Recombineering protocol #4

Design of conditional gene targeting vectors - a recombineering approach

Søren Warming, Ph.D.

The purpose of this protocol is to help you in the gene targeting vector design process. It is not a *practical* recombineering protocol. For practical references, please see other protocols from our website, or the paper by Dr. Pentao Liu *et al.* (Genome Res. 2003 Mar;13(3):476-84).

Overview

Our recombineering approach to making conditional targeting vectors basically involves 4 steps:

(1) "Retrieval" of a 10-15 kb genomic fragment from a BAC by gap repair of a plasmid containing a TK cassette for negative selection in ES cells.
(2) 1st "targeting"; insertion of a loxP flanked (floxed) selection cassette into this retrieved genomic fragment, along with a restriction enzyme for genotyping, 5' to the region (exons) to be floxed.
(3) Cre-mediated "pop-out" of this cassette in bacteria, to leave a single loxP site behind, along with the engineered restriction site
(4) 2nd targeting; insertion of a selection cassette flanked by two Frt sites and a single loxP site, along with one or more engineered restriction sites, 3' of the region to be floxed. This selection cassette stays in the conditional targeting vector for positive selection in ES cells.

After obtaining correctly targeted ES cells, we normally proceed directly to blastocyst injection. Then, after germline transmission of the targeted allele (the "neo" allele), we obtain the conditional knock-out allele ("cko" allele) by crossing to mice expressing *Flpe* from a general promoter (β-actin), and the knock-out ("ko") or null, allele is obtained by crossing to mice expressing *Cre* from a general promoter (also β-actin). Of course, these alleles can also be obtained by expressing *Flp* or *Cre*, respectively, directly in the ES cells, before proceeding to blastocyst injection.

Targeting strategy

This step is crucial and probably the most difficult one. It is really worth spending some time here to avoid problems (including trouble with genotyping) later in the process.The following should be considered:

- What region to target? Has a "straight" knock-out been made before? If so, then it might be a good idea to flox the exon(s) that was/were deleted previously. Any known functional domains that are likely to be crucial for protein function? Alternatively, decide on the exon(s) that will "cause the most damage", a large exon that takes up most of the protein, for example. We don't recommend putting the "lonely" loxP site in the promoter (floxing the first exon), since there is no easy way to predict whether the presence of that loxP site will interfere with promoter function.

- Make sure the absence of the floxed exon(s) will result in an out-of-frame fusion of the remaining exons. This way you are more likely to create a null allele and not just a truncated protein.

- Using a mouse genome browser (like http://genome.ucsc.edu/), download the genomic sequence of interest into a suitable DNA software package (like Vector NTI, MacVector, DNA strider, Gene Construction Kit, or Sequencer)

- Locate suitable restriction sites that can be used for genotyping by Southern blot analysis. Good genomic cutters include: EcoRI, EcoRV, BamHI, BglII, HindIII. Ideally, for both the 5' and the 3' end, the enzyme(s) you choose should cut once *outside* of the region you will use as targeting vector (the flanking site) and again distal to the region you are floxing (the internal site). You can then engineer a new recognition site for this enzyme along with the lonely loxP site (for the 5' enzyme) and a recognition site along with the frt-neo-frt-loxP cassette (for the 3' enzyme) to obtain restriction fragments sizes different from the wildtype bands. The enzyme(s) you choose can be the same for both ends, provided it doesn't cut inside the region, or it can be two different enzymes, one for the 5' end and one for the 3' end.

- Order one or a few BACs from the genomic region. For example, use http://ensembl.org/Mus_musculus, or http://bacpac.chori.org/ to order BACs from a 129 or a C57BL/6 library. It's still being debated whether the DNA source of the targeting vector has to be isogenic with the ES cell. We use a hybrid F1(129xC57BL/6) ES cell line and use targeting vectors made from either 129 or C57BL/6 DNA.

- It is important, when designing the strategy, to include a diagnostic restriction site to be included with the "lonely" loxP site—otherwise there is no way of telling if the loxP site is actually inserted in the genome of the ES cells. Depending on the strategy, this restriction site can be engineered so it remains in the genome after Cre mediated recombination. Alternatively, to aid in genotyping, it may sometimes be an advantage to engineer it such that it disappears after recombination.

