Video Article

Subcloning Plus Insertion (SPI) - A Novel Recombineering Method for the Rapid Construction of Gene Targeting Vectors

Thimma R. Reddy[1], Emma J. Kelsall[1], Léna M.S. Fevat[2], Sarah E. Munson*[3], Shaun M. Cowley*[1]

[1]Department of Biochemistry, University of Leicester

[2]Center for Fisheries, Environment and Aquaculture Sciences

[3]ES Cell Facility, Centre for Core Biotechnology Services, University of Leicester

*These authors contributed equally

Correspondence to: Shaun M. Cowley at smc57@leicester.ac.uk

URL: http://www.jove.com/video/52155
DOI: doi:10.3791/52155

Keywords: Molecular Biology, Issue 95, recombineering, gap-repair, subcloning plus insertion, transgene, knockout, mouse

Date Published: 1/8/2015

Citation: Reddy, T.R., Kelsall, E.J., Fevat, L.M., Munson, S.E., Cowley, S.M. Subcloning Plus Insertion (SPI) - A Novel Recombineering Method for the Rapid Construction of Gene Targeting Vectors. *J. Vis. Exp.* (95), e52155, doi:10.3791/52155 (2015).

Abstract

Gene targeting refers to the precise modification of a genetic locus using homologous recombination. The generation of novel cell lines and transgenic mouse models using this method necessitates the construction of a 'targeting' vector, which contains homologous DNA sequences to the target gene, and has for many years been a limiting step in the process. Vector construction can be performed *in vivo* in *Escherichia coli* cells using homologous recombination mediated by phage recombinases using a technique termed recombineering. Recombineering is the preferred technique to subclone the long homology sequences (>4kb) and various targeting elements including selection markers that are required to mediate efficient allelic exchange between a targeting vector and its cognate genomic locus. Typical recombineering protocols follow an iterative scheme of step-wise integration of the targeting elements and require intermediate purification and transformation steps. Here, we present a novel recombineering methodology of vector assembly using a multiplex approach. Plasmid gap repair is performed by the simultaneous capture of genomic sequence from mouse Bacterial Artificial Chromosome libraries and the insertion of dual bacterial and mammalian selection markers. This subcloning plus insertion method is highly efficient and yields a majority of correct recombinants. We present data for the construction of different types of conditional gene knockout, or knock-in, vectors and BAC reporter vectors that have been constructed using this method. SPI vector construction greatly extends the repertoire of the recombineering toolbox and provides a simple, rapid and cost-effective method of constructing these highly complex vectors.

Video Link

The video component of this article can be found at http://www.jove.com/video/52155/

Introduction

The development of gene targeting technologies has enabled the construction of cell lines and mouse models to investigate different biological systems[1-3]. A key first stage in the modification of a genomic sequence is the design and construction of a gene targeting vector[4]. Targeting vectors are plasmid constructs that carry the allele of interest containing the desired modifications(s) flanked by a selection marker (*e.g., neomycin*), and long genomic regions required for efficient homologous recombination in mammalian cells[5]. Precise gene modification is achieved by the introduction of the targeting vector into embryonic stem (ES) cells[6] or somatic cells[7], whereby homologous recombination between identical stretches of DNA sequence on the targeting vector and the genomic locus results in the transfer of the intended modification to the genome by gene conversion[8]. Such modified ES cells can be injected into mouse blastocysts to produce offspring (chimaeras) that can then transmit these modified alleles via the germline[9]. An alternate route to produce transgenic mice involves the microinjection of a gene expression vector into single celled mouse zygotes, which leads to the random integration of the vector in the mouse genome[10]. Traditional methods of vector construction have relied on conventional 'cut and paste' cloning using restriction enzymes and DNA ligases to clone the different selection marker and genomic fragments into a vector backbone. However, an inherent limiting factor of traditional cloning is the positioning and choice of restriction sites, especially with longer DNA sequences. This often necessitates multiple subcloning steps and also introduces extraneous DNA sequences into the vector, which often leads to a lower efficiency of gene targeting.

