A low degenerate primer pool improved the efficiency of high-efficiency thermal asymmetric interlaced PCR to amplify T-DNA flanking sequences in *Arabidopsis thaliana*

Huchen Zhang¹ · Weiwei Xu² · Zhiyang Feng² · Zhi Hong¹

Received: 16 July 2017 / Accepted: 4 December 2017 / Published online: 11 December 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

We employed hi-TAIL-PCR to identify T-DNA loci in our *Arabidopsis* activation tagging library and only a total of 28 (39%) insertion sites from 72 samples were characterized when the recommended primer pools, C1 and C2 were used. By comparison, we found C1 harboring relatively low degeneracy was more efficient to amplify the flanking sequences of T-DNA insertion than C2. We replaced the degenerate sequences in long arbitrary degenerate (LAD) primers with a piece of 16-bp degenerate sequence originally used in TAIL-PCR, which had the relatively low degeneracy. Our results showed that the new LAD primer pool N increased the valid amplifications and a total of 37 (51%) T-DNA loci were identified, indicating a more effective amplification of T-DNA flanking sequences in *A. thaliana*.

Keywords LAD primers · hi-TAIL-PCR · T-DNA insertion · Flanking sequence

Introduction

T-DNA insertion and transposon-based mutant libraries have been extensively used to explore the gene function in *planta* (Chang et al. 2012; Osborne et al. 1995; Hsia et al. 2017; Kondou et al. 2010; Weigel et al. 2000; Marsch-Martinez et al. 2002; Alonso et al. 2003). To determine the T-DNA loci integrated in the genome, thermal asymmetric interlaced PCR (TAIL-PCR) has been widely used to amplify the flanking sequences adjacent to the insertion sites due

to its simplicity, sensitivity, low cost and high-throughput screening (Miyao et al. 2003; Liu et al. 1995). The efficiency of TAIL-PCR was further improved by simultaneously using several pieces of arbitrary degenerate primers (AD primer pool) during PCR reaction, named as high throughput TAIL-PCR (ht-TAIL-PCR) (Singer and Burke 2003; Sessions et al. 2002). However, both TAIL-PCR and ht-TAIL-PCR tend to produce many false positive and small bands which usually provide little information about the flanking sequences (Sessions et al. 2002; Wu et al. 2015). Based on the suppression-PCR, Liu and Chen (2007) reported a high-efficiency TAIL-PCR (hi-TAIL-PCR) technique by introducing long arbitrary degenerate (LAD) primers and setting alternative cycles of high- and low-stringency (TAIL-cycle), which significantly inhibits the amplification of small and false positive bands and thus increases the PCR efficiency to amplify the target sequences.

We conducted hi-TAIL-PCR to screen T-DNA insertion sites in the home-made activation tagging mutant library. However, we encountered some problems in efficiently amplifying T-DNA flanking sequences. In this study, we replaced the degenerate sequences in long AD primers with 16-bp degenerate sequences originally used in TAIL-PCR (Liu and Whittier 1995; Liu et al. 1995) and the new primers obviously improved the efficiency to amplify the T-DNA flanking sequences in *A. thaliana*.

Huchen Zhang and Weiwei Xu contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-017-1032-y>) contains supplementary material, which is available to authorized users.

✉ Zhiyang Feng
zfeng@njau.edu.cn

✉ Zhi Hong
zhihong@nju.edu.cn

¹ State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing 210023, China

² College of Food Science and Technology, Nanjing Agriculture University, Nanjing 210095, China

Materials and methods

Plant transformation and growth

Genetic transformation of plant through *Agrobacterium tumefaciens* harboring the plasmids (pBASTA-AT2) was conducted by vacuum infiltration as described previously (Clough and Bent 1998). The transgenic lines were screened from the seedlings grown on 1/2 MS medium, containing 1% sucrose, 0.3% gelrite and 10 µg/ml of glufosinate-ammonium (Basta, Sigma), for 10 days (18 h light/8 h dark) and then transferred to soil. The leaves of T₁ seedlings were used for genomic DNA isolation.

