

利用重测序技术获取转基因植物 T-DNA 插入位点

徐纪明, 胡晗, 毛文轩, 毛传澡

浙江大学生命科学院植物生物学研究所, 杭州 310058

摘要: T-DNA 插入位点的获得对于植物功能基因组学研究及转基因植物的筛选鉴定非常重要, 但是目前常用的方法如反向 PCR、半随机引物 PCR 等, 除了操作复杂、消耗时间长外, 特异性较差, 效率也很低。本研究利用全基因组重测序技术, 将 3 份转基因材料基因组 DNA 打包后进行重测序, 利用转基因载体序列作为参考序列进行比对分析, 得到 4 个 T-DNA 插入位点。对 3 份转基因材料进行 PCR 和 Southern blot 验证分析, 成功获得了 3 份转基因材料全部 T-DNA 插入位点, 其中 1 份材料为 2 拷贝插入。本文利用重测序技术建立了一种简单、可靠、高效的获取转基因植物 T-DNA 插入位点的方法, 以为植物功能基因组学及转基因研究奠定基础。

关键词: T-DNA 插入位点; 重测序技术; 转基因植物

Identifying T-DNA insertion site(s) of transgenic plants by whole-genome resequencing

Jiming Xu, Han Hu, Wenxuan Mao, Chuanzao Mao

Institute of Plant Biology, College of Life Science, Zhejiang University, Hangzhou 310058, China

Abstract: The availability of T-DNA insertion sites is very important for plant functional genomics research and the screening and identification of transgenic plants. However, the present protocols for identifying T-DNA insertion sites, like reverse PCR and semi-random primer PCR, are not only complex and time-consuming, but also inefficient. In this paper, a DNA pool of three transgenic plants was sequenced by whole-genome resequencing, and four T-DNA insertion sites were identified by blasting using transgenic T-DNA sequences. After PCR and Southern blot analysis, the T-DNA insertion sites of the three transgenic plants were successfully confirmed, and one of the transgenic plants showed two insertion sites. In conclusion, this study established a simple, reliable and efficient method for obtaining T-DNA insertion sites in transgenic plants.

Keywords: T-DNA insertion site; whole-genome resequencing; transgenic plant

收稿日期: 2018-04-02; 修回日期: 2018-06-11

基金项目: 国家转基因生物新品种培育科技重大专项(编号: 2016ZX08001003-009)和国家自然科学基金项目(编号: 31701983)资助[Supported by the National Transgenic Project (No.2016ZX08001003-009) and the National Natural Science Foundation of China (No. 31701983)]

作者简介: 徐纪明, 博士, 助理研究员, 研究方向: 作物磷高效分子机制。E-mail: xujiming@zju.edu.cn

通讯作者: 毛传澡, 博士, 教授, 博士生导师, 研究方向: 作物养分高效利用的生理及分子机制。E-mail: mcz@zju.edu.cn

DOI: 10.16288/j.ycz.18-080

网络出版时间: 2018/7/26 18:46:03

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20180726.1845.002.html>

T-DNA 插入位点是指转基因植物中载体 T-DNA 序列插入到染色体的位置。在植物功能基因组学研究中,明确 T-DNA 插入位点的信息对相关研究的顺利开展非常重要。不论是使用 T-DNA 激活标签群体还是 T-DNA 插入失活群体筛选材料,获得目标材料后,最关键的一步就是通过得到 T-DNA 插入位点明确目标基因;而在转基因生产应用中,T-DNA 插入位点是每份转基因新材料的标签,转基因材料的筛选鉴定、转基因材料的安全性评估、环境释放申请等都要求提供每份转基因材料的 T-DNA 插入位点。目前获取 T-DNA 插入位点的方法很多,根据原理可以分为 3 类:反向 PCR,外源接头介导 PCR,半随机引物 PCR (如 Tail-PCR)。另外,其他方法如质粒拯救法(plasmid rescue)、PCR-walking 法等,都已经被广泛用于插入位点的鉴定及侧翼序列的获取。但是,这些方法除了操作步骤复杂、消耗时间长等缺点外,还具有很大的不确定性,特异性比较差,无法获得很高的成功率^[1-5]。因此,建立一种简单、快速、成功率高的 T-DNA 插入位点获取方法非常必要。

