

Adaptor Ligation-Based Polymerase Chain Reaction-Mediated Walking

Linus S. Padegimas and Nancy A. Reichert¹

Department of Plant and Soil Sciences, Mississippi State University, Box 9555, Mississippi State, Mississippi 39762

Received February 12, 1998

An improved method of adaptor ligation PCR was developed for isolation of unknown sequences flanking a known DNA sequence. It was determined that the specificity of the adaptor ligation-based walking technique could be significantly enhanced by using uniquely blocked adaptors along with removal of unligated genomic DNA by exonuclease III digestion. This technique was utilized to isolate three novel promoter regions from three different *Zea mays* (maize) peroxidase genes. Sequences encoding a putative maize 6-phosphogluconate dehydrogenase gene were also isolated and confirmed by sequence analysis. The described improvements could be applied to other existing adaptor ligation-based PCR walking techniques. © 1998 Academic Press

Various approaches have been developed for isolation of unknown DNA fragments flanking a known sequence. Generally, PCR-based walking included use of vectorette (1), splinkerette (2), or other adaptors (3, 4) ligated to restriction-digested genomic DNA, followed by PCR amplification with adaptor-specific and gene-specific primers. However, those techniques have been limited by a relatively short walking distance, low specificity, low efficiency, and/or extensive cycling that contributed to a higher error rate in PCR products. A more advanced method was proposed by Siebert *et al.* (5) which employed long adaptors which formed pan-handle structures to help suppress nonspecific, end-to-end amplification by adaptor-specific primer ("suppression PCR"; 5–7). Use of "long and accurate" (LA²) PCR

(8) combined with two nested PCR rounds allowed the specific amplification of up to 6 kb of unknown sequence (5).

We have determined that the specificity of an adaptor ligation-based walking technique could be significantly increased by using uniquely blocked adaptors in concert with removal of unligated genomic DNA strands by exonuclease III (Exo III) digestion. This allowed the development of a precise PCR-mediated walking method which avoided nonspecific, end-to-end amplification.

MATERIALS AND METHODS

Enzymes. Advantage *Tth* polymerase mix (containing *Tth* DNA polymerase as primary polymerase, Vent DNA polymerase as minor component, *Tth*Start antibody, and *Taq*Start antibody) was supplied by Clontech. The following enzymes were supplied: KlenTaq1 DNA polymerase (Ab Peptides); Pfu DNA polymerase (Stratagene); Klenow fragment of *Escherichia coli* DNA polymerase I (Promega); polynucleotide kinase and restriction endonucleases *Eco*RV, *Ssp*I, and *Sca*I (New England Biolabs); *Bam*HI (USB); *Dra*I, T4 DNA ligase; and Exo III (Fermentas, Lithuania). KlenTaqLA-16 polymerase mix was prepared as described (7) with the exception that an equal volume of *Taq*Start antibody (1.1 µg/µl; Clontech) was added. RNase A (Sigma) was boiled for 20 min to remove DNase contamination (9).

Oligonucleotides. All oligonucleotides were supplied by Genosys in desalted form and were used without any additional purification. The following sequences of oligonucleotides (listed from 5' to 3') were used as adaptor DNA: longer strand (LN), CAGCA-GAACGACGCCCCGCGACAAGGACAGGT; shorter strand (Oligo1), ACCTGTCCTGCGAAAGCAsAsAsA ("s" depicts phosphorothioate linkages); and alternative shorter strand (Luk1), ACCTGTCsCsAsAsAA. Also supplied by Genosys were (listed 5' to 3') adaptor

¹ To whom correspondence should be addressed. Fax: (601)-325-8742. E-mail: nreichert@pss.msstate.edu.

² Abbreviations used: LA PCR, long and accurate PCR; Exo III, exonuclease III; RNase A, ribonuclease A; dNTP, deoxynucleotide triphosphate mix; s, depicts phosphorothioate linkages; RACE, rapid amplification of cDNA ends.

specific (APN; CAGCAGAACGACGCCCCGCCGACAA) and gene-specific (3Zea1; GACCGCATCAAGAAGGCG-TACGATAGGAAC) primers.

