Direct Sequencing of Flanking Regions of a Transgene Amplified by Inverted PCR

Takashi Ninomiya, Takeshi Iwabuchi,* Yukiko Soga* and Atsushi Yuki

Research Institute of Life Science, Snow Brand Milk Products Co., Ltd., 519, Shimoishihashi, Ishibashi-machi, Shimotsuga-gun, Tochigi 329-05, Japan
* Applied Biosystems Japan, Inc., 3-3-6, Minamisuna, Koto-ku, Tokyo 136, Japan

Received February 20, 1990

In transgenic mice, recombinant DNAs microinjected into pronuclei of fertilized eggs are generally integrated into host chromosomes. The integration mechanism of microinjected DNA is not fully understood because of limited information concerning the structures of integrated foreign DNAs (transgenes) and their flanking chromosomal DNAs. Transgenes of a few transgenic mice and their flanking regions have been cloned and analyzed. However, these methods included time-consuming steps, such as construction of genomic libraries, cloning, and subcloning.

We have reported that polymerase chain reaction (PCR) and inverted PCR (IPCR) techniques were useful for amplifying transgenes and their flanking regions in transgenic mice. In this study, we extended the research to develop a quick and simple method to sequence them without cloning and subcloning.

In a transgenic mouse line, gE1A/1, only part of the injected human adenovirus type 12 (Ad12) E1A gene was integrated. The flanking region of the transgene could be amplified from the genomic DNA of gE1A/1 by IPCR within a day in a single tube as described previously. Briefly, 100 ng or 20 ng of the genomic DNA was digested with PstI to produce 0.9-kb fragments. These fragments were circularized by ligation, and used as templates for IPCR with primers E1A224 and E1A323. These oligonucleotide primers were synthesized with a 380A DNA synthesizer (Applied Biosystems Japan Inc., Tokyo). Although the products of the IPCR showed smeared bands by gel electrophoresis, an expected band of 850 bp was detected in the products (lanes 1 and 2, Fig. 1).

The products of the IPCR were precipitated with polyethylene glycol to eliminate the primers and dNTPs used for the IPCR. Mineral oil was removed from the reaction mixture by pipetting and 0.6 volumes of 20% aqueous polyethylene glycol 6000 containing 2.5 M NaCl were added to the reaction mixture. After an hour on ice, the aggregates were precipitated by centrifugation at 15,000 × g for 10 min, washed twice with 70% ice-cold ethanol, and dissolved in 100 μl of distilled H2O. A sample corresponding to 0.01 μl of the reaction mixture of IPCR was then used as a template for asymmetric PCR to produce single strand DNAs. The primers E1A224 and E1A323 were designed to prime the IPCR products (Fig. 1). Eighteen bases of -21 M13 sequencing primers (TGTAAGCGAGCCGAGT) were linked to the 5' ends of E1A224 and E1A323 to construct the primers E1A224MP and E1A323MP, respectively. Asymmetric PCR was done either with a combination of E1A224MP and E1A323MP, respectively, with that of E1A323MP (2 pmol) and E1A224 (100 pmol) (lane 4, Fig. 1), or with E1A323MP (2 pmol) and E1A224 (100 pmol) (lane 6, Fig. 1). The volume of the reaction mixture was 100 μl. The products were precipitated with polyethylene glycol as described above and sequenced directly by the dideoxy method with an automated DNA Sequencer, ABI 370 (Applied Biosystems Japan Inc., Tokyo). The details were described in ABI User Bulletin No. 10, December, 1988. The amplified products in 100 μl of reaction mixture were sufficient as templates. Over 500 bases of the products between E1A323MP and E1A224 were sequenced at a time from the primer E1A323MP (Fig. 1). However, only 180 bases of the products between E1A224MP and E1A323 could be identified. In this case, an AT cluster existed 130 bases from the primer E1A224MP, which blocked sequencing beyond this region. To determine the sequence beyond the AT cluster, a new primer, G1PS11MP, was synthesized based on the results obtained, and used for another asymmetric PCR with E1A224. Thus all of the 918 bases of the PstI fragment containing the transgene were sequenced. The PstI fragments contained 368 bases of the 5' flanking region, 424 bases of the Ad12 E1A region, and 126 bases of the 3' flanking region (Fig. 2A). With respect to the transgene, 1178 bases of microinjected Ad12 E1A DNA were deleted from the 5' end, while the sequence of the EcoRI linker at the 3' end was preserved (Fig. 2A). The chromosomal DNA regions flanking the transgene were considerably AT-rich (Fig. 2B). The sequence described above was confirmed by results with another restriction enzyme, BamHI, and also with other primers (data not shown).

