

## Experimental guidelines, suggestions and FAQs

The following information is intended to provide you, a potential user of the Vanderbilt Transgenic Mouse/Embryonic Stem Cell Shared Resource (TMESCSR) with information to help you make effective use of this resource. The information is not intended to be a comprehensive review of the literature on the genetic manipulation of the mouse. However, we have included references to articles and books that go beyond what we provide here.

### Outline

#### I. Design of Gene Targeting Vectors

- A. General considerations
- B. Strain origin of mouse DNA
- C. Recombineering to build targeting vectors
- D. Specific considerations

#### II. Choice of Embryonic Stem Cell Lines

- A. Discussion of available lines
- B. Frequently asked questions about ESCs

#### III. BAC Recombineering and Transgenesis

- A. Bacterial Artificial Chromosomes (BACs) and “Recombineering”: A primer
- B. How to use BACs
- C. Considerations for selecting BACs and handling BAC DNA
- D. How to locate BACs on genome browsers
- E. Web Links
- F. Protocols and Reagents

#### IV. References for Sections I-III.

### I. Design of gene targeting vectors

*Please note: The design of a gene targeting experiment is a scientific issue and is not provided as a service. Some experimental approaches may be conceptually or technically novel, or involve propriety reagents, and thus may necessitate collaboration.*

#### A. General considerations

There are many factors to consider when designing a gene targeting vector, some of which can greatly affect the experimental outcome. First, the type of gene targeting vector also depends on the experimental goals. For instance, the considerations involved in designing a vector to generate a conditional allele are different from those for making a cassette acceptor or knock-in allele. Second, the vector design is also affected by the nature of the gene and the type or structure of the protein it encodes. Third, the location of restriction enzyme sites as well as presence of repetitive DNA sequences need to be taken into account.

If you have never performed a gene targeting experiment and lack the experience and knowledge necessary to make these judgments we strongly recommend that you identify a collaborator to assist you with the steps that are described below. There are several laboratories within the Vanderbilt University Medical Center that have extensive experience in

the design of gene targeting vectors. We are happy to make recommendations. If you do not think collaboration with anyone is necessary but still want some advice you may obtain consultation on fee-for-service basis on the design of gene targeting vectors from Drs. Magnuson or Labosky, and on specific BAC or Red/ET recombineering issues from Dr. Mortlock. There is some written material that can help such as 'Gene Targeting, A Practical Approach', by A.L. Joyner or 'Manipulating the Mouse Embryo' by A. Nagy (Joyner, 1993; Nagy et al., 2003). However, for the most part these texts focus on general issues and often ignore factors that may be critical for the success of a specific type of experiment.

Please keep in mind that even with an 'optimally designed' targeting vector some genetic loci may be difficult or impossible to target. Thus, there is no guaranteed outcome even if hundreds of colonies are screened.

It may not even be necessary to perform a gene targeting experiment to obtain a genetically engineered mouse. There are several repositories of ESCs that have specific retroviral insertions that may have resulted in null mutations. For instance, the Texas Institute for Genomic Medicine (<http://www.tigm.org>) has large libraries of mutant C57BL/6N and 129/SvEvBrd ESCs. These libraries, which together contain over 640,000 mutant cell lines, represent over 13,000 genes. Also, the NIH funded 'Knock-Out Mouse Project (KOMP) is making rapid progress so the KOMP repository ([www.komp.org](http://www.komp.org)) should be searched prior to undertaking any new knock-out experiment.

## **B. Strain origin of mouse DNA used for targeting vectors**

The TMESCSR has the most experience using TL1 mouse ESCs for gene targeting since they reliably give germline transmission (see section II below). Since these cells were derived from 129S6 animals (previously called 129SvEvTac) efficient gene targeting requires DNA that is of 129S6 origin for both the DNA homology arms of the targeting vector. DNA from the very closely related 129S7 (129SvEv) mice is also permitted since it is a nearly identical strain. Even so, it remains possible that small genetic differences could interfere with specific gene targeting experiments since single base mutations can prevent homologous recombination. **Thus, we require that you use either 129S6 or 129S7 DNA to assemble your gene targeting vector.**

Use of PCR to amplify the homology arms is discouraged. Although this may seem like an easy way to avoid having to obtain a BAC clone even single nucleotide mismatches can greatly impair gene targeting efficiency. It may be difficult to determine if base pair mismatches identified during DNA sequence confirmation are due to PCR amplification errors or are simply 129S6 strain specific.

Sources of genomic clones made from 129S6 or 129S7 DNA include:

1. The RPCI-22 BAC genomic DNA library was made from 129S6 ESCs (<http://bacpac.chori.org/mouse22.htm>); however, this particular library was not end-sequenced so it is not indexed in *GenBank* or *Ensembl*; rather, it is available as a spotted array and clones must be identified by hybridization and then characterized to determine whether they contain the full region of interest. Nonetheless, this is the only source we know of for obtaining BAC

DNA that is perfectly isogenic to the TL1 ESCs. It is generally only necessary to screen one segment of the library to obtain several BACs containing your gene of interest.

