High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences

Yao-Guang Liu and Yuanling Chen
South China Agricultural University, Guangzhou, China

BioTechniques 43:649-656 (November 2007)
doi 10.2144/000112601

Isolation of unknown DNA sequences flanked by known sequences is an important task in molecular biology research. Thermal asymmetric interlaced PCR (TAIL-PCR) is an effective method for this purpose. However, the success rate of the original TAIL-PCR needs to be increased, and it is more desirable to obtain target products with larger sizes. Here we present a substantially improved TAIL-PCR procedure with special primer design and optimized thermal conditions. This high-efficiency TAIL-PCR (hiTAIL-PCR) combines the advantages of the TAIL-cycling and suppression-PCR, thus it can block the amplification of nontarget products and suppress small target ones, but allow efficient amplification of large target sequences. Using this method, we isolated genomic flanking sequences of T-DNA insertions from transgenic rice lines. In our tests, the success rates of the reactions were higher than 90%, and in most cases the obtained major products had sizes of 1–3 kb.

INTRODUCTION

DNA tagging by T-DNA and transposon insertions has become an important approach for studying functional genomics in plants. Large numbers of DNA-insertion lines and important mutations have been created in *Arabidopsis* and rice using this approach. To identify the genes tagged by DNA insertions, it is necessary to recover genomic sequences flanking the insertion tags. However, the tagged gene sequences cannot be obtained simply by regular PCR procedures because the genomic flanking sequences are unknown. So far, several PCR-based methods, such as inverse PCR (1,2), adapter-ligation-mediated PCR (3–6), and thermal asymmetric interlaced (TAIL)-PCR (7,8), have been developed for amplification of unknown DNA fragments flanked by known sequences. TAIL-PCR is a PCR-only method and is thus especially suitable for manipulating a large number of samples in manual or automation (7). With the advantages of simplicity and high efficiency, TAIL-PCR and its modified procedures have been widely used in a variety of biological research in various organisms, including large-scale determination of T-DNA and transposon insertion sites in *Arabidopsis* and rice (9,10), and isolation of upstream (promoters) and downstream sequences of the known coding sequences (11,12).

TAIL-PCR utilizes nested known sequence-specific primers with a melting temperature (T_m) >65°C in consecutive reactions together with a short (15–16 nucleotides) arbitrary degenerate (AD) primer with a T_m of about 45°C and 64–256 folds of degeneracy, so that the relative amplification efficiencies of target and nontarget products can be thermally controlled (7,8). In the primary TAIL-PCR of the original method, one low-stringency PCR cycle is conducted to create one or more annealing sites for the AD primer along the target sequence. Target product(s) are then preferentially amplified over nontarget ones that are primed by the AD primer alone by swapping two high-stringency PCR cycles with one that has reduced-stringency (TAIL-cycling). This is based on the principle that, in the high-stringency PCR cycles with high annealing temperatures (65°–68°C) only the specific primer with the higher melting temperature can efficiently anneal to target molecules. The AD primer is much less efficient at annealing due to its lower melting temperature. AD primers with higher degrees of degeneracy, or pooled AD primers (9), may have more chances to bind to the target sequences. However, this tends to produce undesired smaller products. To achieve high success rates in obtaining target sequences with larger sizes, we developed a high-efficiency TAIL-PCR (hiTAIL-PCR) procedure by using specially designed degenerate and known-sequence-specific primers.

MATERIALS AND METHODS

PCR Primers

The primers used for hiTAIL-PCR are shown in Figure 1.

Reagents

Genomic DNAs were prepared from transgenic rice lines as described (10), which were transformed by binary vector constructs based on pCAMBIA1305.1 and pCAMBIA1300 (Cambia, Canberra, Australia). *Ex Taq* DNA polymerase kit with 10× PCR buffer containing 20 mM $MgCl_2$ (Takara-Bio, Dalian, China) was used for the PCR.

