

Mutator 超家族转座子研究进展

从春生¹, 李玉斌^{1,2}

1. 中国农业科学院生物技术研究所, 北京 100081

2. 青岛农业大学农学院, 青岛 266109

摘要: 转座子是一类可以在基因组中不同遗传位点间移动的 DNA 序列, 在其转移过程中有时会伴随自身拷贝数的增加。作为基因组的重要组成部分, 转座子可以通过多种方式影响宿主基因及基因组的结构与功能, 进而在宿主的演化过程中扮演重要角色。目前依据转座过程中间体类型的不同可以将其分为 I 类转座子和 II 类转座子。*Mutator* 超家族转座子是 20 世纪 70 年代在玉米(*Zea mays* L.)中发现的一类特殊的转座子, 其属于 II 类转座子, 广泛存在于真核生物基因组中, 包含遗传特征明晰可分的众多转座子家族。此外, 该超家族转座子转座频率高, 倾向于插入基因富含区及低拷贝序列区, 可快速产生大量新的突变体, 目前已被广泛应用于正向及反向遗传学研究。本文结合近年来相关研究结果, 围绕 *Mutator* 超家族转座子的分类组成、结构特征、转座机制、插入偏好、靶位点重复序列以及玉米自主性 *MULEs* 元件展开综述, 并对转座子研究面临的问题及未来研究方向进行了探讨, 旨在与研究领域内的同行探讨相关研究的可能突破点、未来发展方向及可能产生的重大影响。

关键词: *Mutator* 超家族; *Mu* 家族; *MULEs* 元件; 转座机制; 插入偏好; 靶位点重复序列

Progress on *Mutator* superfamily

Chunsheng Cong¹, Yubin Li^{1,2}

1. Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China

2. College of Agronomy, Qingdao Agricultural University, Qingdao 266109, China

Abstract: Transposable elements (TEs) are fragments of DNA sequence, which can mobile from one locus to another within a genome, often replication in the process. Occupying the main component of the genome, TEs can affect the structure and function of gene and/or genome in a variety of ways, and play an important role in the evolution of the host. Based on the transposition intermediate, eukaryotic TEs can be divided into two classes. The *Mutator* superfamily is found in maize (*Zea mays* L.) in the 1970s. As the member of class II elements, *Mutator* superfamily transposons are found in all eukaryote genomes and contain many families with clearly distinguishable genetic characteristics. In addition, these TEs transpose at high rates and preferentially insert in gene-rich and low-repetitive genomic regions leading to the rapid generation of massive novel mutations, therefore, they are in great use of both forward and reverse genetics researches. In

收稿日期: 2019-09-27; 修回日期: 2019-12-17

基金项目: 国家自然科学基金面上项目(编号: 31871642)资助[Supported by the National Natural Science Foundation of China (No. 31871642)]

作者简介: 从春生, 博士研究生, 研究方向: 生物化学及分子生物学。E-mail: congchunsheng@126.com

通讯作者: 李玉斌, 研究员, 博士生导师, 研究方向: 生物化学及分子生物学。E-mail: liyubin@caas.cn

DOI: 10.16288/j.yczs.19-301

网络出版时间: 2020/1/2 18:26:20

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20191231.1146.001.html>

this review, we summarize the classification, structure characteristic, transposition mechanism, insertion preference and TSD sequence and other autonomous *MULEs* in maize. Moreover, we discuss the problems faced in TEs' research and research directions in the future, with a view to discuss possible breakthroughs, future development directions and significant impacts with colleagues in the related research field..

Keywords: *Mutator* superfamily; *Mu* family; *MULEs* elements; transposition mechanism; insertion preference; target site duplication

20 世纪 40 年代美国遗传学家 Barbara McClintock 在玉米中发现了一些在染色体上可以移动遗传位置的元件,并于 50 年代初提出可移动的遗传基因(即“跳跃基因”)学说,但直到 30 年后这一超越时代的学说才被科学界同行逐渐理解和接受,并将这些可以在基因组中移动的 DNA 序列统称为转座元件或转座子^[1]。现有研究表明,转座子几乎存在于所有生物基因组中,是基因组的主要组成部分。由于其重复特性,转座子曾一度被认为是垃圾 DNA,但越来越多的证据表明转座子是塑造基因组的重要因素^[2-5]。在许多植物基因组中,半数以上的序列属于转座子^[6-12],特别在玉米基因组中转座子序列的占比更高达 85%^[11],并且这些转座子序列通过不同方式影响了玉米的驯化^[13]、传播^[14,15]及优异农艺性状的形成^[16,17]等。此外,转座子在自身序列、蛋白功能域和结构等方面具有十分丰富的变异^[18]。按照转座过程中间体的类型,真核生物转座子可以划分为两大类:I 类转座子(RNA 转座子或反转录转座子)和 II 类转座子(DNA 转座子)^[19]。I 类转座子的转座反应通过 DNA-RNA-DNA 形式介导完成,通常以“copy- and-paste”方式进行转座,按照其长末端重复序列(long terminal repeats, LTR)的有无又可以分为 LTR 反转录转座子和 non-LTR 反转录转座子 2 个亚类;II 类转座子的转座反应通过 DNA-DNA 形式介导完成,绝大部分以“cut-and-paste”方式进行转座,少数转座子通过滚环模型或自我合成途径完成转座反应^[20,21]。根据序列组成和基因结构特征,I 类和 II 类转座子又都可分成不同的超家族^[22],各超家族间既存在共性,同时又各有特性,尽管采用的分析方法不尽相同^[23-25]。*Mutator* 超家族转座子属于 II 类转座子,广泛存在于真核生物基因组中,包含着遗传特征明晰可分的众多转座子家族,并且在转座

子遗传特性方面的研究深入,在功能组学研究中的应用也十分广泛。本文结合近年来的研究结果,围绕 *Mutator* 超家族转座子的分类组成、结构特征、转座机制、插入偏好、靶位点重复序列以及玉米自主性 *MULEs* 元件进行了概述,同时对转座子研究面临的问题及未来研究方向进行了探讨,以便相关科研人员更充分、全面了解 *Mutator* 超家族转座子的研究进展。

1 *Mutator* 超家族转座子分类组成

1978 年,美国爱荷华州立大学(Iowa State University)的 Donald Robertson 博士报道了一份高突变频率玉米材料,其幼苗中的突变频率接近自发突变的 30 倍左右,而这一突变特性的遗传不符合经典的孟德尔遗传定律^[26]并表现出明显的表观沉默^[27]。这一遗传品系由于存在大量一类新型转座元件—*Mutator* 或 *Mu*^[28-30],从而可以发生高频突变,也因此被称之为 *Mutator* 系^[31]。目前这些转座元件同属 *Mutator* 家族或 *Mu* 家族,其中可以编码转座酶并使其自身发生转座的元件称为自主性 *Mu* 转座子—*MuDR* (*Mutator*-Donald Robertson),而在活性 *MuDR* 存在时才能进行转座的元件统称为非自主性 *Mu* 转座子^[32-34]。*MuDR* 与大量的非自主性 *Mu* 转座子组合可形成高效的突变系统,*Mu* 系统是目前被广泛应用的致变能力极强的转座子插入突变体创制系统^[35-37]。

另外, minimal *Mutator* 系是通过筛选获得的只含有单一 *MuDR* 和一个位于颜色基因中的非自主性 *Mu* 的遗传品系,成为研究玉米 *Mu* 转座子系统及调控的理想材料^[38-40]。例如,从 minimal *Mutator* 系中发现了 *Mu killer* (*Muk*)^[41]。作为调控 *MuDR* 的显性遗传性位点 *Muk* 可以沉默一个或多个活性 *MuDR*,

但 *Muk* 并不是维持 *MuDR* 沉默状态所必需的, 在后代分离个体中, 即使缺失 *Muk* 位点, *MuDR* 仍无活性并且可以维持多代^[42]。*Muk* 的发现极大地促进了 *MuDR* 转座子表观沉默的研究, 同时使用于突变体创制的 *Mu* 系统变得更为可控。