Recombineering (recombinogenic engineering) is a DNA engineering technology that overcomes these limitations by using homologous recombination (HR) mediated by phage recombination proteins in *E. coli* cells[11,12]. Since any region of a homologous sequence can serve as a substrate of recombineering, the constraints of availability of restriction sites are removed. Large DNA sequences can be seamlessly modified directly *in vivo,* thus also preserving their structural integrity[13]. Recombineering is very efficient with short homologies (50 bp)[14] and therefore homology arms (HA) can be conveniently incorporated into synthetic oligo sequences. In a typical recombineering experiment, an oligo or a double stranded DNA (dsDNA) fragment containing HA is electroporated into recombineering competent *E. coli* cells containing the target located either on the chromosome or on a plasmid[15]. The recombination potential is conferred by inducible expression of the Red recombination proteins of the phage[16,17] or the RecET proteins of the rac prophage[18]. The Red/RecE exonuclease converts linear dsDNA to a single-stranded

DNA (ssDNA) intermediate, which is then bound by its partner, Red/RecT, a single-stranded annealing protein (SSAP)[19]. The annealing of a long ssDNA or a short oligo to its complementary target sequence occurs on the lagging strand of the replication fork and leads to the incorporation of the sequence at the target site. Lagging strand ssDNA recombination is the basis of the high efficiency of recombineering and can be described by the 'beta' recombination model[20,21].

A typical recombineering workflow to build a gene targeting vector involves either of the two following routes. One route involves subcloning the desired genomic region from a mouse BAC clone into a plasmid followed by the sequential insertion of LoxP recombination sites, a selection marker[22] etc., or the alternative route involves targeting the BAC genomic locus with the different targeting vector elements by multiple rounds of recombineering[10,23] and then subcloning the modified locus into a plasmid by gap repair cloning[24,25]. Variations on this theme have been used in different high-throughput recombineering pipelines as part of large mouse production programs[26,27]. However, these procedures involve complicated and lengthy stages, require the use of specialized vectors and *E. coli* strains (*e.g.,* Cre expressing cells) and utilize one or more intermediate steps of vector DNA purification and re-transformation (**Table 1**). Subcloning plus insertion (SPI) is a novel recombineering technique that combines beta recombination and gap repair cloning into a single process (**Figure 1**). SPI vector assembly is simple, quick and flexible and offers significant improvement on standard recombineering approaches (**Table 1**). Here, we demonstrate the ease and utility of using SPI for different vector construction applications with a particular emphasis on constructing non-standard and challenging vector designs. Test cases included the construction of a fluorescent reporter knockin vector, a dual tagged protein expression knockin vector, a BAC fluorescent reporter vector and a conditional knockout vector. These examined variously the requirement to localize a cell surface receptor, purify a nuclear protein complex or conditionally ablate the expression of a gene.

Protocol

1. Gene Targeting Design

1. Order an appropriate BAC clone covering the genomic region of interest. Ensure that this is isogenic to the type of ES cells to be modified *e.g.,* RPCI-23 and RPCI-24 BAC clones for gene targeting with C57BL/6 ES cells.
2. Apply conventional gene targeting criteria when designing the targeting vector. Key parameters include whether the modification is required to be constitutive or conditional, defining critical exons (CE) for deletion in a gene knockout strategy and spacing and placement of intronic cassettes.
   NOTE: Each of these has been discussed in detail elsewhere[10].
3. Choose genomic regions each of 5-6 kb flanking the target modification site.
   NOTE: The size of the subcloned insert is therefore typically 10-12 kb, although the upper limit can be as high as 80 kb with a low copy subcloning plasmid like p15A, pBR322 etc. and up to 200 kb with a BAC vector.

