A novel method for constructing gene-targeting vectors

Kiyotaka Akiyama*, Hirotaka Watanabe, Shuichi Tsukada and Hitoshi Sasai

Pharmaceutical Frontier Research Laboratories, Central Pharmaceutical Research Institute, Japan Tobacco Inc., 13-2, Fukuura 1-chome, kanazawa-ku, Yokohama, Kanagawa, Japan

Received March 20, 2000; Revised June 2, 2000; Accepted June 28, 2000

ABSTRACT

We developed a simple and rapid method for constructing knockout vectors using inverse-PCR (IPCR). The method consists of three steps: (i) digestion of a target bacterial artificial chromosome with several restriction enzymes (six-base cutters) followed by self-ligation; (ii) IPCR using circular DNAs as templates and two primers which are oriented in opposite directions; and (iii) cloning into a vector containing a positive selection marker, which results in a typical replacement knockout vector. We successfully targeted three mouse genes including the HPRT gene using this method. Compared with the conventional method, this method requires much less time (no more than 3 weeks). Notably, this method requires only small amounts of sequence information (several hundred base pairs such as is available from expressed sequence tags) and can be extended to a systematic mass production of targeting vectors applicable to many organisms, including yeast.

INTRODUCTION

Genome projects of various organisms from bacteria to human are rapidly proceeding. The greatest effort has been spent on the human genome project, mainly to understand the mechanisms of human diseases. Approximately 20 000 genes have been isolated and about one million expressed sequence tags (ESTs) have been registered in public databases, which cover 60 000 unique human genes (http://www-bio.llnl.gov/bbrp/ESTs) and can be extended to a systematic approach to mouse. This method is effective for functional studies of novel genes.

However, using the knockout method is still largely an art. In order to isolate the BAC clone(s) containing the HPRT gene, ‘down to the well mouse ES BAC DNA pool’ (Genome Systems Inc., St Louis, MO) is screened by the primer pair p3A (5′-CTATAGGACCTGAAAGACTTTG-3′) and p3B (5′-TACTTACACAGTAGTCTTC-3′), which amplifies a part of exon 3. The conditions for PCR were as follows: 94°C for 3 min, 50°C for 2 min, (72°C for 1 min, 94°C for 1 min, 50°C for 1 min) for 30 cycles, 72°C for 3 min. The composition of the PCR reaction mixture was 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.025 U/ml Taq (Takara Shuzo, Kyoto, Japan), 0.2 mM of each dNTP mixture, and 1 μM of each primer. The size of the resulting PCR product is 197 bp in the case of one or more positive BAC clones containing the HPRT gene.

MATERIALS AND METHODS

Screening of BAC library

In order to isolate the BAC clone(s) containing the HPRT gene, ‘down to the well mouse ES BAC DNA pool’ (Genome Systems Inc., St Louis, MO) is screened by the primer pair p3A (5′-CTATAGGACCTGAAAGACTTTG-3′) and p3B (5′-TACTTACACAGTAGTCTTC-3′), which amplifies a part of exon 3. The conditions for PCR were as follows: 94°C for 3 min, 50°C for 2 min, (72°C for 1 min, 94°C for 1 min, 50°C for 1 min) for 30 cycles, 72°C for 3 min. The composition of the PCR reaction mixture was 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.025 U/ml Taq (Takara Shuzo, Kyoto, Japan), 0.2 mM of each dNTP mixture, and 1 μM of each primer. The size of the resulting PCR product is 197 bp in the case of one or more positive BAC clones containing the HPRT gene.