PCR

Pre-amplification reactions (20 µL) were prepared, each containing 2.0 µL PCR buffer, 200 µM each of dATP, dCTP, dGTP, and dTTP (dNTPs), 1.0 µM of any one of the LAD primers (in the cases in which two LAD primers were used in single reactions, each was in 1.0 µM), 0.3 µM RB-0a or RB-0b, 0.5 U *Ex Taq*, and 20–30 ng transgenic rice DNA. Each 25-µL primary TAIL-PCR contained 2.5 µL PCR buffer, 200 µM each of dNTPs, 0.3 µM AC1 and RB-1a (or RB-1b), 0.6 U *Ex Taq*, and 1 µL 40-fold diluted pre-amplified product. Each secondary 25-µL TAIL-PCRs

```
                          NotI
LAD1-1: 5′-ACGATGGACTCCAGAGCGGCCGC(G/C/A)N(G/C/A)NNNGGAA-3′
LAD1-2: 5′-ACGATGGACTCCAGAGCGGCCGC(G/C/T)N(G/C/T)NNNGGTT-3′
LAD1-3: 5′-ACGATGGACTCCAGAGCGGCCGC(G/C/A)(G/C/A)N(G/C/A)NNNCCAA-3′
LAD1-4: 5′-ACGATGGACTCCAGAGCGGCCGC(G/C/T)(G/A/T)N(G/C/T)NNNCGGT-3′
AC1:    5′-ACGATGGACTCCAGAG-3′

RB-0a:  5′-GGCAATAAAGTTTCTTAAGATTGAATCCTGT-3′(258bp from RB)
RB-1a:  5′-ACGATGGACTCCAGTCCGGCCTGTTGCCGGTCTTGCGATGATTATCA-3′
RB-1ac: 5′-CCTGTTGCCGGTCTTGCGATGATTATCA-3′(233bp from RB)

RB-2a:  5′-GTAATGCATGACGTTATTTATGAGATGGGTT-3′(161bp from RB)

RB-0b:  5′-CGTGACTGGGAAAACCCTGGCGTT-3′(197bp from RB)
RB-1b:  5′-ACGATGGACTCCAGTCCGGCCCAACTTAATCGCCTTGCAGCACATC-3′
                                              (173bp from RB)
RB-2b:  5′-GAAGAGGCCCGCACCGATCGCCCTT-3′(119bp from RB)
```

Figure 1. Primers used for high-efficiency thermal asymmetric interlaced PCR (hiTAIL-PCR). RB-0a, RB-1a, and RB-2a are specific to pCAMBIA binary vectors (such as pCAMBIA-1305.1) having the Nos terminator sequence adjacent to RB. RB-0b, RB-1b, and RB-2b are specific to pCAMBIA-1300. In the sequence tag of RB-1a (RB-1b), two bases (underlined) were designed to differ from the 3′ end of AC1 in order to avoid the priming of AC1 on this sequence tag in the secondary TAIL-PCR, and in sequencing of the primary TAIL-PCR products from the AC1-end with AC1 as the sequencing primer.

Table 1. Thermal Conditions for hiTAIL-PCR

Pre-amplification			Primary TAIL-PCR			Secondary TAIL-PCR		
Step	Temperature (°C)	Time (min:s)	Step	Temperature (°C)	Time (min:s)	Step	Temperature (°C)	Time (min:s)
1	93	2:00	1	94	0:20	1	94	0:20
2	95	1:00	2	65	1:00	2	68	1:00
3	94	0:30	3	72	3:00	3	72	3:00
4	60	1:00	4	Go to Step 1	1 time	4	94	0:20
5	72	3:00	5	94	0:20	5	68	1:00
6	Go to Step 3	10 Times	6	68	1:00	6	72	3:00
7	94	0:30	7	72	3:00	7	94	0:20
8	25	2:00	8	94	0.20	8	50	1:00
9	Ramping to 72	0.5°C/s	9	68	1:00	9	72	3:00
10	72	3:00	10	72	3:00	10	Go to Step 1	6–7 Times
11	94	0:20	11	94	0:20	11	72	5:00
12	58	1:00	12	50	1:00	12	End	
13	72	3:00	13	72	3:00			
14	Go to Step 11	25 Times	14	Go to Step 5	13 Times			
15	72	5:00	15	72	5:00			
16	End		16	End				

hiTAIL PCR, high-efficiency thermal asymmetric interlaced PCR.

