Cloning of flanking sequence in transgenic plants by restriction site-anchored single-primer polymerase chain reaction

J. Ma, N.N. Wang, S. Ren, Y.P. Fu, S. Lu, Y.P. Wang and P.W. Wang

Biotechnology Center of Jilin Agricultural University, Changchun, Jilin, China

Corresponding author: P.W. Wang
E-mail: davidmedsci@sina.com

ABSTRACT. Determining the insertion position of an exogenous gene in the target plant genome is one of the main issues in the transgenic plant field. This study introduced a simple, rapid, and accurate method to clone the flanking sequences of the transgenic *bar* gene as the anchoring gene in the transgenic maize genome using single-primer polymerase chain reaction (PCR). This method was based on the distribution of restriction sites in the maize genome and adopted the single-primer PCR method. Cloning the flanking sequences with the restriction site-anchored single-primer PCR simplified the experimental procedures by about 70% and reduced the experimental time by more than 80%. In conclusion, the restriction site-anchored single-primer PCR was a simple, rapid method to obtain the unknown flanking sequences in the transgenic plants.

Key words: Single-primer polymerase chain reaction;
Unknown flanking sequences; Transgenic plants; Restriction sites

INTRODUCTION

One of the important issues in the transgenic plant field is determining the integration position of an exogenous gene in the target plant genome or determining the flanking sequences of the inserted gene. Currently, polymerase chain reaction (PCR) techniques are widely used to determine the insertion position of the exogenous gene in the genome. These techniques include adapter-ligation-mediated PCR (Jones and Winistorfer, 1992), thermal asymmetric interlaced PCR (TAIL-PCR) (Liu and Whittier, 1995; Liu et al., 1995; Liu and Chen, 2007), and restriction site extension PCR (Kim et al., 2004; Ji and Braam, 2010); of these, TAIL-PCR is the most commonly used. However, TAIL-PCR requires 3 anchor primers and 4-8 degenerate random primers for the known sequences, and then it obtains the flanking sequences of the known gene through 3 rounds of the TAIL-PCR (Song et al., 2005; Liang et al., 2010). In TAIL-PCR, the primer design is complex, the PCR procedures and operation are complicated, and 8-10 h should be taken to complete an experiment, making it very time-consuming. Restriction site extension PCR requires that the genome is digested prior to PCR, and the technique is complicated (Kim et al., 2004; Ji and Braam, 2010).

The arrangement of restriction sites in the plant genome is fixed, and any exogenous gene will carry some restriction sites. For instance, the *bar* gene carries the *sal*I, *kpn*I, and *sac*II sites (Antal et al., 2004). If the exogenous gene is inserted into the genome of the target plant, it will change the original arrangement of restriction sites of the target plant. Then, if we can identify such changes in the arrangement of restriction sites, we can obtain the flanking sequences of the inserted position (Ma et al., 2011a,b, 2012). Based on the above analysis, the principle of this study is shown in Figure 1 using the *sal*I site in the *bar* gene as an example.

In this study, we used high-generation transgenic maize as the plant material and designed 3 single primers for *bar* gene-based sequences. Using a 33-cycle single-primer PCR, we obtained the flanking sequences of the *bar* gene (as the known gene) in the transgenic maize.
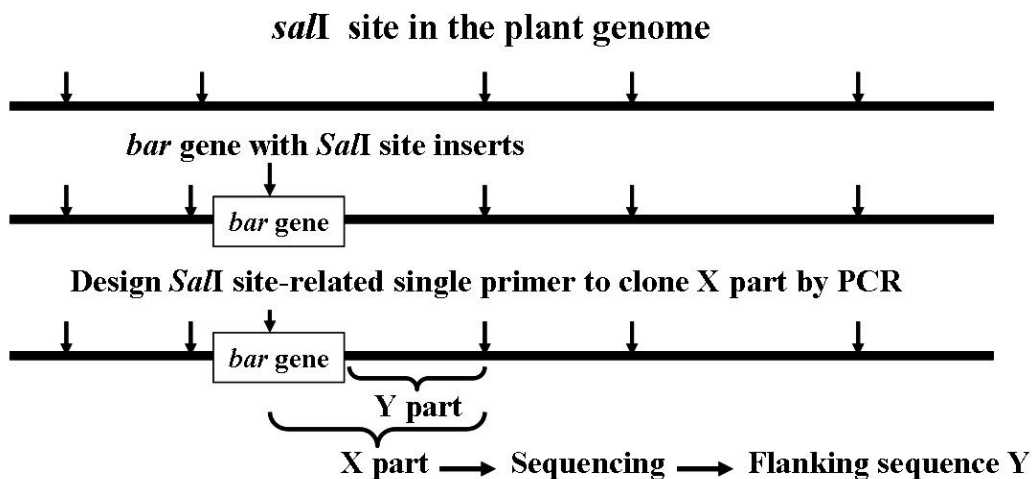

Figure 1. Principle of the study.

