

CDB LARGE GENE TARGETING PROTOCOL

(I) Generation of targeting and control vector

We usually perform positive selection using the neomycin resistance (*neo^r*) gene under the PGK promoter and negative selection with the diphtheria toxin A fragment gene under the MC1 promoter (Yagi et al., 1990, 1993a). Control vector (CV) is the vector used to establish the long-sequence PCR used to identify homologous recombinant colonies after drug (G418) selection (Murata et al., 2004).

Usually we generate targeting vectors (TVs) that have a longer 5' homologous arm and a shorter 3' homologous arm. The identification of homologous recombinant colonies is performed by long-sequence PCR over the short arm. Consequently, the CV is also made with the 3' arm. When one generates a TV that has shorter 5' arm and longer

3' arm, homologous recombinant colonies are to be identified by long-sequence PCR over the shorter 5' arm, and the CV should be made with the 5' arm. Attention should be paid that it is frequently difficult to establish PCR conditions of femtomole (fg) order sensitivity, which is essential for the identification of homologous recombinant colonies (see Setting PCR condition), by PCR over endogenous GC rich promoter regions or promoters, such as that of the PGK-1 gene that directs *neo^r* gene expression.

We recommend to construct the CV first and to test the PCR system in advance of constructing the TV (see Setting up PCR).