Genomic DNA isolation

Extraction of high-quality genomic DNA followed the protocol described previously (Stephen et al. 1983). Briefly, two fresh leaves (1–2 cm²) were used to extract genomic DNA. DNA pellet was air-dried and finally re-suspended in 80 µl of DNase-free double-distilled water for further analysis.

Primer design and hi-TAIL-PCR

The primers used in this study were listed in the Supplementary Table S1. For hi-TAIL-PCR, three nested specific primers, TLB1, TLB2 and TLB3, were designed according to the pBASTA-AT2 left border sequences. TLB2 carried a 14-bp sequence at 5' terminal identical to piece of AC1 (Liu and Chen 2007) to suppress the amplification of small fragments in primary PCR reactions. Four pieces of LAD primers including a AC1 sequence, a *NotI* recognition site, and variable degenerate sequences at 3' terminal. The new degenerate LAD (N-LAD) primers were generated by replacing the 3' terminal degenerate sequences of LAD primers with AD primers used in TAIL-PCR (Liu and Whittier 1995; Liu et al. 1995), without *NotI* sequence. The T_m values were calculated according to the formula, $T_m = 69.3 + 41 \times \text{GC}\% - 650/L$ (L = primer length) (Mazars et al. 1991).

The hi-TAIL-PCR including pre-, primary- and secondary-amplification was performed according to the procedures previously described (Liu and Chen 2007). For pre-amplification, the N-LAD 1, 2, 3 and 4 were used as AD primer pool (pool N). The primers, LAD1 plus LAD3 and LAD2 plus LAD4, were used as control pools and named as pool C1 and pool C2, respectively.

DNA sequencing and mapping of T-DNA insertion sites

The primary and secondary products were separated on 1% agarose gel and those secondary PCR products running

faster than primary PCR products, were isolated for sequencing with primer TLB3. To determine the T-DNA insertion location, the obtained sequences were subjected to BLAST search in NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The positive sequences should contain a short sequence (60–150 bp) of the T-DNA left border and the flanking sequences adjacent to the 3' terminal. The insertion sites were mapped to the *A. thaliana* genome based on the flanking sequences.

Results and discussion

To obtain the flanking sequences of plasmid pBASTA-AT2 integrated in *A. thaliana* genome, we employed the hi-TAIL-PCR to amplify the flanking sequences for mapping T-DNA insertion site. We followed the protocols and long arbitrary degenerate (LAD) primer pools (C1 and C2) as previously described (Liu and Chen 2007) and found that the length of the majority of amplified fragments was between 500 and 1000 bp (Fig. 1a, b) and the PCR products more than 1000 bp in length could only be detected from 20 (~ 28%) and 11 (~ 15%) samples after pre-, primary and secondary amplification using primer pool C1 (LAD1 and LAD3) or C2 (LAD2 and LAD4) from 72 independent samples, respectively (Table 1). Those small PCR products less than 1 kb were mostly non-specific amplification of T-DNA backbone after sequencing.

It has known that the features of LAD primers including the degeneracy, length, GC content and T_m value, were the crucial factors to guarantee the successful amplification of the flanking sequences in hi-TAIL-PCR (Singer and Burke 2003; Liu et al. 1995; Liu and Chen 2007). From the agarose gel analysis, primer pool C1 seemed more efficient to amplify valid products than C2 (Fig. 1a, b and Table 1). We analyzed the primer sequences and found that primers in pool C1 and C2 had similar degeneracy, GC content and T_m value as well (Supplementary Table S1). However, a previous study reported that primer with a thymine at the 3' end was more degenerate than that with an adenine and the presence of a thymine at the 3' end in the primer made more efficient amplification (Kwok et al. 1990). Hence, the LAD primers in pool C2 (GGTT and CGGT at 3' end) should have relative high degeneracy than those in C1 (GGAA and CCAA at 3' end). However, the efficiency of amplifying valid products by pool C1 was better than that of C2 (Fig. 1a, b and Table 1).