contained 2.5 µL PCR buffer, 200 µM each of dNTPs, 0.3 µM AC1 and RB-2a (or RB-2b), 0.5 U *Ex Taq*, and 1 µL 10-fold diluted primary TAIL-PCR product. When necessary, the primary or secondary TAIL-PCRs were scaled up to 50 µL. The PCRs were performed using a PCT-100 PCR cycler (MJ Research, Waltham, MA, USA) with thermal conditions shown in Table 1. The amplified products were analyzed on 1.0% agarose gels, and single fragments were recovered from the gels and purified using a DNA purification kit.

RESULTS AND DISCUSSION

Primer Design

Four relatively longer AD (LAD) primers of 33 or 34 nucleotides were designed for hiTAIL-PCR, which contained four fixed nucleotides at the 3′ ends, followed by degenerate nucleotides at the six or seven positions (2304- or 6912-fold degeneracy) (Figure 1). Of the degenerate nucleotide positions, two or three were designed to have three rather than four nucleotides, to avoid complete self-pairing between the 3′ ends. The arbitrary 3′ four-base sites have a moderate frequency (on average 256 bp a time) to anchor the LADs to the target sequences. A *Not*I restriction site was present in the primers as a choice for cloning of the target TAIL-PCR products in plasmid vectors containing a unique *Not*I site and other unique blunt-end restriction site(s). The different LAD primers shared a common sequence in the 5′ half. A 16-mer primer (AC1) specific to this sequence was prepared, which had a T_m of 52°C as calculated by the formula $T_m = 69.3 + 41 \times GC\% - 650/L$ (L = primer length) (13). Since the sequences of the LAD primers including their 3′ ends were arbitrarily designed, and especially, the four-base sites cannot form specificity for given genomes of organisms, the LAD primers are universally applicable for various organisms.

To isolate T-DNA insertion flanking sequences from transgenic plants

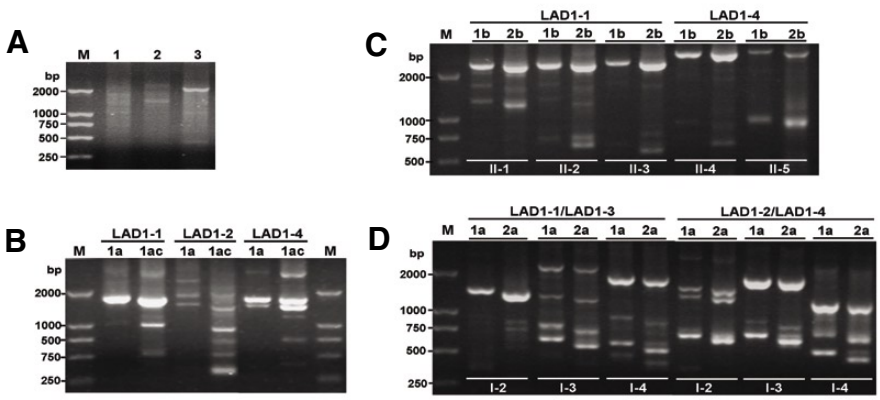

Figure 2. Amplification of T-DNA flanking sequences from transgenic rice plants by high-efficiency thermal asymmetric interlaced PCR (hiTAIL-PCR). (A) Pre-amplification of a transgenic line (I-1) using RB-0a combined with LAD1-1 (lane 1), LAD1-2 (lane 2) and LAD1-4 (lane 3), respectively. (B) Primary TAIL-PCRs of I-1 with RB-1a (1a) or RB-1ac (1ac), showing the effect of using RB-1a on suppressing the amplification of smaller target products. The LAD primers for the pre-amplification are indicated. (C) Analysis of the primary and secondary TAIL-PCR products obtained from the transgenic lines using RB-1b (1b) and RB-2b (2b) for the reactions, respectively. The corresponding primary and secondary products show the expected differential shift (74 bp). (D) hiTAIL-PCRs using pooled LAD primers in the pre-amplification reactions.

