

· 综述 ·

## 植物中活性 MITEs 转座子研究进展

胡冰杰<sup>1</sup>, 周明兵<sup>1,2</sup>

1 浙江农林大学 省部共建亚热带森林培育国家重点实验室, 浙江 杭州 311300

2 浙江农林大学 浙江省竹资源与高效利用协同中心, 浙江 杭州 311300

胡冰杰, 周明兵. 植物中活性 MITEs 转座子研究进展. 生物工程学报, 2018, 34(2): 204–215.

Hu BJ, Zhou MB. Active miniature inverted-repeat transposable elements transposon in plants: a review. Chin J Biotech, 2018, 34(2): 204–215.

**摘要:** MITEs (Miniature inverted-repeat transposable elements) 转座子是一种特殊的转座子, 其既有 DNA 转座子的转座特性——“剪切-粘贴”转座方式, 又有 RNA 转座子的高拷贝特性。目前已被报道的 MITEs 种类和数量虽然很多, 但是关于有转座活性的 MITEs 的报道却甚少。本文总结了近几年来有关活性 MITEs 的相关报道, 发现具有转座活性的 MITEs 种类大都分布在 Tourist 家族, 分别是 mPing、mGing、PhTourist1、Tmi1 和 PhTst-3, 另外还有 Stowaway-like 家族的 dTstu1 和 MITE-39 以及 Mutator 家族的 AhMITE1。文中还分析了这些活性 MITEs 的结构 (TIR 和 TSD)、拷贝数、进化模式以及转座特性等, 为鉴定其他活性 MITEs 以及 MITEs 转座和扩增机制的研究奠定了基础。

**关键词:** 微型颠倒重复转座元件 (MITEs), 转座活性, 末端重复序列 (TIR), 靶位点重复序列 (TSD)

## Active miniature inverted-repeat transposable elements transposon in plants: a review

Bingjie Hu<sup>1</sup>, and Mingbing Zhou<sup>1,2</sup>

1 State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University, Hangzhou 311300, Zhejiang, China

2 Zhejiang Provincial Collaborative Innovation Center for Bamboo Resources and High-efficiency Utilization, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China

**Abstract:** Miniature inverted-repeat transposable elements transposon is a special transposon that could transpose by "cut-paste" mechanism, which is one of characteristics of DNA transposons. Otherwise, the copy number of MITEs is very high, which is one of characteristics of RNA transposons. Many MITE families have been reported, but little about active MITEs. We summarize recent advances in studying active MITEs. Most the MITEs belong to the Tourist-like family, such as mPing, mGing, PhTourist1, Tmi1 and PhTst-3. Additionally, DTstu1 and MITE-39 belong to Stowaway-like family, and AhMITEs1 belongs to Mutator-like family. Moreover, we summarize the structure (terminal inverse repeats and target site

**Received:** May 24, 2017; **Accepted:** September 20, 2017

**Supported by:** Natural Science Foundation of Zhejiang Province, China (No. LR12C16001), National Natural Science Foundation of China (Nos. 31270645, 31470615).

**Corresponding author:** Mingbing Zhou. Tel: +86-571-63731263; E-mail: zmbin@163.com

浙江省自然科学基金 (No. LR12C16001), 国家自然科学基金 (Nos. 31270645, 31470615) 资助。

网络出版时间: 2017-09-25

网络出版地址: <http://kns.cnki.net/kcms/detail/11.1998.Q.20170925.1430.001.html>

duplications), copy number, evolution pattern and transposition characteristics of these active MITEs, to provide the foundation for the identification of other active MITEs and subsequent research on MITE transposition and amplification mechanism.

**Keywords:** miniature inverted-repeat transposable elements, transposon activity, terminal inverse repeats, target site duplications

转座子 (Transposons) 是染色体中一段可以跳跃的 DNA 序列, 可以从基因组的一个位置移动到另一个位置, 引起基因组的改变甚至染色体结构的变异等<sup>[1]</sup>。根据转座子的转座机制不同, 可将转座子分为两大类: RNA 类转座子和 DNA 类转座子。RNA 类转座子又称反转录转座子, 该类转座子先是转录成 RNA, 然后再通过反转录成 DNA 转座到其他位置, 整个过程以“DNA-RNA-DNA”的形式完成转座, 所以 RNA 类转座子是以“复制-粘贴”的方式转座。RNA 类转座子每次转座, 其拷贝数都会相应增加, 一般比 DNA 类转座子的拷贝数要高<sup>[2]</sup>。而 DNA 类转座子是通过“DNA-DNA”的“剪切-粘贴”的形式转座, 是在转座酶的作用下, 从一个位置剪切下来, 插入到另一个位置<sup>[3]</sup>。

DNA 转座子根据能否发生自主转座又可分为自主转座子和非自主转座子。自主转座子由于其结构完整, 具有编码完整转座酶的能力, 所编码的转座酶可以促使自身发生转座; 而非自主转座子与自主转座子相比, 其内部的转座酶编码序列有所缺失, 不能编码转座酶, 必须依靠相应的自主转座子编码的转座酶作用才能转座。

微型颠倒重复转座元件 (Miniature inverted repeat transposable elements, MITEs) 是比较特殊的非自主转座子, 其结构类似于一般的非自主转座子, 但是片段长度要小得多, 一般长度在 100–800 bp 范围。另外, 与一般的非自主转座元件相比, 大多数 MITEs 家族在基因组中具有相当大的拷贝数, 多达上千甚至上万<sup>[4]</sup>, 例如, 在玉米 *Zea mays* 中发现 mPIF 的拷贝数达到 6 000 多个拷贝<sup>[5]</sup>, 在水稻 *Oryza sativa* 基因组有 90 000 多个 MITEs, 占水稻转座子 75% 以上<sup>[6]</sup>。此外, MITEs 的 A/T 含量较高, 可以形成稳定的发夹式二级结构, 且

两端具有高度保守的反向重复序列 (Terminal inverted repeats, TIR, 转座酶的识别位点) 和 2–11 bp 靶位点重复序列 (Target site duplications, TSD, 转座酶识别位点), 倾向于插入基因内部或附近。MITEs 两端的 TIR 可以形成茎环结构, 能够转录加工成 3 种内源小 RNA, 包括 miRNA、endo-siRNA 和 piRNA (PIWI-interacting RNA)<sup>[7–8]</sup>。这些来源于转座子的小 RNA 除了在抑制宿主转座子跳跃、维持基因组的生态环境发挥重要作用外, 还调控着众多宿主基因的表达<sup>[9]</sup>, 如来源于两个 MITEs 的水稻 siRNA441 和 siRNA446 除了参与 MITEs 转座活性的调控外, 还参与了水稻对脱落酸信号和非生物胁迫的响应<sup>[10]</sup>。