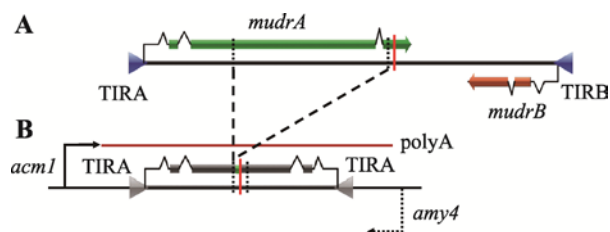
近年来, 伴随着测序技术的发展以及被测物种数量的不断增加, 在植物^[43~47]、真菌^[48,49]、原生动物^[50,51]以及多细胞动物^[52,53]中均发现了与玉米 *MuDR* 序列相类似的转座子, 统称为 *MULEs* 元件 (*Mutator*-like transposable elements)。目前大部分鉴定出来的 *MULEs* 元件都属于非自主性转座子, 它们自身不能编码功能完善的转座酶, 只有极少数 *MULEs* 元件可以进行自主转座, 例如尖孢镰刀菌 (*Fusarium oxysporum*) 中的 *Hop*^[49], 拟南芥 (*Arabidopsis thaliana*) 中的 *AtMu1*^[54], 玉米中的 *Jittery*^[55] 和 *TED*^[56], 水稻 (*Oryza sativa*) 中的 *Os3378*^[57] 以及埃及伊蚊 (*Aedes aegypti*) 中的 *Mutal*^[58] 等。另外, 大量非自主性 *MULEs* 元件内部有时携带着来源于宿主的一个或多个不同基因的片段, 这类元件被特别命名为 Pack-*MULEs*。目前在拟南芥、水稻、玉米、百脉根 (*Lotus japonicus*)、西红柿 (*Solanum lycopersicum*) 及荷花 (*Nelumbo nucifera Gaertn*) 的基因组中都发现了 Pack-*MULEs* 的存在^[59~65], 其中水稻中 Pack-*MULEs* 的数量巨大, 有关研究也更为深入。水稻中有些 Pack-*MULEs* 元件所携带的多个宿主基因片段可形成崭新的开放阅读框并转录出嵌合转录本。氨基酸序列功能分析及蛋白组学研究表明, 捕获的基因片段甚至可能具有特定的功能。结合以上研究结果及 Pack-*MULEs* 在植物中的普遍性, Jiang 等^[63] 推想 Pack-*MULEs* 获取基因片段的方式很可能是高等植株基因进化的一种重要机制。虽然 Pack-*MULEs* 捕获宿主基因组片段的分子机制目前仍不清楚, 但研究发现 Pack-*MULEs* 主要倾向于获得和保留 GC 含量高的序列, 这种选择性捕获使 Pack-*MULEs* 更有可能捕获具有功能性的序列, 进而为新基因的进化及现有基因的修饰提供新的遗传资源^[66~68]。与此同时, 相对于其他超家族转座元件, 水稻中 Pack-*MULEs* 表现出独特的表观遗传学特性, 其插入和表达不仅可以改变水稻染色体的表达模式, 还可以抵消重组对染色体碱基组成的影响, 进而对染色体结

构进化产生影响^[69]。

2 *Mutator* 超家族转座子及其转座酶基本特征

与其他大多数 DNA 超家族转座子相比, *Mutator* 超家族转座子两端具有较长的末端反向重复序列 (terminal inverted repeats, TIR)。TIR 序列中包含有转座酶结合位点^[70], 而携带单一 TIR 的转座子无法正常进行转座^[60,71]。此外, TIR 序列中还含有复杂的启动子序列, 既可以启动转座酶或 TIR 间序列的转录, 也可以调控转座酶在不同组织中的表达^[63,72]。玉米 *Mu* 家族转座子的 TIR 比较保守, 大多长约 215 bp^[73], 根据两端 TIR 间序列的差异, 又划分为不同亚家族 (*Mu1*~*Mu13*)^[74~76]。其中大部分为非自主性转座子, 这些非自主性 *Mu* 转座子是 *MuDR* 内部片段缺失产生的衍生物或者是其他序列点突变导致转座酶功能丧失的 *MuDR* 同源序列 (*MuDR* homologs, h*MuDR*)。h*MuDR* 虽然不能催化转座反应, 但可能在 *MuDR* 表观沉默中发挥增强作用^[31]。相对于玉米 *Mu* 家族转座子, 各种 *MULEs* 元件 TIR 序列变异丰富。有些 *MULEs* 元件 TIR 内含有串联重复序列, 这些串联重复序列可能导致 TIR 自身形成特殊的二级结构, 进而影响转座子的转座行为^[77]。在植物和真菌中大部分 *MULEs* 元件具有较长的 TIR (100~600 bp), 但在拟南芥^[65]、荷花^[61]、玉米^[78]和酵母 (*Yarrowia lipolytica*)^[48] 基因组中鉴定到少数 non-TIR *MULEs* 元件 (TIR<50 bp), 这些元件虽分布较为广泛, 但其与 *MULEs* 元件在进化中的关系仍不清楚。另外, 在玉米、西红柿、水稻和拟南芥基因组中还检测到一些多 TIR *MULEs* 元件, 这些 TIR 大多以串联形式分布, 多 TIR *MULEs* 元件可能更有利于转座子转座和捕获宿主基因组序列^[60]。

MuDR 作为 *Mu* 家族中的自主性转座子, 同时也是 *Mutator* 超家族转座子研究的典型代表。*MuDR* 编码两个转录方向相向的基因: *mudrA* 和 *mudrB*, 各自转录起始于两端的 TIR 内部序列, 两个转录本间没有重叠部分, 在相距 200 bp 处终止转录^[32] (图 1A)。*mudrA* 编码蛋白 MURA (94 kDa), MURA 与原核生物 *IS256* 转座子的转座酶序列相似^[79], 含有

图 1 *MuDR* 和 *Muk* 基因结构组成Fig. 1 The gene structures of *MuDR* and *Muk*

A: 转座子 *MuDR* 及其基因 *mudrA* 和 *mudrB* 的结构; B: *Muk* 及其两翼残存基因的结构。

保守的蛋白结构域^[80], 被认为是转座酶, 催化转座子转座。*mudrB* 编码蛋白 MURB (23 kDa), MURB 并非体细胞组织转座剪切所必需, 可能与生殖类细胞内转座子的重新插入相关^[81,82], MURB 调控方式及其在转座过程中的功能目前还没有更为详尽的报道。如前所述, 在玉米及其他植物、动物、微生物中也已经鉴定到了少数几个自主性 *MULEs* 元件, 但这些新发现的自主性 *MULEs* 元件均只含有 *mudrA* 同源基因, 因此 *mudrB* 基因可能仅存在于玉米的 *MuDR* 中。*Muk* 是 *MuDR* 转座插入并重排形成的 2.2 kb 反向重复序列, 由两段反向加倍的 TIRA 及其下游 *mudrA* 相邻部分序列组成, 不涉及任何 *mudrB* 基因序列(图 1B)。*Muk* 插入位点两翼残存的两个转录方向相向的基因(*acm1* 和 *amy4*)中仅有 *acm1* 启动子起始 *Muk* 转录, 由此产生的发卡状转录本生成小 RNA (主要是 22 nt siRNAs), 然后通过 RNA 介导的 DNA 甲基化方式沉默 *MuDR* 活性^[41,42]。

转座酶是自然界中最丰富、最普遍存在的基因

编码产物^[83]。所有真核生物“cut-and-paste”类型超家族转座子其转座酶均具有 DDE/D 三氨基酸特征结构域^[84], *MuDR* 及其他自主性 *MULEs* 元件转座酶同样具有这样的特征(图 2)。Liu 等^[58]根据保守的 DDE/D 结构域并通过生物信息学方法在埃及伊蚊基因组中发现了自主性 *MULEs* 元件—*Muta1*, 通过定点突变首次证实了 *MULEs* 元件转座酶中 DDE/D 结构域 3 个特征氨基酸的重要性: 其中任何单一氨基酸的改变都足以使转座酶的活性完全丧失。此外, 与另一类 DNA 转座子超家族—*hAT* 超家族的转座酶类似, 大部分 *Mutator* 超家族成员的转座酶在 DDE 结构域的第 2 个 D 和 E 之间还含有一个保守的 CXXH 基序和一个色氨酸^[85](图 2)。CXXH 基序可能参与转座酶对 TIR 的识别, 当 CXXH 基序中的组氨酸突变后, 转座酶催化活性消失殆尽^[58,86]。而色氨酸不仅与转座酶活性相关, 还与转座酶的精确切割或修复相关。当转座酶中色氨酸突变为丙氨酸时, 转座酶催化活性彻底消失; 当其突变为其他芳香族氨基酸时, 转座酶表现出一定活性, 但转座子剪切频率变低, 精确剪切比例也显著下降^[58,85]。

除了 DDE/D 三氨基酸这一特征结构域以外, 在 *Mutator* 超家族转座子的转座酶中还可以鉴定到其他保守结构域(图 3A)。例如, 大部分转座酶的 N 端具有属于 WRKY-GCM1 超家族^[87]的 DNA 结合结构域(DNA binding domain, DBD), 可能通过结合转座子特定区段序列来调控转座酶活性及转座子的转座行为。水稻 *Os3378* 编码的转座酶 DBD 上游特定长度编码序列发生缺失突变后, 转座子的剪切频率显