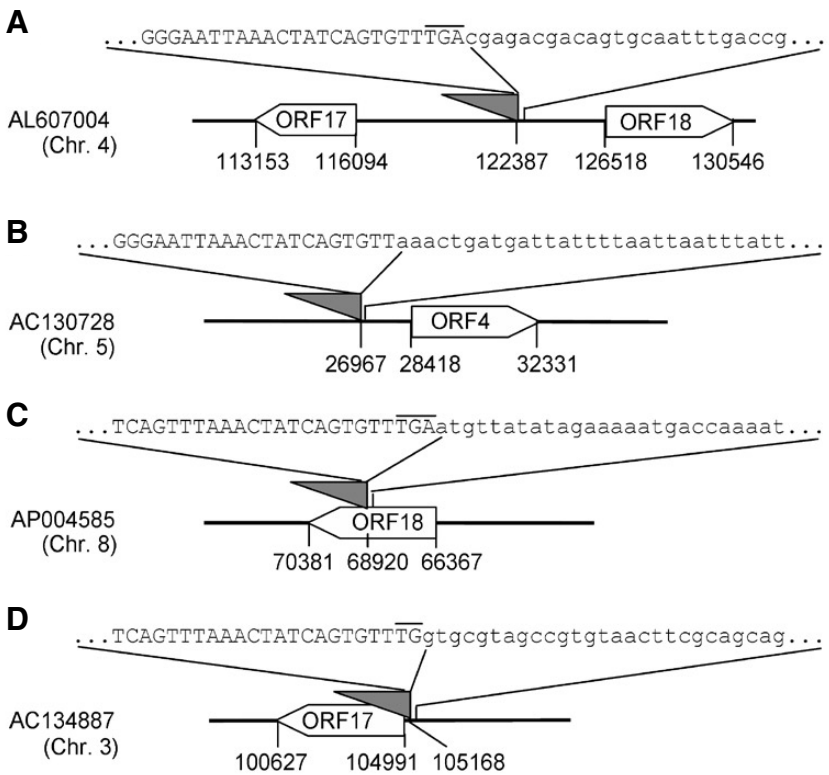

Figure 3. Characterization of T-DNA tagged sites in the rice genome. Diagrams indicate the T-DNA insertion sites in the rice genome of transgenic rice lines (A) I-1, (B) I-3, (C) II-2, and (D) II-4 as determined by sequencing of primary hiTAIL-PCR products using primers RB-2a (panels A and B) or RB-2b (panels C and D). The filled triangles indicate the T-DNA insertions, and the numbers show the positions (in base pairs) of the insertion sites and the predicted open reading frames in the bacterial artificial chromosome clones (BAC) [OSJNba0093F12 (panel A) and OSJNBa0029P07 (panel D)] or P1 artificial chromosome (PAC) clones [P0663C08 (panel B) and P0470B03 (panel C)]. The accession no. of the corresponding rice gene for each clone is indicated to the left of each diagram. The T-DNA sequences of the right border side from pCAMBIA1305.1 (panels A and B) or pCAMBIA1300 (panels C and D) are given in capital letters and the genomic flanking sequences in lower-case letters. Overlines show the retained nucleotides of the right border.

transformed with the pCAMBIA binary vectors, nested specific primer sets (RB-0a/RB-1a/RB-2a, RB-0b/RB-1b/RB-2b) with T_m >68°C were designed according to the sequences adjacent to the T-DNA right border (RB) (Figure 1). In RB-1a (RB-1b) for the primary TAIL-PCR, an additional sequence (21 nucleotides) that was identical (except for two bases) to the 5′ half of the LAD primer was tagged to the 5′ end, in order to produce a suppression-PCR effect (14). As a comparison, a specific primer (RB-1ac) without this sequence tag was prepared for the primary TAIL-PCR. A distance of 88 bp (or 74 bp) between RB-1a and RB-2a (or RB-1b and RB-2b) was set to facilitate the confirmation of the product specificity, if necessary, by the differential shift on agarose gels.