目前已发现上百个不同的 MITEs 家族, 根据 MITEs 两端的 TIR 和 TSD 特性, 可将 MITEs 归为 Stowaway 类 (TSD: TA) 和 Tourist 类 (TSD: TAA 或 TTA) 两个大家族以及其他若干个小家族 (如 CACTA、MUDR 等家族)<sup>[11]</sup>。MITEs 的插入可以通过多种形式调控附近的基因表达, 比如插入使基因失活, 为基因提供顺式调控元件, 引起基因的可变剪接, 提供反义 RNA 或小 RNA 等<sup>[12]</sup>, 这些调控在遗传和表观两个层次都能发挥作用, 因此活性 MITEs 越来越受研究者重视, 以期利用活性 MITEs 开发基因标签, 为大规模研究基因功能和构建饱和突变体库提供新的工具。本文系统收集了目前已鉴定活性 MITEs, 总结了活性 MITEs 的结构 (TIR 和 TSD)、拷贝数、进化模式以及转座特性, 为鉴定其他活性 MITEs 以及 MITEs 转座和扩增机制的研究奠定基础。

## 1 MITEs 的发现及其在植物基因组中的分布

1992 年, Bureau 等<sup>[13]</sup>在研究玉米 *wx-B2* 基因

的插入突变分析中发现一个 128 bp 的短片段元件, 该元件在基因组中分布非常广, 分析其序列发现该元件具有 14 bp 的 TIR (5'-GGCCTTGTTTGGTT-3') 以及 3 bp 的 TSD (TAA/TTA) 保守序列, 与其他带有 TIR 的转座子相比无序列相似性, 且该家族元件的拷贝数非常丰富, 因此将该类转座子另外命名为 Tourist 家族转座子。随后在 Tourist 家族转座子的基础上, Bureau<sup>[11]</sup>又在一些草本植物的基因组中发现另一类转座子, 其 TSD 序列为 TA, 将其命名为 Stowaway 转座子, 与 Tourist 转座子共同称为 MITEs 元件, 自此, 人们开始了对 MITEs 的研究之路。

20 世纪 90 年代中期, 随着基因组测序技术的发展, MITEs 家族种类有暴发式的增长, 人们根据已发现的 MITEs 的一些基本特征, 研发出可以在基因组中鉴定出 MITEs 的计算工具, 如 RSPB<sup>[6]</sup>、FINDMITES<sup>[14]</sup>、MUST<sup>[15]</sup>等, 可以通过 TIR、TSD 等特征结构, 直接在基因组中搜索发现不同种类的 MITEs。这些计算工具的出现, 相比于过去通过实验来检索验证 MITEs, 大大方便了研究人员对 MITEs 的检索。但是由于技术不

成熟, 这些工具鉴定出的 MITEs 假阳性率很高。为了降低其假阳性率, MITE-Hunter<sup>[16]</sup>和 MITE Digger<sup>[17]</sup>相继被开发出来, 可以在全基因组中从头探测, 特别是 MITE Digger, 其假阳性最低, 并且速度快, 是目前应用最为广泛的 MITEs 搜索工具。以水稻基因组中 MITEs 家族为例, MITE-Hunter 的假阳性率为 4.4%–8.3%, FINDMITE 为 85%, MUST 为 86%, 而 MITE Digger 的假阳性率最低, 仅有 1.8%, 并且搜索时间明显比前几种方法短。Ye 等<sup>[18]</sup>基于 MATLAB 采用新颖的数值计算的方法, 开发出一个新的程序——detect MITE, 将其与 MITE-Hunter、MITE Digger 和 RSPB 的工作效率进行比较, 发现 detect MITE 的工作效率是最高的, 尽管 RSPB 搜索出的 MITEs 序列最多, 但大部分都不完整。

根据已报道的一些物种基因组中的 MITEs 含量, 发现虽然 MITEs 在植物基因组中含量丰富, 但在不同的植物基因组中, MITEs 的含量差异较大 (图 1), 例如比较水稻与玉米的 MITEs 含量, 玉米的基因组是水稻的 7.6 倍, 但是其 MITEs 的含量却不到水稻的一半。

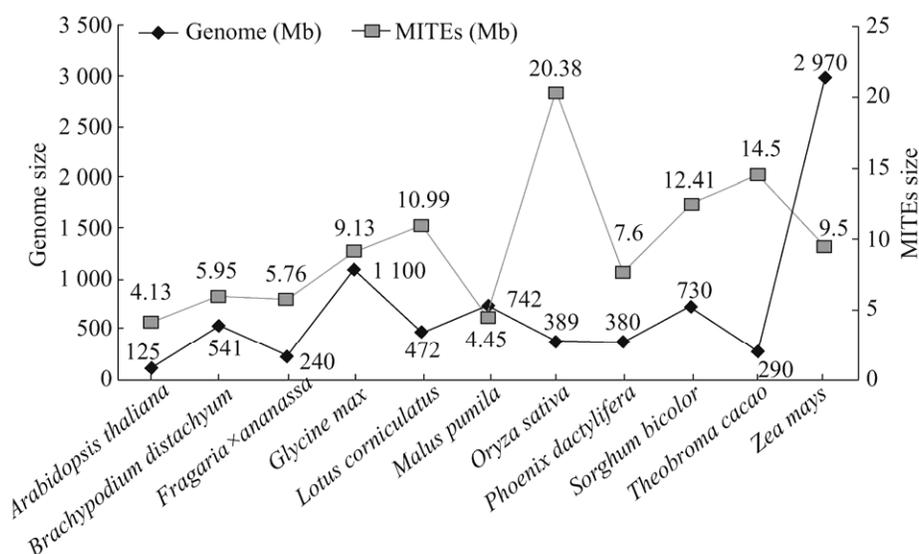


图 1 不同物种基因组大小与其已报道的 MITEs 含量的关系<sup>[19]</sup>

Fig. 1 The relationship between the genomic size of different species and the reported MITEs content<sup>[19]</sup>.