Plant genomic DNA isolation. Plant genomic DNA was isolated from *Zea mays* (maize; nematode resistant line Mp307) leaves using a scaled-up version of the procedure described in (10). Polysaccharides were removed by ethanol precipitation (11). DNA concentration was estimated fluorometrically using Hoechst 33258 dye.

Construction of uncloned genomic DNA libraries. Uncloned genomic DNA "library" construction was performed as described (5) with some modifications. Five libraries were prepared using five different restriction endonucleases: *Bam*HI, *Dra*I, *Eco*RV, *Sca*I, and *Ssp*I. For each library, 10 μ g maize genomic DNA was digested in 200 μ l reaction volume with 100 U of appropriate restriction enzyme for 4 h at 37°C using supplier-recommended buffer which also contained 0.1 mg/ml BSA and 0.3 mg/ml boiled RNase A. Following digestion with *Bam*HI, 8.0 μ l 2.5 mM dNTP and 2.0 μ l (10 U) Klenow fragment of *E. coli* DNA polymerase I were added and then further incubated for 15 min at room temperature. Five microliters of each reaction mix was loaded onto a 1.0% agarose gel to determine digestion efficiency. The remaining DNA was extracted twice with phenol/chloroform (1:1; v/v) and then precipitated by addition of 0.5 vol 6.0 M ammonium acetate, 20 μ g glycogen (Boehringer-Mannheim), and 2.5 vol ethanol. After mixing, the tubes were incubated for 20 min at -80°C and centrifuged at 16,000g in a microcentrifuge for 15 min. Each pellet was washed twice with 85% ethanol, air dried, and dissolved in 20 μ l H₂O.

The shorter strand adaptor oligonucleotide (1600 pmol) was phosphorylated using 40 U polynucleotide kinase in 40 μ l supplier-recommended buffer saturated with 1 mM ATP for 1 h at 37°C and then terminated by 10 min incubation at 75°C. The longer strand adaptor (LN; 1600 pmol) was added, and then the mixture was heated for 1 min at 75°C and incubated for 1 h at room temperature for annealing. Adaptors were extracted twice with phenol/chloroform and then precipitated by addition of 0.1 vol 3.0 M sodium acetate, pH 5.2, 20 μ g glycogen, and 3 vol ethanol following by 20 min incubation at -80°C and 20 min centrifugation in a microcentrifuge at 16,000g. The DNA pellet was washed twice with 85% ethanol, air dried, and dissolved in 20 μ l H₂O.

One-half of the digested genomic DNA (approximately 5.0 μ g) was ligated with 50 pmol adaptor for 20 h at 16°C using 10 U T4 DNA ligase in 20 μ l reaction volume using supplier-recommended buffer. The reaction was terminated by 15 min incubation at 75°C. It was then diluted with 180 μ l H₂O. Ten microliters of each genomic DNA mixture was separated for use later

as control templates for PCR. Exo III (1.0 μ l; 200 U) was added to the remainder and incubated for 20 h at 37°C, and then the reaction was terminated by 15 min incubation at 75°C. One exception was an experiment on aliquots (10 μ l) of the *Ssp*I library, where Exo III incubation times were varied prior to termination. Adaptor-ligated genomic libraries were stored at -20°C.

PCR. Each PCR (50 μ l total volume) contained 1 $\times$ *Tth* PCR buffer [40 mM Tris-HCl (pH 9.3 at 25°C), 15 mM potassium acetate; Clontech], 200 μ M dNTPs (Boehringer-Mannheim), 1.1 mM magnesium acetate, 0.2 μ M each APN and 3Zea1 primers, 1.0 μ l (approximately 25 ng) library DNA, and 1.0 μ l 50 $\times$ Advantage *Tth* polymerase mix. All reactions were carried out in thin-walled tubes. Most PCRs were conducted in the PTC-200 (M. J. Research) thermal cycler, programmed in calculated mode for 7 cycles at 2 s at 93°C and 3 min at 72°C, followed by 32 cycles at 2 s at 93°C, 3 min at 67°C, and 4 min extension for a last cycle. For some GC-rich templates it was necessary to increase denaturation temperature and time up to 95°C for 10–15 s.

