A New Method for Rapidly Generating Gene-Targeting Vectors by Engineering BACs Through Homologous Recombination in Bacteria

Vinícius Cotta-de-Almeida,1,2,3 Susan Schonhoff,4 Tomoyuki Shibata,1,2 Andrew Leiter,4 and Scott B. Snapper1,2,5

1Gastrointestinal Unit and the Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, Massachusetts 02114, USA; 2Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA; 3Department of Ultrastructure and Cell Biology, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Rio de Janeiro, 21045-900, Brazil; 4Division of Gastroenterology, GRASP Digestive Disease Center, New England Medical Center/Tufts University School of Medicine, Boston, Massachusetts 02111, USA

Generating knockout mice is still an expensive and highly time-consuming process. Target construct generation, the first labor-intensive step in this process, requires the manipulation of large fragments of DNA and numerous, and often cumbersome, cloning steps. Here we show the development of a rapid approach for generating targeting constructs that capitalizes on efficient homologous recombination between linear DNA fragments and circular plasmids in Escherichia coli (“recombineering”), the availability of bacterial artificial chromosomes (BACs), and the accessibility of the sequence of the mouse genome. Employing recombineering, we demonstrate with only 1–2 template plasmids, short homologies (40–50 bp) between donor and target DNA, and one subcloning step that we can efficiently manipulate BACs in situ to generate a complicated targeting vector. This procedure avoids the need to construct or screen genomic libraries and permits the generation of most standard, conditional, or knock-in targeting vectors, often within two weeks.

Methods

The generation of knockout mice through gene targeting by homologous recombination in embryonic stem (ES) cells is still an expensive and highly time-consuming process. Knockout construct assembly is the first step in a laborious process that can take up to a year even in the most technically skilled laboratories. Typically, the process requires screening a genomic library to obtain the gene of interest, followed by restriction mapping and numerous cloning steps.

One strategy that can avoid time-consuming subcloning steps and may offer a more efficient approach for generating targeting constructs is the use of homologous recombination in E. coli (Eggleston and West 1996). The basis of the technology is to utilize strains of E. coli that can efficiently carry out homologous recombination between short terminal homology regions on a linear PCR-derived DNA fragment and sequences on a recipient plasmid. This strategy has been referred to as either recombineering—for recombogenic engineering (Copeland et al. 2001; Court et al. 2002)—or recombination cloning (Zhang et al. 2002). This approach has been utilized to alter various DNA targets, including the E. coli chromosome, high copy number plasmids, and bacterial artificial chromosomes (BACs; Zhang et al. 1998, 2002; Angrand et al. 1999; Muyrers et al. 1999; Datsenko and Wanner 2000; Yu et al. 2000; Copeland et al. 2001; Lee et al. 2001; Liu et al. 2003).

While two groups have employed this strategy for the construction of gene targeting vectors, simplifying some of the time-consuming steps for targeting vector construction, these approaches still required gene isolation or the construction of genomic libraries and/or numerous complex gene manipulations. (Angrand et al. 1999; Zhang et al. 2002). We desired a construction strategy that would preclude the need to screen for and subsequently map genomic sequences and, most importantly, would avoid all of the cumbersome cloning steps. To simplify the process, we reasoned that one could identify the genomic sequence of the gene of interest from publicly available databases and then purchase an E. coli clone containing a BAC encompassing these sequences rather than employing standard genomic library screening. With the appropriately constructed template vectors (described below) and the essential recombination machinery, we could then use homologous recombination in E. coli to manipulate directly the gene of interest in the BAC. Finally, in one step, we could remove the manipulated sequence within the BAC by identifying flanking restriction sites from the published sequence, and under appropriate antibiotic selection, generate a targeting construct. Overall, this approach would obviate the need for the numerous time-consuming cloning steps typically required for generating complicated targeting constructs.

RESULTS AND DISCUSSION

We sought to generate a variety of constructs for the conditional targeting of the NOD2/CARD15 gene (the murine homolog of a novel Crohn’s disease gene; Hugot et al. 2001; Ogura et al. 2001). From available sequence databases, we identified the NOD2 gene and purchased a BAC clone containing all 11 coding exons. The overall strategy for generating a conditional targeting vector was to employ two successive recombination steps to introduce a loxP site 5’ to exon 3 and an aminoglycoside phosphotransferase (aph) gene cassette (mediating kanamycin resistance or neomycin resistance when expressed in prokaryotes or eukaryotes, respectively) flanked by loxP sites (“floxed”) 3’ to exon 3 in the BAC clone (Fig. 1). To mediate each homologous recombination step, we introduced a plasmid containing the arabinose-inducible Red
recombination system from bacteriophage λ (Murphy 1998; Datsenko and Wanner 2000) into the BAC containing strain.