2. A BAC genomic DNA library has also been made from AB2.2 ESCs, which were made from 129S7 mice. The library has been end-sequenced and is indexed in *Ensembl* ([www.ensembl.org](http://www.ensembl.org)). Thus, the easiest way to obtain a genomic DNA clone for gene targeting experiments is to identify a BAC clone covering a ~30-40 kb region of your gene of interest using the *Ensembl* browser. You can search the *Ensembl* web site for your gene at the *Ensembl* homepage. To visualize these BAC clones, click on the [ContigView] of your search result, check the box labeled '129S7/AB2.2 clones' under 'DAS Sources' drop-down box under the Detailed View section. The 129S7/AB2.2 BAC clones are displayed as green and pink, which indicates the orientation of the DNA insert in the vector. End reads are shown as grey bars. The clones may be purchased from *Geneservice, Ltd.* online at <http://www.geneservice.co.uk>. They are described on the company's website as Mouse bMQ BACs from the Sanger Institute.

The TMESCSR has also developed ESCs from C57BL/6 blastocysts (see section II below).

### C. Recombineering to build targeting vectors

**It is highly recommended that you assemble your gene targeting vector using BAC or Red/ET recombineering methods.** This is advantageous since longer homology arms can easily be incorporated and the method does not rely on the fortuitous presence of restriction enzyme sites. However, depending on the type of targeting vector being made, not all starting materials are available from online resources.

BAC or Red/ET recombineering can be performed two different ways. The Stewart method has the advantage of not requiring transfer of the BAC DNA into specialized strains of *E. coli*, as does the Copeland method. Otherwise, the methods are similar and success is readily achievable with either approach.

A more detailed description of BAC recombineering and other BAC-related issues is in **Section III**. Section III.B.1 contains information on how and where to obtain reagents for the Stewart or Copeland methods.

### D. Specific considerations for design and use of targeting vectors, and screening targeting events.

1. All replacement-type targeting vectors require two DNA homology arms. **It is highly recommended that the long arm be at least 5 kb in length, and the short arm be at least 2 kb.** This is easy to achieve if BAC recombineering is used to assemble the targeting vector. If possible the homology arms should be free of repetitive DNA sequences. **The absolute minimum size of the DNA homology arms is 4 and 1 kb for the long and short arms, respectively. Gene targeting will not be attempted if the arm lengths do not meet these minimum requirements.** LoxP sites or introduced point mutations that interrupt a DNA homology arm must not be considered when determining the lengths of the DNA sequence.

Due to the exquisitely sensitive nature of homologous recombination any such changes, even a single base, reduce the effective length of the DNA homology regions.

2. When generating a conditional allele, the LoxP flanked excisable region should be kept as short as possible to minimize the probability of homologous recombination occurring in this region and to be sure that Cre-mediated recombination occurs efficiently in the mouse. While there is no firm rule about the length of the LoxP-flanked region it is generally not necessary for it to be longer than 1–2 kb. Also, please keep in mind that Cre-mediated recombination in mice is more efficient when the LoxP sites are not too widely separated.
3. The targeting vector must include at least one chemical selection cassette. Cassettes containing both a phosphoglycerol kinase (PGK) promoter and a neomycin (G418) resistance gene are widely used for so-called positive selection. Recently, use of puromycin resistance has gained favor among some investigators. The TMESCSR can perform positive selection for neomycin, puromycin or hygromycin resistances. Many promoters that function well in somatic cells (CMV, for example) do not function in ESCs and constructs using this promoter to drive a selectable cassette will not be electroporated.
4. It is nearly unthinkable to design a targeting vector without being able to remove the positive selection cassette by site-specific recombination. This is due to numerous reports indicating aberrant effects on the expression of neighboring genes. Thus, the positive selection cassette should be flanked by recombinase recognition sites such as FRT or LoxP. If the goal is to generate a Cre-mediated conditional allele then FRT sites should flank the selection cassette. The use of a triple LoxP strategy is not recommended except in special situations. A dual LoxP/FRT strategy is preferable given the high efficiency of Flp-mediated recombination by FlpE transgenic mice. It is suggested that the positive selection cassette is removed in mice versus in the ESCs. First, the intermediate allele may function as a hypomorphic allele and thus is of possible value. Second, identifying FlpE-mediated recombination in ESCs is not nearly as straightforward as in mice. Third, minimizing *in vitro* manipulations of the targeted cells will improve the likelihood of the ESCs transmitting through the germ line.
5. Incorporation of a negative selection cassette outside the short arm of homology is also highly recommended since this may improve the ratio of targeted clones versus random integrants, thereby decreasing the number of clones that need to be picked and screened. Use of a diphtheria toxin (DT) cassette for negative selection is preferable over a *Herpes Simplex Virus* thymidine kinase (HSV-TK) since it avoids the use of gancyclovir. However, since PGK-HSV-TK cassettes are in wide use the TMESCSR is happy to accommodate this approach.
6. DNA sequence validation of the targeting vector should be performed. At a minimum, DNA sequencing across all cloning or recombination junctions must be performed. To assure the lack of mutations any important open reading frames should also be sequenced, especially if PCR was used at any point in the assembly process.
7. All recombinase recognition sites (e.g. FRT and LoxP) must be sequence confirmed, regardless of origin.

