Simple and straightforward construction of a mouse gene targeting vector using in vitro transposition reactions

Minako Aoyama, Kazuko Agari, Ge-Hong Sun-Wada, Masamitsu Futai and Yoh Wada*

Division of Biological Science, Institute for Scientific and Industrial Research, Osaka University and CREST, Japan Science and Technology Agency, Mihogaoka 8-1, Ibaraki, Osaka 567-0047, Japan and Futai Special Laboratory, Microbial Chemistry Research Center, Microbial Chemistry Research Foundation and CREST, Japan Science and Technology Agency, 3-14-23 Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan

Received February 3, 2005; Revised and Accepted March 7, 2005

ABSTRACT

In a gene targeting experiment, the generation of a targeting construct often requires complex DNA manipulations. We developed a set of cassettes and plasmids useful for creating targeting vectors to modify the mammalian genome. A positive selection marker cassette (PGK/EM7p-npt), which included dual prokaryotic and eukaryotic promoters to permit consecutive selection for recombination in Escherichia coli and then in mouse embryonic stem cells, was flanked by two FRT-loxP sequences. The PGK/EM7p-npt cassette was placed between the minimum regions of a Tn7 transposable element for insertion into another DNA by means of Tn7 transposase in vitro. We also constructed a plasmid having a loxP-Zeo-loxP cassette between the modified Tn5 outer elements. These cassettes can be integrated randomly into a given genomic DNA through the in vitro transposition reaction, thus producing a collection of genomic segments flanked by loxP sites (floxed) at various positions without the use of restriction enzymes and DNA ligase. We confirmed that this system remarkably reduced the time and labor for the construction of complex gene targeting vectors.

INTRODUCTION

Gene targeting is a powerful method for producing genetically modified animals to study gene function in vivo. This powerful technology allows us to introduce designed mutations into any cloned locus, and to analyze generated mice with the corresponding genetic changes. The resulting phenotypes often provide an insight into the functions of genes. However, a gene targeting experiment is laborious and time-consuming, demanding elaborate techniques for the manipulation of DNA, cells and embryos, and it often takes more than one year from the design of constructs to generation of animals.

Replacement of the native gene with a modified gene fragment in the mammalian genome by homologous recombination is performed to introduce genetic alteration in embryonic stem (ES) cell. One of the limiting steps is the generation of gene targeting vectors’, since the design of a targeting vector must fulfill several requirements. First, efficient homologous recombination in the ES genome requires fairly long (>5 kb, >10 kb, if possible) homologous segments flanking the alteration in the targeting construct (1–4). The use of a positive selection marker is mandatory for selecting candidates having homologous replacement, and it should be placed in a proper region and be designed to be removable afterward. For this purpose, the loxP/Cre (4) and FRT/FLPe (5) systems have been used to remove the selective markers in ES cells or animals. The use of these highly specific recombination systems also led to the second generation gene modification strategy for controlling genetic alterations spatio-temporally in animals (6,7). However, the design and construction of floxed alleles is labor-intensive. The loxP sites must be placed at both ends of the genomic segments to be deleted upon Cre recombinase-mediated recombination, but reasonably far from the ends of the construct for efficient homologous recombination in the ES genome. They must be placed in intron regions to conserve the structure and function of the gene products. In addition, genomic regions containing highly repetitive elements should be excluded from the targeting construct, since these elements hamper efficient and accurate detection of homologous recombination on PCR or Southern hybridization analyses.

The mouse genome project allows us to deduce the structures of loci of interest at the nucleotide sequence level in silico. We searched for single restriction sites that fulfill the requirements described above, and then synthesized loxP

*To whom correspondence should be addressed. Tel: +81 6 6879 8481; Fax: +81 6 6875 5724; Email: yohwada@sanken.osaka-u.ac.jp

© The Author 2005. Published by Oxford University Press. All rights reserved.
segments tagged with the recognition sequence at both ends and integrated them into the targeting vector by using DNA ligase. However, it is not always straightforward to find an appropriate unique recognition site within a given genomic segment. Even when a suitable restriction site exists, it still remains difficult to introduce only a single short loxP segment (~40 bp) into a large plasmid (>10 kb) by conventional DNA manipulation relying on T4 DNA ligase.