MATERIAL AND METHODS

Materials

Genomic DNA was extracted from the *bar*-transgenic maize that was subjected to the agrobacterium-mediated method using the pCAMBIA3301 vector. Taq DNA polymerase, 10 mM dNTP, and 10X PCR buffer were purchased from IBM (USA). Agarose was supplied by AMRESCO (USA). The genomic DNA extraction kit and DNA marker were purchased from TaKaRa (China).

Primers

Three *bar*-related single primers were designed, their sequences are shown in Table 1, and the positions of these primers in the T-vector are illustrated Figure 2. The method for designing the P1 single primer was as follows. The *kpn*I restriction site sequence GGTACC (5'-3') in the *bar* gene was found, and then 11 bases toward the 5'-end were incorporated (5'-TTCAGCCTGCC-3'), forming the P1 primer sequence: 5'-TTCAGCCTGCCGGTACC-3'. This single primer could pair with the *bar* gene completely, could be used as the anchor primer, and could also be used as a random primer by considering the genomic distribution of restriction sites. The methods for designing the P2 and P3 single primers were the same as the above.

Table 1. Primer information.

Primer	Sequence (5'-3')	Length (bp)	Distance from the left T-border (bp)
P1 (*kpn*I)	TTCAGCCTGCCGGTACC	17	334
P2 (*sal*I)	TGGACGGCCGAGTCGAC	17	583
P3 (*sac*II)	GATTAGCCCCCCGCGG	16	415
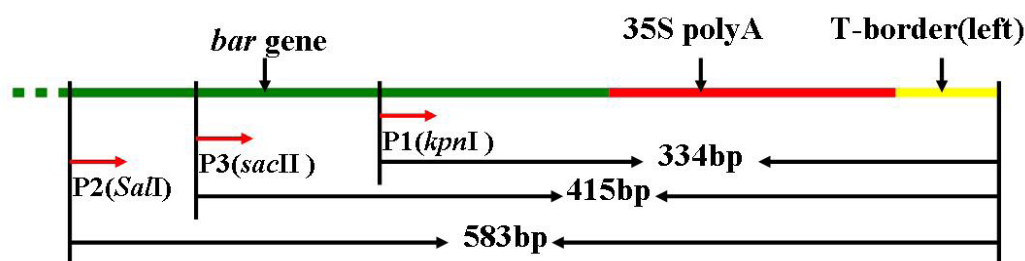
Figure 2. Position of single primers in the T-vector pCAMBIA3301.

PCR procedures and sequencing

Total DNA was extracted from the plant leaves using the Universal Genomic DNA Extract kit (version 3.0) according to manufacturer instructions (TaKaRa Company). The 50-μL reaction mixture for the P1 single-primer amplification contained 5.0 μL 10X buffer, 1.0 μL 10 mM dNTPs, 1.0 μL 20 μM P1 primer, 1.0 μL genomic DNA template (<1 μg), 1.0 μL 2.5

U/µL Taq DNA Polymerase, and 32 µL ddH$_2$O. The PCR conditions were as follows: 94°C for 120 s; 3 cycles of 94°C for 40 s, 59°C for 30 s, and 72°C for 120 s; 30 cycles of 94°C for 40 s and 68°C for 120 s; and 72°C for 5 min. The reaction mixture and cycling conditions for the P2 and P3 single-primer amplifications were the same as those for P1 PCR. Transgenic and non-transgenic plants shared the same PCR procedures. PCR products were separated on 2% agarose gels. The exclusive bands of the transgenic plants, with lengths that were larger than 314, 583, and 415 bp, were extracted routinely, sequenced, and analyzed by Basic Local Alignment Search Tool (BLAST) with published homologous sequences of flanking sequences of the *bar* gene in the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/nuccore/).