近年来,随着测序技术的不断发展,全基因组重测序技术渐趋成熟。全基因组重测序的时间与成本大幅度下降,使得全基因组重测序的应用越来越广泛,除了用于测定不同物种全基因组序列、构建全基因组图谱外,还被用于突变体基因克隆等方面。Nordström 等^[6]使用全基因组重测序技术进行突变基因的克隆,不但缩短了图位克隆的时间,也使得很多杂交后表型难以分离的突变体基因得到克隆。Hu 等^[7]利用全基因组重测序对多个相同表型的突变体直接鉴定突变基因,省去杂交创制分离群体的步骤。为了探讨利用基因组重测序技术是否能够进行 T-DNA 插入位点的鉴定,本文对 3 份转基因材料基因组 DNA 打包后,利用全基因组重测序技术,并使用载体 T-DNA 序列作为模板进行比对分析,成功获

得了 3 份转基因材料的 T-DNA 插入位点。

1 材料和方法

1.1 转基因材料

3 份转基因材料分别为 PT8A-1、PT8A-2 和 PT8A-3,均为本实验室利用农杆菌介导的转基因方法,使用草甘膦作为抗性筛选标记得到的转基因株系,受体亲本为日本晴(*Oryza sativa* L. ssp. *Japonica* cv. Nipponbare, 本实验室保存)。转基因过程参照 Chen 等^[8]方法,转基因载体为 35S-1300-G9A-PT8A (图 1)。

1.2 DNA 提取

转基因材料 PT8A-1、PT8A-2、PT8A-3 和野生型品种日本晴种子萌发露白后,在正常水稻培养液^[9]中培养 30 d,取叶片用液氮磨碎后,参照陈昆松等^[10]方法,提取 3 份转基因材料基因组 DNA (DNA 总量至少 1 μg)。用含有溴化乙锭的 1% 琼脂糖凝胶电泳检测 DNA 的纯度和完整性,并利用 Nanodrop 测定 DNA 浓度。根据转基因载体中草甘膦抗性基因 G9A 序列,使用 Oligo7 软件设计引物,利用 PCR 方法确定 3 份转基因材料为转基因阳性株系。引物由华大基因公司合成,引物信息见表 1。

1.3 基因组重测序

将 3 份基因组 DNA 进行浓度测定后,等量混合。由于重测序需要至少 2 μg 基因组 DNA,因此 3 份转基因材料的 DNA 等量混合时要保证 DNA 总量 > 2 μg。利用非接触式超声破碎仪把 DNA 打断为 200~300 bp 片段,按照 Nextera DNA 文库制备试剂盒 (Illumina 公司,美国)说明书的方法进行建库,并使用 Illumina Hiseq2500 进行常规基因组重测序(由杭州谷禾信息技术有限公司完成),每个 Reads 长度

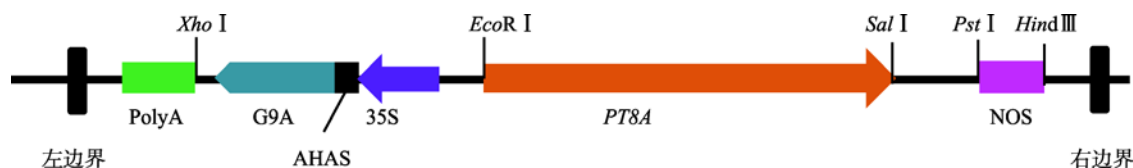


图 1 PT8A 转基因载体 35S-1300-G9A-PT8A 示意图

Fig. 1 Map of 35S-1300-G9A-PT8A vectors used for PT8A transgenes

150 bp, 得到至少 12 Gb 数据, 并保证数据质量指标 $Q30 \geq 80\%$ (即: 测序错误率大于 0.1% 的碱基所占的比例低于 20%)。

1.4 数据分析

根据基因组重测序结果, 利用 Bowtie2 软件以转基因载体 T-DNA 序列作为模板, 与测序得到的全部序列进行序列同源性比较筛选 (Bowtie2, <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>, 默认设置)。将筛选得到的序列进一步拼装筛选, 将序列全部为载体序列的 Reads 去掉, 最终得到一类 Reads 序列, 其特点是一半为基因组序列, 另一半为载体序列。根据得到的基因组序列, 在水稻基因组网站 (http://rice.plantbiology.msu.edu/standalone_blast.shtml) 进行 Blast 序列比对, 得到该基因组序列在基因组上的具体位置, 即为可能的插入位点。