2. Multiplex Recombineering Oligos

1. Design oligos for the plasmid backbone (subcloning plasmid) and the selectable marker (insertion cassette). Design each oligo such that it contains 180 bp HA flanking the genomic target site and 20 bp of specific priming sequence of the insertion cassette or the subcloning plasmid (**Figure 2**).
   NOTE: Homology arms should not contain any repetitive elements, presence of repetitive elements will result in incorrect targeting and subcloning. Repetitive elements can be detected by using web based tools such as the ''Repeat Masker''.
2. Include a unique restriction enzyme (RE) site on one of the vector oligos to linearize the gene targeting vector for ES cell targeting (**Figure 2**).
3. Check oligo parameters using an oligo analyzer program. Take care to avoid secondary structures in the priming regions.
4. Shorter HA (50 bp) of the insertion cassette is tolerated and yields lower number of recombinants but ensure that the subcloning plasmid HA is always longer (180 bp).
5. Determine the orientation of the genomic DNA insert of the particular BAC clone using web based tools like ''cloneDB''. Establish the direction of replication from OriS by checking the map of the BAC plasmid backbone used in the BAC library construction.
   NOTE: OriS is usually opposite to the transcriptional direction of the Chloramphenicol (Chl) marker for all commonly used BAC plasmids.
6. Add two terminal phosphorothioate (PTO) bonds to the 5' end of the oligo that is opposite to the direction of replication on the BAC clone. Add a 5' phosphate modification to the reverse oligo (**Figure 2**).
   NOTE: The asymmetric phosphorothioated PCR cassette upon Red digestion generates an ssDNA intermediate that can prime the lagging strand of the replication fork.
   NOTE: Maximal multiplex recombination frequency is observed with the lagging strand protected cassettes. Leading strand protected or dual protected cassettes may not produce equivalent recombination efficiencies at some loci.
7. Order oligos with PAGE or HPLC purification.

3. BAC Clone Transformation with pSC101 BADgbaA Recombineering Plasmid.

1. To make the *E. coli* strain bearing the BAC recombineering proficient, transform it with the pSC101 BADgbaA plasmid containing the Red genes[16].
   NOTE: The pSC101 BADgbaA plasmid contains the Red and RecA genes under the control of the Arabinose inducible *araC-P_{BAD}* promoter, a tetracycline selection marker and the temperature sensitive pSC101 replicon.
   1. Stab a sterile pipette tip in the BAC agar culture and inoculate 5.0 ml of lysogeny broth (LB) pH 8.0 containing 12.5 µg ml^{-1} Chl. Grow at 37 °C for 5 hr shaking at 200 rpm.
   2. Chill 10% ($^{v}/_{v}$) glycerol solution, microcentrifuge tubes and electroporation cuvettes on ice. Cool a refrigerated large centrifuge and microcentrifuge to 4 °C .
   3. Determine the optical density (OD) of the culture using a spectrophotometer and measure absorbance at 600 nm. Prepare electrocompetent cells (as described below) when an OD_{600} reading of 0.3 to 0.8 is reached.

4. Spin down cells in a 50 ml centrifuge tube in a large centrifuge at 1,216 x g for 5 min at 4 °C .
5. Wash cells with 1 ml of chilled 10% glycerol and spin down cells at 17,949 x g for 20 sec at 4 °C . Perform the wash step a total of 3 times.
6. Resuspend the cells in a total volume of 50 µl of 10% glycerol and add 10-200 ng of the pSC101 BADgbaA recombineering plasmid. Obtain a single cell suspension by pipetting up and down several times and then transfer the cells to a pre-chilled 1 mm gap electroporation cuvette.
7. Electroporate the cells with a setting of 1.8 kv, 25 µF and 200 Ω.
   NOTE: Check the correct settings for each brand of electroporator. A time constant of electroporation less than 4 indicates the presence of salt and other impurities.
8. Immediately recover the cells in 1 ml of LB and transfer the cells to a 50 ml centrifuge tube.
9. Grow the BAC gbaA culture at 30 °C for 2 hr shaking at 200 rpm.
   NOTE: The pSC101 BADgbaA plasmid is lost when cells are grown at 37 °C due to inactivation of the temperature sensitive RepE replication factor. Grow gbaA cells at 30 °C to maintain the recombineering functions.
10. Add 9 ml of LB containing 12.5 µg ml^{-1} Chl and 4 µg ml^{-1} Tetracycline (Tet) to the recovered BAC gbaA culture. Grow the BAC gbaA culture O/N at 30 °C shaking at 200 rpm.
    NOTE: The transformation efficiency of electroporating the supercoiled gbaA plasmid is sufficiently high to permit saturation growth O/N in liquid media.