To overcome these limitations, we applied bacterial transposon systems to simplify the construction of complex targeting vectors. Bacterial transposon exhibits minimal sequence preference for their insertions and single insertions within a several hundred kilo base pairs region of a single molecule due to their long-range cis interactions (8). These features can be reconstituted in vitro (9–11). The nature of the in vitro transposition reactions is very powerful for assembling complex targeting constructs, as reported previously (12). Here, we developed a set of plasmids that significantly reduce the time and labor for the construction of gene targeting constructs. We described a typical experiment for modification of the mouse mVam2/Vps41 gene as a model study.

**MATERIALS AND METHODS**

**Antibiotics and bacteria culture**

*E. coli* was cultured in Terrific Broth. Ampicillin sodium salt (Amp), kanamycin sulphate (Kan), chloramphenicol (Chl), tetracycline HCl (Tet) and Zeocin (Zeo) were used at concentrations of 100, 25, 33, 12.5 and 100 μg/ml, respectively. BAC was cultured in Terrific Broth. Ampicillin sodium salt was prepared with Qiagen large construction kit, and cis several hundred kilo base pairs region of a single molecule preference for their insertions and single insertions within a vectors. Bacterial transposon exhibits minimal sequence poson systems to simplify the construction of complex target-manipulation relying on T4 DNA ligase. However, it is not always straightforward to find an

**FRT sequences** were inserted into unique HindIII (between the upstream loxP and PGKp) and EcoRI (between the down stream loxP and vector backbone) sites in two steps. First, an FRT segment with HindIII sites at both ends was prepared from yeast 2 μm ori containing plasmid YEp24 by PCR with primers HindFLP-A1 and HindFLP-S1, and then inserted at the HindIII site of pLoxPGK7M-Neo to yield an intermediate plasmid. Second, an EcoRI-FRT-EcoRI segment was prepared using primer pair EcoFLP-S1 and EcoFLP-A1, and then inserted at the EcoRI site of the intermediate plasmid. The orientations of the two FRT segments were verified by sequencing, and a correctly configured plasmid, pLoxFRTNeo-1 was obtained. pLoxFRTNeo-1 was digested with Xhol and Smal, polished with T4 DNA polymerase, then subcloned into the SpeI–NotI sites (trimmed with T4 DNA polymerase) of pGPS2.1 (New England Biolabs), introduced into *E. coli* strain EC100D pir-116 (Epictcrease), and then selected on tetracyclin and kanamycin to obtain the final construct, pGPS21loxFRTNeo.

**Assembly of pMODloxZeoΔamp3**

A loxP-Zeo-loxP DNA segment was amplified with a primer pair ERVloxSR-5 and ERVloxSV40-3 using *pVgR XR* (Invitrogen) as a template. The PCR product, having an FRT sequence at the both ends, a mammalian enhancer/promoter, the EM7 promoter and ZeoP, was digested with EcoRV and then introduced into the ClaI/HindIII sites (flushed with T4 DNA polymerase) of pMOD<ele>ΔMCS> (Epictcrease) to yield pMODloxZeo-1. The *bla* gene of the pMODloxZeo-1 was disrupted by digestion with Scal and Sspl, and trimmed with T4 DNA polymerase, and then a part of mammalian transcriptional element was removed by Sphi and PstI digestion to obtain a final plasmid, pMODloxZeoΔamp3.

