



抗病毒转基因大豆事件外源 T-DNA 序列分析及定性检测

刘东波^{1,2}, 邢国杰², 赵倩倩², 杨 静², 牛 陆², 杨向东², 仲晓芳²

(1. 吉林师范大学 生命科学学院, 吉林 四平 136000; 2. 吉林省农业科学院 农业生物技术研究所/吉林省农业生物技术重点实验室, 吉林 长春 130033)

摘 要:转基因作物的 T-DNA 插入及其与宿主基因组的连接序列构成的侧翼区域,是转基因事件检测的关键组成部分,也是转基因作物开发、评估和监管所必需的。前期研究得到 2 个携带大豆花叶病毒 *P3* (soybean mosaic virus *P3*) 基因干扰片段的转基因大豆事件,即 L13 和 L104,可显著提高大豆对大豆花叶病毒、大豆花叶坏死病毒和西瓜花叶病毒的抗性,具有较强的光谱性。为了便于对该转基因事件进行监管以及建立事件特异性检测方法,本研究通过 Illumina Xten 平台,对 L13 和 L104 进行全基因组测序 (whole genome sequencing, WGS),并结合生物信息学分析和 PCR 技术,对转基因事件 L13 和 L104 的旁侧序列进行了分析。通过全基因组测序,每一个事件获得了超过 12.13 Gb 的原始数据,测序深度为 11×,与大豆参考基因组 (Wm82. a2. v1) 和转化载体序列做比较,初步识别出这两个转化事件 T-DNA 的边界及其旁侧序列。转化事件 L13 和 L104 的 T-DNA 分别位于 Chr04 和 Chr11 染色体上。进一步的 PCR 检测和序列测定表明转基因事件 L13 的 T-DNA 为单位点双拷贝整合到大豆基因组 Chr04:46747946 位点;转基因事件 L104 的 T-DNA 为单拷贝插入,整合位点为 Chr11:30461895 位点。结果证明了全基因组测序在识别转基因作物 T-DNA 插入和旁侧序列区域方面的有效性和稳定性。此外,插入位点的鉴定和事件特异性检测的建立将有助于广谱抗病毒转基因大豆品种的应用和发展。

关键词:插入位点;旁侧序列;全基因组测序;抗病毒转基因大豆;时间特异性检测

Efficient Identification of T-DNA Insertions and Qualitative Analysis in Two Virus-resistant Transgenic Soybean Events

LIU Dong-bo^{1,2}, XING Guo-jie², ZHAO Qian-qian², YANG Jing², NIU Lu², YANG Xiang-dong², ZHONG Xiao-fang²

(1. School of Life Science, Jilin Normal University, Siping 136000, China; 2. Agricultural Biotechnology Institute, Jilin Academy of Agricultural Sciences/Jilin Provincial Key Laboratory of Agricultural Biotechnology, Changchun 130033, China)

Abstract: Genomic insertions and flanking regions of transgenes in host genomes constitute a critical component of precise molecular characterization and event-specific detection, which are required in development and assessment for regulatory approval of genetically modified (GM) crops. Previously, we reported two transgenic soybean events, that were L13 and L104, harboring the reverted repeats of soybean mosaic virus *P3* gene, exhibiting significantly enhanced and broad-spectrum resistance to soybean mosaic virus, soybean mosaic necrosis virus, watermelon mosaic virus. To facilitate safety assessment and event-specific detection establishment, we identified the transgenic insertion sites and flanking sequences of the event L13 and L104 using whole genome sequencing (WGS) through Illumina Xten platform combined with bioinformatics analysis and PCR method. More than 12.13 Gb sequence data (11 coverage) were generated of each event through WGS. The sequence reads corresponding to boundaries of inserted T-DNA and associated native flanking sequences were identified by comparison with soybean reference genome (Wm82. a2. v1) and sequence of the transformation vector. Inserted T-DNAs were mapped on Chromosome 04 and 07 for the events L13 and L104, respectively. The putative insertion loci and flanking regions were further confirmed by PCR amplification and sequencing. The T-DNAs of the transgenic event L13 were two copies inserted in Chr04:46747946 site. And the T-DNA of the L104 was inserted in Chr11:30461895 with one copy. Based on the flanking sequences of the inserted T-DNA, the event-specific detection for each event was established with specific PCR primers. Our results demonstrate the efficacy and robustness of whole genome sequencing in identifying the T-DNA insertions and flanking regions of GM crops. Moreover, characterization of insertion loci and establishment of event-specific detection will facilitate the application and development of broad-spectrum virus-resistant transgenic soybean cultivars.