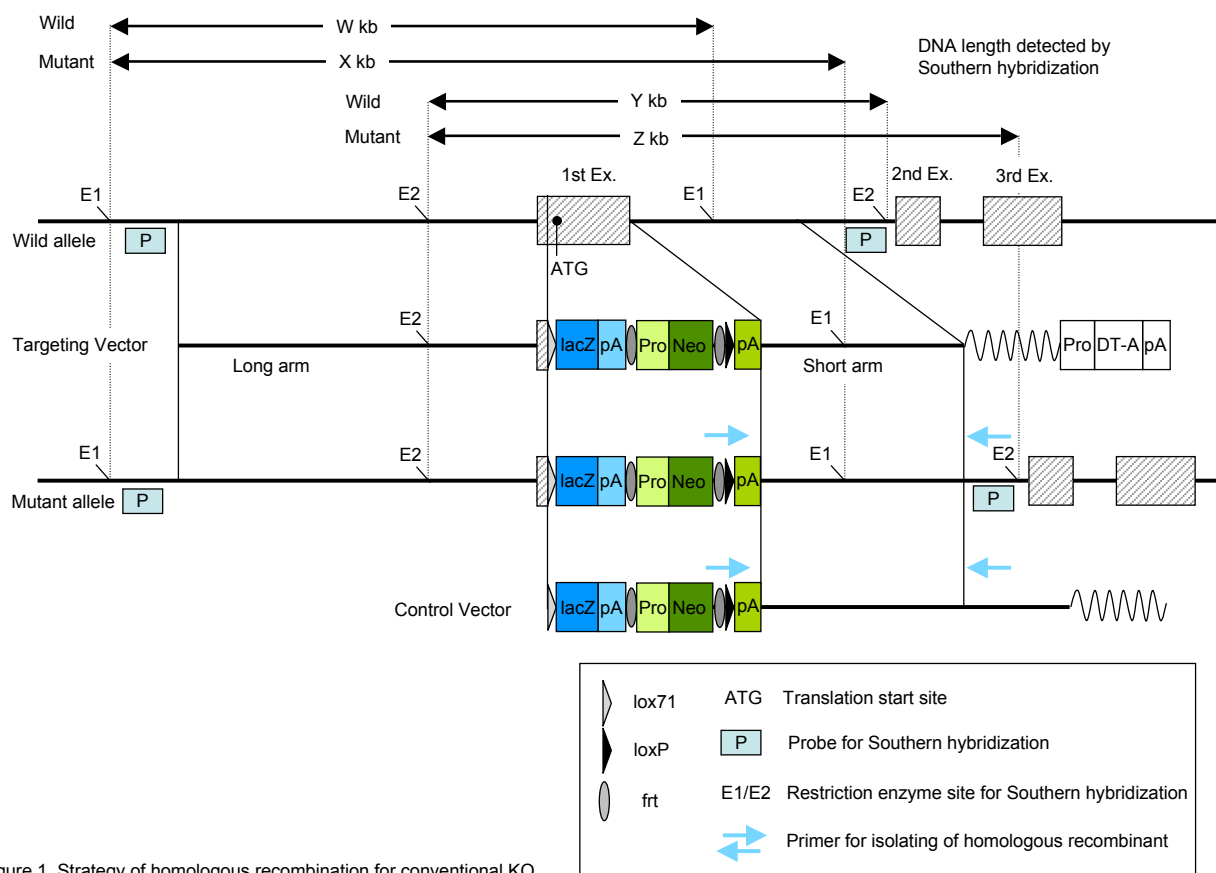


Figure 1. Strategy of homologous recombination for conventional KO

(1) Conventional knock-out and knock-in

The homologous arms of the TV and CV are isolated using the Red/ET recombination cloning system (Gene Bridges GmbH) and bacterial artificial chromosome (BAC) clone. TT2 cells are F1 hybrid between female C57BL/6 and male CBA (Yagi et al., 1993b). We usually obtain BAC clones from the C57BL/6 BAC library. When using C57BL/6 BAC DNA we usually expect homologous mutation to occur in the C57BL/6 allele of the TT2 cells, though it must be confirmed with polymorphic markers to be conclusive. When you want to make the mutation in the CBA allele, you must obtain the BAC clone from the CBA BAC library. The Red/ET recombination system assembles the target genomic DNA of any size and at any chosen position in the BAC clone into the Minimal Vector (MV) in the E.coli.

DT-ApA/Neo and DT-ApA/lacZ/Neo are the cassettes used to construct TVs to simply knock out genes and to knock-in the lacZ gene into each locus, respectively (see LARGE home page, <http://www.cdb.riken.jp/arg/>; Murata et al., 2004). Neo and lacZ/Neo cassettes that have no diphtheria toxin A fragment gene for negative selection are the corresponding cassettes to construct the CVs. The lacZ knock-in is designed so that the lacZ gene is incorporated into the 5' UTR in front of the translational start site of the target gene. Here the construction of the vectors is illustrated with an exam-

ple using the DT-ApA/lacZ/Neo cassette.

Using sequence information provided by you, we will identify the BAC clone by searching NCBI Blast search (<http://www.ncbi.nlm.nih.gov/BLAST/>) or UCSC genome Bioinformatics (<http://genome.ucsc.edu/index.html?org=Mouse>); the BAC clone will be obtained from BACPAC resources (<http://bacpac.chori.org/>).

To obtain the 5' arm two short homologous sequences (SHSs) of 50bp are chosen; one at about 8kb 5' upstream that corresponds to the 5' end of the 5' arm (5'SHS) and the other at the site of mutation that corresponds to the 3' end of the 5' arm (3'SHS) (Fig. 2). In the knock-in mutation 3'SHS corresponds to the site between the 5'UTR and translational start site of the gene. At the 5' end of the 5'SHS a restriction enzyme site for either AscI, AflII or Sal I is generated (we primarily choose SalI) and at the 3' end of the 3'SHS a site for AflII, SalI or NotI (we primarily choose NotI). The 5'SHS is further combined with the 24 bp sequences at the 5' end of the MV, and the 3'SHS with the 22 bp sequences of the MV. The modified MV is amplified by PCR and electroporated into E. coli that has the target BAC clone expressing pRed/ET. In the E. coli the 5' arm is assembled into the MV, and the 5' arm is cloned into the DT-ApA/lacZ/Neo cassette by SalI/NotI digestion (Fig. 2).