Based on the above results, we replaced the degenerate sequences in LAD primers with a piece of 16-bp degenerate sequences originally used in TAIL-PCR to generate N-LAD primers, which reduced the degeneracy of initial LAD primers (Supplementary Table S1). Four pieces of N-LAD primers (N-LAD1–4) composed the primer pool

Fig. 1 Agarose gel analysis of secondary hi-TAIL-PCR products amplified with different primer pools. The secondary products amplified with pool C1 (**a**), C2 (**b**) and N (**c**) were separated on 1% agarose gel and photographs were taken for records. The bands marked with

asterisk indicated the flanking sequences identified by subsequent sequencing. M indicates the DL5000 DNA marker ladder. The size of 1 kb was indicated by dotted lines

Table 1 The amplification efficiency of primer pools C1, C2 and N in hi-TAIL-PCR

Primer pool	Mean of amplified bands	Samples (%) of PCR products over 1 kb can be amplified	Samples (%) that the flanking sequences were identified
C1	1.95	20 (28%)	19 (26%)
C2	1.9	11 (15%)	11 (15%)
C1 and C2	–	28 (39%)	28 (39%)
N	2.5	44 (60%)	37 (51%)
N, C1 and C2	–	61 (85%)	51 (71%)

The percentages and mean number were calculated based on the analysis of PCR products from a total of 72 samples using three different primer pools

N. We then tested the efficiency of the primer pool N in hi-TAIL-PCR. Long PCR products (> 1 kb) could be detected from 44 of 72 samples (more than 60%) and the average number of amplified bands were 2.5, more than those amplified with pool C1 or C2 (2.0 for pool C1 and 1.9 for pool C2) (Table 1). Furthermore, the size of PCR products amplified with primer pool N was usually bigger than that amplified with C1 and C2 (Fig. 2). After sequencing, we confirmed the flanking sequences of 37 (51%) from 72 independent samples could be amplified with pool N. Comparably, 19 (26%) and 11 (15%) samples were identified with C1 and

C2, respectively (Table 1). In addition, we also compared the efficiency between TAIL-PCR using the original AD primers and hi-TAIL-PCR with N-LAD primers. While, only 8 flanking sequences from 24 samples were obtained (Supplementary Fig. S1), even lower than pool C1 and C2 working in hi-TAIL-PCR, indicating that N-LAD primers were more efficient in hi-TAIL-PCR to screen T-DNA insertions in *A. thaliana* genome.

The N-LAD primers greatly improved the efficiency of screening T-DNA insertion. Even so, it could not completely replace pool C1 and C2. As shown in Fig. 1,

Fig. 2 Primer pool N was more efficient to amplify bigger fragments than C1 and C2. The primary (II) and secondary (III) PCR products amplified with different primer pools (C1, C2 and N) were separated

on 1% agarose gel. Four independent samples (No. 25, 26, 27 and 28) were shown. M indicated the DL5000 DNA marker ladder

some clearer bands appeared in samples of 3, 8, 9, 11, 12 and 20 when the pool N was used. Comparably, the pool C1 amplified more distinct bands in samples of 6, 13, 16, and 21 than the pool N. The primers in pool N has lower degeneracy and GC content than those in pool C1. To figure out the amplification differences between two groups of primers, we extracted the DNA fragments from the above samples for sequencing and analyzed the GC content of the flanking sequences adjacent to AD primer binding sites. The results showed that two groups of sequences exhibited the similar GC variation from 20 to 50% (Supplementary Fig. S2), suggesting that the binding preference of C1, C2 and N were not closely related to GC content. In term of the hi-TAIL-PCR results, we obtained 51 (71%) of the T-DNA insertion sites from 72 samples with three primer pools (Table 1).

In brief, we increased the valid amplification of T-DNA flanking sequences via reducing the degeneracy of LAD primers in hi-TAIL-PCR, suggesting the lower-fold degeneracy pool N was more appropriate to screen the T-DNA insertions in *A. thaliana*. However, the relative low degeneracy may also decrease the probability of primer binding to possible sites. Therefore, N-LAD and LAD primers were both necessary to characterize the T-DNA insertion sites. The complementary use of two kinds of primer pools would supply a more comprehensive understanding of the distribution of T-DNA in the genome.

