Exon mapping by PCR

Limin Niu and Gray F.Crouse*

Department of Biology, Emory University, Atlanta, GA 30322, USA

Received November 9, 1992; Revised and Accepted December 7, 1992

Once the cDNA sequence of a mRNA is known and genomic clones of the gene are available, it becomes a fairly easy task to find and sequence the exons of the gene. One can make oligonucleotide primers based on the cDNA sequence and use those primers to find subclones of the gene containing exon sequences. The sequence of a portion of the exon can then be determined with the same oligonucleotide. However, if the entire gene is not to be sequenced, it can prove difficult to determine the precise location of the exon. We present a quick method for determining the precise location of an exon when the sequence of the surrounding DNA is not known.

We have recently sequenced all 26 exons of the mouse *Rep-3* gene [(1); L.N. and G.F.C., manuscript in preparation]. These exons are spread over 134 kb of DNA and because of the large size of the gene, only the exons and the flanking exon/intron junctions were sequenced. Exon-containing fragments of the gene (determined by hybridization with either cDNA or oligonucleotides) were subcloned and then sequenced using an oligonucleotide primer derived from the cDNA sequence. This primer gave the sequence of the exon/intron boundary on one side of the exon; another primer was synthesized to enable sequencing in the opposite direction in order to obtain the other junction sequence.

Although this method allowed a relatively rapid analysis of the exon sequences, we obtained no information about the location of the exon within the cloned DNA fragment. However, all of the genomic fragments had been cloned into Bluescript vectors (Stratagene) and it proved possible to obtain the distance from the exon to the boundary of the restriction fragment by PCR using the gene specific sequencing primer and a primer from the vector, as illustrated in Figure 1. There are several advantages of this method: (1) No other primers had to be synthesized since the primers from the vector were already available, as they are the same primers used for sequencing inserts in the Bluescript vectors. (2) It is not necessary to know the orientation of the exon within the insert; two separate PCR reactions can be done with the sequencing primer and each of the flanking primers. Only a PCR reaction that is correctly oriented will generate a product, so this method can in fact determine orientation of the exon within the restriction fragment. (3) Except for extremely large subclones, the method has an internal check on the value obtained since the distance of the exon to both of the vector boundaries can be obtained by using each of the sequencing primers with the appropriate vector primer. (4) If a restriction fragment contains two or more exons, the distance between those exons can be determined using a sequencing primer from each. (5) This method gives the distance from the exon to the restriction

Figure 1. Structure of various subclones containing *Rep-3* exons in Bluescript plasmids. T3 and T7 refer to the T3 and T7 primer sites of the Bluescript plasmid flanking the inserted DNA fragments. The restriction sites used for cloning are indicated: S = *SaII*, R = *EcoRI*, H = *HindIII*, and the size of the insert is given. The labeled arrows represent the primers used and their orientations; the black rectangles represent the exon with the number of the exon indicated below. The PCR size from the indicated primer pair as determined from the gel in Figure 2 is also given. All primers were 17 bp.

Figure 2. Ethidium-stained 0.8% agarose gel of PCR reaction products. PCR reactions contained 0.2 μ g plasmid DNA and 3 U of Hot Tub DNA polymerase (Amersham) in a volume of 100 μ l with buffer supplied by the manufacturer. After denaturation for 5' at 90°C, 30 cycles of PCR were performed under the following conditions: 1' at 90°C, 1'20" at 48°C, and 5' at 68°C. Reactions were ethanol precipitated and 30% of the reaction was electrophoresed.

* To whom correspondence should be addressed

sites used to generate the subclones. Since it is usually relatively easy to map the location of subcloned fragments within the genome, the location of the exons can then be determined relative to the entire gene. (6) PCR products larger than 5 kb can be generated from cloned restriction fragments; therefore the location and orientation of an exon can be determined by this method even if it sits in the middle of a 10 kb restriction fragment.

Results of several different exon PCR reactions are shown in Figure 2.

ACKNOWLEDGEMENT

This work was funded by a research grant from the US Public Health Service (CA54040).

REFERENCE

1. Linton, J.P., Yen, J.-Y.J., Selby, E., Chen, Z., Chinsky, J.M., Liu, K., Kellems, R.E. and Crouse, G.F. (1989) *Mol. Cell. Biol.* **9**, 3058–3072.