此外, 比较一些物种的基因组与其报道的 MITEs 含量关系, 发现 MITEs 的含量与物种基因组大小无明显的线性关系, 如水稻的基因组比高粱 *Sorghum bicolor* L. 小近一倍, 但是其报道的 MITEs 的含量却比高粱的多出很多 (图 1)。其次, 在亲缘关系较近的物种中, 其拷贝数也存在较大的差异, 如深山南芥 *Arabidopsis lyrata* L. 中的 MITEs 数量大约比拟南芥 *Arabidopsis thaliana* L. 多 4 倍。同样地, 西瓜 *Citrullus lanatus* L. 中的 MITEs 数量大约是甜瓜 *Cucumis melo* L. 的 7 倍<sup>[20]</sup>。

对 MITEs 插入位点分析, 发现大部分 MITEs 倾向插入到基因附近, 也有一部分插入到内含子或者非翻译区<sup>[21]</sup>。如在水稻基因组有 90 000 个 MITEs, 占水稻转座子 75% 以上<sup>[6]</sup>, 其中有 8.2% 位于基因内部, 81.2% 位于基因上游和下游 2 kb 以内区域<sup>[21]</sup>。

## 2 MITEs 的扩增机制

MITEs 和非自主转座子一样, 是缺失编码完整转座酶序列的转座子, 不能自主转座, 但是相比非自主转座子, MITEs 在基因组中的拷贝数却比非自主转座子要高得多, 引起了众多研究者对其扩增机制进行探究。有研究者提出“删除论”: 一些自主的 DNA 转座子通过内部不断删减, 保留了自主转座元件的 TSD、TIR 序列以及中间残留部分, 最终形成了几百 bp 大小的 MITEs, 这些被删减的序列可能就是编码转座酶的序列片段,

而保留的部分可能含有可被转座酶识别的序列, 使编码的转座酶可以识别对应的 MITEs, 从而诱导 MITEs 发生转座<sup>[22]</sup>。2003 年, Feschotte 等<sup>[23]</sup>研究发现, 水稻中 Mariner-like elements (MLE) 自主转座子与 Stowaway-like MITEs 的 TIR 和 TSD 序列高度一致, 推测 MLE 编码的转座酶可能会催化 Stowaway-like MITEs 发生转座。2009 年, Yang 等<sup>[24]</sup>通过在体外构建的 MLE 转座酶序列, 在酵母 *Saccharomyces cerevisiae* 体内催化 Stowaway-like MITEs 转座, 更进一步明确了 MLE 和 Stowaway-like MITEs 之间的关系。由此推测 MLE 在物种进化过程中, 由于内部编码转座酶序列不断删减, 最后成为了缺失编码全部转座酶序列的 Stowaway-like MITEs。另一例子是 mPing, mPing 是第一个发现的具有转座活性的 Tourist-like 家族的 MITEs 转座子<sup>[25-26]</sup>, Yang 等<sup>[27]</sup>又发现了两个自主 DNA 转座子 Ping 和 Pong, 这两个自主转座子的末端与发现的 mPing 具有很高的同源性, 这说明了 mPing 很有可能是 Ping 和 Pong 内部删减的结果。通过拟南芥遗传转化体系, 证实了 Pong 自主转座子编码的转座酶可以促使 mPing 转座子发生转座。这些结果都验证了他们之前所提出的“删除论”。

## 3 活性 MITEs 发现与鉴定

目前在已发现的具有转座活性的 MITEs 中 (表 1), 种类最多的是 Tourist-like 家族, 有 mPing、

表 1 活性 MITEs 特征

Table 1 Features of active MITEs

MITE	Class	Sequence length (bp)	TIR	TSD	Reference
mPing	Tourist-like	430	GGCCAGTCACAATGG	TWA	[25]
mGing	Tourist-like	146	GTTTAGTT	TWA	[28]
PhTourist1	Tourist-like	356	GGCCAGTCTCAATG	TWA	[29]
TMi1	Tourist-like	306-336	GGGGTCGTTT	TAT/TAA	[30]
PhTst-3	Tourist-like	295	GGGCATGTACA	TTA	[31]
dTst1	Stowaway-like	239	CTCCCTCYGTC	TA	[24]
MITE-39	Stowaway-like	260	CTCCCTCCGTTCTTAAATA	TA	[32]
AhMITE1	Mutator-like	201-223	GGTGGATACTACAATGAAGATGGCA	ATGCAATAA	[33]

mGing、PhTourist1、Tmi1 和 PhTst-3; 另外还有属于 Stowaway-like 家族的 dTstu1 和 MITE-39 以及属于 Mutator 家族的 AhMITE1。

### 3.1 Tourist-like 家族活性 MITEs 的发现与鉴定

#### 3.1.1 mPing

水稻作为单子叶植物中的典型模式植物, 其参考基因组序列较为完善。通过生物信息学分析发现, 虽然水稻基因组仅有 430 Mb, 但 MITEs 含量非常高, 并且种类丰富, 几乎涵盖了所有类型的 MITEs, 因此被认为是适合研究 MITEs 转座子的模式材料<sup>[34-35]</sup>。对于这一现象, 有研究者推出一个“cross-mobilization”的假说, 认为 MITEs 的起源和扩增发生在不同的时间段, MITEs 最早起源是由于早期的自主转座子经过长期的内部删除, 经完全删除转座酶编码序列后, 只剩下 TIR 和 TSD 序列以及不编码的小片段, 再经过后面一些“年轻的”自主转座子编码的转座酶作用下发生转座并发生扩增, 并随着时间的积累, 扩增的 MITEs 也不断积累, 从而出现富集的现象<sup>[23]</sup>。

2003 年, 在研究  $\gamma$  射线诱导水稻颖叶表型突变的过程中, 研究人员发现, 经过  $\gamma$  辐射后, 在水稻的 *Rurm 1* 基因中有一个 MITE 片段发生了跳跃, 这是第一个被发现具有转座活性的 MITE 转座子, 遂将其命名为 mPing。比较并分析突变后的水稻与野生型水稻的 *Rurm 1* 基因, mPing 插入到 *Rurm 1* 基因的第 4 个内含子中, 从而改变了 *Rurm 1* 基因的表达, 导致水稻颖叶比野生型的更加细长<sup>[25]</sup>, 并且 mPing 的拷贝数也会明显增加<sup>[36-37]</sup>。同年, 另两个研究发现, mPing 在水稻花药组织培养过程中亦可以被激活并发生转座<sup>[26]</sup>。mPing 除了在 Ping 转座元件编码的转座酶催化下发生转座外, 也会在 Pong 的转座酶作用下被激活<sup>[34]</sup>。2013 年, 有研究发现, 在没有外界胁迫因素的情况下, 当编码泛素蛋白的基因被沉默后, mPing 的转座活