RESULTS

PCR results of each primer

Figure 3 shows the PCR results with the P1, P2, and P3 single primers for the transgenic and non-transgenic maize. The P1 and P3 single primers did not yield exclusive bands of the transgenic plants. However, the P2 single primer obtained the clear band A. The length of band A (about 900 bp) was larger than the distance between the P2 primer and the left T-vector border (583 bp), and band A did not appear in the non-transgenic control.
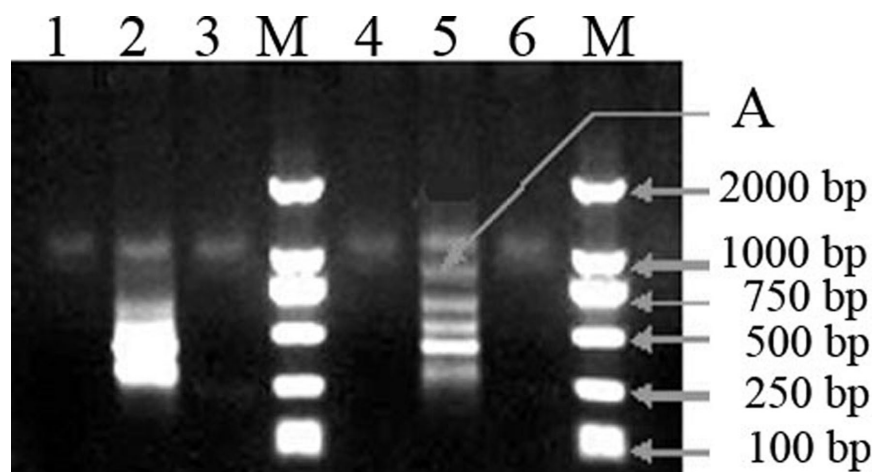
Figure 3. Polymerase chain reaction (PCR) results of each primer for transgenic maize and a non-transgenic control. *Lane M* = DL2000 marker; *lanes 1*, *2*, and *3* = PCR products using the P1, P2, and P3 single primers and the non-transgenic control; *lanes 4*, *5*, and *6* = PCR products using the P1, P2, and P3 single primers and the transgenic maize.

Sequencing

Figure 4 shows the sequencing results of band A. The flanking sequence of the T-DNA insertion site, with a length of 272 bp, was determined. By the NCBI BLAST alignment, the sequences in gray color (Figure 4) were homologous to maize AY664415.1.

GCGAGACACATTCTATAGGGGGCGATTGAGCTGCNCCTTTGGACGGCCGAGT
CGACCGTGTACGTCTCCCCCCGCCACCAGCGGACGGGACTGGGCTCCACG
CTCTACACCCACCTGCTGAAGTCCCTGGAGGCACAGGGCTTCAAGAGCGT
GGTCGCTGTCATCGGGCTGCCCAACGACCCGAGCGTGCGCATGCACGAGG
CGCTCGGATATGCCCCCCGCGGCATGCTGCGGGCGGCCGGCTTCAAGCACG
GGAACTGGCATGACGTGGGTTTCTGGCAGCTGGACTTCAGCCTGCCGGTAC
CGCCCCGTCCGGTCCTGCCCGTCACCGAGATTTGACTCGAGTTTCTCCATAA
TAATGTGTGAGTAGTTCCCAGATAAGGGAATTAGGGTTCCTATAGGGTTTCG
CTCATGTGTTGAGCATATAAGAAACCCTTAGTATGTATTTGTATTTGTAAAAT
ACTTCTATCAATAAAATTTCTAATTCCTAAAACCAAAATCCAGTACTAAAAT
CCAGATCCCCCGAATTAATTCGGCGTTAATTCAGTACATTAAAAACGTCCGC
AATGTGTTATTAAGTTGTCTAAGCGTCAATTTGTTTACACCACAATATATCCT
GCCATAAAATGCCGAGGCCGACCTTCAATCACTCGGCGTTGCCTCTAAAAC
GCCGACCACGTGTTCAATCGCTCGGCATTATTTCTAAAATGCCGACGCCGA
CCTTCAATCACTCGGCGTTGCCTCTAAAAACGCCGACCACGTGTTCAATCGC
TCGGCATTATTTCTAAAATGCCGACGCCGACCTTCAATCACTCGGCGTTGCC
TCTAAAACGCTGACCACGTGTTCAATCGCTCGGCATTACTTCTAAAATGCCG
ACGTCGACTCGGCCGTCCAGCGTCTGTTCAGGGTTCCAAACTCATAGTGGC
AACTAAACACGAGGTAATGG

Figure 4. Sequencing results of band A from the P2 single primer. Blue indicates the *pEASY*-Blunt Simple Cloning Vector that was used for sequencing; green indicates the *bar* gene; red indicates the 35S polyA; yellow indicates the left border of the expression T-vector; and gray indicates the cloned flanking sequences.