Keywords: Insertion site; Flanking sequence; Whole genome sequencing; Virus-resistant transgenic soybean; Time-specific detection

收稿日期:2019-06-19

基金项目:国家转基因生物新品种培育科技重大专项(2016ZX08004-004);国家自然科学基金(31701448)。

第一作者简介:刘东波(1993-),男,硕士,主要从事大豆生物技术育种研究。E-mail:87144326@qq.com。

通讯作者:仲晓芳(1978-),女,博士,副研究员,主要从事大豆生物技术育种研究。E-mail:xfzhong649@163.com。

转基因技术为农业发展提供了巨大的机遇,许多转基因作物已经商业化,同时大量转基因植物正在研发中。目前,对转基因作物的评估需要对插入 T-DNA 进行精确的分子鉴定和完整性描述。此外,在世代稳定性和特定事件检测以及转基因产品的强制标识方面,也需要识别插入位点及其旁侧序列。T-DNA 插入位点及其旁侧序列的特征包括确定插入位点的数量、每个插入位点的 T-DNA 拷贝数、每个插入 DNA 的序列以及每个位点的旁侧序列。在某些情况下,在转化过程中,由于 T-DNA 的插入,插入位点可能随着宿主基因组中整合了某些非预期的质粒 DNA 而发生基因重排,包括缺失、复制、插入细胞器等,或破坏原生宿主基因和易位^[14]。所有这些情况都受到监管机构的监管和批准^[5-7]。Southern 杂交常用来确认在转基因事件中 T-DNA 的整合和拷贝数,而插入位点及其旁侧序列的确定通常采用以 PCR 技术为基础的检测方法^[8-11]和基因组步移或直接测序^[12-14]。联合使用这些方法,可以确定 T-DNA 的插入位点、拷贝数、旁侧序列以及载体骨架的存在与否。然而,当插入位点发生复杂的 DNA 重组,或转基因序列在边界附近与宿主基因组具有相似性(左边界和右边界),这个过程通常是费力而低效的,甚至由于相似序列的误导而难以识别插入位点和旁侧序列^[14]。

下一代基因组测序(next genome sequencing, NGS)技术大大提高了测序通量,可以在短时间内以低成本获得全基因组序列。已有研究证实,如果测序深度足够,就有可能实现复杂基因组的全面覆盖^[15],而且已经证明这种深度测序可以为精确的全基因组研究奠定基础^[16]。为了鉴别外源性片段的插入,开发了几种基于 NGS 的检测方法已应用于 *Arabidopsis thaliana*^[17-18]、水稻^[19-20]、玉米^[21]和大

豆^[22-23]中。与传统的基于 PCR 的方法和基因组步移法相比,利用 WGS 数据结合靶向生物信息学分析和有限的从头装配,为高效、稳健地实现转基因作物的分子表征提供了一种可行的替代方案,无需 Southern 杂交来确定转基因的 T-DNA 拷贝数,但可高效准确地描述转基因插入片段和旁侧序列(确切的插入 DNA 序列和宿主基因组旁侧区域)以及质粒骨架的存在与否,也可以检测到插入位点的 DNA 重排。

前期研究中,采用农杆菌介导法将外源 *SMV-P3* 基因 RNAi 片段导入栽培大豆品种 Williams 82 和沈农 9 号(国审豆 2007015),获得抗病毒转基因大豆 L13 和 L104。抗病毒转基因大豆 L13 和 L104 携带由大豆内源启动子 *GmUBI3* 驱动大豆花叶病毒 *P3* 基因的干扰片段^[24]。在表达载体中,启动子 *GmUBI3* 和位于筛选标记 *bar* 基因下游的终止子 *Tvsp*,与栽培大豆 Williams 82 基因组中的相关序列具有高度相似性(图 1)^[24]。抗性鉴定结果表明,L13 和 L104 对大豆花叶病毒株系(SC7、SC15、SC18 和 SMV-R,中国发现的新种)、大豆花叶坏死病毒 soybean-infecting bean common mosaic necrosis virus (BCMV) 和西瓜花叶病毒 watermelon mosaic virus (WMV)^[24] 抗性水平极显著高于对照品种,表现为较强的广谱抗性,且抗性能够稳定遗传。为进一步评估转基因事件在育种的应用以及加强对其监管,本研究通过基因组重测序技术和 PCR 技术,获得转基因大豆事件 L13 和 L104 外源插入片段左、右边界旁侧序列,并依据其序列特征,建立该转化事件特异性检测方法,为广谱抗花叶病毒转基因大豆事件 L13 和 L104 及其衍生品种或品系商业化应用提供依据。

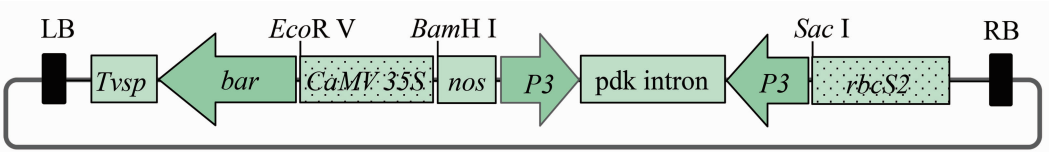


图 1 携带 *P3* 基因干扰片段的植物表达载体示意图

Fig. 1 Schematic map of plant expression vector with *P3*-RNAi fragment

1 材料与方法

1.1 材料

栽培大豆品种 Williams 82、沈农 9 号、吉育 47 和吉育 72 由吉林省农业科学院大豆研究所提供,粳稻品种吉粳 88 号由吉林省农业科学院水稻所提供,

