Nucleic Acids Research, 1993, Vol. 21, No. 9 2263–2264

Direct sequencing of single primer PCR products: a rapid method to achieve short chromosomal walks

Gavin R.Screaton, Charles R.M.Bangham and John I.Bell
Molecular Immunology Group, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

Received March 15, 1993; Revised and Accepted April 2, 1993

One of the limitations of PCR to analyse a region of interest is that it requires knowledge of the flanking sequences in order to design amplification primers. Several adaptations of PCR have been described which allow some unknown sequences adjacent to regions of known sequence to be amplified. Primers directed to repetitive sequences, for instance Alu repeats can be used to pair with a primer specific to the region of interest (1), frequently however the nearest repeat lies too far away to allow PCR to proceed. In another approach known as inverse PCR the target DNA is circularised before amplification, allowing extension to proceed outwards from the region of interest (2).

We have frequently observed a PCR artefact where a single primer generates products when used in a PCR reaction alone. This artefact has been previously reported and used to clone complementary sequences from genomic DNA (3). By combining this technique with direct sequencing of pooled PCR products we have improved the yield of the PCR reactions and by cutting out the cloning steps we have greatly speeded up the method. We have used this technique to make short chromosomal walks in the Human CD44 locus from a yeast artificial chromosome (YAC) (see below), and also in our laboratory to obtain the promoter sequence preceeding the TCRBV2S1 gene in the human T-cell receptor locus from genomic DNA. This technique will have useful applications in sequencing DNA flanking regions of known sequence such as the ends of yeast artificial chromosomes, DNA flanking the insertion of transgenes or retroviruses, and sequencing across intron–exon boundaries and promoters from genomic DNA.

The technique utilises a commonly observed artefact in PCR, mispriming, where an oligonucleotide primer binds to an imperfect complementary sequence, usually under non-stringent conditions, to generate several products of different lengths. It is likely, *a priori*, that some of these products will result from amplification between the true complementary binding site at one end (PI, Figure 1) and a false priming site at the other end (FP1–FP3, Figure 1), in addition to the products generated by false priming at both ends (e.g. FP3 with FP4 in Figure 1). The result of these events is usually the generation of a number of products of different lengths which will cover the region of interest if they contain the true priming site at one end. Mispriming is encouraged by performing the PCR reaction at a number of different annealing temperatures well below the optimal annealing temperature for the given oligonucleotide. Even under these conditions mispriming will be a relatively rare event, and to potentiate the artefact 60 cycles of PCR are performed. A second primer S1 is used as a sequencing primer to provide specificity for those products containing the target sequence. In fact the specificity is such that the products of a number of reactions performed under different annealing conditions

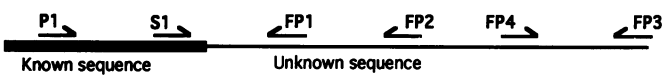

Figure 1. Schematic diagram showing an area of known sequence adjacent to a region of unknown sequence. P1 is the true priming site for the oligonucleotide to be used in the single primer PCR reactions and FP1–FP4 represent potential mispriming sites. S1 is the oligonucleotide used to provide specificity when sequencing the products of the single primer reactions.

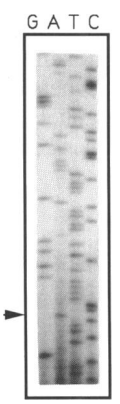

Figure 2. Sequence of the last exon of CD 44 the polyadenylation site is marked with an arrow. 100 ng of total yeast DNA was used in a 50 μl reaction containing Tris–HCl (10 mM pH 8.4), KCl (50 mM), $MgCl_2$ (2.5 mM), dNTPs (250 μM), 2 units of Taq polymerase and the primer 5'GGTGTGGGCAGAAGAAAAA-GCTAG 400 nM. Cycle conditions were 95°C 5 min, 60°C 30 sec, 72°C 2 min for 1 cycle followed by 95°C, 30 sec, 60°C 30 sec, 72°C 2 min for 60 cycles. parallel reactions were performed with annealing temperatures of 45 and 55°C. Products were pooled and passed through a spin column (*Chromaspin*, Clontech, USA) to remove unincorporated primer and nucleotides. The products were then sequenced using a modified cycle sequencing method described by Murray (4). Briefly 3 μl of template was used in each of 4×10 μl reactions containing Tris–HCl (30 mM pH 9), $MgCl_2$ (5 mM), 0.5 pmol of ^{32}P end labelled sequencing primer 5'TTTGTCAGAGGCACGAAAAGGGTTT, 10 μM each of 7 deaza-dGTP, dATP, dTTP and dCTP and either 30 μM dd GTP, 350 μM ddATP, 450 μM ddTTP or 175 μM ddCTP. The cycle parameters were 95°C 5 min, 63°C 30 sec, 72°C 2 min for 1 cycle followed by 95°C 30 sec, 63°C 30 sec 72°C 2 min for 30 cycles. The reactions were stopped with 5 μl of formamide with tracking dyes and run on a 6% denaturing polyacrylamide sequencing gel.

containing multiple PCR bands can be pooled and sequenced together. Pooling of PCR reactions performed under different annealing temperatures is useful as it cuts down the number of sequencing reactions which need to be performed. The yield from any given annealing temperature is not predictable between different oligonucleotides, i.e. one oligonucleotide may work best at 45°C whilst another may work best at 60°C.

The application of this technique to sequence genomic DNA across splice sites and polyadenylation sites in the human CD44 gene is shown in Figure 2. The template used was total yeast DNA from a culture carrying the CD44 gene contained on a YA-C. Three separate PCR reactions were carried out with different annealing temperatures using a single primer. After amplification the reactions were pooled and directly sequenced using an internal primer to give the sequence shown. Details of the reaction and primers are given in the figure legend to Figure 2.

This method has advantages over similar strategies to amplify unknown flanking sequences. These include its speed and the ability to sequence pooled PCR products without the need for cloning and screening to remove spurious bands which make up the majority of products produced by similar methods such as Alu PCR. In addition, the direct sequencing of these PCR products, as opposed to cloning and sequencing, eliminates polymerase errors which would be expected to occur frequently due to the large number of cycles (60) employed in the reactions. In our hands using a YAC as template it gives reliable results, and of the oligonucleotide primers that we used in the single primer PCR reactions, about 50% generated sufficient product, covering the DNA of interest, to allow us to obtain a sequence with an internal primer. Within the CD44 gene 4 out of 5 walks were successful yielding sequence for 2 splice sites and 2 polyadenylation sites. Others in the laboratory have used the technique to walk in the human T-cell receptor locus using genomic DNA generated sequence. The technique will have widespread applicability to sequence the ends of YAC clones as well as splice sites, promoters and unknown insertions in genomic DNA.

ACKNOWLEDGEMENTS

This work was supported by the Wellcome Trust, G.R.S. is an MRC Training Fellow.

REFERENCES

1. Nelson,D.L., Ledbetter,S.A., Corobo,L., Victoria,M.F., Ramirez-Solis,R., Webster,T.D., Ledbetter,D.H. and Caskey,C.T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6686–6690.
2. Silver,J. and Kerrikatte,V. (1989) *J. Virol.* **63**, 1924.
3. Parks,C.L., Cheng,L.-S. and Schenk,T. (1991) *Nucleic Acids Res.* **19**, 7155–7600.
4. Murray,V. (1989) *Nucleic Acids Res.* **17**, 8889.