Isolation, digestion and ligation of BAC DNA

BAC DNAs were isolated using the standard alkaline–SDS method (4). In order to confirm whether an isolated BAC DNA contained the right insert with the right size, we digested it with NotI which resulted in a 7.4-kb fragment (pBeloBAC11 vector sequence) and a large inserted DNA (5). After 2 μg of isolated BAC DNA was digested with several 6 bp-cutters (such as HindIII, BamHI, SfiI, XhoI, XbaI, EcoRI, *To whom correspondence should be addressed. Tel: +81 45 786 7693; Fax: +81 45 786 7692; Email: kiyotaka.akiyama@ims.jti.co.jp
EcoRV and BglII), 0.5 µg of the digested DNA was circularized by self-ligation in 840 µl of reaction mixture at 4°C for 16 h. The reaction mixture was composed of 66 mM Tris–HCl, 6.6 mM MgCl₂, 10 mM DTT, 0.1 mM ATP and 2 U/ml T4 DNA ligase (TOYOBO, Tokyo, Japan). The ligated DNAs were then precipitated with ethanol and dissolved with 20 µl of TE.

Cloning of the fragment amplified with IPCR

We designed the primer pair for the IPCR: their sequences are 5′-ACGCGTTTAATTAACTTCATGACATCTCGAGCAA-GTCTTTGAGT-3′ as pr3-1 and 5′-ACGCGTGCGGCCGC-GCTGACCTGCTGGATTACATTAAAGCACTG-3′ as pr3-2 (Fig. 2). In order to improve the efficiency of cloning the PCR products, we attached the NotI and PacI recognition sequences at the ends of the PCR primers as shown by the underlines. One microliter of the self-ligated DNA was used as a template for IPCR. The IPCR reaction mixture consisted of 0.05 U/ml LA Taq (Takara Shuzo), 2.5 mM MgCl₂, 0.4 mM of each dNTP mixture, and 1 µM of each primer in a total volume of 50 µl. The PCR reaction conditions were: 94°C for 1 min, (98°C for 20 s and 68°C for 10 min) for 30 cycles.

After digesting each end of the PCR product with NotI and PacI, we fractionated the fragment by agarose gel electrophoresis and retrieved it by electroelution using a DNA CELL...
(Daiichi Pure Chemicals Co. Ltd, Tokyo, Japan). The recovered DNA fragment was cloned into p1108 which was digested with NolI and PacI. Then this targeting vector was linearized at the BamHI site prior to transfection.

Cells

E14.1 cells (6) were co-cultured on G418-resistant murine embryonic feeder fibroblasts (EF) (6,7) in high glucose DMEM supplemented with 15% fetal calf serum, glutamine (2 mM), non-essential amino acid (0.1 mM), penicillin (100 U/ml), streptomycin (100 U/ml), 2-mercaptoethanol (50 mM) and leukemia inhibitory factor (1000 U/ml) at 5% CO2 ES cells transfected with the targeted HPRT gene were cultured in the same way except that they were plated onto a gelatin-coated plate instead of onto EF cells to avoid the possibility of introducing HPRT protein to the ES cells via cell-to-cell communication (8).

Introduction of the targeting vector into ES cells

E14.1 cells (1 × 10^6) were transfected with a linearized targeting vector at a concentration of 25 µg/ml in PBS using a Gene Pulser (Bio-Rad, Hercules, CA). We used a cuvette with a 0.4-cm gap between the electrodes and applied a pulse at 200 V, 360 µF. Electroplated cells were plated onto ten 10-cm dishes, which were coated with 0.1% gelatin. After 24 h, the normal medium was exchanged with selection medium containing 250 mg/ml G418 (Gibco BRL, Gaithersburg, MD). At days 7–10, colonies were picked up and grown in the medium containing 5 µg/ml 6-thioguanine and 250 µg/ml G418.

Southern hybridization

Genomic DNA of ES cells was isolated as described by Aldridge et al. (9). Five micrograms of digested DNA was electrophoresed in a 0.7% agarose gel in Tris–acetate–EDTA buffer, transferred to a nylon membrane (GeneScreen Plus: NEN Bioproducts, Boston, MA) and hybridized with the 32P-labeled probe which contains the Amp and Ori sequences of the targeting vector. The hybridization was performed in Rapid Hybridization Buffer (Amersham Pharmacia Biotech, Uppsala, Sweden) with 1 × 10^5 c.p.m./ml of 32P-labeled DNA for 2 h at 65°C. After hybridization, the membranes were washed with 2× SSC plus 0.1% SDS for 15 min at room temperature twice and with 2× SSC plus 0.1% SDS for 30 min at 65°C twice. The membranes were used to expose Kodak XAR-5 film with intensifying screens at –80°C overnight.