8. The orientation of all selection cassettes needs to be known and documented.
9. A complete, assembled DNA sequence file needs to be provided to the TMESCSR prior to electroporation. The DNA sequence file should be in GenBank format and include all annotations present in the gene target. Vector NTI or other DNA applications may be used to generate this type of file. These DNA sequences will be entered into a database and will be retained as confidential information until such time that the results of the study are published by the Principal Investigator.
10. It is recommended that the screening and validation of all homologous recombination events be done by Southern blot hybridization. In our experience it is more reliable than PCR since it also detects potentially mosaic clones. The initial screen is best performed across the short homology arm. However, before any clone is used to produce mice it is essential to confirm homologous recombination at both the 5' and 3' ends of the targeting vector since it is not uncommon for there to be occasional unpredicted changes in the gene locus.
11. If the gene targeting vector will be used to generate a conditional allele then it is also necessary to develop PCR primers that can be used to detect the presence of the LoxP site that is located most distal from the antibiotic selection cassette. The presence of this LoxP site needs to be confirmed prior to blastocyst injection. The PCR primers developed for this purpose can also be used for genotyping of mice.
12. Careful thought should be given to the screening strategy before embarking on vector assembly. Both the cost and digestion efficiency of restriction enzymes vary greatly. It is also important to know the chromosomal location of your gene of interest. Traditional gene targeting efforts have not been successful on the Y Chromosome but some more advanced gene targeting methods can be used (Rohozinski et al., 2002; Simpson et al., 2002). Since germline transmission after blastocyst injection involves passage of the mutation through male germ cells, genes located on the X Chromosome present unique challenges. Moreover, mutations on either sex chromosome may present other issues such as sex-linked lethality. Thus, it may be necessary to use a conditional approach to target genes on the X Chromosome.
13. All DNA probes and restriction digests must be tested by Southern blot hybridization prior to the TMESCSR performing an experiment. It is essential to use DNA isolated from the TL1 ESCs for these blots. The TMESCSR is happy to provide DNA isolated from these cells specifically for this purpose. **Gene targeting experiments will not be initiated until the PI demonstrates the ability to detect the wild type locus by Southern blot hybridization.**
14. The DNA that is used for electroporation of ESCs must be of high quality, and provided to the TMESCSR as linear DNA. Thus, it is important not to forget to include a unique restriction site somewhere outside the homology arms for linearization. Qiagen kits have a proven track record of working very well for purifying vector DNA for this purpose; however other several other brands of kits also provide good results. Also, it should be kept in mind that linear DNA is unstable so once your DNA has been quantified and linearity is verified electrophoretically, the sample should be stored frozen at -20° C and not thawed, even in transit to the TMESCSR.

15. The design of all new targeting vectors will be reviewed by the TMESCSR Management Team prior to initiation of the experiment. The TMESCSR has no obligation to perform a poorly designed experiment. Indeed, this review is performed to avoid wasted effort and expense, and to increase the likelihood of the desired outcome. For this reason, it is highly recommended that the design of a gene targeting vector be reviewed by one of co-directors before effort is expended in assembling the vector.

These guidelines were written by Mark Magnuson with input from Trish Labosky, Doug Mortlock, Jennifer Skelton, Weiping Yuan and Jill Lindner.

Last update: October 27, 2008

## **II. Choice of Embryonic Stem Cell Lines**

### **A. Discussion of available lines**

As noted above, we prefer the use 129S6 ESCs for gene targeting. The main reason for this is that these cells have a high propensity for transmission through the germline. TL1 ESCs, a line that was derived at Vanderbilt University by Trish Labosky in the laboratory of Brigid Hogan in 1994 (Labosky et al., 1994), have been used for years and our experience with these cell lines is extensive. Of the over 160 gene targeting experiments performed by this resource over the past 15 years, the overwhelming majority were performed using this cell line. Indeed, the ease of germline transmission when using TL1 cells has been a key component of the ongoing success of this resource.

Although ESCs are “immortal”, as their passage number increases, they tend to accumulate chromosomal abnormalities and their ability to transmit through the germ line decreases (Longo et al., 1997). Presently, we are using TL1 cells that were frozen at passage 13 and they continue to perform optimally, but in anticipation of the eventual exhaustion of these reserves, we have recently generated new ESCs from 129S6 blastocysts and we have called these cells JDS1. These cells are germ line competent and we have used them in gene targeting experiments in which they perform at or above the level of TL1 cells. We are currently testing this targeted JDS1 cells for germ line competency.