性会提高<sup>[38]</sup>。上述研究表明, mPing 是一个仍然具有转座活性的 MITEs, 并且能在花药组培或辐射等胁迫条件下被激活转座, 而当泛素蛋白被沉默时, 其转座频率会增加。

分析 mPing 的序列发现, 其序列有 430 bp, 包括 15 bp 的 TIR (5'-GGCCAGTCACAATGG-3') 和 3 bp 的 TSD (TTA 或 TAA), 因此该 MITE 家族属于 Tourist 类转座子。在之后的研究中, 研究人员通过同源性比较搜索, 发现了两个自主 DNA 转座子 Ping 和 Pong 的末端与 mPing 具有很高的同源性。mPing 的 TSD 和 TIR 序列与 Ping 一样, 除了 mPing 缺失了 Ping 元件中的两个 ORFs 外, 其他序列几乎一致, 推测 mPing 很可能是 Ping 自主转座元件长期不断的内部删除的结果。通过拟南芥遗传转化体系, 证实了 Pong 与 Ping 两个自主转座子编码的转座酶可以促使 mPing 转座子发生转座<sup>[27,39]</sup>。

进一步研究发现, mPing 序列有 96 个调控保守区, 其中有 1/3 与环境胁迫有关, 对附近基因表达调控起着重要作用。mPing 在基因组的分布广泛, 使水稻 395 个基因具有受冷害、高盐和干旱条件诱导表达的特性<sup>[37]</sup>。当 mPing 插入转录起始位点上游 1-5 kb 范围内, 会上调下游基因的表达; 但是当 mPing 插入到外显子中, 则会下调下游基因的表达; 当 mPing 插入到启动子区, 会影响顺式作用元件和反式作用元件的相互作用, 继而影响下游基因的表达<sup>[40]</sup>。这些结果表明, mPing 的插入为转录因子或其他转录调节蛋白提供一个新的结合位点, 从而影响下游基因的表达丰度。

#### 3.1.2 mGing

Gaijin-like MITEs 属于 Tourist 家族, 最早于 1996 年在水稻基因组中发现<sup>[41]</sup>, 但是当时对 Gaijin 的研究结果仅仅是在水稻基因组中鉴定出 5 个不同位点的 Gaijin-like MITEs 的插入, 并对这几个 Gaijin 元件进行了简单的结构分析, 之后

并没有继续深入研究该家族的其他方面(如转座活性、功能等)。2012年, Dong等<sup>[28]</sup>在水稻 PLRRP (Putative leucine-rich repeat protein) 基因中鉴定出一个新的 Gaijin-like MITEs——mGing, 这也是继 mPing 之后第二个被发现具有转座活性的 Tourist 类转座子。该类 MITEs 大小只有 146 bp, 包括 8 bp 的 TIR (GTTTAGTT) 和 3 bp 的 TSD (TAA 或 TTA) 序列, 并且如其他 Tourist 类转座子一样, 倾向于插入富含 AT 区。以 146 bp 的 mGing 元件作为 query 序列, 在水稻全基因组中比对, 发现其拷贝数达上千, 然而在拟南芥基因组中尚未发现该家族元件。为鉴定 mGing 是否具有转座活性, 选取同一株水稻种子, 用不同剂量的  $\gamma$  射线处理水稻种子并培育, 对不同处理的水稻基因组做转座子显示技术 (Transposon display, TD), 发现 500 gy 处理的水稻基因组出现多处多态性条带, 而未辐射和 300 gy 剂量的水稻条带均一, 并与已报道的 Gaijin-like MITEs 为对照, 均未表现出多态性, 排除了基因重组、基因重排等假阳性, 说明 500 gy 处理后的水稻基因组中, mGing 可能发生转座。通过 PCR 技术扩增含 mGing 的片段并测序, 发现有几个 mGing 的侧翼序列在数据库中并未与 mGing 相连, 这也确定了 mGing 在水稻种子经过辐射处理后的确发生了转座。另外, 比较了 mGing 转座前后的位置序列, 均发现随着 mGing 的剪切或插入, 其侧翼序列都有部分碱基改变, 或者有小片段插入。

### 3.1.3 PhTourist1

2015年, 胡慧等利用已知 mPing 的特点 (TSD 和 TIR 序列结构等), 通过 transpo<sup>[42]</sup>在毛竹 *Phyllostachys edulis* C. 基因组中鉴定出 30 个 mPing-like 的 MITEs——PhTourist1, 比对该 30 条序列发现, 其 TIR 序列是 GGCCAGTCTCAATG, 共 14 bp, 与 mPing 的 TIR (GGCCAGTCACAA TGG) 序列相比 (表 1) 少一个碱基 (GGCCAGT

CACAATGG), 并且伴有一个碱基不同。比较 PhTourist1 家族转座子的侧翼序列发现, 其 TSD 序列为 TWA (W=A/T), 以及插入偏好序列为 (C/T)T(C/A)T(T/A)A(G/T)A(A/C)。通过 PCR 扩增及测序发现, 30 个 PhTourist1 的 MITEs 序列及其侧翼序列均符合毛竹基因组序列, 但是在这 30 个 MITEs 中, 有一个 PhTourist1-3 发现具有多态性, 证明了在毛竹实生苗生长过程中 PhTourist1 发生了转座, 对 PhTourist1-3 转座足迹分析发现, PhTourist1-3 几乎精确切除, 仅在 5'端缺失 1 个碱基。这是首次在毛竹基因组中发现并被鉴定出具有活性的 Tourist-like 的 MITEs。通过定量 PCR 验证发现, 当该位点 PhTourist1-3 转座缺失后, 其下游基因的表达较缺失前明显增加。因为所选取的材料是来自同一棵毛竹的半同胞实生苗, 并且种植环境相同且适宜, 胡慧等推测 PhTourist1 的活性转座可能不是由于外界压力激活的, 很有可能是在种子发芽的过程中发生转座, 也有可能是在授粉过程中, 由于授粉的父本毛竹不同而导致的。PhTourist1 的插入很有可能是改变了顺反式作用元件的作用, 从而抑制了下游基因的表达, 当 PhTourist1 转座切除后, 下游基因的表达量明显增加<sup>[29]</sup>。