1.5 PCR 验证

根据数据分析得到的 T-DNA 插入位点基因组 DNA 序列信息, 以及测序得到的载体序列信息, 使用 Oligo7 软件设计引物 (表 1), 上游引物设计在载体序列上, 下游序列设计在插入位点附近的基因组序列上, PCR 扩增产物长度预期在 250~500 bp 之间, 对 3 份转基因材料基因组 DNA 进行 PCR 验证, 确定每份转基因材料的插入位点是否正确。引物由华大基因公司合成。

表 1 引物序列信息

Table 1 Primers used in this study

引物名称	引物序列 (5'→3')
G9A	F: CCCATCTCGCGGAGCACGTT R: CGCCTCCCGATCTCCGTGTCC
Chr.2	F: ATGAACTGATTGATTCCGTCT R: TTTCATTTGGAGAGAACACGG
Chr.7	F: TAGAGGATCTACCATGGCCAC R: AGCTGCCTTTTGAATGTCAC
Chr.9	F: GGCTTCGCTCTTAAACCCCTC R: CAGTACTAAAATCCAGATCCCC
Chr.11	F: CTGCTCTAGCCAATACGCAAA R: ATGCCCATTTTAACCGGAT

1.6 Southern blot 验证

提取转基因材料 PT8A-1、PT8A-2、PT8A-3 和野生型品种日本晴的基因组 DNA^[10], 每份材料各取 20 μg , 使用 *Hind* III 单酶切后, 利用 G9A 引物扩增出的 DNA 片段作为探针, 按照 Zhou 等^[11]的方法进行 Southern blot 分析。

2 结果与分析

2.1 基因组 DNA 提取与阳性鉴定

分别提取转基因材料 PT8A-1、PT8A-2、PT8A-3 的基因组 DNA, 并进行电泳检测。3 份转基因材料提取的 DNA 结构完整, 没有发生降解 (图 2A)。对 3 份转基因材料进行草甘膦抗性基因 PCR 分析, 结果显示 3 份转基因材料均能扩增出条带大小与阳性对照一致的条带, 而野生型材料没有任何条带 (图 2B), 说明 PT8A-1、PT8A-2 和 PT8A-3 基因组中都含有草甘膦抗性基因, 为转基因阳性株系。

2.2 T-DNA 插入位点分析

对 3 份转基因材料基因组 DNA 等量混合后进行全基因组重测序, 共获得总碱基数约为 26 589 Mb 的数据, 84.40% 的数据可以比对到水稻基因组上, 其中有 10 倍以上覆盖的区域占水稻基因组的 89.24%, 1 倍以上覆盖的区域占水稻基因组的 97.01%, $Q30=84\%$, Mean Quality Score=36.16, 满足 $Q30 \geq 80\%$ 的高通量测序质量要求^[12]。以 35S-1300-G9A-PT8A 载体 T-DNA 序列作为参考模板, 使用 Bowtie2 软件 (默认设置) 对测序数据进行分析, 得到 4 个候选 T-DNA 插入位点, 分别为 2 号染色体 2 078 135 bp 处、7 号染色体 22 297 220 bp 处、9 号染色体 18 991 276 bp 处和 11 号染色体 12 123 278 bp 处 (图 3), 说明至少有一份转基因材料是双拷贝的。

2.3 插入位点 PCR 验证

为确定 3 份转基因材料的 T-DNA 插入位点, 分别用设计的 4 对引物对 3 份转基因材料基因组 DNA 进行 PCR 扩增。结果显示, PT8A-1 材料用 Chr.2 位点引物可以扩增出 PCR 产物, PT8A-2 材料用 Chr.9

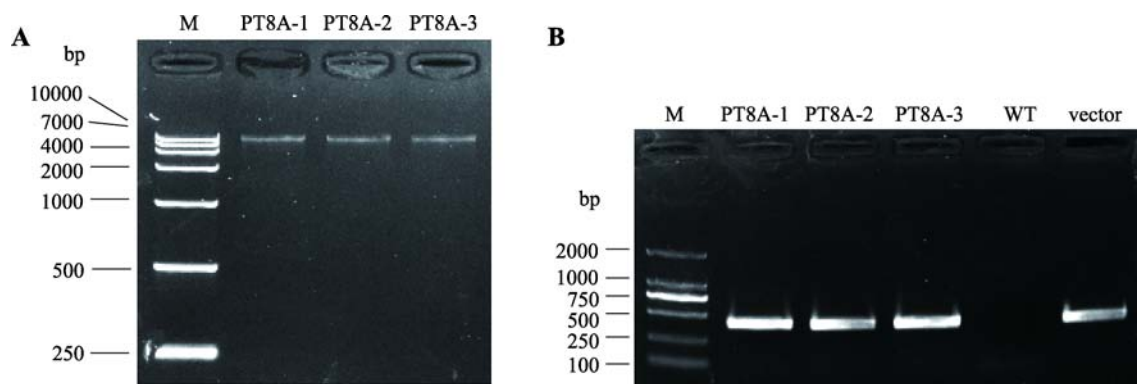


图 2 3 份转基因材料 DNA 和抗性基因 PCR 产物电泳检测结果

Fig. 2 Electrophoresis of the DNA extracted from three transgenic plants and PCR analysis of the resistant gene