- Do a test-Southern analysis using the selected enzymes, to make sure they give the expected wildtype band sizes using the ***flanking*** probes to be used for genotyping. These probes can be amplified with PCR from the ordered BACs. At the same time you test that the probes are actually good enough (specific and with minimal background). The probes should be flanking, so they don't hybridize to the gene targeting vector (otherwise you can't distinguish between random insertion and homologous recombination in the ES cells).

- The Frt-Neo-Frt-loxP cassette. The Neo cassette that remains in the conditional knock-out (cko) vector is flanked be two Frt sites and a single loxP site. It is of course important that the lonely loxP site (5' loxP) and the loxP site next to Neo are in the same orientation (loxP and Frt sites are directional recognition motifs).

- Hypomorphic allele. The orientation of the Pgk-Neo cassette is also important. We sometimes find that when the Pgk-neo cassette downstream of the exon(s) to be floxed is in *opposite orientation* relative to the modified gene, the strong Pgk promoter can interfere with normal transcription/splicing of this gene, resulting in neo/neo homozygous mice with a hypomorphic phenotype (a phenotypic less strong than the null phenotype)—probably do to an insufficient expression level of the gene. This is sometimes a desirable feature, so you might want to consider this. Basically, it's a "free" extra allele that could be useful for your phenotyping analysis. Of course, there is no guarantee that you will always create a hypomorphic allele if Pgk-Neo is oriented in the opposite orientation. On the other hand, there is no guarantee that you will *not* get interference when Pgk-Neo is in the *same* orientation as you gene! If you decide to go for the potential hypomorphic allele, please remember this: Using our vectors (PL451 and PL452) you will have to use both *Flp* and *Cre* to get the "ko" allele, due to the relative position of the Frt and loxP sites. When Pgk-Neo is in the same orientation as your gene, a *Cre* step is all it takes to get the "ko" allele. Alternatively you can make a new vector yourself where the relative position of loxP and Frt is opposite.

- Some general rules for positioning loxP sites: The floxed region should ideally be 2 kb or shorter—that is, up to 2 kb between the two loxP sites. Avoid positioning loxP sites so they interfere with splicing. About 150-200 bp upstream of the splice acceptor (cAG) and 150-200 bp downstream of the splice donor (GT) seems to work. Be extra careful when positioning loxP sites in very small introns.

Retrieval and introduction of loxP sites

- Genomic fragment to be retrieved and turned into a conditional targeting vector. The size of the region to be retrieved depends on your genotyping

strategy, the position of the restriction sites and the position of your flanking probes. Again, this is why it's a good thing to test the probes on a test-Southern first; if you design two probes that map immediately inside of the flanking enzymes, retrieve the region just inside of those probes, and it turns out that the probes are bad, then you have to start all over again…

- Retrieval vector. We use a pBluescript-based vector (PL253 or similar) or a pBR322-based vector that already contains a TK cassette for negative selection in ES cells. Into this vector you clone two "mini-arms", 300-500 bp each. These "mini-arms" are designed so they are homologous to the extreme 5' and 3' end of the region to be retrieved. The easiest way is to PCR amplify these arms from the BAC with primers containing unique restriction sites, and then do a triple ligation into the retrieval vector. For example: the retrieval vector is cut with NotI+ BamHI, the 5' arm is cut with NotI+XhoI, and the 3' arm with XhoI+BamHI. Please note: I use the term "miniarm" for retrieval vector construction, and the term "arm" for the long sequence of homology used for homologous recombination in ES cells. Eventually, when your targeting vector is finished, you will want the "arm" containing the lonely loxP site to be the one initiating homologous recombination in the ES cells, and since this is normally the 5' "arm", you will want to place the unique NotI site at the 5' end of the 5' "miniarm" when designing your retrieval vector. Our vector is designed so the NotI site is proximal to the plasmid backbone. This way, when your final targeting vector is linearized with NotI, the plasmid backbone will protect the TK cassette from degradation in ES cells until homologous recombination takes place. Please remember to check that the restriction sites you add during PCR amplification of the "mini-arms" for subsequent cloning don't cut inside those "mini-arms".