RESULTS

Construction of targeting vector

We obtained three positive BAC clones (64N20, 118E2 and 131H17). 64N20 was used in the following experiment. The mouse HPRT gene is 33 kb long and is split into nine exons (10). We designed a targeting vector for disruption of exon 3. The restriction maps of the mouse HPRT gene are shown in Figure 2a. We used BamHI, EcoRI and HindIII which resulted in fragments which have suitable lengths (7–10 kb) for targeting vectors. The BAC DNA was digested with these three restriction enzymes and circularized by self-ligation. For the IPCR, we used the primer pair (pr3-1/pr3-2) as shown in Figure 2a. Three IPCR reactions resulted in distinct amplifications of the expected sizes, which are 1.3 (I), 7.1 (II) and 7.0 kb (III) (Fig. 2b).

The two amplicons, 7.1 (II) and 7.0 kb (III), were cloned into p1108 which has the neomycin resistance gene regulated by the SV40 promoter (Fig. 3). We designated the two plasmids in which fragments (II) and (III) were cloned as p1209 and p1207, respectively. In order to determine whether these two fragments were amplified correctly by IPCR, restriction maps of the two plasmids were analyzed. Figure 3 shows the maps of two amplicons using eight restriction enzymes (HindIII, BamHI, Sfil, XhoI, XbaI, EcoRI, EcoRV and BglII). In the two overlapping regions, designated as A and B (Fig. 3), seven consecutive restriction sites (EcoRI, Sfil, XhoI and HindIII in the A region, and BamHI, EcoRI and XbaI in the B region) were matched with each other. This indicated that no large deletions or rearrangements were generated during the IPCR.

Transfection to ES cells

We chose p1207 in the following targeting experiment, because the left arm of the homologous region in p1209 is relatively short (600 bp), and might cause a complex rearrangement during a homologous recombination event (11,12). p1207 contains 2.3-kb- and 4.7-kb-long homologous stretches which are long enough to generate homologous recombinations (13). We linearized p1207 with BamHI and transfected it into ES cells.
Transfected ES cells were plated onto a gelatin-coated plate, instead of onto EF cells which might provide HPRT protein to ES cells via cell-to-cell communication. After exposing the ES cells to G418 for 7–10 days, 350 clones were selected. Each clone was picked up and cultured in medium containing 6-thioguanine for an additional 10 days. Two clones, 1207-1 and 1207-2, were grown under these conditions. In order to determine whether a correct homologous recombination between the transfected vector and the mouse genome occurred in these clones, Southern hybridization was performed using six restriction enzymes ($\text{Bam}^\text{HII}$, $\text{Bgl}^\text{II}$, $\text{Eco}^\text{RI}$, $\text{Eco}^\text{RV}$, $\text{Hin}^\text{dIII}$ and $\text{XbaI}$). Based on the structure of the HPRT gene report by Melton et al. (10), we created a map of the disrupted HPRT gene showing predicted restriction sites (Fig. 4a). As a hybridization probe, we used a 2.7-kb $\text{Bss}^\text{HII}$ fragment (shown in Fig. 4a) containing the ampicillin resistance gene and replication origin of p1207. We detected bands that perfectly matched the predicted restriction maps [a 12.5-kb $\text{Bam}^\text{HII}$ band, a 8.6-kb $\text{Bgl}^\text{II}$ band, a 6.8-kb $\text{Eco}^\text{Rl}$ band, a 12.0-kb $\text{Eco}^\text{RV}$ band, a 12.6-kb $\text{Hind}^\text{III}$ band and a 7.8-kb $\text{XbaI}$ band (Fig. 4a and b)]. These results indicated that the events of a precise homologous recombination occurred in the 1207-1 and 1207-2 cell lines.

