be ligated to DNA fragments from any source, e.g., cDNAs obtained after subtractive hybridization or DNA fragments from S1 mapping experiments for analyzing exon-intron boundaries of genomic DNA. Due to the high ligation efficiency of dsPPL to the denatured ssDNA, the subsequent amplification is very efficient. It should be emphasized that PPL can be any palindromic structure that has three random nucleotides at its 3’ end and a ddNTP as its last nucleotide. This makes the method even more widely applicable.

REFERENCES


This work was supported by the China National Natural Science Foundation. Address correspondence to Prof. Sizhong Zhang, Department of Medical Genetics, The First Affiliated Hospital, West China University of Medical Sciences, Chengdu, 610041, PR China. Internet: zszg@wcums.edu.cn

Received 12 November 1997; accepted 28 July 1998.

Cuiying Xiao, Qingjie Xia, Hui Wu and Sizhong Zhang West China University of Medical Sciences Chengdu, PR China

Long-Distance PCR-Based Strategy for Preparing Knock-In Vectors Directly from ES Cell Genomic DNA


Gene targeting is a powerful tool for generating mice with alleles containing mutations as a means to study gene function and the resulting impact of mutations on disease. The first step in this procedure is targeting of the genome in embryonic stem (ES) cells. Success with targeting is positively related to the length of the homologous gene sequence in the vector (5,6) but is negatively related to the degree of sequence variation between the gene and its chromosomal homolog in the ES cell genome (5,10). Therefore, an optimal targeting vector should include a large fragment of the gene with a background highly isogenic to that of ES cell DNA (5). However, a common problem is that the sequence of a gene isolated from a genomic library might not match its gene target in a specific ES cell line if the gene and the ES cell line are not derived from the same mouse strain. The widely used 129 mouse strain was recently found to have a mixed genetic background among its sub-strains and derivative ES cell lines (9,11). About 25% of the genome of the 129/SvJ strain was found to be of non-129 origin (11). The recently developed long-distance polymerase chain reaction (LD-PCR) (1–3) has been increasingly used in isolating genes, mapping genomic structures and screening for mutations (1–4,7). LD-PCR provides an alternative way to make targeting vectors from a large DNA sequence directly isolated from the ES cell line to be used.

Here, we describe a LD-PCR-based strategy for preparing gene-targeting vectors that is simple, flexible and fast. We demonstrate the usefulness of this procedure by generating a knock-in vector for the Asn291Ser mutation in the mouse lipoprotein lipase (LPL) gene.

Embryonic stem cells RI, established from the mouse strain 129/SvX129/Sv-CP F1 (8), were cultured in a 5% CO₂ humidified incubator. The culture media were high-glucose (4.5 g/L) Dulbecco’s modified Eagle medium (DMEM) containing 2 mM L-glutamine (Life Technologies, Burlington, ON, Canada), 150 µM monothioglycerol (MTG) (Sigma Chemical, St. Louis, MO, USA), leukemia inhibitory factor (LIF) (Terry Fox Laboratories, Vancouver, BC, Canada) and 15% selected fetal calf serum. Culture dishes were coated with gelatin. Genomic DNA was extracted from the cultured ES cells by the standard procedure.

The sequences of the primer used for amplification of the mouse LPL gene were taken from the published se-
sequence for mouse strain 129/J (12), which covered the portion of exon/intron boundary sequences (Table1).