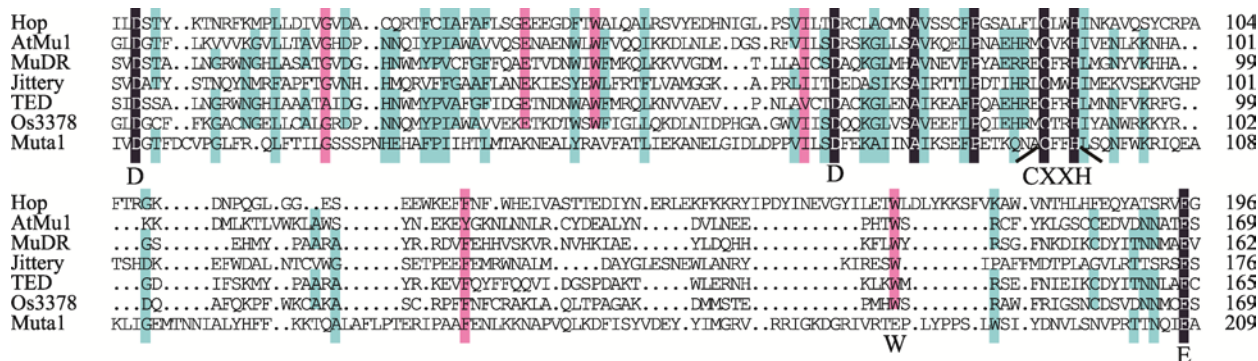


图 2 不同自主性转座酶 DDE 结构域蛋白序列比对分析

Fig. 2 The sequence alignment of DDE domain from different transposases

黑色阴影表示氨基酸完全一致, 粉色阴影表示同源性 $\geq 75\%$, 绿色阴影表示同源性 $\geq 50\%$, 保守的氨基酸及结构标注在序列底部。

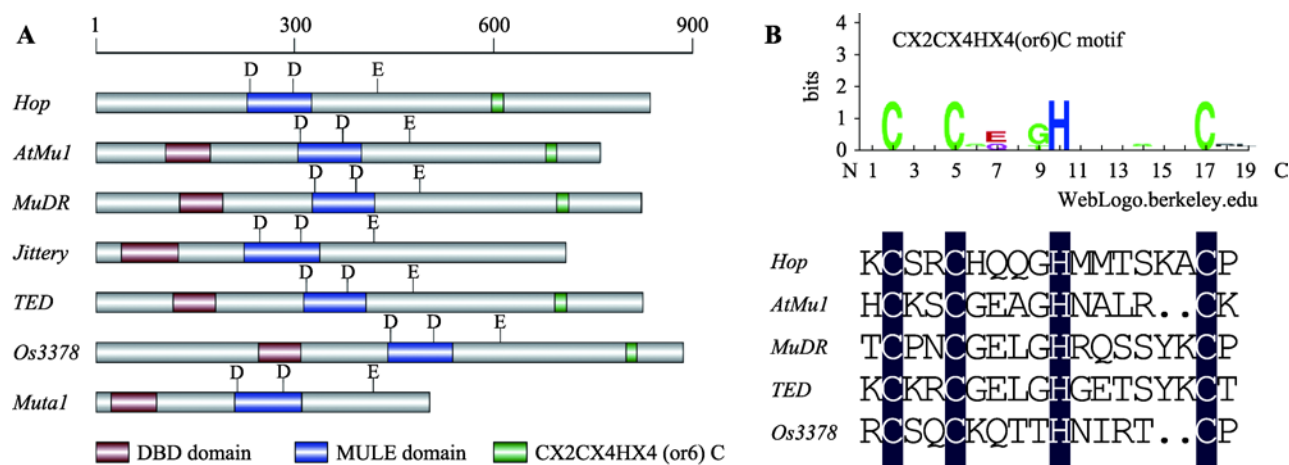


图 3 *Mutator* 超家族转座酶的结构特征

Fig. 3 The characteristic of transposases of *Mutator* superfamily

A: *MuDR* 及其他自主性 *MULEs* 元件转座酶的保守结构域; B: *Mutator* 超家族转座酶 C 端的保守基序。

著增高, 进一步研究发现这部分片段中氨基酸组合的理化特性对 *Os3378* 转座酶活性至关重要^[71]。另外, 大量 MURA 同源蛋白 C 端也具有相对保守的基序, 如在拟南芥、玉米、水稻和甘蔗(*Saccharum* spp.) 中先后鉴定到 CX2CX4HX4 (或 6) C 基序^[47,65], 目前已知的自主性 *Mutator* 超家族转座子的 C 端大多存在这些锌指基序(图 3B), 它们可能通过结合核酸序列(DNA 或 RNA)参与调控转座子的转座行为。

3 *Mutator* 超家族转座子转座机制

利用 *Mu* 家族转座子特有的遗传组成和转座特性, 已经构建了多个玉米突变体资源库(如 TUSC、MTM、RescueMu、UniformMu 和 ChinaMu 等), 为正向遗传学和反向遗传学的研究提供了丰富的突变体遗传材料^[88,89]。然而, 关于 *Mutator* 超家族转座子转座机制的认识仍缺乏直接的证据。基于与其他超家族 DNA 转座子的一些共性的大量转座事件分析, 推测 *Mutator* 超家族转座子的剪切及再次插入可能与某些已知的转座机制存在相似之处。

3.1 转座酶催化作用下的 DNA 双链断裂

真核生物 DNA 转座子的剪切过程一般以转座子两端某一条 DNA 单链的解离为起始, 该过程为亲核裂解反应, 通常 H_2O 作为亲核试剂, 在转座酶的

作用下攻击转座子与侧翼序列连接处的磷酸二酯键而形成断裂口。某些超家族转座子会在转座子末端暴露出自由的 3'-OH, 而其他超家族转座子则在侧翼宿主序列末端暴露出自由的 3'-OH, 随后, 不同类型超家族转座酶催化不同位置的 3'-OH 与不同类型 DNA 底物组合而使第二链断开, 由此形成的 DNA 双链断裂(DNA double strand break, DSB)使转座子最终得以从供体位点释放出来^[90]。最近研究发现, 埃及伊蚊 *Mut1* 第二链断开方式与 *hAT*、*transib* 超家族转座子相类似^[77]。在 *Mut1* 转座酶作用下, 以 H_2O 作为亲核试剂使转座子末端与侧翼 DNA 连接处的磷酸二酯键断开后在侧翼 DNA 的 3'末端暴露出羟基, 3'-OH 进攻另一链而在侧翼 DNA 末端形成发卡结构, 最后释放出转座子(图 4)。剪切位点形成的 DSB 既可以通过非同源末端连接(non-homologous end joining, NHEJ)方式修复, 留下不同类型转座印迹(footprint), 还可能通过同源重组(homologous recombination, HR)方式, 以一条姐妹染色单体或同系物作为模板进行修复。但是, 侧翼 DNA 形成的发卡结构必须在修复前打开, 相关的体外实验表明, 这一过程并不是由转座酶催化完成, 而可能是由宿主自身可以切割类似发卡结构的酶来完成^[86]。目前对于 *Mutator* 超家族转座酶催化作用下的 DNA 双链断裂过程报道较少, 该过程是否是 *Mutator* 超家族转座子的共同遗传特性亟待其他自主性 *Mutator* 超家族转座子相关研究加以验证。

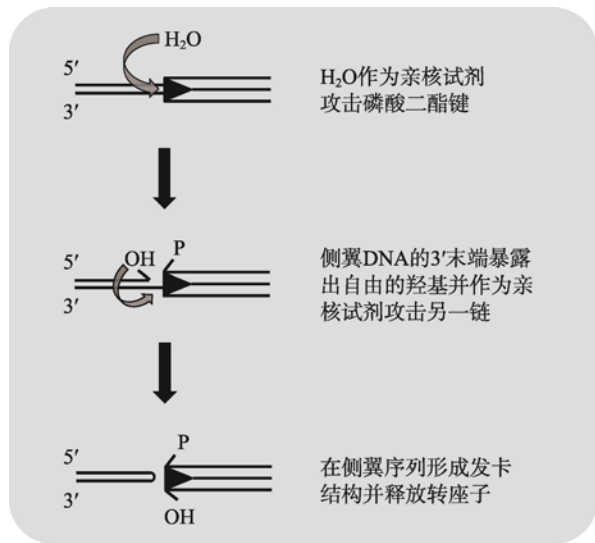


图 4 *Mut1* 转座酶介导的 DNA 双链断裂过程

Fig. 4 The progress of *Mut1* transposase-mediated DSB

3.2 转座子剪切后的 DSB 修复

转座子剪切后导致 DNA 发生双链断裂, 目前认为植物不同组织细胞中 DSB 修复方式不尽相同^[91,92] (图 5)。若剪切发生在生殖类细胞(包括配子体及配子体减数分裂前的有丝分裂细胞)的 S 期或 G₂ 期, 此时转座子已经随着染色体发生了复制, 细胞能够以姐妹染色单体为模板进行精确修复。由于剪切位点被完全修复, 转座反应看似以 “copy-and-paste”