A: 转基因株系 PT8A-1、PT8A-2、PT8A-3 基因组 DNA 电泳图。M: DL10000 marker。B: 转基因株系 PT8A-1、PT8A-2、PT8A-3 PCR 产物电泳图。M: DL2000 marker。WT: 野生型阴性对照; vector: 35S-1300-G9A-PT8A 阳性对照。

位点引物可以扩增出 PCR 产物, 而 PT8A-3 材料用 Chr.7 位点和 Chr.11 位点引物可以扩增出 PCR 产物, 而且所有 PCR 产物大小都与预测一致(图 4A), 说明 PT8A-1 材料在 7 号染色体有 T-DNA 插入, PT8A-2 材料在 9 号染色体有 T-DNA 插入, PT8A-3 材料在 2 号染色体和 11 号染色体有 T-DNA 插入, 为双拷贝插入。

2.4 Southern blot 验证分析

为了确定重测序分析结果是否得到了 3 份转基因株系全部插入位点, 本文对 3 份转基因材料进行 Southern blot 验证分析。结果表明, 野生型(WT)没有检测到任何杂交条带, PT8A-1、PT8A-2 材料基因组 DNA 各有一条杂交条带, 而 PT8A-3 材料基因组 DNA 有 2 条杂交条带(图 4B), 说明 PT8A-1、PT8A-2 材料为单拷贝插入, PT8A-3 材料为双拷贝插入。Southern blot 分析结果与测序分析结果一致, 说明 3 份转基因材料的基因组 DNA 打包后进行重测序分析, 能够得到全部的 T-DNA 插入位点信息。

3 讨 论

T-DNA 插入位点的获得在基因克隆、转基因安全等方面有着非常重要的应用。目前, 获得插入位点比较常用的方法是 Tail-PCR 方法, 由于 Tail-PCR 方法使用的是简并引物, 在实际操作过程中效率并不高。本课题组在使用 Tail-PCR 方法获得插入位点

时, 大多数转基因材料都不能顺利得到插入位点, 而且如果是多拷贝转基因株系, 成功率更低。而其他方法, 如反向 PCR、接头-PCR、SON-PCR 等^[2,13,14], 都需要复杂的操作, 成功率也很低。

本文通过提取 3 份转基因材料基因组 DNA, 等量混合后, 进行全基因组重测序, 使用载体序列作为参考模板进行分析筛选, 顺利得到了 3 份转基因材料的 T-DNA 插入位点, 而且其中有一份是双拷贝插入。本文利用重测序技术建立了一种简单、可靠、高效的获得转基因植物 T-DNA 插入位点的方法, 本方法有以下优点: (1) 操作方法简单。本方法中使用的 DNA 提取、PCR 分析等都是实验室常用的分子生物学技术, 而且随着基因组重测序技术的发展, 基因组重测序的成本下降, 时间加快, 数据质量进一步提高, 使得基因组重测序已经成为了大多数实验室常规操作; (2) 分析工作量小。目前在拟南芥 (*Arabidopsis thaliana*)、大豆 (*Glycine max*) 等转基因植株中有利用基因组重测序获得 T-DNA 插入位点的研究^[12, 15], 但是使用的是全基因组和载体序列同时进行比对, 分析工作量大、消耗时间长。本方法使用载体序列与重测序得到的序列直接进行分析比对, 不需要使用全基因组序列信息, 减少了基因组组装及分析比对的工作量, 减少了数据分析的时间, 而且在一些没有得到全基因组序列的物种上, 使用本方法也可以快速地获得插入位点信息。因此, 本方法相比于已有的 T-DNA 插入位点获得方法, 效率更高, 几乎所有分子生物学实验室都可以使用, 并

A: 2 号染色体插入位点; B: 7 号染色体插入位点; C: 9 号染色体插入位点; D: 11 号染色体插入位点。大写字母为基因组序列, 小写字母为载体序列。