4. Preparation of Insertion Cassettes and Subcloning Plasmids

1. To incorporate HA into the insertion cassette(s) and the subcloning plasmid, perform polymerase chain reaction (PCR) using the long modified oligos as described below.
   1. Optional: to prevent plasmid carryover into the recombineering reaction, use an R6K origin or similar narrow host range plasmid template to amplify the insertion cassette.
   2. Alternatively, linearize the plasmid template using an RE digest (**Table 2**). Choose a RE that cuts outside the PCR amplification region and is heat inactivated. Heat inactivate the RE as recommended by the manufacturer.
   3. Set up polymerase chain reaction (PCR) using a high-fidelity hotstart DNA polymerase system. Prepare a PCR master mix as detailed (**Table 3**). Perform thermal cycling as shown (**Table 3**).
2. Analyze PCR products by agarose gel electrophoresis. Load 1-5 µl of each PCR onto a 1% (w/v) agarose gel containing 0.5 mg ml^{-1} Ethidium Bromide (EtBR).
   NOTE: The presence of non-specific amplification products does not interfere in the recombineering reaction. In some cases however, primer-dimers may decrease recombineering efficiency.
3. Purify PCR products using a PCR purification kit.
   1. Optional: remove the plasmid template from PCRs by treatment with DpnI followed by PCR clean-up. Elute DNA in a minimal volume of sterile deionized water (as recommended by the manufacturer).
      NOTE: Addition of DpnI to unpurified PCR reactions results in reduced efficiency of cleavage of the methylated plasmid DNA template.
4. Quantify PCR amplified DNA by agarose gel analysis against a known set of DNA standards *e.g.*, λ HindIII digest or by using a nanodrop spectrophotometer.

5. Subcloning Plus Insertion

1. Dilute the O/N BAC gbaA culture 50 fold by adding 200 µl in 10 ml LB+Chl+Tet. Grow at 30 °C shaking at 200 rpm for 1 hr 50 min. Include an sample to be used as negative control.
2. Chill all recombineering materials and equipment to 4 °C as described in step 3.1.2.
3. Pour LB agar pH 8 plates containing the correct concentration of the appropriate selective antibiotic(s) that will allow selection of the DNA fragment that is inserted as well as for the selection of the subcloning vector (**Table 4**).
   NOTE: Some antibiotics are pH sensitive. Use LB agar pH 8 as a rule.
4. Prepare a 10% (w/v) solution of L-Arabinose. Filter sterilize through a 0.2 µm syringe filter.
   NOTE: Arabinose induces the expression of the recombineering proteins from the gbaA plasmid.
5. Check the OD of the BAC gbaA culture using a spectrophotometer. Once an OD_{600} of 0.25-0.3 is reached, induce the recombineering proteins as described in the following step.
6. Add 200 µl of the 10% Arabinose solution to 10 ml of the BAC culture to achieve a final concentration of Arabinose of 0.2%. Include an uninduced culture (without Arabinose) to be used as negative control.
7. Transfer the BAC culture to a 37 °C shaking incubator and induce Red expression for 45 min shaking at 230 rpm.
   NOTE: Expression of Red proteins is inefficient at 30 °C .
8. Spin down cells and wash with 10% glycerol 3 times as described in step 3.1.5.
9. Add 600-1,000 ng each of the subcloning plasmid and the insertion cassette(s) to the recombineering reaction. Include a vector only and vector plus single insert controls to check the recombination proficiency and integrity of the vector and the cassettes.
10. Obtain a single cell suspension by pipetting up and down. Perform electroporation as described in step 3.1.7 and subsequent recovery in 1 ml LB at 37 °C for 1 hr for multi-copy plasmids or in 10 ml of LB at 37 °C for 3 hr for BAC vectors.
11. Plate different dilutions of the recovered culture *e.g.,* 90%, 10%, 1% on the dual selection agar plates and grow at 37 °C for 16 hr.