Table 1. List of oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGKEM7S</td>
<td>5’-CCTTCTTCCATCTCCGCCGCCTTTGTACGAAATTAATCATGC-3’</td>
</tr>
<tr>
<td>PGKEM7A</td>
<td>5’-GCAATCTTCCATCTCCGCCGCCTTTGTACGAAATTAATCATGC-3’</td>
</tr>
<tr>
<td>HindFLP-S1</td>
<td>5’-TCAAAAGGCTTTTGAGGTCTCTACTTCT-3’</td>
</tr>
<tr>
<td>HindFLP-A1</td>
<td>5’-CGGAAAAGGCTTTTGAGGTCTCTACTTCT-3’</td>
</tr>
<tr>
<td>EcoFLP-S1</td>
<td>5’-TCAAAAGGCTTTTGAGGTCTCTACTTCT-3’</td>
</tr>
<tr>
<td>EcoFLP-A1</td>
<td>5’-CGGAAAAGGCTTTTGAGGTCTCTACTTCT-3’</td>
</tr>
<tr>
<td>ERVloxSR-5</td>
<td>5’-GGGATATCCTCTCCTGATATGCTCCTATCCGAGAATTTATGCGTCAAAAGTGGGTGCT-3’</td>
</tr>
<tr>
<td>ERVloxSV40-3</td>
<td>5’-GGGATATCCTCTCCTGATATGCTCCTATCCGAGAATTTATGCGTCAAAAGTGGGTGCT-3’</td>
</tr>
<tr>
<td>loxzeo-Clg-5</td>
<td>5’-CCCTAAGTGACACATACTTGCTAT-3’</td>
</tr>
<tr>
<td>loxzeo-Clg-3</td>
<td>5’-CCCTAAGTGACACATACTTGCTAT-3’</td>
</tr>
<tr>
<td>PmeRsr1</td>
<td>5’-GTTGAACCGGACGC-3’</td>
</tr>
<tr>
<td>PmeRsr2</td>
<td>5’-GGTGGGTTTAAAACAGCT-3’</td>
</tr>
<tr>
<td>FseSrf1</td>
<td>5’-CTAGGATCGCCGGCCGGGGCCGGCCGGCG-3’</td>
</tr>
<tr>
<td>FseSrf2</td>
<td>5’-GTAGGATCGCCGGCCGGGGCCGGCCGGCG-3’</td>
</tr>
<tr>
<td>M13 Fwd23</td>
<td>5’-CCAGTACGACGATGTTGAAGAAGACG-3’</td>
</tr>
<tr>
<td>DT-AR</td>
<td>5’-ACTAGTCTAATTGTTTGCTTG-3’</td>
</tr>
</tbody>
</table>
Assembly of pBSDT-AII
A HincII–XhoI segment of the MC-1 promoter and DT-A region was excised from pMC1DT-A (LifeTec) and inserted into Bluescript II SK+ to obtain pBSDT-A. A pair of synthetic oligonucleotides, PmeRsR1 and PmeRsR2 (Table 1), was annealed to obtain a linker, which was then inserted at the SacI–SacII sites of pH1765-A, and then a second linker (a duplex of FseSr1 and FseSr2, see Table 1) was introduced at the BamHI–XbaI sites to generate pBSDT-AII. The structure of multiple cloning site of pBSDT-AII was verified by sequencing.

In vitro transposition reactions
In vitro transposition reaction of pGPS21loxFRTNeo and pMODloxZeoDamp was performed with TnsABC* (New England Biolabs) and EZ::Tn transposase (Epicentre), respectively, following the manufacturer’s instructions. The reaction product was purified by phenol/chloroform extraction and ethanol precipitation, and then introduced into an appropriate E.coli strain by electroporation.

RESULTS AND DISCUSSION
Construction of a transposable KAN/G418 cassette with loxP and FRT sequences
Plasmid pGPS21loxFRTNeo (Figure 1) was designed to introduce a positive selection marker into gene segments by in vitro transposition reaction. This plasmid has several unique features. The Tn5-derived neomycin phosphotransferase gene (npt) was placed downstream of a short fragment of the EM7 promoter (EM7p) and mammalian PGK promoter (PGKp) sequence. This DNA fragment conferred kanamycin resistance (5–50 μg/ml) on E.coli and G418 resistance (100–300 μg/ml) on mouse ES cells, respectively (data not shown), indicating that the PGK/EM7p-npt segment could be used as a positive selection marker in E.coli as well as ES cells.

This PGK/EM7p-npt segment is placed between the right and left borders of the Tn7 transposing element, therefore it can be inserted into other DNAs by incubation with a commercially available Tn7 transposase. Upon Flip- mediated recombination (14), this selectable marker can be removed, leaving one loxP site. The border sequences were placed outside of the two FRT-loxP sequences, therefore transposition introduces a segment of FRT-loxP-PGK/EM7p-npt-FRT-loxP into the target DNA. In vitro transposition events take place almost randomly, and multiple insertions into the same DNA fragment are rare due to the target immunity (8). Plasmid pGPS21loxFRTNeo has R6Kγ replication ori, and common E.coli strains like DH10B and XL1-Blue are not able to maintain this plasmid, but it replicates as a multiple copy plasmid in bacteria expressing the pir-116 protein.

Construction of a transposable cassette with a loxP-Zeo-loxP segment
Another plasmid, pMODloxZeoDamp3 (Figure 2), which has a loxP-Zeo-loxP cassette between modified Tn5 outer elements, was constructed to facilitate the third loxP element insertion. Tn5 and Tn7 require individual machineries for their transposition and both elements can be introduced in the cis configuration. pMODloxZeoDamp3 has a PUC/Col E1 replication origin, a disrupted ampicillin-resistance gene, and a functional EM7-driven Zeocin-resistance gene, and thus confers Zeoγ on common laboratory strains of E.coli.