Despite the years of experience using 129S6-derived mESCs, this strain of mice is not suitable for many types of experiments. A comprehensive discussion of 129 genetic background issues can be found elsewhere (Simpson et al., 1997; Threadgill et al., 1997). For this reason, there has been considerable interest in ES cell lines from different strains of mice. We have recently generated germ line competent ES cell lines from C57BL6/J blastocysts (called JDSB61) and purchased a C57BL/6N line from Primogenix (called PRX B6N and derived by Robin Weselschmidt), so we have several C57BL/6 lines that are germ line competent in our hands. We are also striving to generate albino C57Bl/6 cell lines that would allow germ line transmission to be easily detected by coat color.

### **B. Frequently asked questions about ESCs**

We have generated a short list of FAQs. If you have other questions please email us and we

will add to this document over time. We are hoping to make this a useful resource for the first time user of this technology.

### **1. What DNA should I use to make my targeting vector?**

You should definitely use isogenic DNA to make your vector as this can affect the efficiency of gene targeting by up to 20-fold (te Riele et al., 1992). That means that if you are performing gene targeting in 129S6 ESCs you should use 129S6 genomic DNA. This DNA is available through the TMESCSR or you can purchase 129S6 and 129S7 BACs from commercial vendors as discussed in sections I.B.2 and III.D.6.

### **2. What ESC line should I use for gene targeting?**

First, you should use a line that is germ line competent and that has been proven germ line competent after gene targeting. Next, you need to consider the experiments you are undertaking with the mutant mice that you generate. For example, do these mice need to be on a specific genetic background? Many behavioral studies are conducted on C57BL/6 background and that is why we have generated these lines. While it is also an option to backcross mice to obtain a congenic line (an almost pure genetic background), it is impossible to ever be 100% on a background different from the ES cell line you started in since the mutation was generated in 129S6 DNA. Also, backcrossing mice takes a considerable amount of time. If you use 129S6 as the ESC strain, you should first determine that your chimeras can transmit the mutation through the germ line. After this, while you still have the germ line transmitting chimeras, it is a good idea to cross these chimeras with 129S6 females thereby placing your mutation on a pure genetic background. Once the chimeras are past breeding age this can only be accomplished by backcrossing.

### **3. What issues and risks exist for gene targeting in C57BL/6 cells?**

As long as your targeting vector is generated using isogenic DNA, there should not be notable differences in gene targeting efficiency with C57BL/6 ESCs. However, after you obtain your correctly targeted cell lines, they need to be injected into blastocysts to generate chimeras. It is not desirable to inject C57BL/6 cells into C57BL/6 blastocysts because you will not be able to distinguish chimeras or germ line transmission by coat color. To get around this, we are generating albino C57BL/6 ESCs, but we have not yet determined if our lines are germ line competent. An alternative is to inject C57BL/6 ESCs into albino C57BL/6 blastocysts. This is considerable more expensive since albino C57BL/6 mice are approximately 2.5x the cost of non albinos.

### **4. Are 129-derived ESCs better or worse than C57BL/6 cells?**

Neither are better or worse. People simply have much more experience with 129S6 ESCs, so that is usually the safer way to go.

### **5. Is karyotype analysis helpful in predicting germline transmission?**

Yes. We routinely perform karyotyping analysis on newly generated ES cell lines (fee for service if desired). For some reason, mouse ES cell lines are genetically unstable and tend to lose chromosomes, especially the Y Chromosome. They also become polyploid on occasion. It is highly unlikely that cells with the wrong number of chromosomes will generate normal gametes and transmit through the germ line, the magic number appears to be 50%, if 50-100% of the metaphase spreads contain 40 chromosomes the line is likely to transmit through the germ line, if the line has less than 50% of the metaphase spreads with 40 chromosomes it is not desirable to inject this cell line (Longo et al., 1997). Many factors seem to alter karyotype, the most common is the cell cycle time, cells that are dividing rapidly often have increased chromosomal abnormalities (Liu et al., 1997).

## **6. Should a high passage number concern me?**

Not necessarily. If cells are cultured carefully they can maintain their germ line competence up to and over P30. We (and others) have taken single cell clones of high passage ESCs that would not transmit through the germ line and generated new lines that are now germ line competent. (This is not a preferred approach, however!) But in general, lower passage numbers are better than higher ones (Nagy et al., 2003).

## **7. How long does it take to get a chimera from the time a targeting vector is submitted?**

Currently we can tackle a project within two weeks to one month and get an electroporation under way after the paperwork and DNA are submitted and approved. To obtain chimeric mice for breeding depends on the PI's ability to screen and return results to us, but this averages around 6 months if that screening is done correctly and quickly.

