genome and development of S-SAP molecular marker[J]. *Mol. Biol. Rep.*, 2012, **39**(4):3 409–3 418.
- [54] TAM S M, MHIRI C, VOGELAAR A, *et al.* Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR[J]. *Theor. Appl. Genet.*, 2005, **110**(5):819–831.
- [55] PAUX E, FAURE S, CHOLET F, *et al.* Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat[J]. *Plant Biotechnol. J.*, 2010, **8**(2):196–210.
- [56] VITTE C, ISHII T, LAMY F, *et al.* Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.)[J]. *Mol. Genet Genomics*, 2004, **272**(5):504–511.
- [57] KIM H, TERAOKAMI S, NISHITANI C, *et al.* Development of cultivar-specific DNA markers based on retrotransposon-based insertional polymorphism in Japanese pear[J]. *Breed Sci.*, 2012, **62**(1):53–62.
- [58] DU X Y(杜晓云), ZHANG Q L(张青林), LUO ZH R(罗正荣). Isolation and characterization of RNaseH-LTR sequences of Ty1-copia like retrotransposons in Oriental Persimmon[J]. *Acta Horticulturae Sinica*(园艺学报), 2008, **35**(4):501–508(in Chinese).
- [59] KALENDAR R, SCHULMAN A H. IRAP and REMAP for retrotransposon-based genotyping and fingerprinting[J]. *Nat Protoc*, 2006, **1**(5):2 478–2 484.
- [60] MA J, DEVOS K M, BENNETZEN J L. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice[J]. *Genome Res.*, 2004, **14**(5):860–869.
- [61] WICKER T, STEIN N, ALBAR L, *et al.* Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution[J]. *Plant J.*, 2001, **26**(3):307–316.
- [62] VICIENT C M, KALENDAR R, ANAMTHAWAT-JONSSON K, *et al.* Structure, functionality, and evolution of the BARE-1 retrotransposon of barley[J]. *Genetica*, 1999, **107**(1–3):53–63.
- [63] SANMIGUEL P, GAUT B S, TIKHONOV A, *et al.* The paleontology of intergene retrotransposons of maize[J]. *Nat Genet*, 1998, **20**(1):43–45.
- [64] PETERSON D G, SCHULZE S R, SCIARA E B, *et al.* Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery[J]. *Genome Res.*, 2002, **12**(5):795–807.
- [65] PIEDNOEL M, CARRETE-VEGA G, RENNER S S. Characterization of the LTR retrotransposon repertoire of a plant clade of six diploid and one tetraploid species[J]. *Plant J.*, 2013, **75**(4):699–709.
- [66] FLAVELL A J, DUNBAR E, ANDERSON R, *et al.* Ty1-copia group retrotransposons are ubiquitous and heterogeneous in higher plants[J]. *Nucleic Acids Res.*, 1992, **20**(14):3 639–3 644.
- [67] GUO X, OTHERS. Genetic polymorphism of wheat by IRAP analysis based on retrotransposon Wis2-1 A[J]. *Agricultural Science and Technology*, 2011, **11**.
- [68] KIM H, TERAOKAMI S, NISHITANI C, *et al.* Development of cultivar-specific DNA markers based on retrotransposon-based insertional polymorphism in Japanese pear[J]. *Breed Sci.*, 2012, **62**(1):53–62.