Figure 1. Structure of pGPS21loxFRTNeo. pGPS21loxFRTNeo was designed as a donor plasmid for transposase-mediated in vitro integration of a mammalian/bacterial selection marker and FRT, loxP sequences. Characteristic elements of the plasmid are summarized in (A). The plasmid possesses an R6Kγ replication origin, tetacycline resistance (TetR), and transposable elements having FRT-loxP-PGKp/EM7p-npt-FRT-loxP fragments. The fine structure of the transposable elements is shown in (B), except for the middle part of the KAN/G418 resistance gene (npt). The loxP and FRT sequences are shown in closed and open boxes, respectively. The PGK promoter and 3’-termination sequence are indicated by lowercase letters. The EM7 promoter sequence is indicated by gray uppercase letters.
to create a lethality in the earliest stages of development. We decided that a defect of it may result inBecause it is predicted to be involved in basic cellular function of the mouse homologue (2868 bp) of the VPS41 gene knock-out allele is described in Figure 4. The overall structure and integrity of the subcloned fragment. Then, an appropriate clone with the loxP-Zeo-loxP cassette in a desired position and orientation, 324-4C/DT-A was selected.

by Cre recombinase expressed from cell type- and/or stage-specific promoters or genetically modified adenovirus (22).

We identified a BAC clone containing a part of the mVam2 locus from a mouse 129Sv genomic library by PCR. A PCR primer pair directed toward the 3'-noncoding region of the cDNA (GenBank accession no. ABB028843) gave a strong signal of the expected size upon PCR analysis (data not shown) from BAC clone 191m22 in the ES BAC library (Incyte Genomics). The nucleotide sequences of both ends of the BAC insert were determined and compared with a draft of the mouse genome sequence. BAC 191m22 was not enough to cover the entire region of mVam2, but contained the exon 13 to the last exon of the mVam2 locus (Figure 4). The overall scheme of the construction is summarized in Figure 4.

BAC 191m22 was partially digested with Sau3AI, and then the fractionated 10–15 kb fragments were cloned into the BamHI site of pBSDT-AII (Figure 3). Ninety-six colonies were isolated. Their plasmids were recovered and subjected to direct PCR analysis to obtain three individual clones containing exon 14 of mVam2. Through brief restriction mapping and end-sequencing using primer DT-AR or M13 Fwd23 (Table 1), we selected 324-4C/DT-A as a suitable subclone for the construction.

Next, loxP was introduced into 324-4C/DT-A by means of the in vitro transposition reaction. Plasmid 324-4C/DT-A was incubated with pMODloxZeoAmpl3 in the presence of a modified Tn5-derived transposase (EZ::TN transposase; Epicentre). The reaction products were introduced into E.coli XL1-Blue by transformation, and 96 ZeoR colonies were isolated. Their plasmids were recovered and subjected to sequencing using primer loxzeo-Cla or loxzeo-Nsi (Table 1 and Figure 2B) to determine the position and orientation of each tagging site. The sequence information also verified the structure and integrity of the subcloned fragment. Then, an appropriate clone with the loxP-Zeo-loxP cassette in a desired position and orientation, 324-4CDTAlloxZeo-64, was selected.
The plasmid was propagated in *E.coli* strain BH25.8 expressing Cre recombinase to remove the Zeo sequence and one *lox*P. Cre-mediated excision was very efficient: we found 100% removal in many experiments (data not shown).

The plasmid, 324-4CDTAloxDZeo-64, was recovered from BH25.8 cells and propagated in DH10B cells, because the quality and quantity of plasmids with BH25.8 cells was not satisfactory, partly due to intermolecular recombination between the *lox*P sites.

The final step was introduction of the *FRT-loxP-PGK/EM7p-npt-FRT-loxP* cassette into 324-4CDTAloxDZeo-64. We performed the second round of the *in vitro* transposition reaction in this step. Plasmid pGPS21loxFRTNeo was incubated with 324-4CDTAloxDZeo-64 in the presence of a modified T7 transposase and introduced into *E.coli* DH10B cells, and selected on ampicillin and kanamycin. We picked up 96 Amp<sup>R</sup> Kan<sup>R</sup> colonies and then determined the tagging sites by sequencing from both ends of the *FRT-loxP-PGK/EM7p-npt-FRT-loxP* cassette. The position and orientation of each insertion was determined by comparison with the genome draft and assembled sequence obtained in step 3 (Figure 4).