Acknowledgements The activation tagging vector, pBASTA-AT2, was a kind gift from Professor Jia Li, School of Life Sciences of Lanzhou University. This work was supported by Outstanding Youth Foundation of China (No. 31322008) to Z.H. and the National Natural Science Foundation of China (No. 31370088) to Z.F.

Author contribution HZ and WX did the experiments and contributed equally to this work. ZF and ZH designed the project and completed the manuscript.

Compliance with ethical standards

Funding This work was supported by Outstanding Youth Foundation of China (No. 31322008) to Z.H. and the National Natural Science Foundation of China (No. 31370088) to Z.F.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301(5633):653–657. <https://doi.org/10.1126/science.1086391301/5633/653>
- Chang Y, Long T, Wu C (2012) Effort and contribution of T-DNA Insertion mutant library for rice functional genomics research in China: review and perspective. *J Integr Plant Biol* 54(12):953–966. <https://doi.org/10.1111/j.1744-7909.2012.01171.x>
- Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16(6):735–743
- Hsia MM, O'Malley R, Cartwright A, Nieu R, Gordon SP, Kelly S, Williams TG, Wood DF, Zhao Y, Bragg J, Jordan M, Pauly M, Ecker JR, Gu Y, Vogel JP (2017) Sequencing and functional

- validation of the JGI *Brachypodium distachyon* T-DNA collection. Plant J. <https://doi.org/10.1111/tpj.13582>
- Kondou Y, Higuchi M, Matsui M (2010) High-throughput characterization of plant gene functions by using gain-of-function technology. Annu Rev Plant Biol 61:373–393. <https://doi.org/10.1146/annurev-arplant-042809-112143>
- Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sninsky JJ (1990) Effects of primer template mismatches on the polymerase chain-reaction—human-immunodeficiency-virus type-1 model studies. Nucleic Acids Res 18(4):999–1005. <https://doi.org/10.1093/Nar/18.4.999>
- Liu YG, Chen Y (2007) High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. Biotechniques 43(5):649–650, 652, 654 passim
- Liu YG, 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(3):674–681
- Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8(3):457–463
- Marsch-Martinez N, Greco R, Van Arkel G, Herrera-Estrella L, Pereira A (2002) Activation tagging using the En-I maize transposon system in Arabidopsis. Plant Physiol 129(4):1544–1556. <https://doi.org/10.1104/pp.003327>
- Mazars GR, Moyret C, Jeanteur P, Theillet CG (1991) Direct sequencing by thermal asymmetric PCR. Nucleic Acids Res 19(17):4783. <https://doi.org/10.1093/nar/19.17.4783>
- Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinzuka Y, Onosato K, Hirochika H (2003) Target site specificity of the Tos17 retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15(8):1771–1780
- Osborne BI, Wirtz U, Baker B (1995) A system for insertional mutagenesis and chromosomal rearrangement using the Ds transposon and Cre-lox. Plant J 7(4):687–701
- Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell 14(12):2985–2994
- Singer T, Burke E (2003) High-throughput TAIL-PCR as a tool to identify DNA flanking insertions. Methods Mol Biol 236:241–272. <https://doi.org/10.1385/1-59259-413-1-241>
- Stephen L, Dellaporta JW, Hicks James B (1983) A plant DNA mini-preparation: version II. Plant Mol Biol Rep 1:19–21
- Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. Plant Physiol 122(4):1003–1013
- Wu L, Di DW, Zhang D, Song B, Luo P, Guo GQ (2015) Frequent problems and their resolutions by using thermal asymmetric interlaced PCR (TAIL-PCR) to clone genes in Arabidopsis T-DNA tagged mutants. Biotechnol Biotechnol Equip 29(2):260–267. <https://doi.org/10.1080/13102818.2014.998161>