BENCHMARKS

Benchmarks are brief communications that describe helpful hints, shortcuts, techniques or substantive modifications of existing methods.

Extender PCR: A Method for the Isolation of Sequences Regulating Gene Expression from Genomic DNA

BioTechniques 26:804-806 (May 1999)

The isolation and characterization of non-coding genomic DNA sequences flanking coding regions is critical for the analysis of gene transcriptional control mechanisms. In this report, we describe a new polymerase chain reaction (PCR)-based method for "walking" into previously uncloned regions of genomic DNA that negates the need for synthesis of double-stranded vector linkers (2), blocked oligonucleotide adaptors (3) or genomic DNA library construction and screening. The method relies on both the ligation of a single-stranded oligonucleotide adaptor to restriction enzyme-digested genomic DNA and the blocking of nonspecific replication of the adaptor-complementary strand by incorporation of a dideoxynucleotide. Figure 1A summarizes the *Taq* DNA polymerase-catalyzed extension from internal gene-specific sites that generates the complementary strand of the ligated adaptor sequence, thus producing the primer annealing sites necessary for the amplification of the desired target sequences. Extender PCR can be used to clone both upstream (5') and/or downstream (3') regulatory regions using antisense or sense internal gene-specific

primers, respectively.

A large molar excess (200:1) of a single-stranded 'adaptor' oligonucleotide, 5'-TGCGAGTAAGGATCC-TCACGCAAGGAATTCCGACCAG-ACANNNN-3', in which NNNN denotes nucleotides complementary to the restriction site used, is annealed to the overhanging ends of genomic DNA fragments that had been produced by single digests with restriction enzymes that leave unique 4-bp 3' overhanging ends (e.g., *Pst*I, *Nsi*I, *Aat*II or *Sph*I; Figure 1Bi). Annealing restriction digested genomic DNA with adaptor primer is performed in ligation buffer by slow cooling in a thermal cycler, followed by overnight ligation with T4 DNA ligase. Unligated adaptors are then removed by an isopropanol precipitation of the DNA. Using this technique, individual 'libraries' for each restriction digest are generated that provide DNA for use in gene-specific PCR amplifications.

Restriction enzymes generating 5' overhanging ends can also be used with this technique (Figure 1A, Bii). Double-stranded genomic DNA digested with any appropriate enzyme (e.g., *Eco*RI or *Bam*HI) is annealed to an excess of a 1:1 mixture of the adaptor oligonucleotide and a 12-mer that has both a 5' end complementary to the 4-bp 5' overhang and a 3' end complementary to the 3' end of the adaptor sequence (e.g. 5'-AATTTGCATGTC-3' for *Eco*RI-digested DNA; see Figure 1Bii). Temporary annealing of the 12-mer to the restriction enzyme-generated 5' overhang produces a complementary site for adaptor sequence annealing (note: this 12-mer is not ligated to the genomic DNA because synthetic oligonucleotides with non-phosphorylated 5' ends are used). Ligation can then occur between the 3'-hydroxyl group of the adaptor sequence and the 5'-phosphate groups of the genomic DNA.

A key feature of this technique is the use of a novel 'blocking' procedure to prevent the production of adaptor primer annealing sites at the ends of all genomic DNA fragments during the initial PCR cycle. A dNTP-ddATP mixture is added to the adaptor-ligated genomic DNA in *Taq* DNA polymerase buffer, and nucleotide incorporation is performed using *Taq* DNA polymerase

until addition of the first ddATP molecule prevents further chain extension (Figure 1Biii). After removal of unincorporated nucleotides, by isopropanol precipitation, the blocked DNA is used as template DNA in PCR using a gene-specific primer, usually derived from a cDNA sequence, and the adaptor primer P1 (5'-TGCGAGTAAGGATC-CTCACGCA-3'). In the initial cycle of PCR, only 3'-hydroxyl groups of the annealed gene-specific primer are available for extension, ensuring that P1 priming sites are synthesized only at the ends of the DNA fragments of interest. Successive rounds of PCR between the adaptor primer P1 and the gene-specific primer amplify the desired sequence. Gene-specific primers of either orientation can be used for the amplification of DNA located 5' or 3' to the known sequence. The option of a second PCR using nested gene-specific primers and the adaptor primer P2 (5'-CGCAAGGAATTCCGACCAGACA-3'), allows further specific amplification of the desired flanking DNA sequences.