We used the Expand™ Long Template PCR System (Boehringer Mannheim Canada, Laval, QC, Canada) to amplify the mouse LPL gene fragments and to introduce the Asn291Ser mutation. This mixture of Taq and Pwo DNA polymerases reduces the PCR error rate and facilitates long-range amplification. Briefly, according to the manufacturer’s recommendation, amplifications were set up in a total volume of 50 µL, containing 160–200 ng ES cell DNA, 350 µM of each dNTP, 240 nM of each primer, with No. 1 PCR Buffer (from the Expand Long Template PCR System) and 0.75 µL enzyme mixture in 0.2-mL, thin-walled PCR tubes. Amplifications were preceded by a denaturation at 93°C for 3 min, followed by 32 cycles of 92°C for 10 s, 65°C for 1 min, 68°C for 6 min with a final extension at 68°C for 10 min on the GeneAmp® PCR System 9600 (PE Applied Biosystems, Mississauga, ON, Canada). The amplified genomic fragments were blunted for cloning with 1.5 U T4 DNA Polymerase and then treated with 5 U Polynucleotide Kinase (Life Technologies). These samples were then purified on Amicon® Centricon® 100 Spin Columns (Millipore, Bedford, MA, USA) and cloned into the

Figure 1. LD-PCR-based strategy to construct knock-in targeting vector. (A) The boxes represent exons, and the lines represent introns. [1] The LPL gene fragments containing the desired mutation were individually amplified. [2] Overlap PCR linked fragments A1-B1 and C1-D1. The mutation at the end of separately amplified sequence in step [1] was moved in the central region of a linked larger genomic sequence. [3] Another LPL gene section was amplified and cloned to be used to extend the length of the LPL gene sequence in the vector and to include the selection gene PGK-neo in the unique restriction site (EcoRI). [4] The knock-in vector pKSAsn291Ser. (B) Agarose gel electrophoretic analysis of PCR products (right) and the knock-in vector pKS Asn291Ser digested with Xhol (X) and BgIII (B) (left).
Benchmarks

Table 1. Oligonucleotide Primers for Amplification of the Mouse LPL Gene

<table>
<thead>
<tr>
<th>No.</th>
<th>Oligonucleotidesa (5′→3′)</th>
<th>Size</th>
<th>Genomic Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>tccccctaagGGTTTGAGCCAGCT</td>
<td>24-mer(S)</td>
<td>intron 4/exon 5</td>
</tr>
<tr>
<td>B1</td>
<td>TCTCTTGCTCTCGAGGATCTCATA</td>
<td>30-mer(AS)</td>
<td>exon 6</td>
</tr>
<tr>
<td>C1</td>
<td>TATGAGTAGCTGCAAGGTCCAGGAAAGAGA</td>
<td>30-mer(S)</td>
<td>exon 6</td>
</tr>
<tr>
<td>D1</td>
<td>GGAACGACgttgagagagatgacacaa</td>
<td>30-mer(S)</td>
<td>exon 7/intron 6</td>
</tr>
<tr>
<td>E1</td>
<td>tcttttccagACGTGGACCAGCTGGTGAA</td>
<td>30-mer(S)</td>
<td>intron 6</td>
</tr>
<tr>
<td>F1</td>
<td>gtctctacTTTTTCTGACTCTCTCGGCTT</td>
<td>30-mer(AS)</td>
<td>exon 8/exon 8</td>
</tr>
</tbody>
</table>

aUppercase letters represent exonic nucleotides, and lowercase letters represent intronic nucleotides; changed nucleotides for the mutations Asn291Ser are underlined; S in the parentheses represents sense, and AS represents antisense.

pBluescript® II KS+ vector (Stratagene, La Jolla, CA, USA) at the SmaI site. Sequencing was carried out using a Model 373A Automated DNA Sequencer (PE Applied Biosystems).

The mouse LPL gene contains 10 exons (100–250 bp in length) and 9 introns ranging from 0.5–7.7 kb in size (12). The flexibility of the LD-PCR approach allows any large section of the LPL gene to be amplified based on the known sequences of exons. For our purpose, a 4–7-kb section comprising exon 6 was required (Figure 1).

First, two pairs of primers were used to separately amplify fragments A1-B1 (exons 5 and 6) and C1-D1 (exons 6 and 7) directly from the ES cell genomic DNA. The mutagenic nucleotides were included in primers B1 and C1 for creating the Asn291Ser mutation in exon 6 (Table 1 and Figure 1).