Agarose gel analysis of PCR products. PCR products were analyzed using agarose (1.0–1.2%) gel electrophoresis in 1 $\times$ TAE buffer containing 0.1 μ g/ml ethidium bromide.

DNA molecular weight markers were supplied by Fermentas (Lithuania). The 1-kb ladder contained 12 fragments (kb: 0.5–4.0 kb in 0.5-kb increments plus 5.0, 6.0, 8.0, and 10 kb) with 11 in the 100-bp ladder (80 bp plus 100–1000 bp in 100-bp increments).

Subcloning and sequencing of PCR products. PCR products were isolated from the agarose gel as described (12) and ligated with pBluescript SK(-) (Stratagene) vector prepared as described (13).

Sequencing was performed using the Thermo Sequase radiolabeled terminator cycle sequencing kit (Amersham) following the supplier's instructions.

RESULTS

We previously determined that an amino group used to block the lower adaptor strand was not effective when the LA PCR technology was used. This could easily be explained by 3' $\rightarrow$ 5' proofreading activity of the minor component (usually *Pfu* or *Vent* DNA polymerase) present in polymerase mixes. However, phosphorothioate bonds had been determined to be relatively resistant to 3' $\rightarrow$ 5' proofreading activity of *Vent* and *Pfu* DNA polymerases (14). In addition, it had been demonstrated that splinkerettes–vectorettes containing a hairpin structure on the bottom strand could significantly enhance walking specificity (2). Therefore, we combined those two features in our walking adaptor design.

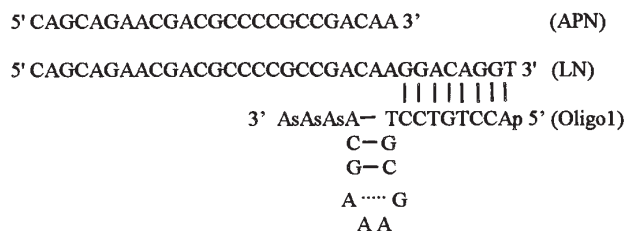


FIG. 1. Sequence and putative secondary structure of walking adaptor; adaptor-specific primer (APN) sequence.

The walking adaptor (used in this investigation; Fig. 1) contained an upper strand [LN; 33 nucleotide (nt) length] and lower strand (Oligo1; 21 nt length). The 5' end of the lower (shorter; Oligo1) adaptor strand was phosphorylated to enable ligation with genomic DNA. This adaptor strand was composed of three discrete regions (from 5' to 3'): an area (8 nt) complementary to the 3' region of the upper (longer; LN) adaptor strand, an area (10 nt) which formed an extremely stable hairpin structure (15), and an oligoA tail (4 nt; not complementary to upper adaptor strand) joined by phosphorothioate linkages which also prevented annealing to the upper adaptor strand. This 3' end could not be extended by DNA polymerase activity and therefore should be resistant to 3' → 5' proofreading activity as well as to Exo III activity (16). If the oligoA tail was removed during PCR (resistance is never absolute), the hairpin structure should initiate elongation along the lower strand. That resulting long hairpin-structured DNA would be not able to support PCR as a template.

In earlier experiments we prepared uncloned genomic DNA libraries as described under Materials and Methods, but with the exception that Exo III was not used. Unfortunately, one round of PCR produced a smear of nonspecific products regardless of the library used.

The following scenario could explain these results. Some library DNA fragments could contain adaptors which were only partially ligated (only one strand ligated). If only the upper adaptor (LN) strand was ligated, the genomic DNA could be elongated, followed by nonspecific amplification with APN (Fig. 2).

To determine if this scenario was plausible, we removed partially ligated and unligated genomic DNA by digestion with Exo III prior to PCR. It had been determined that Exo III digested in processive and synchronous (when added in excess) manners (17) and was demonstrated to remove up to 14.4 kb DNA sequence during 32 min incubation at 37°C (18). However, phosphorothioate linkages were resistant to Exo III (16) which should have ensured the “safety” of Oligo1-ligated template DNA.