In the first recombination step, intronic BAC sequences 5’ to exon 3 were targeted with PCR products containing a loxP site and an frt-flanked aph gene (Fig. 1a). These PCR products were amplified with 60-bp primer pairs consisting of 40 bp of sequence homologous to the intronic region flanking the targeted exon, and 20 bp complementary to the template plasmid containing the DNA to be inserted (Fig. 1a). The template plasmid, pKD4Lox, was derived from pKD4 (Datsenko and Wanner 2000), a conditional replicon (oriR) that requires the pir gene product (absent in most E. coli strains that carry BACs) for replication. Transformation with the loxP/aph-containing PCR fragment resulted in several kanamycin-resistant (KnR) recombinants (Fig. 1b). Following transformation with a temperature-sensitive replicon that expresses the flp recombinase (Cherepanov and Wackernagel 1995), which recognizes frt sites, the aph gene was excised by site specific recombination (Fig. 1c).

The second step was to manipulate BAC intronic sequence 3’ to exon 3. Kanamycin-sensitive strains isolated above (Fig. 1c) were targeted with a PCR product containing a floxed aph cassette generated from pSBS150. This pKD4-based template contains the aph gene under the control of both prokaryotic and eukaryotic promoters to allow selection in bacterial and ES cells (Fig. 1d). From the genomic sequence, we identified a restriction enzyme site (KpnI) flanking the manipulated exon that would yield a fragment of sufficient length for homologous recombination in ES cells (Fig. 1e). To isolate the construct, the recombinant BAC was digested with KpnI, and the fragments were subcloned into a multi-copy cloning vector (e.g., pBluescript or a thymidine kinase [TK]-containing plasmid; Fig. 1f), and selected for ampicillin resistant (ApR) and KnR colonies. The presence of the aph gene within the region of the BAC of interest allowed, upon kanamycin selection, enrichment of clones containing only the desired DNA fragments. PCR (Fig. 1g), restriction enzyme, and sequencing analyses confirmed the desired products in all steps.

We also developed a similar method for a “one-step” targeting strategy that would be applicable for either conditional knockout targeting construct generation or for the “knock-in” of a mutated exon. We first generated pSBS156, another conditional replicon. This plasmid contained a unique EcoRI site to allow the cloning of a wild-type or “mutated” exon of interest flanked by a loxP sequence and a floxed aph gene cassette (Fig. 2a). Subsequently, we developed a NOD2 “knock-in” template by introducing into pSBS156 a mutated exon 10 that corresponded to the mutant genotype in CD (Fig. 2b, pSBS158; Hugot et al. 2001; Ogura et al. 2001). Homologous recombination between linear PCR fragments and the BAC generated KnR recombinants containing the mutated exon and a 3′ floxed aph gene, but not the 5′ loxP. To circumvent this recombination that occurred within the exon (which would be incompatible with a conditional targeting construct), we increased the homology arm in the sense primer to 50 bp, and selected a target sequence located 100 bp upstream of the mutated exon as the 5′ region for homologous

Figure 1  Conditional knockout vector construction. (a–f) Sequential steps leading to a NOD2 exon 3 recombinant BAC and targeting vector (see text for details). (g) Gel electrophoresis depicting PCR products confirming the recombinants schematized in diagrams: (a) lane 1 (WT-BAC), (b) lane 2 (BAC loxP/ft/aph/ft), and (c) lane 3 (BAC loxP), generated with primers 1 (P1) and 2 (P2); (d) lane 4 (BAC loxP) and (e) lane 5 (BAC loxP-loxP/aph/loxP), generated with P3 and P4; and (f) lanes 6, 7 (targeting construct) using respectively P3 and P4, and P1 and P4. Circled “F”, frt site; large arrowhead, loxP site.
recombination. Appropriate recombinants were identified and confirmed (Fig. 2c,e). Restriction analyses of the recombinant BACs revealed the absence of unwanted internal rearrangements (data not shown). A targeting construct was generated by digesting the recombinant BAC with NheI, followed by subcloning into a TK-containing vector and selection on ampicillin/kanamycin plates (Fig. 2d).