**DISCUSSION**

In this experiment, we have shown that our novel method of constructing a targeting vector works efficiently in the case of knocking out the HPRT gene. This method has several advantages compared with a conventional method. First, it is relatively quick. Typically, the novel method requires several weeks, while the conventional method requires several months. This is because the conventional method requires many time-consuming steps: screening of a phage library, construction of restriction maps, determination of exon/intron boundaries and cloning of homologous fragments into a vector. Also, when designing a targeting vector, there is limited availability of preferred restriction sites. However, we are free from this limitation, because we simply use IPCR and a one-step cloning of an ampiclon into a vector.

Second, this method requires much less information about the target than does the conventional method. We can start with only a small piece of sequence information such as that in ESTs. In theory, we need a consecutive sequence of only 80 bp, because we can screen the BAC and do IPCR by using primers that are based on a short sequence. Until now, targeting experiments have been done after characterizing full-length cDNAs and the genomic structures of the genes of interest. These days, enormous ESTs, which are partial sequences of cDNAs, are available in public databases. Now ESTs are available on gene chips and on micro-arrays for differential display, and are utilized for identifying disease-associated genes. Using this database and the power of bioinformatics, we have been able to find DNA sequences with interesting motifs or expression profiles. Using this novel method, we can prepare a batch of targeting vectors for ESTs of interest before they have been fully analyzed.

Third, this method makes it easy to insert any sequence into any desirable region in a target gene. In the conventional method, site-directed insertions or deletions are often difficult due to the limited availability of preferred restriction sites. This sometimes causes a serious problem if there is no restriction site for modifying the gene. Since we obtain the homologous region by IPCR in our novel method, we can design PCR primers for any region in a gene.
specific modification of BACs in Escherichia coli of target vectors (14, 15). In short, Stewart and colleagues reported an efficient method for modifying BAC DNA, which can be used for the construction of targeting vectors for expression. In this paper, we have shown the feasibility of this method when the region on which IPCR primers are based is closed to the cloning site of the vector, because this might result in amplification of an undesirable fragment derived from a vector sequence. A YAC clone can also be utilized as a template for IPCR because Triglia et al. (3) reported that the DNA segment lying inside the YAC arm was amplified by IPCR.

It is known that PCR products contain mutations due to replication errors of Taq polymerase. Riele et al. (17) showed that mismatches between a targeting vector and a genomic DNA seriously reduce the efficiency of homologous recombination. On the other hand, Randolph et al. (18) and Sedivy and Dutriaux (19) showed that a mismatch at the level of 10⁻³–10⁻⁴ does not affect homologous recombination. In this experiment, the targeting frequency was 0.6%, which is consistent with the efficiency reported in other papers (20, 21). Since the replication error in the LA PCR system used in this study has been reported to be 1.6 × 10⁻⁴, it is plausible that this mutation rate is sufficiently low to allow homologous recombination. In case there is a reduced efficiency of homologous recombination due to mismatches, we could reduce the number of mutations in the PCR products by using fewer cycles in the PCR reaction.

Using this novel method, we can easily make vectors for promoter assays. These vectors can be used for drug screening, determination of transcriptional promoter/suppressor regions, screening of transcriptional regulation factors and so on. Usually, the reported genomic sequence of the 5’ region of genes is too short to construct vectors for these purposes. Conventionally, we construct the vectors from cloned genomic DNAs using restriction enzymes, which is tedious and laborious. Our new method makes this step much easier and faster than the conventional method, by simply amplifying and cloning the 5’ flanking region using IPCR, even if the 5’ flanking sequences are unknown. However, care should be taken to disrupt the coding region rather than the untranslated region.