### 3.1.4 TMi1

骆红梅等<sup>[30]</sup>在烟草 *Nicotiana tabacum* L. 中发现一个具有转座活性的 MITEs——TMi1, 这也是首次在烟草基因组中发现具有转座活性的 MITEs。TMi1 的 TSD 序列为 TAT/TAA, 属于 Tourist-like 家族转座子, 其序列长度在 306–336 bp 不等, TIR 序列为 GGGGTCGTTT, 共 17 bp。该研究利用 TD 技术初步评估出 TMi1 家族在若干个不同品种的烟草中存在着明显的多态性, 初步判断烟草 TMi1 家族的 MITEs 可能仍具有转座活性; 将具有扩增差异的条带胶回收并测序, 分析测序结果发现, 回收的条带大部分都是带有 TMi1 元件的

序列, 随后利用 Blastn 将上述测序结果与数据库中的序列比对, 发现部分测序结果中的序列多出 TMi1 元件, 这些结果说明 TMi1 家族在若干个不同烟草品种间发生了跳跃, 推测烟草中 TMi1 家族转座元件目前仍具有转座活性。

### 3.1.5 PhTst-3

随着毛竹的基因组测序的完成, 我们在毛竹基因组中鉴定出 489 592 个 MITEs, 分别可以归纳为 362 个家族, 占据毛竹全基因组的 4.74%。根据这些 MITEs 的 TIR 和 TSD 序列的保守性, 可将其分成 6 个超家族, 分别是 Stowaway-like MITEs、hAT-like MITEs、Tourist-like MITEs、Mutator-like MITEs 和 CACTA-like MITEs, 而对于一些 TIR 或 TSD 序列特征不明显的 MITE 暂时归为未知超家族类别<sup>[43]</sup>。在对毛竹基因组中的 MITEs 家族进化、插入时间、分布等分析之后, 我们进一步对 MITEs 的转座活性以及转座后对附近基因的影响进行研究, 在胡慧等<sup>[29]</sup>研究发现 PhTourist1 家族具有转座活性之后, 又先后在毛竹基因组中鉴定了两个具有转座活性的 Tourist-like 和 Stowaway-like MITEs。陈昂<sup>[31]</sup>在不同毛竹实生苗中, 发现有一个 MITE 在 PH01002699G0010 基因的第 6 个内含子中存在插入多态性, 初步判断该类 MITEs 具有转座活性。通过生物信息学的方法, 分析发现该类 MITEs 在毛竹基因组中具有 94 个拷贝, 全长 295 bp, 其 TIR 序列为 GGGCATGTACA, TSD 序列为 TTA, 属于 Tourist-like MITEs, 命名为 PhTst-3。通过荧光定量 PCR 发现, 与含有 PhTst-3 插入的基因相比, 未插入的 PH01002699G0010 基因的表达量明显有所上升, 表明 PhTst-3 插入内含子的作用相当于一个沉默子, 下调了基因的表达。

### 3.1.6 Tourist-like 家族活性 MITEs 进化模式分析

将 mPing、mGing、PhTourist1、Tmi1 和 PhTst-3 共 5 个 Tourist-like 家族的 MITE 序列与 Ping 和 Pong 序列通过 DNAMAN 比对 (图 2A), 发现除

了 Tmi1 家族 MITE 的 TSD 序列为 TAW (T/A), 其余几个家族的 TSD 均为 TW (A/T)A。Tmi1 家族 MITEs 的 TIR 序列也与其他 MITEs 的 TIR 差异明显。Ping、Pong、mPing、PhTourist1 的 TIR 序列一致性较高, 特别是 mPing 与 Ping 和 Pong, 除 TIR 序列一致以外, 其内部序列与 Ping 的内部序列相似度很高。

为了进一步分析 Tourist-like 家族活性 MITEs 的进化特征, 利用 MEGA 6.0 将 Ping、Pong 与活性 Tourist-like MITEs (mPing、mGing、PhTourist1、Tmi1 和 PhTst-3) 进行聚类分析 (图 2B), 构建 ML 进化树。结果显示, PhTourist1、mPing 与 Ping 和 Pong 亲缘关系较近, 推测 PhTourist1、mPing 与 Ping 和 Pong 转座子可能来自同一祖先。而 PhTst-3、Tmi1 和 mGing 与前四者亲缘关系较远, 说明这 3 个 MITEs 家族与前四者 MITEs 起源的祖先不一样。

## 3.2 Stowaway-like 家族活性 MITEs 的发现与鉴定

### 3.2.1 dTstu1

Yang 等<sup>[24]</sup>通过将水稻中一个 Stowaway-like 的 MITE 序列和 Mariner-like 自主转座子转座酶编码序列扩增出来, 分别构建到表达载体中, 共同转化到酵母中, 检测发现该 MITE 在转座酶的作用下发生转座, 整合到酵母基因组中。然而这只是通过体外试验间接说明了 Stowaway 类的 MITEs 有转座活性, 并未在植物基因组中证实。次年, 研究人员在马铃薯 *Solanum tuberosum* L. 表皮颜色变化的研究中发现一个新的具有转座活性的 MITE, 其不同品种的马铃薯中存在插入多态性<sup>[44]</sup>, 将其命名为 dTstu1。经过分析发现, dTstu1 序列有 239 bp, 其中 A/T 含量是 67%, 两端 TIR 序列为 CTCCTCYGTC, TSD 序列与 Stowaway 类转座子的序列一致, 为 TA 序列<sup>[11]</sup>。这是第一次在植物体内发现具有转座活性的 Stowaway-like 家族的 MITEs。