To demonstrate this principle, 5 µg of genomic DNA from the mantis shrimp *Gonodactylus oerstedii* was restriction-digested with 10 U of *Nsi*I at 37°C for 12 h, purified and resuspended in 10 µL of ligation buffer. The resuspended DNA was mixed with 740 pmol of the appropriate adaptor primer (containing a 3' end complementary to the *Nsi*I restriction overhang: 5'-TGCA-3') heated to 55°C for 10 min and cooled to 10°C over 1.5 h. Ligation was performed overnight at 14°C with 40 U T4 DNA ligase (Igi Fermentas, Sunderland, UK). The DNA was precipitated with isopropanol, resuspended in *Taq*Supreme™ (Igi Fermentas) buffer [50 mM Tris-HCl (pH 9.1), 16 mM (NH₄)₂SO₄, 3.5 mM MgCl₂, 150 µg/mL bovine serum albumin (BSA)] containing a dNTP-ddATP mixture (200 µM each of dCTP, dGTP, dTTP and ddATP), 25 U *Taq*Supreme, and the 3' ends were blocked (68°C, 5 min) by ddATP incorporation. The blocked, adaptor-ligated template DNA was phenol-extracted, precipitated and the resulting DNA pellet was dissolved in 50 µL TE buffer. Amplification of target sequences was carried out using 0.2 µg of adaptor-ligated template DNA,

20 pmol of adaptor primer P1 combined with 20 pmol of an opsin-specific primer, OP1 (5'-AGTTCAGTATAGT-GCGTATT-3'), 1 U *Taq*Supreme and 0.08 U *Pfu* DNA polymerase. Reaction conditions were 94°C for 5 min, 94°C for 30 s, 58°C for 30 s and 68°C for 6 min for a total of 30 cycles with a final cycle of 68°C for 10 min. A single band of 1.9 kb (Figure 2A) was excised from the gel, cloned into pTAG (R&D Systems, Abingdon, UK) and subjected to

automated sequencing. The 101 bp of sequence directly upstream of the opsin primer site was identical to that predicted from the 5' end of the *G. oerstedii* opsin cDNA sequence (GenBank® Accession No. AF057015), followed by approximately 1.8 kb of novel sequence that included a TATA box sequence 26 bp upstream of the first nucleotide of the cloned cDNA, indicating that this was likely to be the promoter region of the gene.

A second upstream walk was performed, this time using *Hae*II, *Aar*II and *Sph*I-restriction-digested DNA samples, primer P1 and gene-specific primer OP2 (5'-ACTGGCTACTCC-AGGATTTACCAC-3'); followed by PCR using the nested-adaptor primer P2 and gene primer OP3 (5'-TTA-GGGCAGTGCTATCGACT-3' under identical conditions except that 1 µL of a 1/100 dilution of the primary PCR products was used as the template. This

Figure 1. An outline of the extender PCR protocol. (A) Genomic DNA is digested with restriction enzymes that produce double-stranded DNA fragments with either 5'- or 3'-overhanging ends (parallel black lines). The restriction fragments are ligated to both a 45-mer linker oligonucleotide containing sequences that are identical to primers P1 (diagonal striped bar) and P2 (stippled bar) and a 3' end complementary to the restriction site of the enzyme used. In the case of DNA with 5'-overhanging ends a 12-mer linker/adaptor (black bar) is used to produce a temporary 3'-overhanging end to allow the adaptor to anneal. Incorporation of ddATP (asterisks) by *Taq* DNA polymerase blocks the 3' ends of all the DNA fragments. This DNA is then used as a template for PCR with adaptor primer P1 and a gene-specific primer. During the first PCR cycle, only extension from the unblocked 3' ends of the gene-specific primer can occur (zig-zag line), thus creating the complementary sites for P1 primer annealing exclusively on the ends of the target sequences. Further amplification with a second, nested gene-specific primer and the nested P2 adaptor primer can be performed to enhance PCR specificity. (B) Schematic diagrams detailing the ligation and 3' blocking steps outlined in Panel A. (Bi) Restriction fragments of genomic DNA digested with enzymes that generate 3'-overhanging ends (e.g., *Nsi*I) anneal to the complementary end of the extender oligonucleotide (black bar). Positions of P1 and P2 primer sequences are given above. Ligation occurs between the phosphate-bearing 5' end of the genomic DNA and the hydroxyl-bearing 3' end of the adaptor oligonucleotide. (Bii) Restriction fragments of genomic DNA digested with enzymes that generate 5'-overhanging ends (e.g., *Eco*RI) anneal to the adaptor linker, which in turn anneals to the 3' end of the extender oligonucleotide. Ligation occurs between the 5' end of the genomic DNA and the 3' end of the extender oligonucleotide, but not between the 3' end of the genomic DNA and the unphosphorylated 5' end of the adaptor linker. (Biii) All 3' ends are blocked by the incorporation of ddATP during incubation of the ligation products from Bi and Bii with *Taq* DNA polymerase and a dNTP:ddATP mixture. *Taq* DNA polymerase extends the 3' ends of the genomic DNA restriction fragments until the first ddATP is incorporated (H3'A).