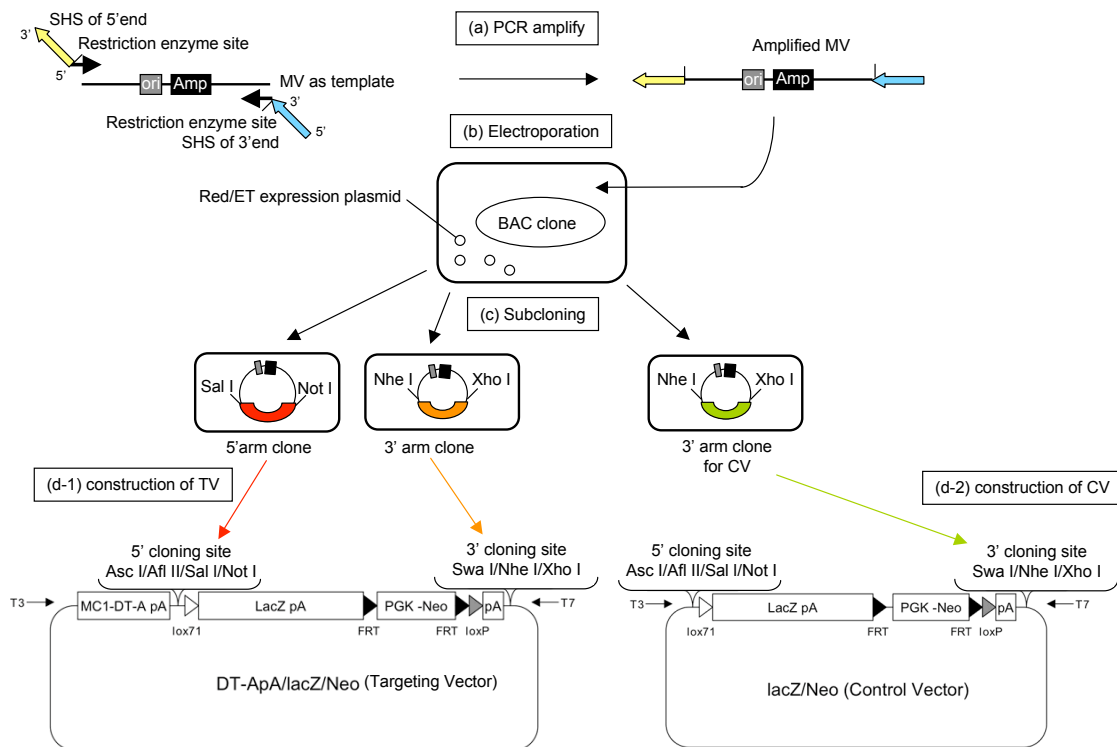


Figure 2. Cloning of homologous region for conventional KI

The 3' arm of the TV is obtained similarly, using the 5' SHS at the site of the mutation with *SwaI* or *NheI* sites (we primarily choose *NheI*) and 3' SHS at about 4kb 3' downstream with *NheI* or *XhoI* sites (we primarily choose *XhoI*) and cloned into the DT-ApA/*lacZ*/Neo cassette at the *NheI* and *XhoI* sites. The 3' arm of the CV must be longer than the 3' arm of the TV (we usually make it about 2kb longer), so that the primer used to identify homologous recombinant ES colonies is set up in the 2kb extra region (Fig. 1). The 3' arm of the CV is

cloned similarly into the *lacZ*/Neo cassette; alternatively it can be obtained by long-sequence PCR as described previously (Murata et al., 2004).

(2) Conditional KO

For conditional KO, three homologous regions must be cloned separately; the 5' arm, the floxed-out body, and the 3' arm (Fig. 3). The floxed-out body region is sandwiched by *loxP* sites and excised by the Cre protein in the targeted allele.

(II) Setting up PCR

After positive and negative selection with a TV, each surviving colony is composed of about 4000 cells. We identify homologous recombinant colonies with half of these cells (about 2000); the other half of the cells is passaged into one of 24-well dish. 2000 cells have about 10 ng DNA in which one copy of the vector corresponds to a few fg.