- Mini-targeting cassettes for inserting loxP and Frt-Neo-Frt-loxP.

  (1) For the 1st targeting step, the fastest way is probably to use PCR to amplify a neo/kanamycin cassette driven by a procaryotic promoter and flanked by two loxP sites. You can use the PL452 plasmid as template and just amplify the em7-neo part (921 bp), and then add homology+loxP sequences + recognition site for a restriction enzyme. The primers (100 bp, most companies don't make longer oligos) should look like this (annealing temperature should be 60°C):

  5'42bp-homology_restr.enz(6 bp)_loxP(34bp)_cgacctgcagcctgttga3'
  5'48bp-hom.(reverse)_reverse-loxP(34bp)_gtcgaggctgatcagcga3'

  (2) Same goes for the 2nd targeting step, although for this step you need to include the eucaryotic Pgk promoter and the polyA signal since this cassette is going to be used for positive selection in ES cells. The

primers should be designed like this (if you need to introduce a restriction site, subtract from the length of the homology arm, so you have a total length of 100 bp):

5'48bp-hom_gaagttcctattctctagaaagtataggaacttcaggtctgaagaggagttt3'
5'50bp-hom_ ataacttcgtatagcatacattatacgaagttatattatgtacctgactg3'

For the PCR approach, use 1-2 ng plasmid template, and after the PCR reaction (25 cycles) add 1-2 □l DpnI directly to the reaction, incubate for 1 hour @ 37°C and gel-purify o/n. DpnI destroys the plasmid template but leaves the PCR product intact (cuts only dam-methylated DNA)—this way there will be no background due to residual template. The primers listed here work well with an annealing temperature of 60°C.

(3) Alternatively, and a bit more time-consuming, but often resulting in more efficient recombineering, you can use PCR to amplify 300-500 bp "mini-arms" just as for the retrieval vector, and clone these arms into PL451 and PL452, using SalI-EcoRI and BamHI-NotI. This can be done in a single step using 4-way ligation, or in two steps, by cloning one arm at a time. You then release the "mini-targeting" vector from the plasmid backbone by digesting with SalI-NotI, and gel-purification o/n. Background due to residual plasmid backbone can be minimized by (a) cutting with a third enzyme that disrupts the amp gene (ScaI), and (b) by digesting less vector (about 250 ng and no more than 500 ng).

Other considerations

- Efficiency of 1st and 2nd targeting steps is highest if you introduce your targeting vector(s) into heat-shocked and electrocompetent bacteria already containing the plasmid with the retrieved fragment. This way all bacteria have the target, so efficiency is increased. This is true especiallly when using the PCR strategy, since the homology arms are relatively short. Alternatively, co-transform the target plasmid and the targeting vector. In both situations, and especially if the plasmid was present in the bacteria first, you might observe a "funny" restriction digest pattern resulting from a mixed "population" of un-modified retrieval plasmid and targeted plasmid. You will see some bands of roughly half the intensity as the other. This is because some restriction fragments don't change size after recombination, and you get these both from the un-modified and the targeted plasmid, whereas the "new" bands, derived from the introduction of the selection cassette, only come from the targeted plasmid, so they are "under-represented" in the population. Often, however, you should be able to find a clone that only has the targeted molecule (a pure clone). If not,

you can either (a) re-transform the plasmid (select for kanamycin resistance), or (b) linearize with NotI (unique cutter), column-purify the digest to get rid of the enzyme, dilute, self-ligate, and transform (select for kanamycin resistance).

- When your cko vector is finished and ready, you should do a functional test of the loxP and Frt sites. This is easy: simply transform your vector into arabinose-induced and electrocompetent EL250/SW105 (for test of functional Frt sites) or EL350/SW106 (for test of loxP sites), and plate on amp (kanamycin resistance should be gone). Test a few colonies by miniprep and compare to the original cko vector. The changes in band sizes tells you whether everything looks o.k. and works all right.