方式进行, 但事实上是转座子剪切后又被重新修复的结果, 而并非转座子的简单加倍。在 DSB 的修复过程中, 由于模板内存在一些长短不一、散落分布的微同源序列, 修复复制链发生位置滑移, 便形成了大小有别、序列组成不同的多种缺陷型转座子。而在体细胞发育过程晚期, 转座子剪切可能发生在细胞 S 期之前或者剪切形成的 DSB 主要通过易错易突变的 NHEJ 方式进行修复, 结果导致转座子原插入位点产生多种类型的 footprint 序列^[93,94]。

目前已有多方面证据支持上述假说。在 *MuDR* 和 *TED* 生殖类细胞转座研究中均可以检测到缺陷型转座子, 并且这些缺陷型转座子在缺失序列的两翼存在微同源序列^[56,91]。此外, 研究还发现在一些缺陷型转座子序列内部含有填充序列(filler DNA), 并且这些 filler DNA 均来自缺失位点附近序列^[56,95]。这正是由于修复模板内存在多组微同源序列, 修复复制时发生了多次复制链滑移造成的。由此可见, 在生殖类细胞中转座子剪切后是通过依赖模板的方式进行修复的。另外, 研究发现在玉米 RAD51 突变体中生殖类细胞内 *Mu* 转座子转座行为异常。RAD51 在细胞减数分裂的 DSB 修复过程发挥重要作用, 玉米中存在两个 *rad51* 同源基因, 在含有活性 *MuDR* 的 RAD51 双突变材料中, 生殖细胞转座反应中 *MuDR* 内部及侧翼序列缺失的频率比野生型材料高

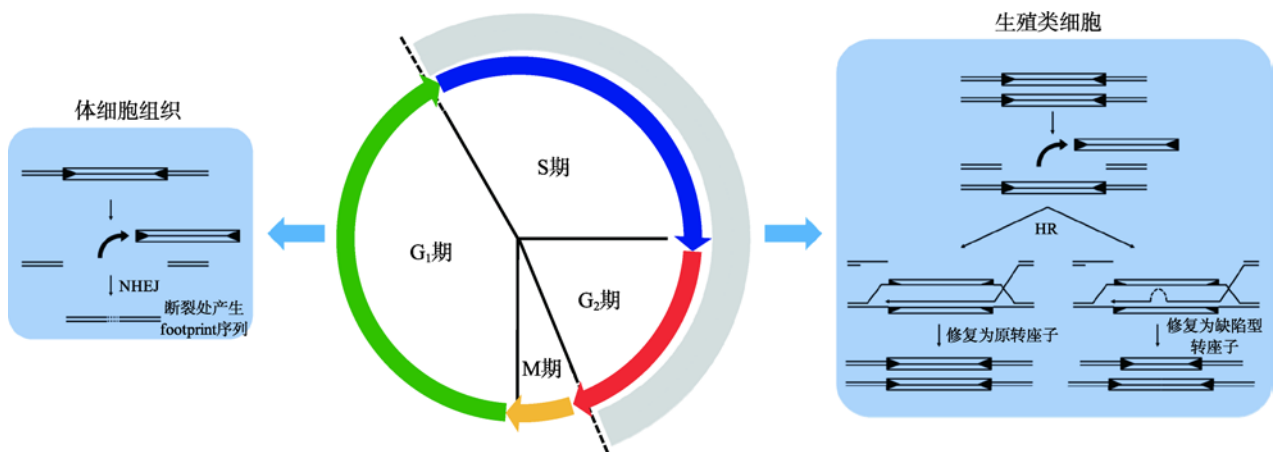


图 5 转座子剪切后的 DNA 双链断裂修复

Fig. 5 DSB repairing after transposon excision

中间圆形表示细胞周期; 左图表示转座子剪切发生在体细胞的 S 期前, 转座子剪切后形成的 DSB 通过 NHEJ 方式修复, 产生不同类型 footprint 序列; 右图表示转座子剪切发生在生殖类细胞的 S 期或 G₂ 期, 转座子剪切后形成的 DSB 通过 HR 方式以姐妹染色单体为模板进行修复, 在断裂处修复为原转座子或由于微同源序列导致修复链滑移, 修复为缺陷型转座子。

出数 10 倍,这也表明玉米生殖类细胞内 *Mu* 剪切后需要 RAD51 介导的 HR 进行修复^[96,97]。与生殖类细胞剪切修复相比,在体细胞组织转座过程中,转座子剪切后往往会形成多种类型 footprint 序列,大量 *Mu* 及 *TED* 体细胞组织转座事件研究已经证实了这一点^[56,93,94]。但体细胞组织内转座子剪切部位绝大多数不存在微同源序列,因此,体细胞组织内转座子剪切后更可能是通过非模板修复方式进行修复的。

3.3 染色体外环形结构

原核 IS256 转座元件通过闭合环形结构介导转座过程^[98],而早期研究发现,在携带活性 *MuDR* 的玉米材料中,*Mu* 转座子也能够以染色体外共价闭合的环形结构形式存在,并且这些环形 *Mu* 的出现依赖于活性 *MuDR* 的存在,因此,其可能是 MURA 作用下的转座中间体或是 *Mu* 剪切后的产物^[99]。但由于环形 *Mu* 序列信息的缺乏,难以明确其在转座反应中的作用及生物学意义。Li 等^[56]对玉米 *TED* 的研究发现,在含有活性 *TED* 的玉米体细胞中可以检测到共价闭合的环形 *TED* 或环形缺陷型 *TED* 结构;序列分析证实,这些环形结构的确是转座子两端共价连接的产物;同时酶切实验表明,所有检测的环形结构并不是两侧 TIR 末端完美的“头顶头(Head-to-Head)”共价连接,而可能是其他更为复杂的序列组成;除了预期大小的扩增产物,该研究还检测到一些其他扩增产物,克隆测序发现多数产物序列在连接点处缺失转座子单侧或两侧末端序列,缺失长度不等(<100 bp 至 >2 kb),有些涉及编码转座酶的区段。这些染色体外环形结构可能与转座子某些特性相关,例如转座子剪切后在非连锁位点的再次插入。目前,类似的转座子共价闭合环形结构在其他转座子研究中也有所报道^[54,100,101],但不同转座子的染色体外环形结构是否参与转座反应,它们如何发挥作用及其生物学意义仍不明确。

4 插入偏好

Mutator 超家族转座子在基因组内的转座并非随机插入而具有一定的偏好性。目前通过多个 *Mu*

突变系已经获得了数万份玉米 *Mu* 插入突变体,这些 *Mu* 插入遍布整个玉米基因组^[89,102,103],与 *Ac/Ds* 插入突变体不同,*Mu* 新插入位点与原初插入位点并不连锁。值得注意的是,虽然玉米基因组大部分为反转录转座子序列,但绝大部分 *Mu* 插入在基因组低拷贝区域的基因内部或基因附近^[103~105]。进一步研究发现,*Mu* 更倾向于插入基因的 5'末端,且插入区段序列 GC 含量较高,这与玉米及其他单子叶植物基因 5'末端 GC 含量略高相对应^[106]。另外,*Mu* 插入位点与开放染色质表观遗传标记(如 DNA 甲基化和组蛋白修饰)紧密相关^[107]。近期的研究也表明,真核生物转座子插入位点的选择受染色质结构影响^[108]。在玉米 W22 基因组中,*Ds* 插入位点与两翼序列染色质开放性无明显差异,而 *Mu* 插入位点染色质开放性显著增加。此外,*Mu* 和 *Ds* 都倾向于插入 CG 和 CHG 甲基化程度极低的区域,但 *Ds* 插入位点通常与 CG 和 CHG 高度甲基化区域相距较远,而 *Mu* 插入位点与这些高度甲基化区段距离较近^[109],这些特征有助于理解 *Mu* 更多插入在基因 UTR 区而 *Ds* 更倾向于插入基因编码区。

5 靶位点重复序列

不同超家族转座子插入基因组后会在其两侧形成一定长度的正向重复序列,这些序列来自插入位点,被称为靶位点重复序列(target site duplication, TSD)。通常同一家族的转座子重新插入后形成相同长度甚至固定组成的 TSD,因此,TSD 序列长度、固定的序列组成也是进行转座子分类的依据之一^[20,21]。*Mutator* 超家族转座子转座主要形成长度为 9 bp 的 TSD,并且这些 TSD 无明显的序列组成规律。近期研究表明,TSD 与 DNA 转座子的转座行为之间关系密切。例如,在异源酵母系统中研究水稻 *Os3378* 的转座遗传特征时发现,当改变一侧 TSD 中紧邻 *Os3378* 的前 3 个碱基后,转座子的剪切频率显著下降,对于较长的非自主性转座子这种影响更为明显。并且不一致的 TSD 还会影响非自主性转座子 *Os3378NA* 剪切位点的精确修复,而对转座子重新插入频率并无显著影响^[71]。另外,在异源酵母系统中,TSD 同样影响埃及伊蚊 *Mutal* 的转座行为。