The PCR should be set up so that the long-sequence PCR over more than 4kb works at a sensitivity of a few femtograms between the two primers: one primer in the *neo^r* insert and the other primer at the 3' site that is included in the CV, but not in the TV (Fig. 1). Sometimes the difficulty to achieve fg sensitivity can be overcome by changing primers and PCR condition. We do not recommend to make a second PCR with the second inner primers, since this usually generate a high number of false positives. Sometimes, probably due to the sequence composition, the PCR condition at fg sensitivity cannot be established using the 3' arm in the design, and as a result the TV must be redesigned. For this reason, we recommend generating the CV first and testing the PCR condition in advance of constructing the TV.

PCR setups vary by equipment. For the LA PCR Kit (TaKaRa) we use the following:

PCR buffer: 10xLA PCR Buffer II (Mg²⁺ free, TaKaRa)

MgCl₂: final 1.5 – 2.5 mM

DNA polymerase: LA Taq (TaKaRa)

Pretreatment: at 96 °C for 5 min

PCR: 45 cycles (denaturation, 1 min at 96 °C; annealing and extension 5 min at 68 °C)

Post-treatment: at 72 °C for 15 min.

First PCR conditions are tested with 0.5-10 fg CV DNA. When the sensitivity is obtained at a few fg order, 10ng ES cell DNAs are added to 0.5-10 fg CV DNAs, and the sensitivity is confirmed. The PCR condition must be readjusted if the fg sensitivity is lost with the ES DNAs.

Subsequently, 1000 wild type ES cells are added to 0.5 -10 fg CV DNAs and the fg sensitivity is re-confirmed.

When you construct the vector, you have to make these tests to establish the PCR condition. Without this, we cannot begin to isolate homologous recombinants.

Finally, the CV is linearized with an appropriate restriction enzyme and electroporated into the ES cells (see (III) Isolation of homologous recombinant ES cells <Preparation of DNAs>). After selection against G418, twenty colonies that randomly integrated the CV are arbitrarily picked up using a 200 µl Gilson pipette, and placed into trypsin solution, after which PCR is performed using half of the cells (about 2000 cells) of each colony. About 2000 trypsin-suspended cells of each colony are washed free of serum with PBS buffer, lysed in water at 95 °C for 5 min, treated with protease K at 55 °C for 2 h and heat-treated at 95 °C for 10 min. The extract is subjected to PCR. The PCR condition is considered to have been established when more than 80% of colonies are PCR-positive.

Even when you construct the vector, we (LARGE) perform this electroporation upon receipt of the CV, and we will send you about twenty colonies of the cells for you to make the PCR.