6. Analysis of Recombinants

1. Optional: Pick 6 to 12 colonies and perform colony PCR using a HA flanking primer and an insert specific primer. Include the subcloning plasmid and the insertion cassette plasmid as well as the parent BAC clone as negative controls.
    1. Perform agarose gel analysis of the colony PCRs. Identify positive clones by the presence of a bright band at the expected size.
    NOTE: The high efficiency of the SPI cloning process generates mostly correct recombinants.
2. Pick colonies into 5 ml LB pH 8 containing the selective antibiotic and grow O/N at 37 °C .
3. Prepare DNA minipreps using a column purification kit as per the manufacturer's instructions.
4. Prepare BAC miniprep DNA using standard phenol-chloroform isolation or a similar protocol.
5. Perform RE digests on the miniprep DNA. Separate DNA by agarose gel electrophoresis. Analyze RE patterns to identify the clones containing the expected fragments sizes of the correct targeting vector. Choose an RE that clearly discriminates between the vector lacking insert(s) and the vector containing the insert(s).
6. Optional: to obtain cells that are devoid of the BAC, which is still present in the *E. coli* after recombineering, use the miniprep plasmid DNA to transform DH5alpha or DH10B *E. coli* cells.
7. Perform DNA sequencing across the HA and insertion cassette to check oligo synthesis errors.

Representative Results

Knockin Vectors

Knockin targeting vectors reflect the need to introduce a novel sequence feature in the genome including single base pair substitution in a protein coding region, the fusion of a fluorescent marker or an affinity tag to a protein or the integration of a gene expression cassette. To test the application of SPI in knockin vector construction strategies, two different test cases were examined. *Dnttip1* encodes the deoxynucleotidyltransferase, terminal, interacting protein 1A (TDIF1) that together with class I histone deacetylase (HDAC) form a mitotic deacetylase complex (MiDAC)[28]. To investigate the role of *Dnttip1* in cell division, a tandem-affinity tagging approach was taken to isolate the TDIF1 interacting proteins. Previous attempts to subclone a portion of the *Dnttip1* gene using a p15A vector containing long homology regions resulted in low gap repair efficiency and frequent aberrant recombination products (data not shown). Thus, the construction of a knockin vector at the *Dnttip1* locus provided a challenging recombineering exercise. A SPI strategy was designed to subclone a 12 kb section of the *Dnttip1* gene spanning the last exon (exon 13) into a low copy p15A vector and to simultaneously insert a dual affinity tag selection cassette into exon 13, replacing the stop codon (**Figure 3A**). The 2X FLAG-calmodulin binding protein (CBP) linked FRT-PGK-em7-Neomycin (Neo)-BGhpA-FRT cassette (2.0 kb) was amplified from an RE linearized plasmid using dual PTO modified oligos that contained 120 bp HA flanking the *Dnttip1* stop codon. A p15A *zeoDnttip1* subcloning vector (1.7 kb) was constructed that contained 200 bp regions homologous to the ends of the *Dnttip1* sequence to be subcloned. The *Dnttip1* subcloning plasmid was RE linearized and PCR amplified using 20 bp modified oligos that generated a leading strand protected vector. The PCR products were DpnI treated, purified and co-electroporated into recombineering competent *Dnttip1* BAC *E. coli* cells as well as uninduced control cells. The SPI reactions were plated on Zeocin (Zeo) and Kanamycin (Kan) containing agar plates (**Figure 3B**). SPI produced the correctly modified *Dnttip1* knockin vector in all of the 12 recombinants analysed (**Figure 3C**). DNA sequencing verified the absence of any errors resulting from oligo synthesis or PCR amplification.