鲁棉研 15 号为市售转基因抗虫棉花。

1.2 方法

1.2.1 基因组 DNA 提取和重测序 取纯合株系的转基因事件 L13 和 L104 的幼嫩叶片 1~2 mg,采用改良的 CTAB 法提取总 DNA^[25],然后通过 Nano-Drop 2000 spectrophotometer (Thermo Fisher Scientif-

ic, USA)测定 DNA 的浓度和纯度。

检测合格的样品基因组 DNA 委托北京百迈客生物科技有限公司对转基因大豆进行重测序分析。采用二代高通量测序 Xten 平台进行测序。对测序得到的原始数据进行质量评估,然后参考基因组序列(Wm82. a2. v1, http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Gmax)进行比对。通过比对定位所得序列在参考基因组上的位置,统计各样品的测序深度、基因组覆盖度等信息^[26-27]。

将转基因大豆基因组重测序数据分别比对到参考基因组和外源插入序列,采用 Burrows-Wheeler-Alignment Tool(bwa)^[28] 比对参考基因组,选取能比对外源插入序列的全部序列,进行局部组装。组装的序列使用 blastn 分别比对外源插入序列和参考基因组,获得外源插入片段插入位置信息^[26-27]。

1.2.2 转基因大豆事外源插入片段左、右边界旁侧序列获得 选取参考基因组 *Glycine max* Wm82. a2. v1 (<https://phytozome.jgi.doe.gov/pz/portal.html>) 上插入位点上下游约 2 kb 的片段以及 T-DNA 序列,设计引物,使扩增产物融合一部分大豆基因组序列和一部分 T-DNA 序列。

以转基因事件 L13 和 L104 基因组 DNA 为模板,利用上述引物分别进行 PCR 扩增。PCR 反应体系(25 μ L)为:10 \times PCR 缓冲液 2.5 μ L,10 mmol \cdot L⁻¹ dNTPs 0.5 μ L,5 U \cdot μ L⁻¹ *Taq* 酶 0.5 μ L,样品 DNA 1.0 μ L(约 50 ng),10 μ mol \cdot L⁻¹ 正向引物 0.5 μ L,10 μ mol \cdot L⁻¹ 反向引物 0.5 μ L, ddH₂O 19.5 μ L。PCR 反应条件为:95 $^{\circ}$ C 5 min;94 $^{\circ}$ C 45 s,60 $^{\circ}$ C 30 s,72 $^{\circ}$ C 3 min,共 35 个循环;72 $^{\circ}$ C 15 min。利用 1% 琼脂糖凝胶电泳检测 PCR 扩增产物。然后利用胶回收试剂盒纯化 PCR 产物,并连接到 GENSTAR 公

司的 EZ-T 克隆载体。委托上海生工进行测序验证,并将测序结果与外源 T-DNA 序列和参考基因组序列比对,最终获得转基因大豆外源插入片段左、右边界旁侧序列。根据获得的 T-DNA 在基因组中的具体位置,选取整合位点左右两侧各约 2 kb 的基因组 DNA 序列,与 NCBI 数据库进行比对,分析 T-DNA 是否插入到基因区。

1.2.3 转基因大豆事件特异性 PCR 检测 根据转基因大豆外源插入片段左边界和右边界旁侧序列,分别设计特异性检测引物。分别提取转基因大豆植株根、茎、叶、花和种子 DNA 样品,DNA 提取方法与 1.2.1 相同。以受体非转基因大豆品种沈农 9 号、常规大豆品种吉育 47、吉育 72 及水稻和棉花作为对照,进行 PCR 扩增。PCR 反应体系(25 μ L)为:10 \times PCR 缓冲液 2.5 μ L,10 mmol \cdot L⁻¹ dNTPs 0.5 μ L,5 U \cdot μ L⁻¹ *Taq* 酶 0.5 μ L,DNA 样品 1.0 μ L,10 μ mol \cdot L⁻¹ 正向引物 0.5 μ L,10 μ mol \cdot L⁻¹ 反向引物 0.5 μ L,ddH₂O 19.5 μ L。PCR 反应条件为:95 $^{\circ}$ C 5 min;94 $^{\circ}$ C 30 s,60 $^{\circ}$ C 30 s,72 $^{\circ}$ C 45 s,共 30 个循环;72 $^{\circ}$ C 5 min。PCR 产物经 1% 琼脂糖凝胶电泳分离,以鉴定是否存在特异性扩增条带。

2 结果与分析

2.1 转基因事件 T-DNA 整合位点的获得

将转基因大豆重测序数据分别与参考基因组 Williams 82 序列和外源 T-DNA 序列进行比对,选取能比对外源 T-DNA 序列的全部 reads,进行局部组装。组装、比对以及 BWA 分析获得的外源 T-DNA 片段整合位置信息结果表明:L13 的 T-DNA 拷贝整合位点为 Chr04 染色体的 46747946 位点;L104 T-DNA 整合位点为 Chr11 染色体的 30461895 位点(表 1)。

表 1 重测序数据
Table 1 Data of genome re-sequencing

转化事件 Transformation event	原始序列 Draw sequence	过滤序列 Filter sequence	测序深度 Sequencing depth	整合位点 Integration location
L13	48336495	46704340	12 \times	Chr04;46747946
L104	50901392	49700850	12 \times	Chr11;30461895