As mentioned previously, each library (approximately 5.0 µg genomic DNA ligated with walking adaptor) had been digested with 200 U Exo III at 37°C for 20 h.

Figure 5 demonstrated that digestion with Exo III for only 30 min significantly reduced nonspecific background, whereas digestion for 2 or more hours removed nonspecific background completely (Fig. 3).

Each of the five adaptor-ligated libraries produced a single DNA band following PCR with APN and 3Zea1 primers (Fig. 4). These bands were determined to be overlapping products by restriction analysis (Fig. 5). To confirm the walking specificity of the generated PCR products, two (*EcoRV* and *SspI* products) were subcloned into the pBluescript SK(–) vector and completely sequenced. Both cloned sequences corresponded to the 3' end of a maize 6-phosphogluconate dehydrogenase gene which also included flanking non-coding sequences.

DISCUSSION

We have developed an adaptor ligation PCR approach to provide an improved method for isolation of unknown sequences flanking a known sequence. The

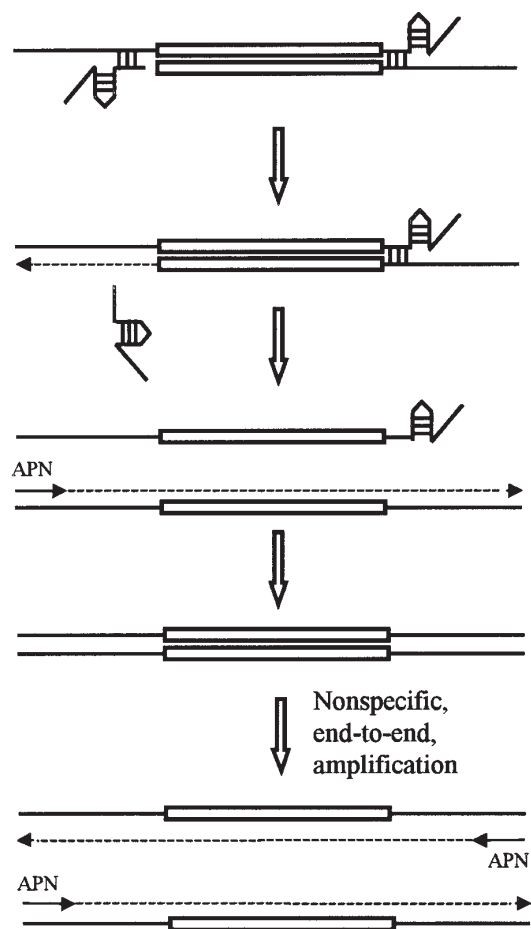


FIG. 2. Schematic diagram of nonspecific, end-to-end amplification as proposed in the scenario described in text.

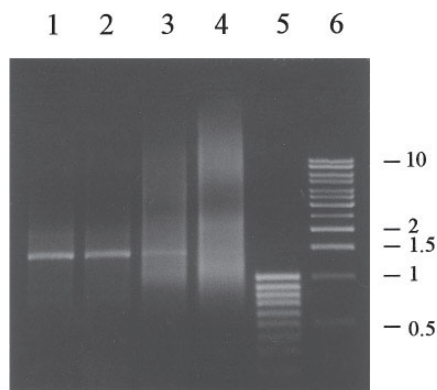


FIG. 3. Dependence of PCR walking specificity on Exo III digestion times. PCR walking was performed as described under Materials and Methods using a 1.0- μ l *Ssp*I library treated with Exo III for 20 h (lane 1), 2 h (lane 2), 0.5 h (lane 3), and 0 h (lane 4). Lanes 5 and 6 contained 100-bp and 1-kb DNA ladders, respectively.

powerfulness of this method can be illustrated by the following examples. Based on sequence information obtained by using the gene expression screen technique (19), we recovered the 5' and 3' ends of a putative maize 6-phosphogluconate dehydrogenase gene. In total, 3.99 kb was sequenced which included coding plus flanking 5' and 3' sequences³ (manuscript in preparation). Additional flanking noncoding sequences were obtained, but sequence determinations were not conducted. In the second case, we wanted to isolate peroxidase promoters from maize for use in transgenic plant analyses. Therefore, we designed a gene-specific primer based on 5' sequence homology of known plant peroxidase gene sequences. We isolated three PCR products which contained 5' coding sequences plus attached promoter regions from three different maize peroxidase genes⁴ (manuscript in preparation).

