Identification and Cloning of Integration Site of DNA by PCR


Exogenous DNA can be inserted into chromosomes by retroviral integration, transposon jumping, intrachromosomal translocations, as well as gene transfer techniques such as cell transfection assays and the preparation of transgenic animals. The sites of such insertions of exogenous DNA are usually identified by preparing genomic libraries and isolating clones containing both the exogenous DNA and flanking sequences (4,6,7). The task of isolating the appropriate genomic clone is usually time-consuming and tedious. In some situations, the flanking sequences can be isolated rapidly by plasmid rescue in Escherichia coli (1) or by inverse polymerase chain reaction (IPCR) (5). These rescue procedures have limitations in that either the inclusion of plasmid sequences is required (1) or the knowledge of restriction enzymes sites close to the site of integration (5) is essential. Though IPCR is known as being a simple procedure, on the other hand it involves several steps such as digestion of genomic DNA with different restriction enzymes, followed by recircularization and digestion with another enzyme. Another limitation is that amplification of DNA may become difficult where the enzyme sites are separated by large DNA sequences. In this report we present a simple PCR-based method to clone the flanking chromosomal DNA at the site of insertion of foreign DNA.

In the first and key step, the chromosomal DNA at the junction of transgene insertion is enriched by repeated cycles of primer extension using a DNA polymerase and a primer specific for the transgene, thus extending it into the flanking chromosomal DNA (see Figure 1). This is followed by amplification of the DNA by PCR using several arbitrary primers, each separately and the same transgene-specific primer. The chromosomal:transgene junction fragments are identified by Southern blot analysis. The junction fragment is further amplified in a second round of PCR using a nested primer and the specific arbitrary primer, followed by cloning and sequencing.

Genomic DNA from a transgenic mouse that contained 1–2 copies of a human minigene for α1 collagen type I (COL1A1) (2) was subjected to 5 cycles of primer extension by Taq DNA Polymerase (Perkin-Elmer, Norwalk, CT, USA) at 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min, with 0.5 μM concentration of primer S1 situated about 200 bp 3'-end and to the 5'-end of the transgene (primer S1, Figure 1).

The primer-extended DNA was then amplified by PCR using the same primer (S1) and commercially available arbitrary primers (10-mers; Operon Technologies, Alameda, CA, USA) at 94°C for 30 s, 36°C for 30 s, and 72°C for 30 s for 40 cycles in a PEC9600 thermal cycler (Perkin-Elmer). The amplified fragments were separated by electrophoresis on 1.5% NuSieve® (FMC BioProducts, Rockland, ME, USA) and the 1.0% agarose gel (Life Technologies, Gaithersburg, MD, USA). Forty arbitrary primers were tested, each separately, for the amplification.
cation of DNA with primer S1. Each arbitrary primer, apart from amplifying the transgene-specific fragment, may also amplify several random fragments of the genomic DNA. Figure 2A shows the gel electrophoresis pattern of amplified fragments with 15 separate arbitrary primers. To identify fragments that are specific for the transgene, the DNA was transferred onto nitrocellulose filters and analyzed by Southern blot analysis using 32P-labeled transgene-specific probe (probe A, Figure 1) following the method described by Lee et al. (3). The blots were washed for the amplification of DNA with 15 fragments. Figure 2B shows the data from another transgenic mouse line with multiple copies of transgene identified sequences spanning the junction of two copies of the transgene, as well as sequences at the junction of transgene with chromosomal DNA (data not shown).

The method described here involves techniques that are routinely used in any molecular biology laboratory. The unique sequences identified by this method can be used directly to scan the DNA databases or to screen the genomic libraries to identify novel genes. Alternatively, the sequence can be used to identify further upstream or downstream flanking sequences by repeating the whole procedure with the primers specific for these sequences. Since this procedure involves an initial step of enrichment of sequences by primer extension at the chromosomal junction site, the problems associated with the detection of single-copy sequences are reduced. Furthermore, the procedure also provides the flexibility of cloning the flanking chromosomal sequences at either end of the foreign DNA by using primers for both ends of the transgene (5). Although not tested here, the method may be applicable for cloning the sites of viral integration and transposable elements. Another important application may be to identify the junctions of chromosomal translocations, where partial information of the translocated gene is available.

REFERENCES


We thank Ms. Machiko Arita for technical assistance. This work was supported by NIH Grant AR38188. Address correspondence to Jaspal S. Khillan, Department of Biochemistry and Molecular Biology, Jefferson Medical College, Thomas Jefferson University, 233 South Tenth Street, Philadelphia, PA 19107, USA.

Received 24 May 1995; accepted 20 September 1995.

Shi-Wu Li, Kaye Marie Nembhard, Darwin J. Prockop and Jaspal S. Khillan
Jefferson Medical College
Thomas Jefferson University
Philadelphia, PA, USA

Figure 3. Sequence of the 270-bp PCR-amplified fragment. Unique chromosomal sequences are shown in bold letters. The junction of the chromosomal and foreign DNA is indicated by an arrow. The positions of arbitrary primer A-7 and the transgene-specific primer S2 are underlined.