Rationale of hiTAIL-PCR

The RB-0a (RB-0b) and a LAD primer (or pooled LAD primers) are applied to the pre-amplification reaction. After 10 cycles of linear amplification of target sequences primed by RB-0a (RB-0b), which increases the copy number of target molecules, a single cycle with a low annealing temperature (25°C) is carried out. The higher degree of the primer degeneracy and the low annealing temperature allow the LAD primer to bind to the target sequence with a higher probability, thus efficiently creating one or more annealing sites for the AC1 primer. Then the primary TAIL-PCR is carried out using the nested specific primer RB-1a (RB-1b) and AC1. Since the DNA strands of specific products primed by RB-1a (RB-1b) and AC1 contain complementary ends (with a Melting temperature of 63°C), the relatively small ones (<500 nucleotides) tend to form a hairpin structure, which suppresses the PCR amplification, while amplification of larger products is less affected, because the longer distances between the complementary ends decrease the potential for the hairpin structure to form. The amplification of nontarget products primed by AC1 alone is of low efficiency during the TAIL-

cycling (7,8). Nontarget products from nonspecific priming by the specific primer RB-0a (RB-0b) alone, if any, are diluted and cannot be amplified in the following TAIL-PCRs using the nested specific primers. On the other hand, new nontarget products cannot be generated and amplified to visible levels from such diluted (approximately 1000-fold) templates by RB-1a (RB-1b) alone in the primary TAIL-PCR.

Effectiveness of hiTAIL-PCR

The hiTAIL-PCR procedure was used to isolate T-DNA insertion flanking sequences from transgenic rice lines. We first tested the efficiency of using single LAD primers in hiTAIL-PCRs. Six lines transformed with pCAMBIA1305.1 (I-1 to I-6) and nine lines with pCAMBIA1300 (II-to I9), which had single T-DNA insertion, were selected for the test in combination with the four LAD primers. The pre-amplification reactions produced smear bands (Figure 2A), which contained a large number of randomly amplified products from the rice genomic DNA. In some cases, the pre-amplified products were not detectable on the agarose gels (not shown), but target products could be obtained in the primary TAIL-PCR. In most hiTAIL-PCR, the obtained major target products had sizes of about 1–3 kb, and no products smaller than about 0.5 kb were produced (Figure 2, B–D), whereas the original and modified TAIL-PCR procedures usually produce fragments of approximately 0.2–1.5 kb (7–9). In comparison, the control reactions without the suppression-PCR effect produced some smaller products (Figure 2B). Of the 60 reactions, 56 produced target products, giving an average success rate of 93.3%. This is much higher than that (50%–70%) of the original TAIL-PCR procedure (8). All the four LAD primers worked well in our tests. Most primary TAIL-PCR reactions produced detectable target products. Therefore, in general the secondary TAIL-PCR can be omitted. In a few cases the primary TAIL-PCR products were at relatively low levels (data not shown); however, they could be amplified to high concentrations in the secondary TAIL-PCRs.

We also used LAD primer pools, LAD1-1/LAD1-3 and LAD1-3/LAD1-4, in the pre-amplification reactions to test if the amplification efficiency can be further increased. The results showed that 37 (97.4%) of the 38 reactions tested succeeded to produce target products. The reactions using the pooled LAD primers produced an average of 2.5 products, while those using single LAD primers gave an average of 1.8. The sizes of the products were similar to or somewhat smaller than those obtained by using single LAD primers, but those smaller than about 0.5 kb were blocked effectively (Figure 2D). Therefore, it is a good choice to use the LAD primer pools in the application of hiTAIL-PCR.