Although not described here, construction of simple targeting vectors that aim to replace a portion of a gene with an *aph* gene cassette can be generated using pSBS150 as template. PCR primers homologous to the sequence flanking the deleted regions can be used to generate a linear fragment for homologous recombination with the BAC. It was previously shown that primers designed in this way efficiently loop out intervening sequence with replacement by the selectable marker (Zhang et al. 1998; Angrand et al. 1999; Datsenko and Wanner 2000; Yu et al. 2000).

Two groups previously described methods for the construction of targeting vectors utilizing recombinase (Angrand et al. 1999; Zhang et al. 2002). Both of these approaches required screening genomic libraries and did not utilize the availability of BACs. Here we describe a simple and rapid method of generating complex gene knockout targeting constructs using recombinase engineering requiring only one or two plasmid templates (Figs. 1, 2). This approach exploits the commercial availability of mouse genomic BAC clones spanning known sequences in the mouse genome. Since they can accommodate 5-10-fold larger inserts than bacteriophage, the use of BACs ensures that sufficient sequence is available to generate large regions of homology required for efficient homologous recombination in ES cells. Furthermore, this strategy avoids library construction and most subcloning steps that are characteristic of conditional targeting construct generation, and can be applied to the rapid generation of most standard, conditional, or knock-in targeting vectors.

While our work was being finalized for publication, another recombinase-based approach that manipulated BACs for the generation of targeting vectors was reported (Liu et al. 2003). Those authors employ gap-repair for the subcloning of a BAC fragment into a recipient plasmid, followed by further targeting of this plasmid to construct a conditional targeting vector. Although both approaches utilize recombinase and BAC DNA sequences for vector construction, there are unique and notable distinctions. Because we did not use Cre recombinase in the BAC targeting process, the presence of *loxP* sites in the BAC vectors has not posed any constraints to the insertion of further *loxP* sites for conditional targeting vector construction. On the other hand, the presence of three *loxP* sites leads to three possibilities of recombinants when the Cre recombinase is expressed in ES cells; the utilization of additional sequences (e.g., *frt* sites) which orient site-specific deletion can further simplify this process (Liu et al. 2003). In addition, in our approach, the requirement for unique restriction sites in the BAC that are located outside the regions of homology is one potential obstacle. Although we have been able to generate three different targeting constructs using this approach without difficulty, one can circumvent this obstacle by subcloning via gap repair (Liu et al. 2003).

Most importantly, in our studies, we manipulated BACs in situ in the original host, which avoids numerous subcloning steps and permits additional uses of this technology. Furthermore, we utilized a plasmid, rather a prophage, to encode the lambda Red recombination system and achieved efficient recombination utilizing only two templates for all targetings and very short terminal homologies (40–50 bp). Short homologies did not pose any constraint to our targetings. We achieved a specific targeting efficiency of 65% ± 30% for eight unique targetings within four different genes (average total number of colonies tested per targeting: 13 ± 9; average number of correct colonies identified per targeting: 7 ± 3). Aberrant recombination in the alternative approach employing gap repair may have resulted from short homologies between the donor sequences and repetitive sequences in the target DNA (Liu et al. 2003). It is likely that this potential problem can be avoided by surveying the BAC sequence prior to designing primers, to avoid using repetitive sequences for the site of recombination. Finally, manipulating BACs directly can permit the use of the entire BAC as a targeting construct in ES cells—avoiding constraints on the size of the homology arms.
Recombineering BACs for Gene Knockouts

METHODS

Construction of Template Plasmids

The template plasmids pKD4Lox, pSBS150, and pSBS158 were derived from pKD4, and each maintained in pir+ bacteria (Datsenko and Wanner 2000). The use of pKD4 as the backbone was essential to prevent replication of the template plasmid in transformed strains.

pKD4Lox was constructed by cloning a single loxP site upstream of the flanked TNS-aph segment in pKD4. pSBS150 was constructed by cloning a floxed aph gene (driven by both TNS and PGK promoters) in the pKD4 vector double-digested with BamHI/Ncol. To allow single targeting of BACs (see below), pSBS156 was generated from pSBS150 by cloning a loxP site upstream of the floxed aph gene. pSBS158 was constructed by cloning the mutated CARD15 exon (see text) into a unique EcoRI site located between the loxP sequence and the floxed aph gene cassette in pSBS156 (see Fig. 2b).