Figure 2. Amplification of a *G. oerstедii* opsin promoter sequence by extender PCR. (A) Agarose gel of a 1.9-kb PCR product (lane 1) amplified from *NsiI*-digested, extender-adapted genomic DNA using primer P1 and gene-specific primer OP1. Marker lane (M) is *HindIII*-restriction-digested λ phage DNA. (B) Agarose gel of PCR products generated by initial PCR amplification using P1 and gene-specific primer OP2, followed by nested PCR with P2 and OP3 from *HaeII*, *SphI* and *AatII* (Igi Fermentas; lanes 1–3, respectively) restriction-digested extender-adapted genomic DNA. The marker (M) is *HindIII*-restriction-digested λ phage DNA. (C) Restriction map of the upstream untranscribed DNA sequences (open bar) of a *G. oerstедii* opsin gene (black bar), generated by extender PCR. The putative transcriptional start site is shown as +1. Positions of restriction sites used to generate four fragments of the promoter region are shown above (A, *AatII*; S, *SphI*; H, *HaeII*; N, *NsiI*). Gene-specific primer sites are shown below (labeled arrows). Each PCR product is labeled with the primers used to generate it and the lane to which it corresponds in parts A and B.

generated products of approximately 0.4, 1.0 and 0.7 kb, respectively (Figure 2B, lanes 1–3) that extended the 5' sequence by an additional 1.0 kb to give a total of approximately 2.8 kb of sequence upstream of the transcriptional start site (Figure 2C).

The same adaptor-ligated DNA has also been used to isolate both the proximal promoter elements from two other *G. oerstедii* opsin genes. In additional experiments, we have also successfully isolated the 5'-untranslated region and initiation codon of an opiate receptor-like gene from *Lymnaea stagnalis*, whereas conventional approaches (e.g., 5' rapid amplification of cDNA ends [RACE]) had failed. If used in conjunction with long-range PCR protocols (1),

this technique could also rapidly amplify fragments of several kilobase pairs in length, negating the need to construct and screen genomic DNA libraries. It will be particularly useful for the isolation of promoter regions from information contained within expressed sequence tag (EST) databases. This technique is also compatible with direct dye-terminator sequencing of gel-purified, P1-PCR-amplified products using the nested-adaptor primer P2.

REFERENCES

1. Barnes, W.M. 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. USA* 91:2216-2220.
2. Rosenthal, A. and D.S.C. Jones. 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucleic Acids Res.* 18:3095-3096.
3. Siebert, P.D., A. Chenchik, D.E. Kellogg, K.A. Lukyanov and S.A. Lukyanov. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23:1087-1088.

Address correspondence to Dr. Alastair J.H. Brown, Novartis Institute of Medical Research, 5 Gower Place, London, England, UK, WC1E 6BN. Internet: alastair.brown@pharma.novartis.com

Received 26 August 1998; accepted 17 February 1999.

Alastair J.H. Brown, Stephen J. Perry, Susan E. Saunders and Julian F. Burke
University of Sussex
Brighton, England, UK