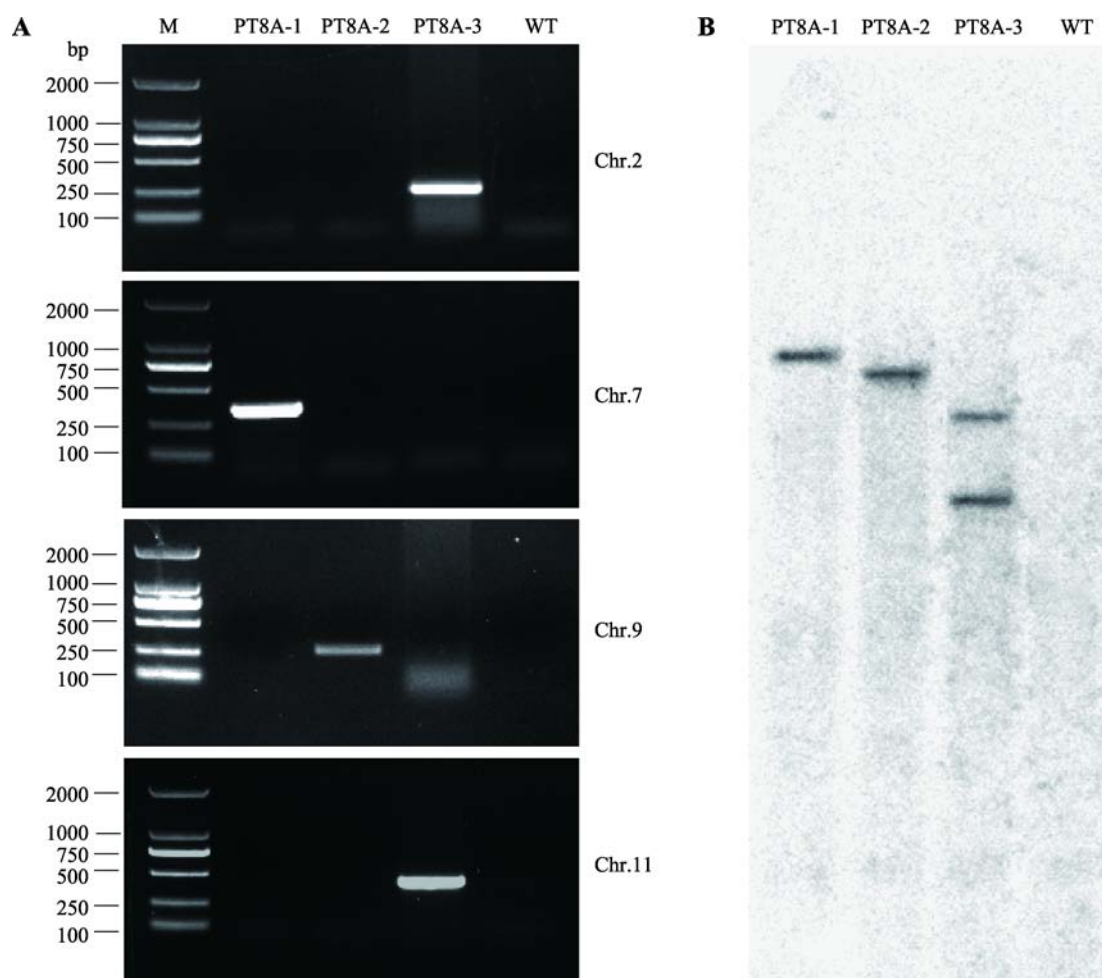


图 4 转基因株系 T-DNA 插入位点验证分析

Fig. 4 Verification of T-DNA insertion sites for transgenic plants

A: PCR 验证分析结果; B: Southern blot 验证分析结果。M: DL2000 marker。

因材料混合进行全基因组重测序, 由于测序深度不够, 导致有 1 份材料未得到插入位点信息。本文将 3 份转基因材料的基因组 DNA 混合后进行重测序, 共得到大约 25 Gb 的数据, 经分析, 3 份转基因材料都得到了准确的插入位点。因此, 多份转基因材料等比例混合后测序, 在建库均匀和保证测序深度的情况下都可以得到插入位点。Polko 等^[15]认为多至 20 份材料混合后也可以成功得到插入位点信息。但是, 如果混合的材料过多, 会不会由于目前方法的局限性, 导致建库不均匀, 进而导致得不到插入位点或者多拷贝株系的插入位点信息不完整, 还需要进一步的研究去验证。