Another example of SPI involved the in-frame insertion of an enhanced yellow fluorescent protein (*eYFP*) linked selection cassette at the *P2rx1* gene. The *P2rx1* gene encodes a G-protein-coupled receptor that functions as an ATP-gated ion channel[29]. Linkage to YFP fluorescence permits the tracking of the P2x1 receptor on the cell surface in various functional assays. *P2rx1* is another difficult locus that has presented significant problems with conventional recombineering methodologies (data not shown). Using SPI, a 12 kb segment of the *P2rx1* gene encompassing the terminal exon (exon 12) was subcloned into a p15A *zeo* vector and modified with the concerted insertion of an *eYFP*-LoxP flanked Neo cassette replacing the stop codon in exon 12 to construct the *P2rx1-eYFP* knockin vector (**Figure 3D**). In this example, the lagging strand protected p15A vector (1.7 kb) contained 230 bp HA and the *eYFP* cassette contained 50 bp HA and was also left unmodified. The *eYFP* insertion cassette (2.7 kb) was assembled by splicing overlap PCR of the *eYFP* gene, PCR amplified from pEYFP-C1, and the LoxP- PGK-em7-Neo-BGhpA-LoxP cassette, PCR amplified from pL452 (NCI, Frederick). The SPI reaction produced hundreds of colonies (data not shown). RE analysis of DNA minipreps prepared from 11 clones showed the majority contained the correctly assembled *P2rx1-eYFP* knockin vector (**Figure 3E**). Additional validation by DNA sequencing showed error free insertion of the *eYFP* cassette. However, some of the clones showed profiles of the untagged gap repaired and the tagged gap repaired plasmids in the same cell (lanes 6-9). A few samples also contained incorrectly gap repaired (lane 2) or mistargeted plasmids (lanes 4 and 5). The failure of concurrent subcloning and targeting in some SPI recombinants highlights the limits of efficient SPI cloning when using large cassettes (> 3 kb), which arise from the constraints on Red processivity of long DNA fragments[30], or if using short HA. Indeed, increasing the HA of the *eYFP* cassette to 200 bp increased SPI efficiency and the correct targeting of the cassette (data not shown). Gene targeting with the *P2rx1-eYFP* knockin vector in JM8.N4 mouse ES cells (C57BL/6 strain) generated 6 positive clones out of 96, which contained the correctly targeted *P2rx1-eYFP* sequence (**Figure 3F**).

BAC Reporter Vectors

A BAC clone of the target gene often contains all the requisite upstream and downstream regulatory elements *e.g.,* enhancers, UTRs etc. as well the endogenous promoter to drive gene expression at natural levels[31]. A BAC reporter vector is therefore the preferred vehicle to recapitulate the endogenous expression pattern of a gene[32]. However, the large size of a BAC plasmid (up to 200 kb) presents significant practical problems with transfecting an intact BAC into cells[33]. Reducing the size of the BAC genomic insert through BAC trimming[34,35], while retaining the necessary regulatory elements of a gene, enables easier handling of the BAC and results in a more efficient transfection. Current BAC engineering technology involves multiple rounds of recombineering to achieve this goal[35] . To demonstrate the utility of SPI in BAC trimming, a pBeloBAC11 BAC vector was used to subclone a 30 kb genomic sequence including the full length *P2rx1* gene from the 168 kb *P2rx1* BAC together with the simultaneous insertion of an *eYFP* cassette in the *P2rx1* gene (**Figure 4A**). The pBeloBAC11 *zeo* vector backbone (6.5 kb) containing 180 bp HA was PCR amplified with modified oligos that generated a lagging strand protected vector. The same *eYFP* Neo cassette (3 kb), used in the previous *P2rx1* knockin vector construction, was PCR amplified using lagging strand protected oligos containing 180 bp HA and targeted

the *P2rx1* exon 12 replacing the stop codon. Following combined electroporation of the *eYFP* cassette and the pBeloBAC11 *zeo* vector into *P2rx1* BAC cells expressing the gbaA recombineering proteins, the culture was recovered in 10 ml LB pH 8 for 3 hr at 37C to segregate the BAC plasmids. Pure recombinants were selected on Zeo and Kan agar plates and were observed at a frequency of 2 x 10^{-6}. Colony PCR genotyping analysis revealed successful BAC trimming in the 3 out of 6 clones that were analysed (**Figure 4B**). Long range PCR amplification across the *eYFP* insertion site confirmed the correct cassette incorporation in the three positive clones (**Figure 4B**). The three clones lacking the *eYFP* insert were also incorrectly gap repaired at the 5' end, though the cause of mistargeting of the *eYFP* cassette in these clones may be separate to the correct closure of the 5' BAC end. The three *eYFP* positive BAC clones were further analysed with RE digests and showed the expected pattern of the correctly trimmed *eYFP* recombinant BAC (**Figure 4C**). Sequence analysis revealed absence of errors in the *eYFP* cassette in only 1 clone out of the 3 positives.