当非自主性转座子 *Mutator* 携带有 8 bp 或 9 bp TSD 时, 剪切频率较无 TSD 情况下显著提高, 但与 *Os3378* 不同的是, 当携带 TSD 时, 相应转座子重新插入的频率也有所提高。此外, TSD 的缺失同样影响相应 *Mutator* 转座子剪切位点的精确修复。当携带有 8 bp 或 9 bp TSD 时, 90% 的回复突变均为精确剪切; 当无 TSD 时, 精确剪切频率仅占有回复突变的 10%。而对于 TSD 序列组成的研究表明, 不同的 TSD 序列组成对于转座子的剪切和重新插入均无显著影响^[58]。由此可见, TSD 序列的有无、一致性及其长度对 *MULEs* 元件转座行为的影响更大, 而 TSD 的序列组成对于转座行为的影响较小。在植物中 *Mutator* 超家族转座子的转座反应受到严格调控, 而异源酵母系统中开展的研究可能不足以完全涵盖和揭示 TSD 在植物转座过程中的作用和调控机制。

6 玉米自主性 *MULEs* 元件

除了 *MuDR* 以外, 目前玉米中还克隆了另外两个自主性 *MULEs* 元件: *Jittery*^[55] 和 *TED*^[56]。这两个转座子与 *MuDR* 间存在一些共性, 例如都具有 *mudrA* 同源序列, 含有较长的 TIR, 插入位点形成 9 bp TSD。而系统进化分析表明, 这两个转座子各自作为独立的自主性转座子已经存在了数百万年^[73]。与 *MuDR* 相比, *Jittery* 和 *TED* 共享某些特性: (1) 玉米中 *Jittery* 和 *TED* 的拷贝数很低; (2) 它们在生殖类细胞内发生回复突变的频率均高于 *MuDR*; (3) 两者都不含有 *mudrB* 同源序列^[55,56]。*Jittery* 更为特殊, 在生殖类细胞和体细胞组织内剪切后的修复都不留有任何 footprint 序列。此外, *Jittery* 的自主性略显欠缺, 目前只检测到玉米 *bronze* 位点的 *Jittery* 可以发生剪切, 并未检测到其重新整合到基因组中, 虽然不排除可能与缺少 *mudrB* 同源基因相关^[55], 但更可能是 *Jittery* 两端 TIR 序列微小差异影响了转座剪切后的再次插入。*Jittery* 3'端 TIR 比 5'端 TIR 在末端少了 4 个核苷酸(GCTC), 生物信息学分析发现, 在其他已测序的玉米材料中, *Jittery*-like 序列两侧 TIR 中均含有这 4 个核苷酸。因此, *Jittery* 很可能原本两侧 TIR 序列一致, 在转座到 *bronze* 位点过程中 3'端 TIR 发生了序列丢失进而影响到 *Jittery* 的转座反应, 导

致其剪切后不能重新插入到基因组中。另外, 其他已鉴定的自主性 *MULEs* 元件同样不含有 *mudrB* 同源基因, 但转座后均可以重新插入到基因组其他位点。因此, 即使 *mudrB* 基因确实与生殖类细胞内 *MuDR* 转座剪切后的重新插入相关, 目前鉴定到的这些 *MULEs* 元件可能在转座剪切后的重新插入方面进化出了不同的机制, 不再需要 *MURB* 功能蛋白。玉米中除了以上 3 个 *Mutator* 超家族转座子外, 遗传学实验还鉴定到另外几个自主性 *MULEs* 元件, 但这些转座子完整的基因组序列目前仍未被克隆, 如玉米中 *Mrh* 家族转座子^[110]。已有研究表明, 缺陷型 *Mrh* (*rMrh*) 的 TIR 与 *Jittery* TIR 的序列在前 50 bp 高度同源, 并且 *Jittery* 可以使 *rMrh* 发生移动。但关于自主性 *Mrh* 转座特性及其与 *Jittery* 间相互关系目前仍不清楚, 尚需进行深入研究。

7 结语与展望

转座子在真核生物基因和基因组的结构及进化过程中扮演着重要角色, 众多农作物在其驯化过程中优异农艺形状和优良品质的形成以及对生物胁迫和非生物胁迫的不断适应的遗传基础都与转座子引发的变异密不可分。*Mutator* 超家族转座子作为 II 类转座子研究的重要方面, 是转座子遗传学及功能基因组学的主要研究对象, 而转座子研究中仍有许多科学问题亟待解决, 转座子的开发应用更有待加强。因此, 继续深入转座子基础遗传学研究并不断开发利用转座子资源必将发挥重要的学术及应用价值。随着高通量测序、生物信息学分析及机器深度学习等新技术的发展, 转座子深入研究的成果势必更好地服务和推动生命科学的发展。

参考文献(References):

- [1] Ravindran S. Barbara McClintock and the discovery of jumping genes. *Proc Natl Acad Sci USA*, 2012, 109(50): 20198–20199. [DOI]
- [2] Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, Mager DL, Feschotte C. Ten things you should know about transposable elements. *Genome*

- Biol*, 2018, 19(1): 199. [DOI]
- [3] Huang CR, Burns KH, Boeke JD. Active transposition in genomes. *Annu Rev Genet*, 2012, 46: 651–675. [DOI]
- [4] Liu Z, Xu JH. The application of the high throughput sequencing technology in the transposable elements. *Hereditas(Beijing)*, 2015, 37(9): 885–898.
刘振, 徐建红. 高通量测序技术在转座子研究中的应用. *遗传*, 2015, 37(9): 885–898. [DOI]
- [5] Li SF, Li S, Deng CL, Lu LD, Gao WJ. Role of transposons in origin and evolution of plant XY sex chromosomes. *Hereditas(Beijing)*, 2015, 37(2): 157–164.
李书粉, 李莎, 邓传良, 卢龙斗, 高武军. 转座子在植物 XY 性染色体起源与演化过程中的作用. *遗传*, 2015, 37(2): 157–164. [DOI]
- [6] Li MM, Zhang DF, Gao Q, Luo YF, Zhang H, Ma B, Chen CH, Whibley A, Zhang YE, Cao YH, Li Q, Guo H, Li JH, Song YZ, Zhang Y, Copsey L, Li Y, Li XX, Qi M, Wang JW, Chen Y, Wang D, Zhao JY, Liu GC, Wu B, Yu LL, Xu CY, Li J, Zhao SC, Zhang YJ, Hu SN, Liang CZ, Yin Y, Coen E, Xue YB. Genome structure and evolution of *Antirrhinum majus* L. *Nat Plants*, 2019, 5(2): 174–183. [DOI]
- [7] Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, Bayer M, Ramsay L, Liu H, Haberer G, Zhang XQ, Zhang Q, Barrero RA, Li L, Taudien S, Groth M, Felder M, Hastie A, Šimková H, Staňková H, Vrána J, Chan S, Zhou M, Amatriain M, Ounit R, Wanamaker S, Bolser D, Colmsee C, Schmutz T, Aliyeva-Schnorr L, Grasso S, Tanskanen J, Chailyan A, Sampath D, Heavens D, Clissold L, Cao S, Chapman B, Dai F, Han Y, Li H, Li X, Lin C, Mccooke JK, Tan C, Wang P, Wang S, Yin S, Zhou G, Poland JA, Bellgard MI, Borisjuk L, Houben A, Dolezel J, Ayling S, Lonardi S, Kersey P, Langridge P, Muehlbauer GJ, Clark MD, Caccamo M, Schulman AH, Mayer FXK, Platzer M, Close TJ, Scholz U, Hansson M, Zhang G, Braumann I, Spannagl M, Li C, Waugh R, Stein N. A chromosome conformation capture ordered sequence of the barley genome. *Nature*, 2017, 544(7651): 427–433. [DOI]
- [8] Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, Wright J, Borrill P, Kettleborough G, Heavens D, Chapman H, Lipscombe J, Barker T, Lu FH, McKenzie N, Raats D, Ramirez-Gonzalez RH, Coince A, Peel N, Percival-Alwyn L, Duncan O, Trösch J, Yu G, Bolser DM, Namaati G, Kerhornou A, Spannagl M, Gundlach H, Haberer G, Davey RP, Fosker C, Palma FD, Phillips AL, Millar AH, Kersey PJ, Uauy C, Krasileva KV, Swarbreck D, Bevan MW, Clark MD. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res*, 2017, 27(5): 885–896. [DOI]
- [9] Li FG, Fan GY, Lu CR, Xiao GH, Zou CS, Kohel RJ, Ma ZY, Shang HH, Ma XF, Wu JY, Liang XM, Huang G, Percy RG, Liu K, Yang WH, Chen WB, Du XM, Shi CC, Yuan YL, Ye WW, Liu X, Zhang XY, Liu WQ, Wei HL, Wei SJ, Huang GD, Zhang XL, Zhu SJ, Zhang H, Sun FM, Wang XF, Liang J, Wang JH, He Q, Huang LH, Wang J, Cui JJ, Song GL, Wang KB, Xu X, Yu JZ, Zhu YX, Yu SX. Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat Biotechnol*, 2015, 33(5): 524–530. [DOI]
- [10] Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA. Genome sequence of the palaeopolyploid soybean. *Nature*, 2010, 463(7278): 178–183. [DOI]
- [11] Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higinbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M,

- Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, Mccann MC, Sanmiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK. The B73 maize genome: complexity, diversity, and dynamics. *Science*, 2009, 326(5956): 1112–1115. [DOI]
- [12] Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otiillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, Mccann MC, Ming R, Peterson DG, Mehboob-Ur-Rahman, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. The Sorghum bicolor genome and the diversification of grasses. *Nature*, 2009, 457(7229): 551–556. [DOI]
- [13] Studer A, Zhao Q, Ross-Ibarra J, Doebley J. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet*, 2011, 43(11): 1160–1163. [DOI]
- [14] Yang Q, Li Z, Li WQ, Ku LX, Wang C, Ye JR, Li K, Yang N, Li YP, Zhong T, Li JS, Chen YH, Yan JB, Yang XH, Xu ML. CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc Natl Acad Sci USA*, 2013, 110(42): 16969–16974. [DOI]
- [15] Castelletti S, Tuberosa R, Pindo M, Salvi S. A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL *Vgt1*. *G3 (Bethesda)*, 2014, 4(5): 805–812. [DOI]
- [16] Wang C, Yang Q, Wang WX, Li YP, Guo YL, Zhang DF, Ma XM, Song W, Zhao JR, Xu ML. A transposon-directed epigenetic change in *ZmCCT* underlies quantitative resistance to *Gibberella* stalk rot in maize. *New Phytol*, 2017, 215(4): 1503–1515. [DOI]
- [17] Zhang ZH, Zhang X, Lin ZL, Wang J, Liu HQ, Zhou LN, Zhong SY, Li Y, Zhu C, Lai JS, Li XR, Yu JM, Lin ZW. A large transposon insertion in the *stiff1* promoter increases stalk strength in maize. *Plant Cell*, 2019, DOI: 10.1105/tpc.19.00486. [DOI]
- [18] Kojima KK. Structural and sequence diversity of eukaryotic transposable elements. *Genes Genet Syst*, 2018, DOI: 10.1266/ggs.18–00024. [DOI]
- [19] Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet*, 1989, 5(4): 103–107. [DOI]
- [20] Kapitonov VV, Jurka J. A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet*, 2008, 9(5): 411–412, 414. [DOI]
- [21] Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, Sanmiguel P, Schulman AH. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*, 2007, 8(12): 973–982. [DOI]
- [22] Bao WD, Kojima KK, Kohany O. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA*, 2015, 6: 11. [DOI]
- [23] Xu HE, Zhang HH, Han MJ, Sheng YH, Huang XZ, Xiang ZH, Zhang Z. Computational approaches for identification and classification of transposable elements in eukaryotic genomes. *Hereditas(Beijing)*, 2012, 34(8): 1009–1019.
- 许红恩, 张化浩, 韩民锦, 沈以红, 黄先智, 向仲怀, 张泽. 真核生物转座子鉴定和分类计算方法. *遗传*, 2012, 34(8): 1009–1019. [DOI]
- [24] Hou XG, Zhang X, Guo DL. Identification and analysis methods of plant LTR retrotransposon sequences. *Hereditas (Beijing)*, 2012, 34(11): 1491–1500.
- 侯小改, 张曦, 郭大龙. 植物 LTR 类反转录转座子序列分析识别方法. *遗传*, 2012, 34(11): 1491–1500. [DOI]
- [25] Shen D, Chen C, Wang SS, Chen W, Gao B, Song CY. Research progress of Tc1/Mariner superfamily. *Hereditas (Beijing)*, 2017, 39(1): 1–13.
- 沈丹, 陈才, 王赛赛, 陈伟, 高波, 宋成义. Tc1/Mariner 转座子超家族的研究进展. *遗传*, 2017, 39(1): 1–13. [DOI]
- [26] Robertson DS. Characterization of a mutator system in maize. *Mutat Res*, 1978, 51(1): 21–28. [DOI]

- [27] Robertson DS. Genetic studies on the loss of *Mu* mutator activity in maize. *Genetics*, 1986, 113(3): 765–773. [DOI]
- [28] Strommer JN, Hake S, Bennetzen J, Taylor WC, Freeling M. Regulatory mutants of the maize *Adh1* gene caused by DNA insertions. *Nature*, 1982, 300(5892): 542–544. [DOI]
- [29] Bennetzen JL. Transposable element *Mu1* is found in multiple copies only in Robertson's Mutator maize lines. *J Mol Appl Genet*, 1984, 2(6): 519–524. [DOI]
- [30] Taylor LP, Walbot V. Isolation and characterization of a 1.7-kb transposable element from a mutator line of maize. *Genetics*, 1987, 117(2): 297–307. [DOI]
- [31] Lisch D, Jiang N. *Mutator* and *MULE* transposons. In: Bennetzen JL, Hake S, eds. *Handbook of Maize: Genetics and Genomics*. New York, NY: Springer New York, 2009, 277–306. [DOI]
- [32] Hershberger RJ, Warren CA, Walbot V. Mutator activity in maize correlates with the presence and expression of the *Mu* transposable element *Mu9*. *Proc Natl Acad Sci USA*, 1991, 88(22): 10198–10202. [DOI]
- [33] Chomet P, Lisch D, Hardeman KJ, Chandler VL, Freeling M. Identification of a regulatory transposon that controls the *Mutator* transposable element system in maize. *Genetics*, 1991, 129(1): 261–270. [DOI]
- [34] Qin MM, Robertson DS, Ellingboe AH. Cloning of the *Mutator* transposable element *MuA2*, a putative regulator of somatic mutability of the *a1-Mum2* allele in maize. *Genetics*, 1991, 129(3): 845–854. [DOI]
- [35] Mccarty DR, Settles AM, Suzuki M, Tan BC, Latshaw S, Porch T, Robin K, Baier J, Avigne W, Lai J, Messing J, Koch KE, Hannah LC. Steady-state transposon mutagenesis in inbred maize. *Plant J*, 2005, 44(1): 52–61. [DOI]
- [36] May BP, Liu H, Vollbrecht E, Senior L, Rabinowicz PD, Roh D, Pan XK, Stein L, Freeling M, Alexander D, Martienssen R. Maize-targeted mutagenesis: A knockout resource for maize. *Proc Natl Acad Sci USA*, 2003, 100(20): 11541–11546. [DOI]
- [37] Raizada MN. *RescueMu* protocols for maize functional genomics. *Methods Mol Biol*, 2003, 236: 37–58. [DOI]
- [38] Qian YX, Cheng X, Liu Y, Jiang HY, Zhu SW, Cheng BJ. Reactivation of a silenced minimal *Mutator* transposable element system following low-energy nitrogen ion implantation in maize. *Plant Cell Rep*, 2010, 29(12): 1365–1376. [DOI]
- [39] Lisch D, Chomet P, Freeling M. Genetic characterization of the *Mutator* system in maize: Behavior and regulation of *Mu* transposons in a minimal line. *Genetics*, 1995, 139(4): 1777–1796. [DOI]
- [40] Lisch D, Freeling M. Loss of *Mutator* activity in a minimal line. *Maydica*, 1994, 39(4): 289–300. [DOI]
- [41] Slotkin RK, Freeling M, Lisch D. *Mu killer* causes the heritable inactivation of the *Mutator* family of transposable elements in *Zea mays*. *Genetics*, 2003, 165(2): 781–797. [DOI]
- [42] Slotkin RK, Freeling M, Lisch D. Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nat Genet*, 2005, 37(6): 641–644. [DOI]
- [43] Stawujak K, Startek M, Gambin A, Grzebelus D. *MuTAnT*: A family of *Mutator*-like transposable elements targeting TA microsatellites in *Medicago truncatula*. *Genetica*, 2015, 143(4): 433–440. [DOI]
- [44] Yan L, Gu YH, Tao X, Lai XJ, Zhang YZ, Tan XM, Wang HY. Scanning of transposable elements and analyzing expression of transposase genes of sweet potato [*Ipomoea batatas*]. *PLoS One*, 2014, 9(3): e90895. [DOI]
- [45] Gbadegesin MA, Wills MA, Beeching JR. Diversity of LTR-retrotransposons and *enhancer/suppressor mutator*-like transposons in cassava (*Manihot esculenta* Crantz). *Mol Genet Genomics*, 2008, 280(4): 305–317. [DOI]
- [46] van Leeuwen H, Monfort A, Puigdomenech P. *Mutator*-like elements identified in melon, *Arabidopsis* and rice contain ULP1 protease domains. *Mol Genet Genomics*, 2007, 277(4): 357–364. [DOI]
- [47] Rossi M, Araujo PG, de Jesus EM, Varani AM, Van Sluys MA. Comparative analysis of *Mutator*-like transposases in sugarcane. *Mol Genet Genomics*, 2004, 272(2): 194–203. [DOI]
- [48] Neuveglise C, Chalvet F, Wincker P, Gaillardin C, Casaregola S. *Mutator*-like element in the yeast *Yarrowia lipolytica* displays multiple alternative splicings. *Eukaryot Cell*, 2005, 4(3): 615–624. [DOI]
- [49] Chalvet F, Grimaldi C, Kaper F, Langin T, Daboussi MJ. *Hop*, an active *Mutator*-like element in the genome of the fungus *Fusarium oxysporum*. *Mol Biol Evol*, 2003, 20(8): 1362–1375. [DOI]
- [50] Lopes FR, Silva JC, Benchimol M, Costa GG, Pereira GA, Carareto CM. The protist *Trichomonas vaginalis* harbors multiple lineages of transcriptionally active *Mutator*-like elements. *BMC Genomics*, 2009, 10: 330. [DOI]