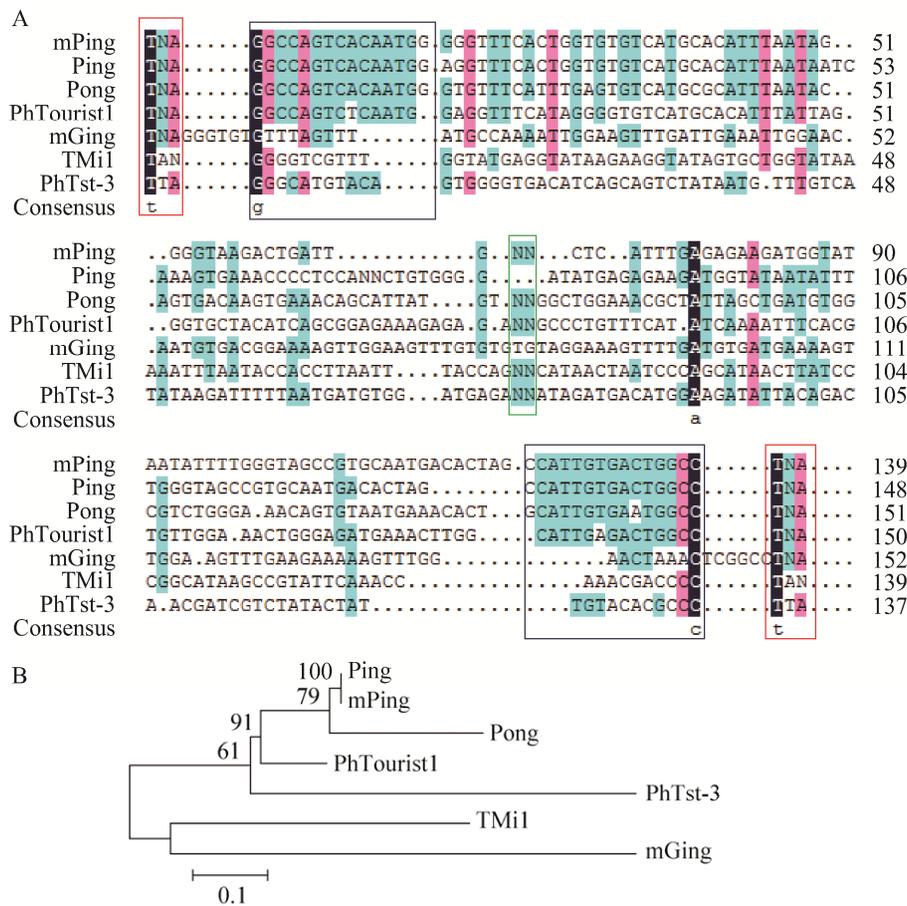


图 2 活性 Tourist-like MITEs 与 Ping 和 Pong 的进化关系. (A) 活性 Tourist-like MITEs (mPing、mGing、PhTourist1、Tmi1 和 PhTst-3) 与 Ping 和 Pong 的 TSD 与 TIR 比对图 (红框: TSDs; 蓝框: TIRs; 绿框: 省略的部分序列). (B) 活性 Tourist-like MITEs (mPing、mGing、PhTourist1、Tmi1 和 PhTst-3) 与 Ping 和 Pong 转座子的聚类分析

Fig. 2 Evolution of active Tourist-like MITEs, Ping and Pong. (A) The comparison of TSD and TIR of active Tourist-like MITEs (mPing, mGing, PhTourist1, Tmi1 and PhTst-3) with Ping and Pong (Red frame: TSDs; blue frame: TIRs; green frame: partially omitted sequence). (B) Clustering analysis of Tourist-like MITEs (mPing, mGing, PhTourist1, Tmi1 and PhTst-3), Ping and Pong.

随后,为了研究 dTstu1 对马铃薯表皮颜色变化的调控机制, Momose 等从三倍体栽培品种 72218 (块茎表皮为红色) 的叶子上获取原生质体,通过体外细胞培养获得其组培苗<sup>[45]</sup>。三倍体栽培的 72218 品种和组培苗 JKR 品种的基因组中存在 F3'5'H 基因,该基因的第三个外显子中有 dTstu1 插入,该 MITE 的插入导致一个终止子 (GTA) 编码,产生一个只有 24 个氨基酸的截短蛋白;而在组培苗 JKP 品种的该基因位置,

dTstu1 发生缺失,使得缺失 dTstu1 的 F3'5'H 基因编码一个含 510 个氨基酸的完整蛋白<sup>[46]</sup>。dTstu1 的跳跃导致 F3'5'H 基因合成不同的蛋白,致使后期块茎表皮细胞中二氢山奈酚中羟基基团数不同,合成的花青素也不一样:72218 品种和 JKR 品种的二氢山奈酚合成天竺葵色素,使块茎表皮表现为红色;而 JKP 品种块茎表皮细胞中二氢山奈酚比另两个品种中的多出两个羟基,合成牵牛花色素,表现为块茎紫色。整个研究过程表明双

子叶植物中存在着具有活性的 Stowaway 类转座子, 并且该 MITE 的插入和缺失会导致下游基因的表达, 改变表型。

### 3.2.2 MITE-39

周倩倩<sup>[32]</sup>通过 RepeatMasker (www.repeatmasker.org) 软件, 在毛竹基因组数据库中搜索出启动子中的 MITEs。根据搜索出的序列设计引物, 通过实验发现并验证, 在毛竹基因 PH01003704G0280 的启动子有一个 MITE 元件, 该 MITE 在不同野生实生苗中存在插入多态性。经分析发现, 该 MITE 全长 260 bp, TIR 为 CTCCTCCGTTCTTA AATA, TSD 序列为 TA, 属于 Stowaway-like MITE, 命名为 MITE-39。利用定量 PCR 研究该 MITE 在启动子中的作用, 发现当该 MITE 插入 PH01003704G0280 基因的上游启动子时, 会下调该基因的表达量; 而该 MITE 转座离开后, 该抑制会被解除, 基因表达水平得以恢复。由于实验中所用的实生苗为半同胞苗, 是同一母本但不一定是同一父本, 而实生苗生长环境均统一且适宜, 作者推测认为 MITE-39 的转座可能是因为来源的父本基因不同导致的。

## 3.3 其他 MITE 家族的活性转座子

### 3.3.1 Mutator 家族的 AhMITE1

2004 年, Patel 等<sup>[33]</sup>在花生 *Arachis hypogaea* L. 的脂肪酸合成研究中发现一类 MITEs 转座子, 发现当这个 MITE 插入编码合成脂肪酸的基因 (*ahFAD2B*) 中, *ahFAD2B* 基因发生突变, 不能正常合成脂肪酸。比对该家族 MITEs 序列发现其含有 25 bp 的 TIR 和 9 bp 的 TSD, 既不属于 Stowaway 家族也不属于 Tourist 家族, 而是属于 Mutator 家族。2012 年, Shirasawa 等<sup>[47]</sup>对几个花生品种中 AhMITE1 进一步研究, 首先从富含 MITEs 的 *A. hypogaea* 花生品种的基因组数据库中比对获取 504 条 AhMITE1, 大小在 201–223 bp 不等, 平均 GC 含量 30.1%。将 504 条 AhMITE1 的 5'和 3'端