Conditional Knockout (cko) Vectors

Conditional ablation of gene expression is an important tool to investigate developmental processes or to study biological systems at a particular time point. A conditional gene knockout strategy typically involves the placement of LoxP recombination sites surrounding a critical exon (CE). The deletion of a CE upon Cre expression or activation produces a frameshift and a premature stop codon, resulting in degradation of the mRNA due to nonsense mediated decay (NMD). The construction of a conditional gene targeting vector is a complex task and involves several steps of subcloning, targeting and transformation[22]. SPI offers a convenient route to simplify this process. As a test case, a conditional allele of the *Zrsr2* gene was constructed using the SPI methodology (**Figure 5A**). The *Zrsr2* gene encodes a splicing factor and a single copy is located on the X chromosome. The conditional status of gene deletion is particularly important in this instance to control for the possibility of cells adapting to the lack of *Zrsr2* during ES cell selection in a constitutive gene deletion targeting strategy. SPI was performed to subclone a 10 kb portion of the *ZrSr2* gene with concurrent insertion of two different LoxP flanked selection cassettes. The FRT-PGK-em7-Neo-FRT-LoxP cassette (2 kb) was PCR amplified from RE digested pL451 using lagging strand protected oligos that contained 180 bp HA targeting the *ZrSr2* intron 2. A second cassette containing Rox-PGK-em7-Blasticidin (Bsd)-Rox-LoxP (2 kb) was PCR amplified from an R6K plasmid with lagging strand protected oligos containing 180 bp HA identical to a downstream region in intron 3. The two LoxP sites flanked exon 3, the CE whose deletion upon conditional Cre activation results in a frameshift and introduces a premature stop codon. The *Zrsr2* subcloning plasmid (1.6 kb) was PCR amplified from an RE linearised p15A *zeo* plasmid using lagging strand protected oligos containing 180 bp HA matching the ends of the 10 kb *Zrsr2* sequence. SPI reactions were performed as described before and plated on Zeo+Neo+Bsd plates. The *ZrSr2* conditional targeting vector was successfully assembled in most of the 12 recombinants examined (**Figure 5B**). Further DNA sequencing analysis showed correct insertion of both the selection markers in the *ZrSr2* cko vector.