Homology searches by GENETYX Ver. 6.30 (Software Development Co., Ltd., Tokyo) with the DNA database (EMBL-GDB, Ver. 9.00) showed that the 5' flanking region was highly homologous (82.8%) to the mouse immunoglobulin kappa germline v region: v16 (Fig. 2A). This suggested that microinjected DNA was integrated among the immunoglobulin kappa gene families. However, it was likely that further rearrangement or deletion took place after the initial integration, since the 3' flanking region was not homologous to the rest of the kappa gene or any other sequences registered in EMBL-GDB.

A short direct repeat is usually produced at junctions of...
Fig. 1. Amplification of Flanking Regions of the Transgene by IPCR and Asymmetric PCR.

Transgene; ———, mouse chromosomal DNA. Arrowheads (→) indicate the locations and directions (5' to 3') of the primers. Squares indicate 21M13 sequences. The primers and their sequences are: E1A208, AACCTGGTTTCGGTTGAAGTGTCTTG; E1A224, GCCGTGGTTA AACAGGGGATAT; E1A224MP, TGTTAAACGACGGCCAGTGCGTGGTTAAACAGGGA; E1A305, CAAAGTCCATTAATTACATCTAGG; E1A323, CCACAGCACATCTACGCCTCCCAGT; G1PS11MP, TGTAAAACGACGGCCAGTGCGTGGTTAAACAGGGA; E1A224, GCCGTGGTTA AACAGGGGATAT; E1A224MP, TGTTAAACGACGGCCAGTGCGTGGTTAAACAGGGA; E1A305, CAAAGTCCATTAATTACATCTAGG; E1A323, CCACAGCACATCTACGCCTCCCAGT; G1PS11MP, TGTAAAACGACGGCCAGTGCGTGGTTAAACAGGGA. PCR was done by repeating 35 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 2 min and extension at 72°C for 2 min with 2.5 units of Taq polymerase (Perkin Elmer-Cetus, Connecticut). For IPCR, 100 pmol each of E1A208 and E1A305 was added to 100 μl of reaction mixture containing 100 ng (lane 1) or 20 ng (lane 2) of templates. Results of symmetric (lanes 3 and 5) and asymmetric (lanes 4 and 6) PCR are shown. The primers used are: lane 3, E1A323(100 pmol)/E1A224MP(100 pmol); lane 4, E1A323(100)/E1A224MP(2); lane 5, E1A224(100)/E1A323MP(2); lane 6, E1A224(100)/E1A323MP(2). Fifteen μl of the reaction mixture was directly electrophoresed in a composite gel of 3% NuSieve and 1% Seakem agarose (FMC BioProducts, Maine). DNA in the gel was stained by EtBr. M, size marker; HaeIII digests of φX174 DNA.

Arrows in the figure of asymmetric PCR indicate the sequences identified from the primers.
Fig. 2. A. Nucleotide Sequences of the Transgene and Its Flanking Regions of a Transgenic Mouse, gElA/1. The transgene integrated in gElA/1 is shown in a box with nucleotide numbers of Ad12 E1A gene.\(^{15}\) The underlined sequences are the linkers used for cloning E1A gene into plasmid pSV2.\(^{16}\) The nucleotide sequence of the mouse immunoglobulin kappa germline v region: v16\(^{17}\) is shown aligned under the 5' flanking region. Only bases non-identical to the flanking region are shown in small letters. △ indicates an insertion of the base(s) below it.

B. A-T Content Distribution Diagram of the Transgene and Its Flanking Regions. AT content(%) per every 20 nucleotides was plotted. ■, transgene; ——, mouse chromosomal DNA.
a transposon\(^\text{18}\) or a retrovirus (provirus).\(^\text{19}\) However, no such a repeat was found at the junctions of the transgene.

This method requires only 4 to 5 days to amplify and to sequence a transgene and its flanking regions, while general methods take months to clone, subclone, and sequence. Therefore this direct method would help in analyzing structures of other transgenes and their flanking regions. These data would promote elucidation of the integration mechanism. This method could also be used as a rapid alternative method in gene walking.

References