The primary innovations of the method described above are specific blockage of the lower adaptor strand plus the Exo III digestion step. In subsequent investigations, we have determined that the hairpin structure included in the lower adaptor strand is not absolutely required. We have obtained similar success, but with slightly less specificity, when we utilized an alternative lower adaptor strand, Luk1 (5'-ACCTGTCsCsAsAs-AA-3').

The Exo III digestion step is essential. Despite the demonstrated success with a shorter digestion time (Fig. 3), prolonged incubation (longer than 2 h) with excess enzyme is frequently required. Even trace amounts of partially ligated DNA with the adaptor could cause significant nonspecific background. Therefore, we recommend use of prolonged (20 h) incubation with an excess of Exo III.

³ GenBank Accession Nos. AF037036 and AF037037.

⁴ GenBank Accession Nos. AF037031, AF037032, and AF037035.

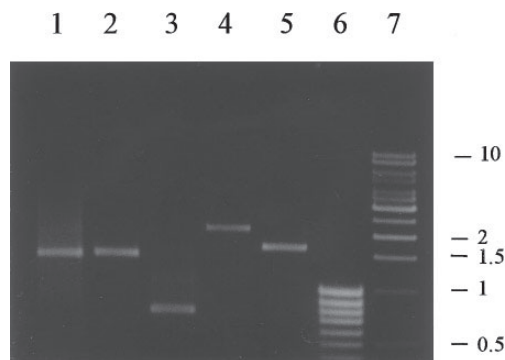


FIG. 4. PCR walking in *Ssp*I, *Bam*HI, *Dra*I, *Sca*I, and *Eco*RV adaptor-ligated libraries (lanes 1–5, respectively) treated with Exo III for 20 h. Lanes 6 and 7 contained 100-bp and 1-kb DNA ladders, respectively.

We have determined that use of Advantage *Tth* polymerase mix is also essential. Use of KlenTaqLA-16 polymerase (8) saturated with Advantage *Taq* antibody gave poor results compared to *Tth* polymerase mix. We have also tried to use the KlenTaqLA-16 polymerase with PC2 (8) and *Tth* buffers which contained 1.1 or 3.5 mM magnesium acetate without success. This reinforces the requirement of PCR optimization for each polymerase mix.

In the majority of cases we obtained a single specific band, but sometimes nonspecific bands would also be present. In all the investigated cases, sequence analysis revealed that it was a different kind of nonspecificity, i.e., not end-to-end amplification by an adaptor-specific primer (APN). Nonspecific bands resulted from mispriming of the gene-specific primer which resulted in incorrect walking. In this case we recommend using a second round of PCR for 25–30 cycles (5) employing a

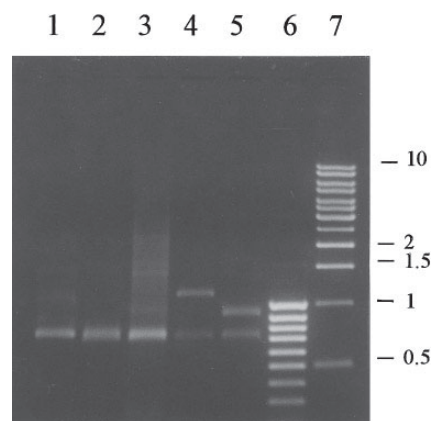


FIG. 5. Walking products obtained from *Bam*HI, *Ssp*I, *Sca*I, and *Eco*RV adaptor-ligated libraries (lanes 2–5, respectively) after digestion with *Dra*I restriction endonuclease. Lane 1 contained walking product obtained from the *Dra*I library. Lanes 6 and 7 contained 100-bp and 1-kb DNA ladders, respectively.

nested gene-specific primer along with APN. An alternative gene-specific primer may also be used instead of nested PCR.