Second, these two fragments were linked by LD-PCR overlap amplification using primers A1 and D1 at the same conditions. Centricon-100-purified fragments A1-B1 and C1-D1 (at the ratio ca. 1:1) were diluted 200-fold in water and used as the template. Electrophoresis of PCR product revealed the expected 4-kb band for the fusion fragments of A1-B1 plus C1-D1, indicative of successful overlap linking by the LD-PCR amplification.

In the final PCR step, a 5-kb LPL gene fragment (E1-F1) containing exons 6–8 was amplified and cloned as described above. The integrity of this and the other amplified gene sections was verified by restriction mapping, which yielded the expected fragment lengths except for the size of intron 6, which was found to be 2.8 kb with a single EcoRI site (Figure 1), instead of 3.6 kb with two EcoRI sites, as depicted in the published sequence (12).

To create the knock-in vector, a 1.2-kb DNA from the PCR-mutagenized and cloned fragment A1-D1 was removed by HindIII and EcoRI digestion and replaced by a 3-kb DNA from fragment E1-F1. Then, a 1.7-kb phosphoglycerate kinase (PGK) promoter-neo gene flanked by two direct LoxP sites was inserted into the unique EcoRI site in this region. The PGK-neo gene is used to select positive ES cells for the homologous recombination. Figure 1, A and B both show the final construct, pKS (Asn291Ser) knock-in vector containing a 6-kb LPL gene sequence including exons 5, 6, 7 and part of exon 8.

PCR-generated DNA fragments are inherently at risk of containing PCR errors, particularly after two rounds of amplification. These errors were sought in our final construct by sequencing all exons, splicing junctions and approximately 100 bp of flanking intronic sequence. Three single-base substitutions (two T→C and one A→G) were found in exons 5 and 6 when these data were compared with that of the published sequence. Thus, these data yield an error rate of approximately one in 1500 bases. However, the actual frequency might be less because one alteration was detected in each of the triplet
clones sequenced and could represent a polymorphism. The other two substitutions, which are close to the overlap junctions, are most likely PCR errors because they were not detected in each of the triplet clones sequenced. Significantly, none of these changes altered the codon specificities. The 0.8-kb DNA deviation we detected in intron 6 might reflect a LPL gene polymorphism. The other two substitutions, which are close to the overlap junctions, could represent a genomic sequence directly isolated from ES cells. In the future there could be a variety of possible modifications to this method. Although the risk of introducing undesired mutations by LD-PCR might be still higher than that of the conventional library construction and screening, this approach is an attractive choice, especially when a specific gene fragment could not be obtained by the library screening. Using this strategy we have successfully constructed other targeting vectors containing up to 9 kb of the gene sequence of interest.

REFERENCES


Address correspondence to Dr. Lin-Hua Zhang, Centre for Molecular Medicine and Therapeutics, 950 West 28th Avenue, Vancouver, BC, Canada, V5Z 4H4. Internet: linhua@ulam.generex.ca

Received 23 February 1998; accepted 27 July 1998.

Lin-Hua Zhang1 and Howard E. Henderson1,2
1University of British Columbia
Vancouver, BC, Canada
2University of Cape Town
Cape Town, South Africa

ADDENDUM


Recently, it was brought to my attention that some papers describing the essence of this method have been previously published. All of these articles described similar approaches to address slightly different questions. For example, the 1996 paper by Ailenberg and Silverman emphasized a way to overcome the low efficiency of digesting PCR products and their 1997 article emphasized PCR-based mutagenesis, while my method focused on solving the presence of internal restriction sites of PCR products. I am correcting this inadvertent oversight by acknowledging these prior publications (1–4).

REFERENCES


Gang Zeng
National Cancer Institute
National Institutes of Health
Bethesda, MD, USA
Internet: gzeng@pop.nci.nih.gov

Vol. 25, No. 5 (1998)