## **8. How many clones do you suggest injecting?**

This is totally PI dependent. If you are targeting a locus that has never been targeted before and the null phenotype has not been described, we suggest having two independent cell lines/mouse lines that give the same result. If this is a well-characterized locus, one line might be acceptable for publication.

## **9. How many days per clone do you suggest injecting?**

Again, this is PI and money dependent. We usually suggest 2 days per ESC clone because the cells have to be brought up for injection anyhow and a second injection day usually only requires one additional passage. This also covers the possibility inherent in natural matings that there may be few or no blastocysts recovered for injection on a given day.

## **10. I generated a heterozygous ES cell line using the Core several years ago and would like to have a homozygous line to use in my research. What is the best way to go about this?**

There are several ways to generate a homozygous ESC line starting with the heterozygous mouse or ESC line. The first way is easiest; you can culture the heterozygous lines in an

increased concentration of G418 and some of the cells will convert to homozygous (Mortensen et al., 1992). The problem with this approach is that often the ESCs convert back to heterozygous after a while in culture and the increased selection can often lead to chromosomal defects that may confound study of the phenotype. However, this is a simple approach that can be taken in the PI's laboratory and detailed protocols are available (Mortensen, 2001). A second approach is to target the remaining wild type allele *in vitro*. If the *neo* selection cassette is flanked by LoxP sites, it can be removed using a transient transfection with Cre. Then the original targeting vector can be used to target the second allele. The final approach, and the one least impacted by increased culture time, is to generate an ESC line *de novo* from homozygous mutant blastocysts. We have done this successfully and if there is interest we might be offering this as a fee-for-service technology.

### III. BAC Recombineering and Transgenesis

#### A. Bacterial Artificial Chromosomes (BACs) and BAC “Recombineering”: A primer

##### 1. What are BACs?

BACs are bacterial cloning vectors that can carry large (150-250 kb or more) inserts of foreign genomic DNA ligated to a 7-12 kb vector. They are propagated in bacteria as circular molecules at 1-2 copies per cell.

##### 2. What is BAC recombineering?

Recombineering is the engineering of BAC clones by homologous recombination. This is performed inside bacterial cells, using *E. coli* cells that have been engineered to permit homologous recombination, via controlled expression of the analogous Red or ET protein systems. This enables insertion, deletion or replacement of DNA sequences in the BAC, or subcloning (retrieval) of fragments out of the BAC into plasmids. Since homologous recombination is used, the engineering is not limited by location of restriction sites used for traditional ligation steps.

##### 3. Why use BACs?

**a. BAC library resources:** Excellent BAC libraries exist for the genomes of mouse, human and other species. BAC clones that span a gene of interest can usually be quickly identified on the web, using genome browsers (see below). BACs containing mouse DNA from the C57/BL6 strain or the 129/Sv strain can be identified this way. If you can identify a BAC clone on a genome browser, you can usually buy a bacterial glycerol stock of the clone from an outside vendor. BAC DNA can be engineered purified using methods that are accessible to most labs that use routine molecular cloning methods.

**b. Specific uses of BACs:** For users of the transgenic core, BACs are primarily useful for **(1) building targeting vectors** for ESC-mediated gene targeting experiments, or **(2) generating transgenes intended for pronuclear injection**. In the case of building targeting vectors, BACs

can be very useful for obtaining the homology arms that are needed for proper targeting of vectors in ESCs. Also, by using recombineering methods many steps in building the targeting vector can be done using the BAC itself as a scaffold. This can eliminate many of the often-difficult ligation steps needed to make targeting vectors.

In the case of pronuclear injection, BACs can be very useful for making transgenes that drive expression of desired genes (e.g. CRE, LacZ, GFP). This is primarily due to the large size of BACs as compared to traditional plasmid-based transgenes. Because of the large size of BAC inserts, they are more likely to contain all cis-regulatory elements for the gene in question. Unlike smaller transgenes, which are fairly prone to position-mediated silencing and ectopic expression following transgene integration, BAC transgenes tend to be more resistant to these problems.

#### **4. What are key issues to know about using BACs?**

BACs are large as compared to most standard plasmid cloning vectors (150-250 kb vs. 3-15 kb). Therefore, purified BAC DNA is more susceptible to shearing and enzymatic degradation than most plasmids, so not all methods and prep kits that work for plasmids can be used for BACs. However, standard methods and kits for BAC preps are widely available. They are low copy vectors, so the DNA yield is low compared to high-copy plasmids.