As displayed in this report, the construction of adaptor-ligated libraries is not limited to use of blunt-end-producing restriction endonucleases. We have successfully used *Bam*HI (see Fig. 4, lane 2) and *Hind*III (data not shown) restriction endonucleases followed by filling in the recessed 3'-termini prior to adaptor ligation.

The method described above should be amenable to a number of research goals. It can be used to isolate promoters as well as full gene sequences, based on limited sequence information obtained by differential display (20) or other gene-screening techniques. This method should also be extremely useful in genome-sequencing programs.

We have found that the use of a specially blocked adaptor followed by removal of unligated DNA strands by Exo III is more effective compared to suppression PCR (5, 6). Therefore, we believe that the innovations described above could also improve various adaptor ligation-based PCR techniques such as rapid amplification of cDNA ends (RACE; 21) and PCR subtraction (22, 23).

ACKNOWLEDGMENTS

Funds were provided by the Mississippi Agricultural and Forestry Experiment Station (MAFES). We gratefully acknowledge receipt of these funds under MAFES' Adaptive Biotechnology and Enhanced Feed Grains Initiatives. This paper was approved for publication as MAFES Journal Article No. J9334.

REFERENCES

1. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., and Markham, A. F. (1990) *Nucleic Acids Res.* **18**, 2887–2890.
2. Devon, R. S., Porteous, D. J., and Brookes, A. J. (1995) *Nucleic Acids Res.* **23**, 1664–1645.
3. Fors, L., Saavedra, R. A., and Hood, L. (1990) *Nucleic Acids Res.* **18**, 2793–2799.
4. Rosenthal, A., and Jones, D. S. C. (1990) *Nucleic Acids Res.* **18**, 3095–3096.
5. Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A., and Lukyanov, S. A. (1995) *Nucleic Acids Res.* **23**, 1087–1088.
6. Lukyanov, S. A., Gurskaya, N. G., Lukyanov, K. A., Tarabykin, V. S., and Sverdlov, E. D. (1994) *J. Bioorg. Chem. (Russia)* **20**, 386–388.
7. U.S. Patent 5,565,340.
8. Barnes, W. M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220.
9. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *in* Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10. Padegimas, L., Shulga, O. A., and Skryabin, K. G. (1993) *Mol. Biol. (Russia)* **27**, 583–585.
11. Michaels, S. D., John, M. C., and Amasino, R. M. (1994) *BioTechniques* **17**, 274–275.
12. Zhen, L., and Swank, R. T. (1993) *BioTechniques* **14**, 894–898.
13. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1990) *Nucleic Acids Res.* **19**, 1154.
14. Skerra, A. (1992) *Nucleic Acids Res.* **20**, 3551–3554.
15. Sandusky, P., Wooten, E. W., Kurochkin, A. V., Kavanaugh, T., Mandecki, W., and Zuiderweg, E. R. P. (1995) *Nucleic Acids Res.* **23**, 4717–4725.
16. Putney, S. D., Benkovic, S. J., and Schimmel, P. R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7350–7354.
17. Wu, R., Ruben, G., Siegel, B., Jay, E., Spielman, P., and Tu, C. D. (1976) *Biochemistry* **15**, 734–740.
18. Henikoff, S. (1984) *Gene* **28**, 351–359.
19. Wang, Z., and Brown, D. D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11505–11509.
20. Liang, P., and Pardee, A. B. (1992) *Science* **257**, 967–971.
21. Chenchik, A., Diatchenko, L., Moqadam, F., Tarabykin, V., Lukyanov, S., and Siebert, P. D. (1996) *BioTechniques* **21**, 526–534.
22. Diatchenko, L., Lau, Y.-F. C., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., and Siebert, P. D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 6025–6030.
23. Gurskaya, N. G., Diatchenko, L., Chenchik, A., Siebert, P. D., Khaspekov, G. L., Lukyanov, K. A., Vagner, L. L., Ermolaeva, O. D., Lukyanov, S. A., and Sverdlov, E. D. (1996) *Anal. Biochem.* **240**, 90–97.