2.2 转基因事件 T-DNA 整合位点旁侧序列 PCR 鉴定

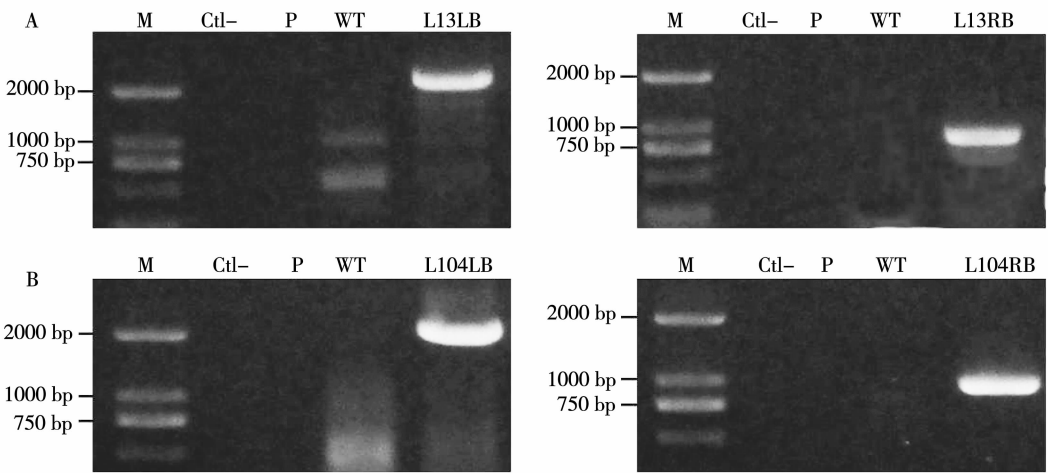
为了准确获得插入到基因组中 T-DNA 序列及其旁侧序列,以整合位点上下游 2 kb 基因组序列及 T-DNA 序列设计引物,使扩增产物含有部分大豆基

因组序列和部分载体序列,引物序列如表 2 所示。电泳检测结果显示,引物可分别在转基因事件 L13 和 L104 中扩增得到相应的左右边界旁侧序列片段,产物片段大小与预期相符,而阳性质粒和野生型沈农 9 号均不存在该片段(图 2)。

表 2 整合位点左、右侧翼序列检测引物

Table 2 Primer pairs used for left and right flanking sequences amplification

引物名称 Primer name	引物序列 Primer sequence	目的片段 Target fragment	长度 Length/bp
L13LB-F1	5'-CCCTCACTCCATTTGTCCTCT-3'	L13 左边界	2101
P3-R1	5'-TCCCACATACTTCCTCCCTCT-3'		
P3-F1	5'-TACTTCCTCCCTCTTCAGCACC-3'		
L13RB-R1	5'-GCTAGCGATGTGGACCTTAGCCTT-3'	L13 右边界	845
L104LB-F1	5'-GCATTATGTTTGAGGGAGACAAGC-3'	L104 左边界	2078
P3-R2	5'-AAGAGGGAGGAAGTATGTGGGAG-3'		
P3-F2	5'-CTTTCTTCTGAGTTACATCTTTGTCTG-3'		
L104RB-R1	5'-ATACACAAATGGAGGCTACAACG-3'	L104 右边界	897



M:2 kb DNA 标准片段;Ctl-:ddH₂O 对照;P:阳性质粒;WT:沈农 9 号;L13LB 和 L13RB 分别为转基因事件 L13 左、右边界;L104LB 和 L104RB 分别为转基因事件 L104 左、右边界。

M: 2 kb DNA Marker; Ctl-: ddH₂O control; P: Positive plasmid; WT: Shennong 9; L13LB and L13RB: Left and right flanking sequences of L13, respectively; L104LB and L104RB: Left and right flanking sequences of L104, respectively.

图 2 L13 (A) 和 L104 (B) 转基因事件旁侧序列 PCR 验证

Fig. 2 PCR verification of side sequence of transgenic events of L13 (A) and L104 (B)

扩增片段测序和分析结果能够说明转基因事件 L13 和 L104 T-DNA 的整合位点信息。L13 和 L104 T-DNA 在基因组中的整合方式为两种模式:第一种为转基因事件 L13 的 T-DNA 为单位点双拷贝模式,整合位点为 Chr04:46747946,两个拷贝左边界相连(图 3);第二种为转基因事件 L104 为单拷贝,整合位点为 Chr11:30461895,左边界存在部分 T-DNA 片段(图 3)。

将以上两个转基因事件 T-DNA 整合位点的左右两侧的大豆基因组序列约 4 kb 在 NCBI 数据库比对,结果表明这两个转化事件的 T-DNA 整合位点均位于染色体的非编码区。但是,转基因事件 L13 在 T-DNA 整合过程中发生了 5 个碱基的缺失 (Chr: 46747941 ~ 46747945) (图 4A);L104 由于 T-DNA 插入,发生了碱基替换,由 AATCGAT 替换了 CGAG-CAAGAG (Chr:30461885 ~ 30461894) (图 4B)。