### **B. How to use BACs**

#### **1. Recombineering methodologies: an overview**

“Recombineering” refers to various methods that all use homologous recombination in *E.coli* to engineer modifications into vectors (primarily, BACs). Of these, the most widely used systems are based on controlled expression of proteins that permit homologous recombination at reasonable efficiencies. These allow recombination through homologies as small as 42 bp. Since most *E.coli* strains do not have this capability, the methods require either the use of engineered strains in which the recombination genes have been inserted into the *E.coli* chromosome (e.g. Neal Copeland method), or alternatively, the recombination genes are temporarily transferred into the cells via a plasmid vector (e.g. Francis Stewart method). In either case, the recombination proteins are under inducible control so they are not constitutively expressed. They are induced transiently to permit recombination, and the desired recombinant clones are usually identified with antibiotic selection.

a. For the Copeland method, a very useful website exists that describes available protocols, reagents and instructions for requesting reagents: <http://recombineering.ncifcrf.gov/> Descriptions have been published on how to use these tools to engineer BACs (Copeland et al., 2001; Lee et al., 2001; Warming et al., 2005) and for making knockout vectors (Liu et al., 2003).

b. For the Francis Stewart method, descriptions of use (Muyrers et al., 1999; Zhang et al., 1998) and a review (Muyrers et al., 2001) are available. A kit containing the necessary material and information can be purchased from <http://www.genebridges.com/>.

## 2. Constructing gene targeting vectors using BACs

First, read **Section I** above for important considerations that apply to all targeting vectors.

Vectors for ESC-mediated gene targeting need to have several components: sufficiently long homology arms to the target locus, positive and negative selection cassettes, and often other components such as LoxP and FRT sites. Recombineering frees the researcher from needing to depend on restriction sites and ligations for cutting and pasting these together in vitro. Essentially, recombineering allows direct subcloning of a targeting fragment from a BAC into the plasmid in order to build the targeting vector. Recombineering can be used either before or after this step (that is, to modify the BAC first, or the plasmid subclone) to insert selection cassettes, LoxP sites, or other modifications into the targeting fragment, depending on your specific design goal. Articles describing these approaches have been published (Chan et al., 2007; Liu et al., 2003).

A typical workflow for building targeting vector from a BAC might be as follows:

- a. Transfer BAC into strain permitting homologous recombination.
- b. Subclone a segment of the BAC containing the targeting cassette into a plasmid that contains a negative selection vector (e.g. HSV-Tk).
- c. Recombine other components as needed, e.g. PGK-NeoR, FRT or LoxP sites, into desired locations in the BAC. LoxP-flanked and FRT-flanked Neo cassettes exist to facilitate building many types of vectors.

Note, the precise order of recombineering steps and whether they occur before or after subcloning from the BAC may vary depending on your design strategy.

## 3. Pronuclear injection of BACs

BAC DNA can be successfully used for pronuclear injection. BAC DNAs are usually injected at 0.5-1.5 ng/ $\mu$ l concentration, which is slightly lower than what is typically used for plasmid DNAs (~3 ng/ $\mu$ l). The lower concentration helps reduce needle clogging problems that are more common with injecting BAC DNA.

BAC DNAs can be injected as uncut, circular molecules or linearized and both have been used successfully to make transgenic mice. For both BACs and plasmid DNAs, a rapid process of breakage, religation and recombination between transgene isomers occurs that tends to generate concatamers of transgene DNA before integration into a random genomic location (Bishop and Smith, 1989). This is often sufficient for the scientific goals of the project. While plasmid transgenes are almost always injected as linear fragments, BACs may be injected as circular or linear molecules with similar integration frequencies. Most BAC vectors have *NotI* sites near the vector/insert junctions, so many mammalian BAC inserts can be cut free of the

BAC vector using NotI. For projects where the insertion of single, linear, intact BACs is absolutely required, linearization and purification steps may be needed.

A typical workflow of recombineering and injecting BACs might be as follows:

- a. Transfer BAC into strain permitting homologous recombination.
- b. Recombineer the BAC as desired.
- c. Purify BAC DNA by cesium chloride gradient or BAC column kit (CsCl method is strongly recommended.)
- d. Verify concentration and integrity of purified BAC DNA by agarose gel analysis.
- e. Dilute BAC DNA prior to injection.

The Mortlock lab has made available a BAC DNA purification protocol that has been reliably used to generate transgenic mice (Chandler et al., 2007a; Chandler et al., 2007b):

[http://web.mac.com/mortlock1/iWeb/Public%20lab%20site/Protocols%20and%20reagents\\_file/BAC%20CsCl%20prep%20Oct07.doc](http://web.mac.com/mortlock1/iWeb/Public%20lab%20site/Protocols%20and%20reagents_file/BAC%20CsCl%20prep%20Oct07.doc)

### C. Considerations for selecting BACs and handling BAC DNA

**1. Gene structure issues:** Most, but not all mammalian genes can fit entirely within BACs. Existing mouse and human BAC libraries have fairly deep coverage and at least some clones exist for almost every region of the genome without gaps. However, individual BAC clones will vary in the extent of 5' and 3' coverage around a gene of interest. Genome browsers (see below) are invaluable for selecting BAC clones that best match your needs. BACs can contain multiple genes, so if the BAC is intended for pronuclear injection, overexpression of linked genes may have undesired phenotypic effects or even lethality. In this case, recombineering may be used to trim other genes out of the BAC, or a BAC with a closer breakpoint might be a better choice.