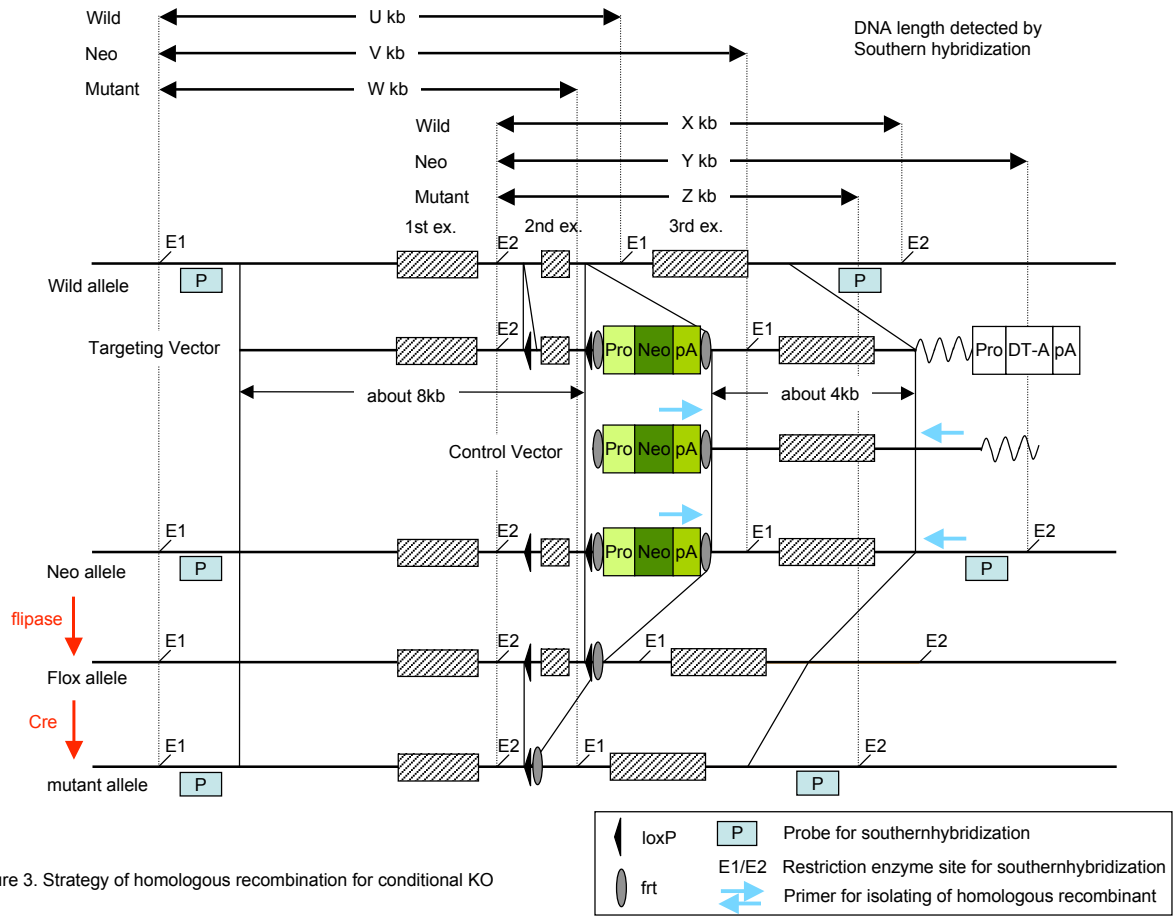


Figure 3. Strategy of homologous recombination for conditional KO

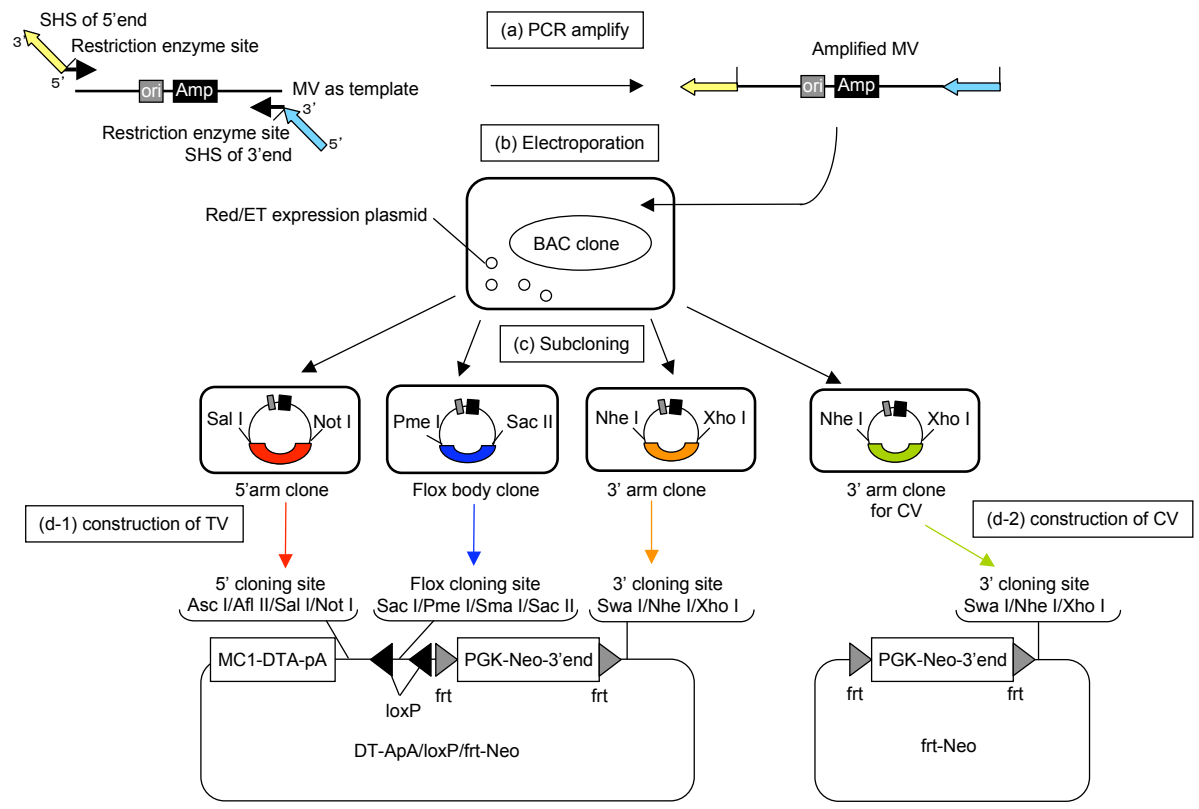


Figure 4. Cloning of homologous region for conditional KO

(III) Isolation of homologous recombinant ES cells

Once the PCR is set up with the CV as described above, it is the time to start the isolation of homologous recombinants with the TV. The TV is

linearized with the appropriate restriction enzyme digestion and electroporated into TT2 ES cells (Yagi et al., 1993b).

<Preparation of DNAs>

TV and CV DNAs are purified by equilibrium centrifugation with CsCl-ethidium bromide gradient; we do not recommend purification using QIAGEN columns. Purified DNAs are linearized with an appropriate restriction enzyme so that DTA-pA locates at the 3' end of TV or plasmid sequences at the 3' end of CV. The completeness of the linearization must be checked by gel electrophoresis. The linearized DNAs are purified by sequentially treating with phenol once, with phenol/chloroform twice, and with chloroform once. After the last chloroform treatment, the supernatant fluid must be maintained under sterile conditions. DNAs are precipitated with ethanol and twice rinsed with 70% ethanol.

The DNA precipitates are then dried moderately and dissolved in HBS buffer at about 60nM concentration. For 10kb DNA, 45µg in 100µl corresponds to about 60nM. Prepare more than 500µl of 60 nM TV DNA and 300µl of 60 nM CV DNA and send them to us either at ice-cold temperature or frozen. Indicate your name and the name of the vector on the label.