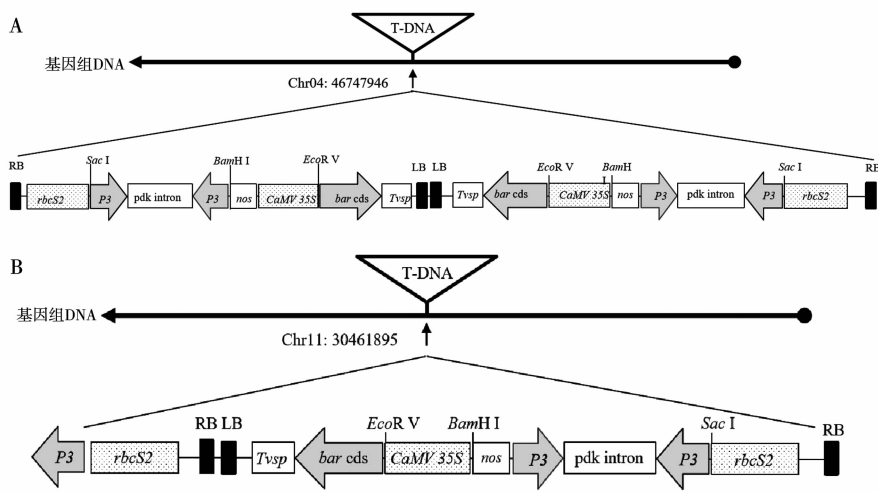


图3 转基因事件 L13(A) 和 L104(B) T-DNA 插入序列在染色体的位置

Fig. 3 Position of transgenic events L13 and L104 T-DNA insertion sequences on chromosomes

A

ACAGCTGAATACGTGAAGACACAGCTAAAGATGTGCTAAACCGAAGCATATCTCACC TTCCA
TCTCTCTTATTTT TTGTACCATAGATAGAATGTTATTAATCATCTTTT TTATAGAAAATCT
ATTAGTTTATATGTAACCAACCTTTATTTATTTGTAAAAACCTGGCTGATCACATGACCTC
TCTTCC TTCTTTTGTGATCTTCTCTCTCTCTATATGCTGAATTTATCTCTGAGATATGAAGTT
AAACATGAGAATTTCATTAAGCAATTTAAATCTAATCTATCTGACCAATGTTTCAACATACCTT
AAACTGATATGGTTTCTTCTCTTTATCTGG

caaaccaatgattgttttaaactgaagcgggaaacgaacaactctga*** aaaaaaaacatgtattg
aagtg ggtatggaatcatagaattatataaaaggaacacattgggcatctagcatgattggtg
gcagaaataatcaacagaatacaattttcaaatctactaacacacacactatagaacacctctacat
atgcagccaatctattctctcttttaactatt

467/4946

GTGCGAAGATGAATAAAAAGGTGAGATATATCTTTGATCTTTACTTGGCTGGCTTTAAATGG
GTGTGAATGTTGATGTATGCATGTCTATGTCACGCTGATGAGCAGATAGAAGCTCCGTATGTG
AAGCAACACTCTCTATATCTTACTACTCTACTATCTAGTAAGTAGGCTGATGATGTCTG
TACCTGACGACGAGGATGGCATGGCATGTTTCAAAAAGTGGCTGGACATACATTTCT
TGAAGTGTTTAGGGAAAGAAATGAAATAAAGAGTGAATCAAAATAAAATTTTGAATTAAGTA

B

TCATATGCATATGTTTGTGGGGAAGCAAGCTTACGATGGATGAGGAATATATATGAGTGAAGAAT
TTCTGTGTGGAAGAAATGAAATTAATTTCTTCAATTAATTTAAGCGGAATATTTATGCTGAATCT
TTCCACGTGGAAATTTACCAAAATATAATATGAAAAACACATATTCCATGTGGAAGACG
TTTCAATAAAGATGAATCAAAAATTAACACGAATTTTCTTTGTGTAAGAAAGATGAATGAAT
TTTATGCAATGCTGAAGTGAATGATCTGTATGAATATTTCTGTGCTATGACCAAGTGAATGAC
GTACATACAGCTTTCTTGGCTCCCAAAAGAAATGAT (CGAGCAAGAG)
caattggggttggtggtgtcttctctctctctgattggtgacaaactttcttgcgctgtgatt
gactgtggcgctgaggacagctctctagaatgctcatgaatctctcagaagttctgaaatcat
***aacatctctctctctctcagacactctcgaatccacgaatgctgaatcatgctcagctg
tctctgtgtaagtgtgattcgtgcacatccacacacatcagcgggaagcatcaagtgc
taaaagcgtggggtgctaatgagtgagctacaatcaattatgagctgctgcactgcac

30461895

CGATCTTATAAAGTAAATTTGAAGACTTTTTTAATATATGATATAAACAACCTACGTTTTCAC
AGAGTAGTCAAGCAATGTTT TTATTTTGTCAATCAAAAAATACATAAAAGTGTGTGA
TCAATTTCAAAAAATAAAAAATCATTTAAAAAAGAAAGCTGTTCTAATATGATGTTTCTCC
ATTGGCTCATTTTGGCTGTGTAGGCTCAATTTGTTATGATATGATATCTTTAATATTTGGGCTG
TGTGTTCTTCTTCAATTTCTAGCTTTTCTGTGATAATATGAAGAATAATTTTCCAAAGTAAT

A 中双划线为 L13 T-DNA 整合位点基因组的缺失序列;B 中单划线为 L104 T-DNA 整合位点替换碱基;B 中虚线为替换碱基;B 中星号为省略的载体序列。

Double underline of A indicates deletion genomes sequence in integrated site of L13; Underline of B indicates substitute bacterial DNA inserted site of L104; Dotted line of B indicates replaced bases; Asterisk of B indicates omitted vector sequence.