- [51] Pritham EJ, Feschotte C, Wessler SR. Unexpected diversity and differential success of DNA transposons in four species of *Entamoeba* protozoans. *Mol Biol Evol*, 2005, 22(9): 1751–1763. [DOI]
- [52] Jacinto DS, Muniz Hdos S, Venancio TM, Wilson RA, Verjovski-Almeida S, Demarco R. Curupira-1 and Curupira-2, two novel *Mutator*-like DNA transposons from the genomes of human parasites *Schistosoma mansoni* and *Schistosoma japonicum*. *Parasitology*, 2011, 138(9): 1124–1133. [DOI]
- [53] Marquez CP, Pritham EJ. *Phantom*, a new subclass of *Mutator* DNA transposons found in insect viruses and widely distributed in animals. *Genetics*, 2010, 185(4): 1507–1517. [DOI]
- [54] Singer T, Yordan C, Martienssen RA. Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation (DDM1)*. *Genes Dev*, 2001, 15(5): 591–602. [DOI]
- [55] Xu NZ, Yan XH, Maurais S, Fu HH, O'Brien DG, Mottinger J, Dooner HK. *Jittery*, a *Mutator* distant relative with a paradoxical mobile behavior: excision without reinsertion. *Plant Cell*, 2004, 16(5): 1105–1114. [DOI]
- [56] Li YB, Harris L, Dooner HK. *TED*, an autonomous and rare maize transposon of the *Mutator* superfamily with a high gametophytic excision frequency. *Plant Cell*, 2013, 25(9): 3251–3265. [DOI]
- [57] Gao DY. Identification of an active *Mutator*-like element (MULE) in rice (*Oryza sativa*). *Mol Genet Genomics*, 2012, 287(3): 261–271. [DOI]
- [58] Liu K, Wessler SR. Functional characterization of the active *Mutator*-like transposable element, *Mut1* from the mosquito *aedes aegypti*. *Mob DNA*, 2017, 8: 1. [DOI]
- [59] Ming R, Vanburen R, Liu Y, Yang M, Han Y, Li LT, Zhang Q, Kim MJ, Schatz MC, Campbell M, Li J, Bowers JE, Tang H, Lyons E, Ferguson AA, Narzisi G, Nelson DR, Blaby-Haas CE, Gschwend AR, Jiao Y, Der JP, Zeng F, Han J, Min XJ, Hudson KA, Singh R, Grennan AK, Karpowicz SJ, Watling JR, Ito K, Robinson SA, Hudson ME, Yu Q, Mockler TC, Carroll A, Zheng Y, Sunkar R, Jia R, Chen N, Arro J, Wai CM, Wafula E, Spence A, Han Y, Xu L, Zhang J, Peery R, Haus MJ, Xiong W, Walsh JA, Wu J, Wang ML, Zhu YJ, Paull RE, Britt AB, Du C, Downie SR, Schuler MA, Michael TP, Long SP, Ort DR, Schopf JW, Gang DR, Jiang N, Yandell M, Depamphilis CW, Merchant SS, Paterson AH, Buchanan BB, Li S, Shen-Miller J. Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biol*, 2013, 14(5): R41. [DOI]
- [60] Ferguson AA, Jiang N. *Mutator*-like elements with multiple long terminal inverted repeats in plants. *Comp Funct Genomics*, 2012, 2012: 695827. [DOI]
- [61] Holligan D, Zhang XY, Jiang N, Pritham EJ, Wessler SR. The transposable element landscape of the model legume *Lotus japonicus*. *Genetics*, 2006, 174(4): 2215–2228. [DOI]
- [62] Hoen DR, Park KC, Elrouby N, Yu Z, Mohabir N, Cowan RK, Bureau TE. Transposon-mediated expansion and diversification of a family of *ULP*-like genes. *Mol Biol Evol*, 2006, 23(6): 1254–1268. [DOI]
- [63] Jiang N, Bao ZR, Zhang XY, Eddy SR, Wessler SR. Pack-MULE transposable elements mediate gene evolution in plants. *Nature*, 2004, 431(7008): 569–573. [DOI]
- [64] Lisch D. *Mutator* transposons. *Trends Plant Sci*, 2002, 7(11): 498–504. [DOI]
- [65] Yu Z, Wright SI, Bureau TE. *Mutator*-like elements in *Arabidopsis thaliana*: structure, diversity and evolution. *Genetics*, 2000, 156(4): 2019–2031. [DOI]
- [66] Wang J, Yu Y, Tao F, Zhang JW, Copetti D, Kudrna D, Talag J, Lee S, Wing RA, Fan CZ. DNA methylation changes facilitated evolution of genes derived from *Mutator*-like transposable elements. *Genome Biol*, 2016, 17(1): 92. [DOI]
- [67] Ferguson AA, Zhao D, Jiang N. Selective acquisition and retention of genomic sequences by Pack-*Mutator*-like elements based on guanine-cytosine content and the breadth of expression. *Plant Physiol*, 2013, 163(3): 1419–1432. [DOI]
- [68] Jiang N, Ferguson AA, Slotkin RK, Lisch D. Pack-*Mutator*-like transposable elements (Pack-MULEs) induce directional modification of genes through biased insertion and DNA acquisition. *Proc Natl Acad Sci USA*, 2011, 108(4): 1537–1542. [DOI]
- [69] Zhao D, Hamilton JP, Vaillancourt B, Zhang W, Eizenga GC, Cui Y, Jiang J, Buell CR, Jiang N. The unique epigenetic features of Pack-MULEs and their impact on chromosomal base composition and expression spectrum. *Nucleic Acids Res*, 2018, 46(5): 2380–2397. [DOI]
- [70] Benito MI, Walbot V. Characterization of the maize *Mutator* transposable element MURA transposase as a DNA-binding protein. *Mol Cell Biol*, 1997, 17(9): 5165–5175. [DOI]
- [71] Zhao D, Ferguson A, Jiang N. Transposition of a rice