**2. Fragility:** DNA molecules that are over 30-50 kilobases are much more prone to shearing and degradation than smaller molecules. Standard techniques for plasmid DNA handling such as vortexing or repeated pipetting to dissolve DNA pellets are not appropriate for BACs and will shear the DNA. Minimize pipetting steps and use wide-bore tips when possible. Keep purified BAC DNAs on ice or at 4°. Store at 4° and avoid freeze-thawing, which will cause shearing.

**3. Validation:** BACs can usually be propagated very stably. However, keep in mind you are trying to engineer the BAC clone. PCR is usually not sufficient to verify completely that successful engineering has occurred for two reasons: first, false positives can occur if the clone is not truly clonally pure (i.e. a mixture of recombineered and unmodified cells); second, PCR usually can't tell you if your BAC is still intact. Restriction digest of BAC DNA to observe predicted alterations, and lack of unwanted rearrangements, is highly recommended. Prior to pronuclear injection, BAC DNA should be run on a pulsed-field gel to verify it is intact and of correct size. Standard agarose gel electrophoresis cannot distinguish well between intact and sheared BAC DNA.

**4. Purification Kits:** Standard plasmid DNA prep kits usually cannot purify intact BAC DNA. This is because almost all kits use DNA-binding columns with wash buffers optimized for low-molecular-weight plasmid DNAs. Different buffers are needed for large molecules such as BACs, so use BAC purification kits. However, standard alkaline lysis / ethanol precipitation is fine for isolating BAC DNA in sufficient amounts for simple PCR and gel analysis.

**5. Direct BAC sequencing** is best performed on BAC DNA made by cesium chloride method or BAC prep kits. Note, for sequencing, shearing is not as big a concern as purity.

#### **D. How to locate BACs on genome browsers:**

Several mouse, human and other BAC libraries have been deeply end-sequenced, meaning that single sequence reads were generated from both ends of the insert for most clones. If both end sequences for one clone can be aligned to unique positions in the genome so that they are “facing each other” and within a few hundred kilobases, their aligned positions probably reflect the actual ends of the BAC insert. These predictions can be displayed on genome browsers. Most of them are right, but when you get your BAC it is important to check that it contains your gene by PCR.

**1. C57BL/6 strain BACS:** The **UCSC genome browser** has extensive BAC maps for the *Mus musculus domesticus* (C57BL/6 strain) BAC libraries RPCI-23 and RPCI-24. To visualize them on the browser:

- a. Log on to UCSC genome browser gateway page  
<http://genome.ucsc.edu/cgi-bin/hgGateway>
- b. Select mouse genome and search for your gene or genomic region of interest
- c. On the browser window page, scroll down below window to “Mapping and sequencing Tracks”. Click and change the BAC end pairs pulldown tab setting to “full”.
- d. Click the “refresh” button. BAC positions should be visible. BACs from the C57BL/6 strain will have the prefix “RP23” or “RP24”. Note: BACs with the prefix “MSMg01” are from a subspecies *Mus musculus molossinus* BAC library.

**2. 129S7 strain BACs:** The **Ensembl genome browser** has BAC maps for the bMQ library that was made from this strain. See part I.2.b above for a detailed description on how to locate a BAC for your gene.

#### **3. Where to get BACs**

a. *Mus musculus domesticus* (C57/BL6 strain) BAC libraries RPCI-23 and RPCI-24:  
Weblink to: [BACPAC Resources Center \(BPRC\)](#)  
(The BPRC is at the Children’s Hospital Oakland Research Institute (CHORI), Oakland, CA.)

b. *Mus musculus domesticus* (129S7 strain) BAC library bMQ  
Weblink to: [Geneservice, Ltd.](http://www.geneservice.co.uk)  
<http://www.geneservice.co.uk>

c. *Mus musculus domesticus* (129S6 strain) BAC library RPCI-22: Clones and library filters are available from the [BACPAC Resources Center \(BPRC\)](#). Clones are also available from [Invitrogen](#) (call to inquire about availability of pooled plate DNAs to facilitate screening the library by PCR).

Note: Invitrogen and Open Biosystems also provide clones from the mouse CITB/CTB BAC libraries; however, these are derived from the 129SV substrain, which has been shown to be genetically contaminated with DNA from a non-129 strain. In other words, clones from this strain may or may not be isogenic to 129S6 or 129S7 ESCs and we discourage their use to make targeting vectors.