图4 转基因事件 L13(A)和 L104(B) T-DNA 插入导致基因组序列发生缺失或替换

Fig. 4 Deletion and substitution caused by T-DNA insertion in transgenic events L13 (A) and L104 (B)

2.3 事件特异性 PCR 检测

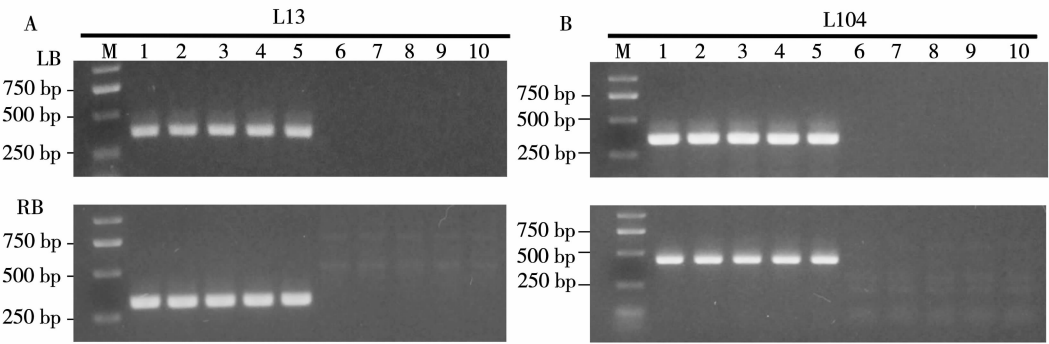
基于以上测序结果,设计 L13 和 L104 转基因事件特异性 PCR 检测引物,扩增产物仍然包含大豆基因组序列和 T-DNA 序列(表 3)。转基因事件 L13 和 L104 用特异性引物进行 PCR 扩增时,非转基因大豆品种沈农 9 号、常规大豆品种吉育 47、吉育 72

及水稻和棉花样品同样无扩增条带,只有转基因大豆 L13 和 L104 样品包括根、茎、叶、花和种子产生特异性扩增条带(图 5A 和 B)。本研究表明,利用外源插入片段侧翼序列特异性引物进行 PCR 分析可以特异性地检测样品中是否含有来源于转基因事件 L13 和 L104 的成分。

表3 事件特异性 PCR 检测引物

Table 3 Primer pairs used for event-specific PCR analysis

引物名称 Primer name	引物序列 Primer sequence	目的片段 Target fragment	长度 Length/bp
L13LB-F2	5'-ACGCAAGCAGATACTTCACCT-3'	L13 左边界特异引物	355
P3-R3	5'-TCAGATTGTCGTTTCCCGCC-3'		
P3-F3	5'-TCCACACAACATACGAGCCG-3'	L13 右边界特异引物	316
L13RB-R2	5'-GCATACATCACATTACACCCA-3'		
L104LB-F3	5'-GAAAAAACACATTATTCATGTGGAGAAG-3'	L104 左边界特异引物	342
P3-R4	5'-GCTGGCGTAATAGCGAAGAG-3'		
P3-F4	5'-TTCCACACAACATACGAGCCG-3'	L104 右边界特异引物	426
L104RB-R3	5'-TGGGGAAAAATACATATTAGAACACTG-3'		



M:2 kb DNA 标准序列;1~5:分别为转基因大豆的根、茎、叶、花和种子;6:沈农 9 号;7:吉育 47;8:吉育 72;9:非转基因水稻吉梗 88;10:转基因棉花鲁棉研 15。

M: 2 kb DNA marker; 1-5: Root, stem, leaf, flower and seed, respectively; 6: Wild type Shennong 9; 7: Soybean cultivar Jiyu 47; 8: Soybean cultivar Jiyu 72; 9: Rice cultivar Jijing 88; 10: Transgenic cotton Lumianyan 15.