In our laboratory, many kinds of knockout mice are currently being developed. Some of the results of these studies are summarized in Table 1. Over a short period, we succeeded in the disruption of two genes in ES cells. One was the erythropoietin receptor (EpoR) gene and another was an EST, AU067162. These results indicate that this method is applicable to the targeting of any gene or EST, and the targeting efficiency for EpoR gene was almost as good as the targeting efficiency that we obtained for the HPRT gene. We believe that our new method can promote the gene knockout process from a type of art to an easier technology.

As the genome projects of various organisms proceed, we need more functional studies of many genes, partly by gene knockouts, in simple organisms as well as in mammalian cells. In yeast, gene targeting has been used for a long time to analyze gene function, since the efficiency of homologous recombination is high. The whole genomic DNAs of yeast might be used for a starting material for IPCR, because the size of the yeast genome is sufficiently small (3). Now we believe that this new, efficient method of gene targeting is possible in any organism, and that it will help to fill the gap between what efficient promoters, such as MC1 and PGK, to reduce the number of passages that are required (16).

Although we used a BAC clone as a starting material, we can choose phages or cosmid clones provided that enough targeting region is included. However, we must be careful when the region on which IPCR primers are based is closed to the cloning site of the vector, because this might result in amplification of an undesirable fragment derived from a vector sequence. A YAC clone can also be utilized as a template for IPCR because Triglia et al. (3) reported that the DNA segment lying inside the YAC arm was amplified by IPCR.

In this paper, we have shown the feasibility of this method when the region on which IPCR primers are based is closed to the cloning site of the vector, because this might result in amplification of an undesirable fragment derived from a vector sequence. A YAC clone can also be utilized as a template for IPCR because Triglia et al. (3) reported that the DNA segment lying inside the YAC arm was amplified by IPCR.

### Results of the genomic Southern hybridization of 1207-1 and 1207-2.

Figure 4. Results of the genomic Southern hybridization of 1207-1 and 1207-2. (a) Restriction maps and structure of the targeted HPRT locus. The fragments that hybridized with the probe are indicated by bold lines with their predicted sizes. Genomic Southern hybridization of two clones, 1207-1 (b) and 1207-2 (c).

primes for any region we choose. This method is especially powerful when a reporter gene, such as the luciferase gene, is placed so that it is strictly under the control of an endogenous promoter.

Stewart and colleagues reported an efficient method for modifying BAC DNA, which can be used for the construction of targeting vectors (14, 15). In short, they demonstrated site-specific modification of BACs in Escherichia coli via homologous recombination using a targeting vector harboring homologous DNA fragments as small as 50 nt. While Stewart’s method uses more steps for cloning vectors than ours, it avoids the risk of PCR error that might cause a low efficiency of homologous recombination. However, both methods require only a small amount of sequence information to create a complex targeting vector.

In this paper, we have shown the feasibility of this method using simple vectors that contain only the elements that are indispensable for gene targeting. However, the vector can be modified in several ways. For example, the Cre–lox recombination system can be used to avoid a possible artificial effect caused by putting loxP sequences at both ends of the neomycin cassette. Although we used the SV40 promoter for expression of the neomycin-resistant gene, it is advisable to use more
we know about gene sequences and what we know about gene functions.

ACKNOWLEDGEMENTS

We express our thanks to Ms Seiko Kubo and Ms Suguri Niwa for their excellent experimental techniques and Mr Hideki Kodera for encouragement throughout the preparation of this manuscript. We deeply thank Dr Yoshisuke Nishi and Dr Takashi Matsumoto for encouraging us to carry out this study.

REFERENCES


Table 1. The efficiency of the homologous recombination with two targeting vectors

<table>
<thead>
<tr>
<th>Gene</th>
<th>Size of 5′ flanking region (kb)</th>
<th>Size of 3′ flanking region (kb)</th>
<th>Number of pick-up colony</th>
<th>Number of H.R. colony</th>
<th>Targeting efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epo receptor</td>
<td>7.0</td>
<td>1.1</td>
<td>192</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>EST (AU067162)</td>
<td>3.0</td>
<td>8.0</td>
<td>713</td>
<td>1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 1. The efficiency of the homologous recombination with two targeting vectors