As described in the original TAIL-PCR (7,8), the differential shift between the primary and secondary products on agarose gels (Figure 2, C and D) is a good indicator of the product specificity. The specificity also was tested in some cases by control (primary or secondary) reactions with RB-1a (RB-1b) or RB-2a (RB-2b) alone and AC1 alone, which did not generate detectable products (data not shown). We further purified several primary hiTAIL-PCR products for direct sequencing. The result showed that all of the sequenced products contained the inserted T-DNA and its flanking genomic sequences of the rice genome. Several examples of the T-DNA tagged sites in the rice genome are shown (Figure 3).

Conclusion

We described a substantially improved TAIL-PCR procedure, hiTAIL-PCR, by special design of the arbitrary degenerate and specific primers. This method combined the advantages of the TAIL-cycling and suppression-PCR, thus greatly increased the success rate but avoided to produce small target fragments. This new version of TAIL-PCR thus can replace the original one in molecular biology research of various organisms.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Science and Technology of China.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Ochman, H., A.S. Gerber, and D.L. Hartl.** 1988. Genetic applications of an inverse polymerase chain reaction. Genetics *120*:621-623.
2. **Triglia, T., M.G. Peterson, and D.J. Kemp.** 1988. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res. *16*:8186.
3. **Mueller, P.R. and B. Wold.** 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. Science *246*:780-786.
4. **Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner, S. Powell, R. Anand, J.C. Smith, and A.F. Markham.** 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res. *18*:2887-2890.
5. **Rosenthal, A., D. Stephen, and C. Jones.** 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. Nucleic Acids Res. *18*:3095-3096.
6. **Jones, D.H. and S.C. Winistorfer.** 1992. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. Nucleic Acids Res. *20*:595-600.
7. **Liu, Y.-G. and R.F. Whittier.** 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.
8. **Liu, Y.-G., N. Mitsukawa, T. Oosumi, and R.F. Whittier.** 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. *8*:457-463.
9. **Sessions, A., E. Burke, G. Presting, G. Aux, J. McElver, D. Patton, B. Dietrich, P. Ho, et al.** 2002. A high-throughput *Arabidopsis* reverse genetics system. Plant Cell *14*:2985-2994.
10. **Miyao, A., K. Tanaka, K. Murata, H. Sawaki, S. Takeda, K. Abe, Y. Shinozuka, K. Onosato, et al.** 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*:1771-1780.
11. **Uozumi, N., E.J. Kim, F. Rubio, T. Yamaguchi, S. Muto, A. Tsuboi, E.P. Bakker, T. Nakamura, et al.** 2000. The

Arabidopsis HKT1 gene homolog mediates inward Na^+ currents in *Xenopus laevis* Oocytes and Na^+ uptake in *Saccharomyces cerevisiae.* Plant Physiol. *122*:1249-1259.

12. **Zhao, Y., S.E. Blumer, and G.W. Sundin.** 2005. Identification of *Erwinia amylovora* genes induced during infection of immature pear tissue. J. Bacteriol. *187*:8088-8103.
13. **Mazars, G.-R., C. Moyret, P. Jeanteur, and C.G. Theillet.** 1991. Directing sequencing by thermal asymmetric PCR. Nucleic Acids Res. *19*:4783.
14. **Diatchenko, L., Y.-F.C. Lau, A.P. Campbell, A. Chenchik, F. Mogadam, B. Huang, S. Lukyanov, K. Lukyanov, et al.** 1996. Suppression subtracive hybridization, a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA *93*:6025-6030.

Received 8 June 2007; accepted 11 September 2007.

Address correspondence to Yao-Guang Liu, Key Laboratory of Plant Functional Genomics and Biotechnology of Education Department, Guangdong Province, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China. e-mail: ygliu@scau.edu.cn