图 5 L13(A)和 L104(B)事件特异性 PCR 检测

Fig. 5 Event-specific PCR analysis of L13 (A) and L104 (B)

3 讨论

对于转基因作物的安全评价以及转化事件的追踪来讲,外源 T-DNA 及其旁侧序列的详细分子特征是非常重要的^[29]。基于 PCR 技术的 TAIL-PCR 和染色体步移结合 Southern 杂交和 Sanger 测序是常用来确定外源插入序列整合位点和与宿主基因组连接区域的方法^[5]。由于大豆基因组中存在大量的重复序列,如果 T-DNA 整合方式复杂、发生重组、存在缺失/插入或者单个碱基替换,那么使用这些方法确定整合位点以及旁侧序列往往难以奏效,即使可行,也会耗时耗力。随着高通量测序技术的出现以及不断发展,全基因组序列获得已经又快又省钱,并且已经成为分析基因组变异的有力工具,包括重组、基因融合以及 DNA 结构变异等^[30-32]。NGS 技术结合生物信息学,不仅应用于农业生物技术领域^[17,20,33],目前也开始应用于转基因作物分子特征的获得以及安全评价方面^[23,34-36]。本研究中,应用该技术对转基因事件 L13 和 L104 的 T-DNA 整合位点、整合方式以及旁侧序列进行分析,证明 L13 的 T-DNA 整合方式为单位点双拷贝,整合位点为大豆基因组 Chr04:46747946;L104 的 T-DNA 整合方式为单拷贝并残余部分右边界序列,整合位点为大豆基因组 Chr11:30461895。这两个转基因事件里,T-DNA 的整合方式复杂,运用传统方式均难以获得其分子特征,而使用重测序方法获得了 T-DNA 的整合位点以及旁侧序列,并建立了特异性检测方法,为这两个转基因事件的安全评价以及衍生品系的追踪评价提供了技术支持。

与传统的基于 PCR 技术方法相比较,全基因重测序结合生物信息学分析,在分析转基因作物分子特征方面既高效又省时省力,不需要大量的 DNA 和繁琐的人工操作,更容易发现遗传变化。随着技术

的发展,检测物质用量也会进一步降低。同时,全基因组测序还易于发现基因组中存在的 SNPs、小的 InDels 等序列变异^[36]。这些详细的序列信息也可应用于转基因作物的毒理以及致敏性分析中。

4 结论

用重测序和 PCR 结合的方法分离获得了转 SMV-P3 基因干扰片段的大豆 L13 和 L104 株系 T-DNA 整合位点旁侧序列。转基因事件 L13 的 T-DNA 整合位点为大豆基因组 Chr04:46747946,综合位点为单位点双拷贝;L104 T-DNA 整合位点为 Chr11:30461895,整合方式为单拷贝。据旁侧序列特征设计了针对 L13 和 L104 株系的事件特异性检测引物,采用特异性 PCR 扩增方法可以准确地将不同的转化株系区别开来,也可追踪不同转化株系是否来自同一转化体。

参考文献

[1] Latham J R, Wilson A K, Steinbrecher R A. The mutational consequences of plant transformation[J]. Journal of Biomedicine and Biotechnology, 2006(2): 25376.

[2] Nacry P, Camilleri C, Courtial B, et al. Major chromosomal rearrangements induced by T-DNA transformation in *Arabidopsis*[J]. Genetics, 1998, 149(2): 641-650.

[3] Wilson A K, Latham J R, Steinbrecher R A. Transformation-induced mutations in transgenic plants; Analysis and biosafety implications[J]. Biotechnolnoly and Genetic Engineering Reviews, 2006, 23(1):209-238.

[4] Windels P, Sylvie D B, Erik V B, et al. T-DNA integration in *Arabidopsis* chromosomes. Presence and origin of filler DNA sequences[J]. Plant Physiology, 2003, 133(4): 2061-2068.

[5] Codex Alimentarius Commission. Guide line for the conduct of food safety assessment of foods derived from recombinant-DNA plants[S]. FAO/WHO Joint Publications, 2003:1-14.