综上所述, 本文提供了一种简单、可靠、高效的获得转基因插入位点的方法, 利用该方法可以快

速得到多份转基因株系的插入位点, 而且不受插入拷贝数的影响。

参考文献(References):

- [1] Pu YB, Guo C, Ping SZ. Application progress on amplified methods to clone flanking sequence. *Biotech Bull*, 2016, 32(11): 72-79.
蒲远波, 郭翠, 平淑珍. 基因侧翼序列扩增技术的应用进展. *生物技术通报*, 2016, 32(11): 72-79. [\[DOI\]](#)
- [2] Chen L, Tu ZM, Hussain J, Cong L, Yan YJ, Jin L, Yang GX, He GY. Isolation and heterologous transformation analysis of a pollen-specific promoter from wheat (*Triticum aestivum* L.). *Mol Biol Rep*, 2010, 37(2): 737-744. [\[DOI\]](#)

- [3] Tsaftaris A, Pasentzis K, Argiriou A. Rolling circle amplification of genomic templates for inverse PCR (RCA-GIP): A method for 5'- and 3'-genome walking without anchoring. *Biotechnol Lett*, 2010, 32(1): 157. [DOI]
- [4] Shang CH, Shi L, Ren A, Qin L, Zhao MW. Molecular cloning, characterization, and differential expression of a lanosterol synthase gene from *Ganoderma lucidum*. *Biosci Biotechnol Biochem*, 2010, 74(5): 974–978. [DOI]
- [5] Wang Z, Ye S, Li J, Zheng B, Bao M, Ning G. Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning. *BMC Biotechnol*, 2011, 11: 109. [DOI]
- [6] Nordström KJ, Albani MC, James GV, Gutjahr C, Hartwig B, Turck F, Paszkowski U, Coupland G, Schneeberger K. Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *Nat Biotechnol*, 2013, 31(4): 325–330. [DOI]
- [7] Hu H, Wang WT, Zhu ZX, Zhu JH, Tan DY, Zhou ZP, Mao CZ, Chen X. GIPS: A software guide to sequencing-based direct gene cloning in forward genetics studies. *Plant Physiol*, 2016, 170(4): 1929–1934. [DOI]
- [8] Chen SY, Jin WZ, Wang MY, Zhang F, Zhou J, Jia QJ, Wu YR, Liu FY, Wu P. Distribution and characterization of over 1000 T-DNA tags in rice genome. *Plant J*, 2003, 36(1): 105–113. [DOI]
- [9] Yoshida S, Forno DA, Cock JH, Gomez KA. Laboratory Manual for Physiological Studies of Rice, Ed 3. The International Rice Research Institute: The Philippines, 1976. [DOI]
- [10] Chen KS, Li F, Xu CJ, Zhang SL, Fu CX. An efficient macro-method of genomic DNA isolation from *Actinidia chinensis* leaves. *Hereditas(Beijing)*, 2004, 26(4): 529–531.
陈昆松, 李方, 徐昌杰, 张上隆, 傅承新. 改良 CTAB 法用于多年生植物组织基因组 DNA 的大量提取. *遗传*, 2004, 26(4): 529–531. [DOI]
- [11] Zhou J, Jiao FC, Wu ZC, Li YY, Wang XM, He XW, Zhong WQ, Wu P. OsPHR2 is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants. *Plant Physiol*, 2008, 146(4): 1673–1686. [DOI]
- [12] Guo BF, Guo Y, Hong HL, Qiu LJ. Identification of genomic insertion and flanking sequence of *G2-EPSPS* and *GAT* transgenes in soybean using whole genome sequencing method. *Front Plant Sci*, 2016, 7: 1009. [DOI]
- [13] Antal Z, Rascle C, Fèvre M, Bruel C. Single oligonucleotide nested PCR: A rapid method for the isolation of genes and their flanking regions from expressed sequence tags. *Curr Genet*, 2004, 46(4): 240–246. [DOI]
- [14] Guo H, Xiong J. A specific and versatile genome walking technique. *Gene*, 2006, 381: 18–23. [DOI]
- [15] Polko JK, Temanni MR, Zanten M, Workum W, Iburg S, Pierika R, Voeseneka LA, Peeters AJ. Illumina sequencing technology as a method of identifying T-DNA insertion loci in activation-tagged *Arabidopsis thaliana* plants. *Mol Plant*, 2012, 5(4): 948–950. [DOI]

(责任编辑: 吴为人)