DISCUSSION

Determining the integration position of the exogenous gene in the genome of the target plant is an important task in transgenic plant research. Currently, TAIL-PCR is the most commonly used method (Liu and Whittier, 1995; Liu et al., 1995; Song et al., 2005; Liu and Chen, 2007; Liang et al., 2010). However, the primer design of TAIL-PCR is complex, and the PCR procedures are complicated and time-consuming. This study designed single primers with restriction sites of the exogenous gene for single-primer PCR, and the flanking sequence was identified based on the distribution of restriction sites in the plant genome. Three short single primers were designed with the transgenic *bar* gene as the anchor gene, which was inserted into the transgenic maize genome. A 33-cycle single-primer PCR obtained the flanking sequences of the known gene, the *bar* gene, in the transgenic maize. With this PCR method, designing the primers was simple, the PCR procedure and operation were simple, and PCR took only 2-3 h. In PCR procedures, 3-6 cycles should be at a low-annealing temperature; the annealing temperature should be equal to or slightly greater than the optimum primer annealing temperature, but the difference should not exceed 3°C. The subsequent 30 cycles at the high-annealing temperature were sufficient for amplification. The best method was to use the 2-point method for PCR.

Certainly, the genome distribution of restriction sites is different between various species and varieties of the same species. This is just the theoretical basis of amplified fragment length polymorphism. According to the completely random arrangement principle, a 6-bp restriction site should appear once in every 4096 bases in the genome (Antal et al., 2004). In practice, the

genome distribution of restriction sites does not completely follow the random arrangement principle. Two identical restriction sites may be separated by only a short distance sometimes, but they may be very far away other times. The rearrangement of the different restriction sites is also different after the insertion of a foreign gene. Hence, not all single primers that are designed according to the genome distribution of restriction sites will yield clearly distinguishable PCR products (Ma et al., 2011a,b, 2012). For example, in this study, the P1 and P3 primers did not yield a differential band in the transgenic maize in comparison with the non-transgenic control. Therefore, it is necessary to use a non-transgenic control. The obtained band that is sequenced, with a length that is larger than the distance between the single primer and the left border of the T-vector, should be amplified exclusively in the transgenic plants.

In conclusion, single-primer PCR can be used to obtain the flanking sequences of the T-DNA insertion position based on the distribution of restriction sites in the plant genome. This method excludes the complex primer design and experimental procedures of TAIL-PCR and restriction site extension PCR, and it saves time and money. It is a simple, fast method to determine the flanking sequences of the T-DNA insertion position.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Antal Z, Rascle C, Fevre M and Bruel C (2004). Single oligonucleotide nested PCR: a rapid method for the isolation of genes and their flanking regions from expressed sequence tags. *Curr. Genet.* 46: 240-246.

Ji J and Braam J (2010). Restriction site extension PCR: a novel method for high-throughput characterization of tagged DNA fragments and genome walking. *PLoS One* 5: e10577.

Jones DH and Winistorfer SC (1992). Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. *Nucleic Acids Res.* 20: 595-600.

Kim YJ, Kwak CI, Gu YY, Hwang IT, et al. (2004). Annealing control primer system for identification of differentially expressed genes on agarose gels. *Biotechniques* 36: 424-6, 428, 430.

Liang CZ, Zhang R, Sun GQ and Meng ZG (2010). Cloning of stress-related transcription factors gene from cotton by optimized TAIL-PCR. *Cotton Sci.* 22: 195-201.

Liu YG and Whittier RF (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25: 674-681.

Liu YG and Chen Y (2007). High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques* 43: 649-50, 652, 654.

Liu YG, Mitsukawa N, Oosumi T and Whittier RF (1995). Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* 8: 457-463.

Ma J, Guan SC, Yao D, Wei YF, et al. (2011a). Problems with and a system to eliminate single-primer PCR product contamination in simple sequence repeat molecular marker-assisted selection in soybean. *Genet. Mol. Res.* 10: 1659-1668.

Ma J, Wang PW, Yao D, Wang YP, et al. (2011b). Single-primer PCR correction: a strategy for false-positive exclusion. *Genet. Mol. Res.* 10: 150-159.

Ma J, Guan SC, Zhang Z and Wang PW (2012). Single- and double-SSR primer combined analyses in rice. *Genet. Mol. Res.* 11: 1032-1038.

Song DF, Han N, Bian HW and Zhu MY (2005). Analysis of cbfl flanking sequences in transgenic tobacco by TAIL-PCR. *Acta Agron. Sin.* 31: 1377-1379.