[6] European Food Safety Authority Panel on Genetically Modified Or-

- ganisms. Guidance on the environmental risk assessment of genetically modified plants[J]. European Food Safety Authority Journal, 2010, 8(11):1879.
- [7] Kok E J, Pedersen J, Onori R, et al. Plants with stacked genetically modified events: To assess or not to assess[J]. Trends Biotechnology, 2014, 32(2): 70-73.
- [8] Ochman H, Gerber A S, Hartl D L. Genetic applications of an inverse polymerase chain reaction[J]. Genetics, 1998, 120(3): 621-623.
- [9] Liu Y G, Mitsukawa N, Oosumi T, et al. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR[J]. Plant Journal, 8(3):457-463.
- [10] O'Malley R C, Ecker J R. Linking genotype to phenotype using the *Arabidopsis* unimutant collection[J]. Plant Journal, 61(6): 928-940.
- [11] Ji J B, Braam J. Restriction site extension PCR: A novel method for high-throughput characterization of tagged DNA fragments and genome walking[J]. PLoS ONE, 2010, 5: e10577.
- [12] Spalinskas R, Bulcke M V D, Eede G V D, et al. LT-RADE: An efficient user-friendly genome walking method applied to the molecular characterization of the insertion site of genetically modified maize MON810 and rice LLRICE62[J]. Food Analytical Method, 2013, 6(2): 705-713.
- [13] Babekova R, Funk T, Pecoraro S, et al. Development of an event-specific real-time PCR detection method for the transgenic Bt rice line KMD1[J]. European Food Research and Technology, 2009, 228(5): 707-716.
- [14] Milcamps A, Rabe S, Cade R, et al. Validity assessment of the detection method of maize event Bt10 through investigation of its molecular structure[J]. Journal of Agricultural and Food Chemistry, 2009, 57(8):3156-3163.
- [15] Wang J, Wang W, Li R, et al. The diploid genome sequence of an Asian individual[J]. Nature, 2008, 456(7218): 60-65.
- [16] Ajay S S, Parker S C, Abaan H O, et al. Accurate and comprehensive sequencing of personal genomes[J]. Genome Research, 2011, 21(9): 1498-1505.
- [17] Lepage E, Zampini E, Boyle B, et al. Time and cost-efficient identification of T-DNA insertion sites through targeted genomic sequencing[J]. PLoS One, 2013, 8(8): e70912.
- [18] Inagaki S, Henry I M, Lieberman M C, et al. High-throughput analysis of T-DNA location and structure using sequence capture[J]. Plos One, 2015, 10(10): e0139672.
- [19] Wahler D, Schauser L, Bendiek J, et al. Next-generation sequencing as a tool for detailed molecular characterisation of genomic insertions and flanking regions in genetically modified plants: A pilot study using a rice event unauthorised in the EU[J]. Food Analytical Methods, 2013, 6(6): 1718-1727.
- [20] Park D, Kim D, Jang G, et al. Efficiency to discovery transgenic loci in GM rice using next generation sequencing whole genome re-sequencing[J]. Genomics & Informatics, 2015, 13(3):81.
- [21] Williams-Carrier R, Stiffler N, Belcher S, et al. Use of illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy mutator lines of maize[J]. The Plant Journal, 2010, 63(1):167.
- [22] Kovalic D. The use of next generation sequencing and junction sequence analysis bioinformatics to achieve molecular characterization of crops improved through modern biotechnology[J]. Plant Genome, 2012, 5(3):149-163.
- [23] Guo B, Yong G, Hong H, et al. Identification of genomic insertion and flanking sequence of *G2-EPSPS* and *GAT* transgenes in soybean using whole genomes equencing method[J]. Frontiers in Plant Science, 2016, 7: 1009.
- [24] Yang X D, Niu L, Zhang W, et al. RNAi-mediated SMV *P3* cis-tron silencing confers significantly enhanced resistance to multiple *Potyvirus* strains and isolates in transgenic soybean[J]. Plant Cell Reports, 2018, 37(1): 103-114.
- [25] Porebski S, Bailey L G, Baum B R. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components[J]. Plant Molecular Biology Reporter, 1997, 15(1): 8-15.
- [26] 马阔, 仲晓芳, 牛陆, 等. 耐盐转基因大豆 FA8015 旁侧序列分离及定性 PCR 检测[J]. 大豆科学, 2018, 37(6):854-859. (Ma K, Zhong X F, NIU L, et al. Sequence separation and qualitative PCR detection of salt-tolerant transgenic soybean FA8015[J]. Soybean Science, 2018, 37(6): 854-859.)
- [27] 仲晓芳, 杨静, 贺红利, 等. 基于基因组重测序的高含量油酸转基因大豆 T-DNA 旁侧序列分析及事件特异性 PCR 检测[J]. 农业生物技术学报, 2018, 26(12):2017-2026. (Zhong X F, Yang J, He H L, et al. Sequence analysis and event specific PCR detection of high content oleic acid transgenic soybean T-DNA based on genomic resequencing[J]. Journal of Agricultural Biotechnology, 2018, 26(12): 2017-2026.)
- [28] Langmead B and alzberg S L. Fast gapped-read alignment with Bowtie 2[J]. Nature Methods, 2012, 9:357-359.
- [29] Yang L T, Wang C M, Jensen A H, et al. Characterization of GM events by insert knowledge adapted re-sequencing approaches[J]. Scientific Reports, 2013, 3:2839.
- [30] Campbell P J, Stephens P J, Pleasance E D, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing[J]. Nature Genetics, 40(6):722-729.
- [31] Hormozdiani F, Hajirasouliha I, Mcpherson A, et al. Simultaneous structural variation discovery among multiple paired-end sequenced genomes[J]. Genome Research, 2011, 21(12): 2203-2212.
- [32] DuBose A J, Lichtenstein S T, Narisu N, et al. Use of microarray hybrid capture and next-generation sequencing to identify the anatomy of a transgene[J]. Nucleic Acids Research, 2013, 41(6):e70.
- [33] Kovalic D, Garnaat C, Guo L, et al. The use of next generation sequencing and junction sequence analysis bioinformatics to achieve molecular characterization of crops improved through modern biotechnology[J]. The Plant Genome, 2012, 5(3): 149-163.
- [34] Urbanski D F, Malolepszy A, Stougaard J, et al. Genome wide *LOREI* retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*[J]. The Plant Journal, 2012, 69(4):731-741.
- [35] Daniela W, Leif S, Joachim B, et al. Next-generation sequencing as a tool for detailed molecular characterization of genomic insertions and flanking regions in genetically modified plants: A pilot study using a rice event unauthorized in the EU[J]. Food Analytical Method, 2013, 6(6):1718-1727.
- [36] Pauwels K, DeKeersmaecker S C J, DeSchrijver A, et al. Next-generation sequencing as a tool for the molecular characterization and risk assessment of genetically modified plants: Add value or not? [J]. Trends in Food Science and Technology, 2015, 45(2):319-326.