的 9 bp 的 TSD 序列与已报道的花生突变体中的 AhMITE1 的 TSD 比较, 发现有 286 条 AhMITE1 的 TSD 与之前报道的完全匹配, 而其余 218 条出现 1 到 9 bp 不等的错配, 这 218 条序列有可能在 TSD、TIR 序列或转座子内部发生突变。

为了研究 AhMITE1 的活性, 取 4 种花生品种 Nakateyutaka、YI-0311、Satonoka 和 Kintoki 的种子。取一部分 Nakateyutaka 种子, 经过  $\gamma$  射线处理后发苗, 培育两代, 取第二代 M<sub>2</sub> 叶子进行 PCR 扩增, 与原 4 个花生品种的序列比对, 发现经过  $\gamma$  射线处理后, 至少在 AhTE0433、AhTE0426、AhTE0121 三个位点的 AhMITE1 具有活性, 发生转座, 并伴有一到两个碱基插入或突变。

## 4 总结与展望

30 年前, 在转座子发现初期, 转座子一直被认为是没有用的 DNA 片段<sup>[48]</sup>, 随着人们对转座子研究的不断深入, 发现转座子的插入、切除和扩增改变着基因组的结构和大小, 继而影响了基因的表达, 促使物种多样性的形成<sup>[49]</sup>。而 MITEs 转座子作为特殊的 DNA 转座子, 由于其拷贝数高, 插入多态性高, 而受研究者青睐。本文总结了近几年鉴定具有转座活性的 MITEs, 发现已鉴定的活性 MITEs 大部分属于 Tourist 家族。

目前已被报道具有转座活性的 MITEs 共有 8 个家族: 包括水稻中的 mPing、mGing 两个家族, 毛竹中的 PhTourist1 家族、PhTst-3 和 MITE-39 三个家族, 还有烟草中的 TMi1 家族、土豆中的 dTstu1 家族以及花生中的 AhMITE1 家族。本文汇集了近几年来这些被发现并鉴定具有转座活性的 MITEs, 总结了这些活性 MITEs 的 TIRs、TSDs、插入偏好性等特性, 以及这些活性 MITEs 转座后对附近基因的影响, 例如 mPing 影响水稻种子颖叶的长细, dTstu1 影响马铃薯表皮颜色调控基因, PhTst-3 插入基因启动子下调下游基因的表达等,

这些结果表明, MITEs 的转座很可能会调控附近植物基因表达, 甚至可能影响植物的表型, 对植物基因组结构、进化或植物物种进化等有重要作用。

毛竹和水稻同属于禾本科植物, 是我国东南部地区的重要经济竹种, 其基因组中含有丰富的 MITEs<sup>[43]</sup>, 在毛竹表观遗传学研究中具有重要的意义。我们通过研究毛竹基因组中的活性 MITEs, 分析其转座和扩增机制。在胡慧等<sup>[29]</sup>发现并鉴定具有转座活性的 PhTourist1 后, 陈昂<sup>[31]</sup>、周倩倩<sup>[32]</sup>先后分别发现了另外两个具有转座活性的 MITEs 家族 PhTst-3 和 MITE-39, 发现这两个 MITEs 在基因的启动子区发生转座跳跃, 比较转座前后该位置基因的表达, 发现这两类 MITEs 的存在抑制沉默了下游基因的表达, 而转座后该抑制被解除, 下游基因表达水平恢复。这些发现为后续更多毛竹活性 MITEs 的研究发现、鉴定以及功能分析等作了基础性的铺垫。

## REFERENCES

- [1] McClintock B. Chromosome organization and genic expression. Cold Spring Harb Symp Quant Biol, 1951, 16: 13–47.
- [2] Wicker T, Keller B. Genome-wide comparative analysis of *copia* retrotransposons in Triticeae, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual *copia* families. Genome Res, 2007, 17(7): 1072–1081.
- [3] Finnegan DJ. Eukaryotic transposable elements and genome evolution. Trends Genet, 1989, 5(4): 103–107.
- [4] Kidwell MG, Lisch D. Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci USA, 1997, 94(15): 7704–7711.
- [5] Zhang XY, Feschotte C, Zhang Q, et al. *P* instability factor: an active maize transposon system associated with the amplification of *Tourist*-like MITEs and a new superfamily of transposases. Pro Natl Acad Sci USA, 2001, 98(22): 12572–12577.
- [6] Jiang N, Feschotte C, Zhang XY, et al. Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). Curr Opin Plant Biol, 2004, 7(2): 115–119.
- [7] Piriyaopongsa J, Jordan IK. Dual coding of siRNAs and miRNAs by plant transposable elements. RNA, 2008, 14(5): 814–821.
- [8] Saito K, Siomi MC. Small RNA-mediated quiescence of transposable elements in animals. Develop Cell, 2010, 19(5): 687–697.
- [9] Blumenstiel JP. Evolutionary dynamics of transposable elements in a small RNA world. Trends Genet, 2011, 27(1): 23–31.
- [10] Yan YS, Zhang YM, Yang K, et al. Small RNAs from MITE-derived stem-loop precursors regulate abscisic acid signaling and abiotic stress responses in rice. Plant J, 2011, 65(5): 820–828.
- [11] Bureau TE, Wessler SR. Stowaway: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell, 1994, 6(6): 907–916.
- [12] Kuang HH, Padmanabhan C, Li F, et al. Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs. Genome Res, 2009, 19(1): 42–56.
- [13] Bureau TE, Wessler SR. Tourist: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell, 1992, 4(10): 1283–1294.
- [14] Nene V, Wortman JR, Lawson D, et al. Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science, 2007, 316(5832): 1718–1723.
- [15] Chen Y, Zhou FF, Li GJ, et al. MUST: a system for identification of miniature inverted-repeat transposable elements and applications to *Anabaena variabilis* and *Haloquadratum walsbyi*. Gene, 2009, 436(1/2): 1–7.
- [16] Han YJ, Wessler SR. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acids Res, 2010, 38(22): e199.
- [17] Yang GJ. MITE Digger, an efficient and accurate