Generation of BAC Targeting Constructs

PCR products for BAC targeting were generated with 60–74-bp primer pairs using Advantage 2 polymerase (Clontech). Within each primer, the first 40–54 bp (homology arm) was homologous to the BAC sequence, and the next 20 bp was complementary to the template plasmid of interest. The 60-bp sequences to prime pKD4Lox, pSBS150, and pSBS158 (and further BAC targeting) were, respectively: CCTTGTTTAACCATATCAAAGTTCCCG GCTGTATCTCCTGCTCTTGAGCGATTGTGTAGGC (forward) and GGACCCAAGGGGTCTCC TAAACCTGGTTCTTCCCCTG CAA ACATAGTGAATCTCCTCCTT (reverse), ATT CCACCTCTGC CACCCCTTCATA CAACCCACGGCCAGCCAGCCGGATCAT A TTCAATA ACCC (forward) and GCTATAGTATTTGTCTT CAAACTAAGGCCCAAGAGGG CTA GCTACCTCGT CAAATAA (reverse), AGGCCTGCAATGAGCTTGGCCTCCT CAGCCCTTCTTCTTGTGGGTATC ACCGGGATCTT (forward) and ATACAAAATGTG TCTGA A AACATGATCTTCTAATCCACCTTACGT CTA CGCTTCGTA AAA (reverse), QAGCGTTCGCAATGAGCTTGGCCTCCT CAGCCCTTCTTCTTGTGGGTATC ACCGGGATCTT (forward) and ATACAAAATGTG TCTGA A AACATGATCTTCTAATCCACCTTACGT CTA CGCTTCGTA AAA (reverse), QAGCGTTCGCAATGAGCTTGGCCTCCT CAGCCCTTCTTCTTGTGGGTATC ACCGGGATCTT (forward) and ATACAAAATGTG TCTGA A AACATGATCTTCTAATCCACCTTACGT CTA CGCTTCGTA AAA (reverse), QAGCGTTCGCAATGAGCTTGGCCTCCT CAGCCCTTCTTCTTGTGGGTATC ACCGGGATCTT (forward) and ATACAAAATGTG TCTGA A AACATGATCTTCTAATCCACCTTACGT CTA CGCTTCGTA AAA (reverse), QAGCGTTCGCAATGAGCTTGGCCTCCT CAGCCCTTCTTCTTGTGGGTATC ACCGGGATCTT (forward) and ATACAAAATGTG TCTGA A AACATGATCTTCTAATCCACCTTACGT CTA CGCTTCGTA AAA (reverse).

Two-Step BAC Gene Targeting

Bacteria carrying the target BAC and pKD46 (λ Red recombination system) were grown in 100-ml LB cultures with chloramphenicol (12.5 µg/ml), ampicillin (100 µg/ml), and L-arabinose (150 µg/ml, Sigma) at 30°C to an OD600 of ~0.6. Electrocompetent bacteria were transformed with 100 ng of the pKD4Lox-amplified and aph gene-containing PCR construct. Chloramphenicol-resistant (CmR)/KmR transformants were first replated to achieve correct resistance, and then screened by PCR (using P1 and P2 primers, AGCCCTGCCCCTCTATTI and TCACACG GGACCTACACAG, respectively; see Fig. 1), and the PCR products were sequenced to confirm the desired targeting. Positive colonies were grown overnight at 42°C to cure pKD46 (confirmation was performed by testing ampicillin sensitivity). In order to remove the aph gene, a single CmR/KmR and ampicillin-sensitive colony was transformed with pCP20 (an ApR and Tn10-containing vector, and the desired transformants were selected by growth on ampicillin/kanamycin plates.

ACKNOWLEDGMENTS

We gratefully acknowledge members of the Snapper laboratory for their support and the critical reading of the manuscript by Drs. Daniel Podolsky, Vijay Yajnik, Antonio Oliveira-dos-Santos, Sheila Thomas and Anna Lyubimova. These studies were supported by the National Institutes of Health (AI50950 and HL59561—S.B.S.) and by CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Brazil—V.C.-A.).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES


Genome Research 2193 www.genome.org


Received March 20, 2003; accepted in revised form June 10, 2003.