**4. A description of BAC clone nomenclature can be found at:**

<http://www.ncbi.nlm.nih.gov/genome/clone/nomen.html>

## **E. Web Links**

Very useful NCI web site with information, protocols and instructions for requesting reagents:

<http://recombineering.ncifcrf.gov/>

UCSC genome browser gateway page

<http://genome.ucsc.edu/cgi-bin/hgGateway>

Ensembl genome browser

<http://www.ensembl.org>

BACPAC Resources Center (BPRC)

<http://bacpac.chori.org/>

[Geneservice, Ltd.](#)

<http://www.geneservice.co.uk/home/>

## **F. Protocols and Reagents**

1. Protocols and reagent request forms for Copeland method strains can be found at

<http://recombineering.ncifcrf.gov/>

A kit with reagents for the Stewart recombineering method can be obtained from Genebridges at <http://www.genebridges.com/>

2. [Mortlock lab BAC DNA CsCl prep protocol](#)

3. [Mortlock lab BAC or plasmid DNA miniprep protocol](#)

#### IV. References Cited in Sections I-III:

- Bishop, J. O. and Smith, P.** (1989). Mechanism of chromosomal integration of microinjected DNA. *Mol Biol Med* **6**, 283-98.
- Chan, W., Costantino, N., Li, R., Lee, S. C., Su, Q., Melvin, D., Court, D. L. and Liu, P.** (2007). A recombineering based approach for high-throughput conditional knockout targeting vector construction. *Nucleic Acids Res* **35**, e64.
- Chandler, K. J., Chandler, R. L., Broeckelmann, E. M., Hou, Y., Southard-Smith, E. M. and Mortlock, D. P.** (2007a). Relevance of BAC transgene copy number in mice: transgene copy number variation across multiple transgenic lines and correlations with transgene integrity and expression. *Mamm Genome* **18**, 693-708.
- Chandler, R. L., Chandler, K. J., McFarland, K. A. and Mortlock, D. P.** (2007b). Bmp2 transcription in osteoblast progenitors is regulated by a distant 3' enhancer located 156.3 kilobases from the promoter. *Mol Cell Biol* **27**, 2934-51.
- Copeland, N. G., Jenkins, N. A. and Court, D. L.** (2001). Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* **2**, 769-79.
- Joyner, A. L.** (1993). Gene targeting: A Practical Approach. New York: IRL Press at Oxford University Press.
- Labosky, P. A., Barlow, D. P. and Hogan, B. L.** (1994). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. *Development* **120**, 3197-204.
- Lee, E. C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D. A., Court, D. L., Jenkins, N. A. and Copeland, N. G.** (2001). A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56-65.
- Liu, P., Jenkins, N. A. and Copeland, N. G.** (2003). A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res* **13**, 476-84.
- Liu, X., Wu, H., Loring, J., Hormuzdi, S., Disteché, C. M., Bornstein, P. and Jaenisch, R.** (1997). Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission. *Dev Dyn* **209**, 85-91.
- Longo, L., Bygrave, A., Grosveld, F. G. and Pandolfi, P. P.** (1997). The chromosome make-up of mouse embryonic stem cells is predictive of somatic and germ cell chimaerism. *Transgenic Res* **6**, 321-8.
- Mortensen, R.** (2001). Production of a homozygous mutant embryonic stem cell line (double knockout). *Curr Protoc Mol Biol Chapter 23*, Unit 23 6.
- Mortensen, R. M., Conner, D. A., Chao, S., Geisterfer-Lowrance, A. A. and Seidman, J. G.** (1992). Production of homozygous mutant ES cells with a single targeting construct. *Mol Cell Biol* **12**, 2391-5.
- Muyrers, J. P., Zhang, Y. and Stewart, A. F.** (2001). Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA. *Trends in Biochemical Sciences* **26**, 325-31.
- Muyrers, J. P., Zhang, Y., Testa, G. and Stewart, A. F.** (1999). Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Research* **27**, 1555-7.
- Nagy, A., Gertsenstein, M., Ventersten, K. and Behringer, R.** (2003). Manipulating the Mouse Embryo: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

- Rohozinski, J., Agoulnik, A. I., Boettger-Tong, H. L. and Bishop, C. E.** (2002). Successful targeting of mouse Y chromosome genes using a site-directed insertion vector. *Genesis* **32**, 1-7.
- Simpson, E. M., Johnson, K. A., Shirley, B. J., Fang, G. Y., Bayleran, J. K. and Lerner, C. P.** (2002). Novel Sxr(a) ES cell line offers hope for Y chromosome gene-targeted mice. *Genesis* **33**, 62-6.
- Simpson, E. M., Linder, C. C., Sargent, E. E., Davisson, M. T., Mobraaten, L. E. and Sharp, J. J.** (1997). Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat Genet* **16**, 19-27.
- te Riele, H., Maandag, E. R. and Berns, A.** (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc Natl Acad Sci U S A* **89**, 5128-32.
- Threadgill, D. W., Yee, D., Matin, A., Nadeau, J. H. and Magnuson, T.** (1997). Genealogy of the 129 inbred strains: 129/SvJ is a contaminated inbred strain. *Mamm Genome* **8**, 390-3.
- Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. and Copeland, N. G.** (2005). Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* **33**, e36.
- Zhang, Y., Buchholz, F., Muyrers, J. P. and Stewart, A. F.** (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genetics* **20**, 123-8.