We chose 324DTLN-F05 as the final targeting vector. It has a 4.2 kb uninterrupted homologous left arm and a short right arm of 1.3 kb. The final construct contains *FRT-loxP-PGK/EM7p-npt-FRT-loxP* in intron 12, and the third *loxP* sequence in intron 16. The expression of Flpe recombinase will remove the positive selection marker, *npt*, leaving two *loxP* sites in introns 12 and 16. This *floxed* allele can be inactivated by the expression of Cre recombinase in particular tissues or developmental stages.

The plasmid was purified by alkaline-lysis and CsCl ultracentrifugation, linearized with *PmeI* at the multiple cloning site of the pBSDT-AII backbone and then electroporated into R1 embryonic stem cell (23). We identified 4 ES clones that underwent homologous recombination out of 264 G418R colonies on PCR analysis. These homologous recombinants were introduced into blastocysts of C57Bl/6 mice and we obtained germ line transmitting chimeric mice.

The idea of the transposon-generated gene targeting constructs was reported first by Westphal and Leder (12). We applied the idea to create conditional knock-out constructs that are indispensable for studying the function of genes involved in basic cellular functions including organelle assembly (24) or acidification (25). The described plasmids and transposons are components of the system applicable to any gene of interest, and we could save time and effort in our several ongoing gene targeting projects.

**Figure 3.** Structure of pBSDT-AII. Plasmid pBSDT-AII functions as the backbone for mouse targeting vectors. It has a unique multiple cloning site (MCS) and a negative selection marker, diphtheria toxin-A fragment (DT-A), to be expressed in mammalian cells through the MC-1 promoter (MC-1p). This plasmid can be propagated in common *E.coli* strains in the presence of ampicillin (Amp). The nucleotide sequence around the MCS and a part of the DT-A coding region are shown.
Although most of mouse genome sequence has been determined, there remain ambiguous segments scattering all over the genome. The \textit{mVam2} locus, as well as the other genes of our interests, is not assembled into contiguous sequence in the public databases, but several contigs separated by highly repetitive sequences. We avoid repeat sequences in the short homologous arm of the final targeting vector, because they are practically impossible to amplify by PCR thus compromise
the screening of ES recombinants. We determined the position and orientation of tagging sites by sequencing the resultant plasmids at steps 3 and 6 (Figure 4) using primers designed to be annealed to the ends of transposons. The 96 clones gave ~90 reliable sequences of 400–500 bp, in each direction, that cover almost all (10–20 kb) of the cloned segments in both strands, thus we could verify the genomic structure at the nucleotide sequence level.

In the case of the mVam2 gene targeting described here, we first subcloned a genomic segment from BAC into a plasmid, pBSDT-AII, by 'shot-gun' strategy involving restriction digestion, size fractionation and ligation. Several groups developed alternative highly efficient strategies for retrieving a large genomic DNA by homologous recombination in *E. coli* cells having a bacterial or phage-derived recombination system (26,27), and many excellent applications for engineering a large genomic DNA has been developed to date (28–31). We found that the use of such recombining system eliminates *in vitro* DNA manipulations. In current projects, we employ bacterial *in vivo* homologous recombination for step 1 of Figure 4, by using a PCR generated pBSDT-AII fragment flanked by 75 bp arms homologous to the ends of genomic segment to be retrieved. However, we prefer the transposon-mediated approach for insertion of *FRT*, *loxP* and *neo* gene to the recombining vector, because we can obtain the genomic structure of the locus as discussed above. In addition, because many potential candidates having the *neo* cassette with *FRT* and *loxP* at different position can be prepared simultaneously (Figure 4, step 5), there are more choices for assembling different targeting vectors without intensive work. This versatility is an obvious advantage if one construct fails to undergo homologous recombination in ES cells, which is not a rare event in our experience.

While this study was under review, Zhang et al. (32) reported a similar strategy to introduce *loxP* and selectable markers into targeting constructs by the use of Mu-based transposon. Combination of their and our sets of transposons will be possible because Mu, Tn5 and Tn7 can be integrated into one contiguous DNA segment, adding further layer of gene modification strategy.

**ACKNOWLEDGEMENTS**

We wish to thank Drs Osamu Minowa and Tetsuo Noda for providing the *loxP*-Kan plasmid, and Drs Andreas Nagy, Yujiro Higashi and Hiato Kondoh for ES cell line R1. We also thank Dr Neal G. Copeland for providing the *E. coli* homologous recombination system. This study was supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and by grants from the Noda Foundation and the Japan Foundation for Applied Enzymology. Funding to pay the Open Access publication charges for this article was provided by MEXT.

**Conflict of interest statement.** None declared.

**REFERENCES**