HBS buffer:

25 mM HEPES

137 mM NaCl

5 mM KCl

0.7 mM Na₂HPO₄ · 12H₂O

6 mM Dextrose

Adjust to PH 7.0 and sterilize by filtration.

-to be continued-

<Electroporation >

<Selection and harvest of surviving colonies>

<Identification of homologous recombinants by long PCR>

<Propagation of potential homologous recombinant clones>

(IV) Generation of mutant mice

<Generation of chimeric mice>

<Generation of heterozygous mutants and breeding of mutant mice for homozygous phenotype analysis>

References

Murata, T., Furushima, K., Hirano, M., Kiyonari, H., Nakamura, M., Suda, Y., and Aizawa, S. (2004). ang is a novel gene expressed in early neuroectoderm, but its null mutant exhibits no obvious phenotype. *Gene Expr. Patterns* 5, 171-178.

Yagi, T., Ikawa, Y., Yoshida, K., Shigetani, Y., Takeda, N., Mabuchi, I., Yamamoto, T., and Aizawa, S. (1990). Homologous recombination at c-fyn locus of mouse embryonic stem cells with diphtheria toxin A fragment gene in negative selection. *Proc. Natl Acad. Sci. USA* 87, 9918-9922.

Yagi, T., Nada, S., Watanabe, N., Tamemoto, H., Kohmura, N., Ikawa, Y., and Aizawa, S. (1993a). A novel negative selection for homologous recombinants using diphtheria toxin A fragment gene. *Anal. Biochem.* 214, 77-86.

Yagi, T., Tokunaga, T., Furuta, Y., Nada, S., Yoshida, M., Tsukada, T., Saga, Y., Takeda, N., Ikawa, Y., and Aizawa, S. (1993b) A novel ES cell line, TT2, with high germline-differentiating potency. *Anal. Biochem.* 214, 70-76.