- algorithm for genome wide discovery of miniature inverted repeat transposable elements. *BMC Bioinform*, 2013, 14(1): 186.
- [18] Ye CT, Ji GL, Liang C. *detectMITE*: a novel approach to detect miniature inverted repeat transposable elements in genomes. *Sci Rep*, 2016, 6: 19688.
- [19] Fattash I, Rooke R, Wong A, et al. Miniature inverted-repeat transposable elements: discovery, distribution, and activity. *Genome*, 2013, 56(9): 475–486.
- [20] Chen JJ, Hu Q, Zhang Y, et al. P-MITE: a database for plant miniature inverted-repeat transposable elements. *Nucleic Acids Res*, 2014, 42(D1): 1176–1181.
- [21] Oki N, Yano K, Okumoto Y, et al. A genome-wide view of miniature inverted-repeat transposable element (MITEs) in rice, *Oryza sativa* ssp. *japonica*. *Genes Genet Syst*, 2008, 83(4): 321–329.
- [22] Feschotte C, Jiang N, Wessler SR. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet*, 2002, 3(5): 329–341.
- [23] Feschotte C, Swamy L, Wessler SR. Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). *Genetics*, 2003, 163(2): 747–758.
- [24] Yang GJ, Nagel DH, Feschotte C, et al. Tuned for transposition: molecular determinants underlying the hyperactivity of a *Stowaway* MITE. *Science*, 2009, 325(5946): 1391–1394.
- [25] Nakazaki T, Okumoto Y, Horibata A, et al. Mobilization of a transposon in the rice genome. *Nature*, 2003, 421(6919): 170–172.
- [26] Kikuchi K, Terauchi K, Wada M, et al. The plant MITE *mPing* is mobilized in anther culture. *Nature*, 2003, 421(6919): 167–170.
- [27] Yang GJ, Zhang F, Hancock CN, et al. Transposition of the rice miniature inverted repeat transposable element *mPing* in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*, 2007, 104(26): 10962–10967.
- [28] Dong HT, Zhang L, Zheng KL, et al. A *Gaijin*-like miniature inverted repeat transposable element is mobilized in rice during cell differentiation. *BMC Genomics*, 2012, 13(1): 135.
- [29] Hu H, Zhou MB, Yang P, et al. Cloning and analysis of miniature inverted repeat transposable elements *PhTourist1* from *Phyllostachys edulis*. *Sci Silv Sin*, 2015, 51(5): 127–134 (in Chinese).  
胡慧, 周明兵, 杨萍, 等. 毛竹微型颠倒重复序列转座子 *PhTourist1* 的克隆与分析. *林业科学*, 2015, 51(5): 127–134.
- [30] Luo HM, Wang JJ, Ma WG, et al. One tourist-like miniature inverted repeat transposable element (MITE) is active mobilized in *nicotiana tabacum* genome. *J Agric Biotech*, 2015, 23(9): 1157–1166 (in Chinese).  
骆红梅, 王菁菁, 马文广, 等. 烟草基因组一类微小倒置重复转座元件(MITE)目前仍具转座活性. *农业生物技术学报*, 2015, 23(9): 1157–1166.
- [31] Chen A. Miniature inverted-repeat transposable elements (MITEs) identification and the effect on host gene expression in Bamboo[D]. Hangzhou: Zhejiang Agriculture and Forestry University, 2016 (in Chinese).  
陈昂. 毛竹微型反向重复转座子 (MITEs) 鉴定及对宿主基因表达的影响[D]. 杭州: 浙江农林大学, 2016.
- [32] Zhou QQ. Distribution and functions of the transposons in *Phyllostachys pubescens* (edulis) promoters[D]. Hangzhou: Zhejiang Agriculture and Forestry University, 2016 (in Chinese).  
周倩倩. 转座子在毛竹基因组启动子区域的分布及功能分析[D]. 杭州: 浙江农林大学, 2016.
- [33] Patel M, Jung S, Moore K, et al. High-oleate peanut mutants result from a MITE insertion into the *FAD2* gene. *Theor Appl Genet*, 2004, 108(8): 1492–1502.
- [34] Jiang N, Bao Z, Zhang X, et al. An active DNA transposon family in rice. *Nature*, 2003, 421(6919): 163–167.
- [35] Turcotte K, Srinivasan S, Bureau T. Survey of transposable elements from rice genomic sequences. *Plant J*, 2001, 25(2): 169–179.
- [36] Naito K, Cho E, Yang G, et al. Dramatic amplification of a rice transposable element during recent domestication. *Proc Natl Acad Sci USA*, 2006, 103(47): 17620–17625.
- [37] Naito K, Zhang F, Tsukiyama T, et al. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature*, 2009,

- 461(7267): 1130–1134.
- [38] Tsukiyama T, Teramoto S, Yasuda K, et al. Loss-of-function of a ubiquitin-related modifier promotes the mobilization of the active MITE *mPing*. *Mol Plant*, 2013, 6(3): 790–801.
- [39] Shan XH, Liu ZL, Dong ZY, et al. Mobilization of the active MITE transposons *mPing* and *Pong* in rice by introgression from wild rice (*Zizania latifolia* Griseb.). *Mol Biol Evol*, 2005, 22(4): 976–990.
- [40] Yasuda K, Ito M, Sugita T, et al. Utilization of transposable element *mPing* as a novel genetic tool for modification of the stress response in rice. *Mol Breed*, 2013, 32(3): 505–516.
- [41] Bureau TE, Ronald PC, Wessler SR. A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. *Proc Natl Acad Sci USA*, 1996, 93(16): 8524–8529.
- [42] Santiago N, Herráiz C, Goñi JR, et al. Genome-wide analysis of the emigrant family of MITEs of *Arabidopsis thaliana*. *Mol Biol Evol*, 2002, 19(12): 2285–2293.
- [43] Zhou MB, Tao GY, Pi PY, et al. Genome-wide characterization and evolution analysis of miniature inverted-repeat transposable elements (MITEs) in moso bamboo (*Phyllostachys heterocycla*). *Planta*, 2016, 244(4): 775–787.
- [44] Momose M, Abe Y, Ozeki Y. Miniature inverted-repeat transposable elements of stowaway are active in potato. *Genetics*, 2010, 186(1): 59–66.
- [45] Okamura M. Somaclonal variation and variety improvement: development of potato new variety ‘‘JAGA KIDS’’. *Tissue Cult*, 1991, 17: 207–212.
- [46] Jung CS, Griffiths HM, De Jong DM, et al. The potato *P* locus codes for flavonoid 3', 5'-hydroxylase. *Theor Appl Genet*, 2005, 110(2): 269–275.
- [47] Shirasawa K, Hirakawa H, Tabata S, et al. Characterization of active miniature inverted-repeat transposable elements in the peanut genome. *Theor Appl Genet*, 2012, 124(8): 1429–1438.
- [48] Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. *Nature*, 1980, 284(5757): 601–603.
- [49] Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*, 2007, 41: 331–368.

(本文责编 郝丽芳)