- Mutator*-like element in the yeast *Saccharomyces cerevisiae*. *Plant Cell*, 2015, 27(1): 132–148. [DOI]
- [72] Raizada MN, Benito MI, Walbot V. The *MuDR* transposon terminal inverted repeat contains a complex plant promoter directing distinct somatic and germinal programs. *Plant J*, 2001, 25(1): 79–91. [DOI]
- [73] Lisch D. *Mutator* and *MULE* transposons. *Microbiol Spectr*, 2015, 3(2): A3–A32. [DOI]
- [74] Tan BC, Chen ZL, Shen Y, Zhang YF, Lai JS, Sun SSM. Identification of an active new *Mutator* transposable element in maize. *G3 (Bethesda)*, 2011, 1(4): 293–302. [DOI]
- [75] Dietrich CR, Cui F, Packila ML, Li J, Ashlock DA, Nikolau BJ, Schnable PS. Maize *Mu* transposons are targeted to the 5' untranslated region of the *gl8* gene and sequences flanking *Mu* target-site duplications exhibit nonrandom nucleotide composition throughout the genome. *Genetics*, 2002, 160(2): 697–716. [DOI]
- [76] Bennetzen JL, Springer P, Cresse AD, Hendrickx M. Specificity and regulation of the *Mutator* transposable element system in maize. *Crit Rev Plant Sci*, 1993, 12(1–2): 57. [DOI]
- [77] Liu K, Wessler SR. Transposition of *Mutator*-like transposable elements (MULEs) resembles *hAT* and *Transib* elements and V(D)J recombination. *Nucleic Acids Res*, 2017, 45(11): 6644–6655. [DOI]
- [78] Wang QH, Dooner HK. Remarkable variation in maize genome structure inferred from haplotype diversity at the *bz* locus. *Proc Natl Acad Sci USA*, 2006, 103(47): 17644–17649. [DOI]
- [79] Eisen JA, Benito MI, Walbot V. Sequence similarity of putative transposases links the maize *Mutator* autonomous element and a group of bacterial insertion sequences. *Nucleic Acids Res*, 1994, 22(13): 2634–2636. [DOI]
- [80] Hua-Van A, Capy P. Analysis of the DDE motif in the *Mutator* superfamily. *J Mol Evol*, 2008, 67(6): 670–681. [DOI]
- [81] Raizada MN, Walbot V. The late developmental pattern of *Mu* transposon excision is conferred by a cauliflower mosaic virus 35S-driven MURA cDNA in transgenic maize. *Plant Cell*, 2000, 12(1): 5–21. [DOI]
- [82] Lisch D, Girard L, Donlin M, Freeling M. Functional analysis of deletion derivatives of the maize transposon *MuDR* delineates roles for the MURA and MURB proteins. *Genetics*, 1999, 151(1): 331–341. [DOI]
- [83] Aziz RK, Breitbart M, Edwards RA. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res*, 2010, 38(13): 4207–4217. [DOI]
- [84] Yuan YW, Wessler SR. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. *Proc Natl Acad Sci USA*, 2011, 108(19): 7884–7889. [DOI]
- [85] Hickman AB, Ewis HE, Li X, Knapp JA, Laver T, Doss AL, Tolun G, Steven AC, Grishaev A, Bax A, Atkinson PW, Craig NL, Dyda F. Structural basis of *hAT* transposon end recognition by Hermes, an octameric DNA transposase from *Musca domestica*. *Cell*, 2014, 158(2): 353–367. [DOI]
- [86] Zhou L, Mitra R, Atkinson PW, Hickman AB, Dyda F, Craig NL. Transposition of *hAT* elements links transposable elements and V(D)J recombination. *Nature*, 2004, 432(7020): 995–1001. [DOI]
- [87] Babu MM, Iyer LM, Balaji S, Aravind L. The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Res*, 2006, 34(22): 6505–6520. [DOI]
- [88] Mccarty DR, Meeley RB. Transposon resources for forward and reverse genetics in maize. In: Bennetzen JL, Hake S, eds. *Handbook of Maize: Genetics and Genomics*. New York, NY: Springer New York, 2009, 561–584. [DOI]
- [89] Liang L, Zhou L, Tang Y, Li N, Song T, Shao W, Zhang Z, Cai P, Feng F, Ma Y, Yao D, Feng Y, Ma Z, Zhao H, Song R. A Sequence-Indexed mutator insertional library for maize functional genomics study. *Plant Physiol*, 2019, 181(4): 1404–1414. [DOI]
- [90] Hickman AB, Chandler M, Dyda F. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Crit Rev Biochem Mol Biol*, 2010, 45(1): 50–69. [DOI]
- [91] Hsia AP, Schnable PS. DNA sequence analyses support the role of interrupted gap repair in the origin of internal deletions of the maize transposon, *MuDR*. *Genetics*, 1996, 142(2): 603–618. [DOI]
- [92] Donlin MJ, Lisch D, Freeling M. Tissue-specific accumulation of MURB, a protein encoded by *MuDR*, the autonomous regulator of the *Mutator* transposable element family. *Plant Cell*, 1995, 7(12): 1989–2000. [DOI]
- [93] Britt AB, Walbot V. Germinal and somatic products of *MuI* excision from the *Bronze-1* gene of *Zea mays*. *Mol Gen Genet*, 1991, 227(2): 267–276. [DOI]
- [94] Doseff A, Martienssen R, Sundaresan V. Somatic excision of the *MuI* transposable element of maize. *Nucleic Acids Res*, 1991, 19(3): 579–584. [DOI]

- [95] Raizada MN, Nan GL, Walbot V. Somatic and germinal mobility of the *RescueMu* transposon in transgenic maize. *Plant Cell*, 2001, 13(7): 1587–1608. [DOI]
- [96] Li J, Harper LC, Golubovskaya I, Wang CR, Weber D, Meeley RB, McElver J, Bowen B, Cande WZ, Schnable PS. Functional analysis of maize RAD51 in meiosis and double-strand break repair. *Genetics*, 2007, 176(3): 1469–1482. [DOI]
- [97] Franklin AE, McElver J, Sunjevaric I, Rothstein R, Bowen B, Cande WZ. Three-dimensional microscopy of the Rad51 recombination protein during meiotic prophase. *Plant Cell*, 1999, 11(5): 809–824. [DOI]
- [98] Loessner I, Dietrich K, Ditttrich D, Hacker J, Ziebuhr W. Transposase-dependent formation of circular IS256 derivatives in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *J Bacteriol*, 2002, 184(17): 4709–4714. [DOI]
- [99] Sundaresan V, Freeling M. An extrachromosomal form of the *Mu* transposons of maize. *Proc Natl Acad Sci USA*, 1987, 84(14): 4924–4928. [DOI]
- [100] Gorbunova V, Levy AA. Analysis of extrachromosomal *Ac/Ds* transposable elements. *Genetics*, 2000, 155(1): 349–359. [DOI]
- [101] Gorbunova V, Levy AA. Circularized *Ac/Ds* transposons: formation, structure and fate. *Genetics*, 1997, 145(4): 1161–1169. [DOI]
- [102] Settles AM, Holding DR, Tan BC, Latshaw SP, Liu J, Suzuki M, Li L, O'Brien BA, Fajardo DS, Wroclawska E, Tseung CW, Lai J, Hunter CT, Avigne WT, Baier J, Messing J, Hannah LC, Koch KE, Becraft PW, Larkins BA, Mccarty DR. Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. *BMC Genomics*, 2007, 8: 116. [DOI]
- [103] Fernandes J, Dong QF, Schneider B, Morrow DJ, Nan GL, Brendel V, Walbot V. Genome-wide mutagenesis of *Zea mays* L. using *RescueMu* transposons. *Genome Biol*, 2004, 5(10): R82. [DOI]
- [104] Cresse AD, Hulbert SH, Brown WE, Lucas JR, Bennetzen JL. *MuI*-related transposable elements of maize preferentially insert into low copy number DNA. *Genetics*, 1995, 140(1): 315–324. [DOI]
- [105] Hanley S, Edwards D, Stevenson D, Haines S, Hegarty M, Schuch W, Edwards KJ. Identification of transposon-tagged genes by the random sequencing of *Mutator*-tagged DNA fragments from *Zea mays*. *Plant J*, 2000, 23(4): 557–566. [DOI]
- [106] Wong GK, Wang J, Tao L, Tan J, Zhang J, Passey DA, Yu J. Compositional gradients in *gramineae* genes. *Genome Res*, 2002, 12(6): 851–856. [DOI]
- [107] Liu S, Yeh CT, Ji T, Ying K, Wu H, Tang HM, Fu Y, Nettleton D, Schnable PS. *Mu* transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet*, 2009, 5(11): e1000733. [DOI]
- [108] Sultana T, Zamborlini A, Cristofari G, Lesage P. Integration site selection by retroviruses and transposable elements in eukaryotes. *Nat Rev Genet*, 2017, 18(5): 292–308. [DOI]
- [109] Springer NM, Anderson SN, Andorf CM, Ahern KR, Bai F, Barad O, Barbazuk WB, Bass HW, Baruch K, Ben-Zvi G, Buckler ES, Bukowski R, Campbell MS, Cannon E, Chomet P, Dawe RK, Davenport R, Dooner HK, Du LH, Du C, Easterling KA, Gault C, Guan JC, Hunter CT, Jander G, Jiao Y, Koch KE, Kol G, Köllner TG, Kudo T, Li Q, Lu F, Mayfield-Jones D, Mei W, Mccarty DR, Noshay JM, Portwood JN, Ronen G, Settles AM, Shem-Tov D, Shi J, Soifer I, Stein JC, Stitzer MC, Suzuki M, Vera DL, Vollbrecht E, Vrebalov JT, Ware D, Wei S, Wimalanathan K, Woodhouse MR, Xiong W, Brutnell TP. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat Genet*, 2018, 50(9): 1282–1288. [DOI]
- [110] Shepherd NS, Rhoades MM, Dempsey E. Genetic and molecular characterization of a-*mrh-Mrh*, a new mutable system of *Zea mays*. *Dev Genet*, 1989, 10(6): 507–519. [DOI]

(责任编辑: 严建兵)