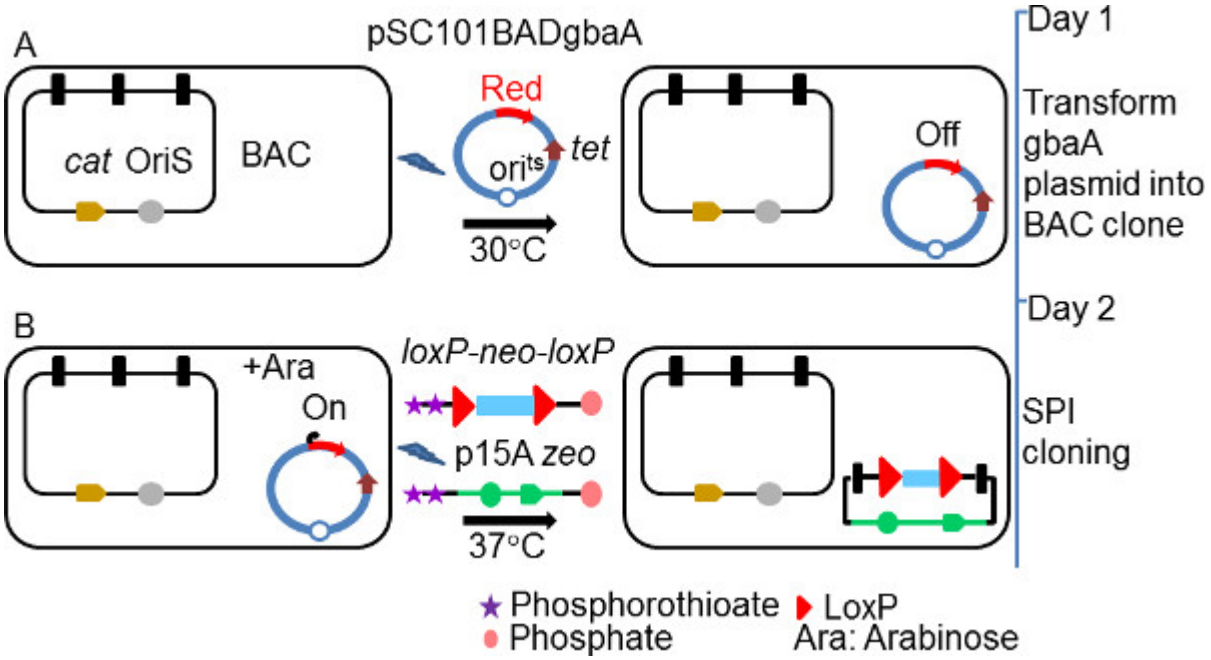

Figure 1. **SPI cloning.** Overview of the SPI cloning process combining cassette insertion and subcloning in a single step. (A) The BAC clone is transformed with the pSC101 BADgbaA plasmid and grown at 30C. (B) Following arabinose induction to express the Red proteins, the asymmetric modified insertion cassette and subcloning plasmid are introduced into the BAC clone and selected with both the insertion and subcloning markers to generate the final vector. Please click here to view a larger version of this figure.

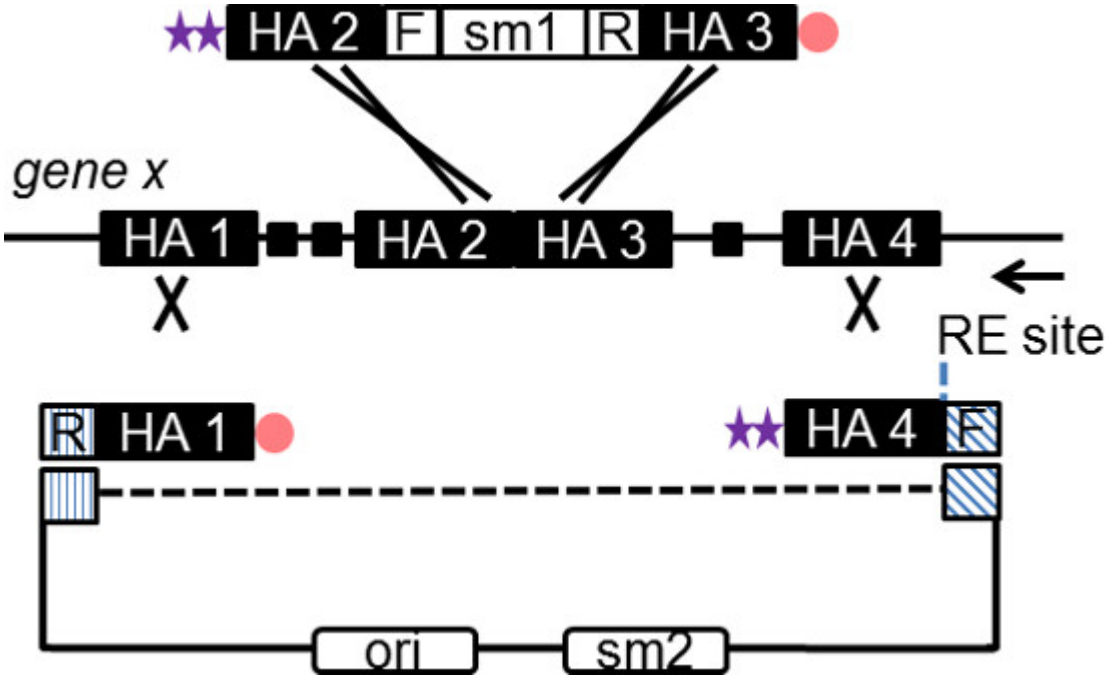

Figure 2. Schematic illustrating the design of SPI recombineering oligos. The subcloning plasmid and insertion cassette are both generated by PCR using a combination of terminal PTO and phosphate modified oligos. The PCR fragment upon Red digestion *in vivo*, produces an ssDNA intermediate, which anneals to the lagging strand of the replication fork. Gene specific HA of 50-180 bp is incorporated into each oligo as shown. The arrow indicates the direction of DNA replication across a candidate gene. Dashed line represents subcloning plasmid sequence not incorporated into the PCR product. RE site, restriction enzyme site of linearization of final vector; F, forward sequence (20 bp) specific to the subcloning plasmid or insertion cassette; R, reverse complement sequence (20 bp) of the plasmid or cassette; sm, selection marker. Please click here to view a larger version of this figure.

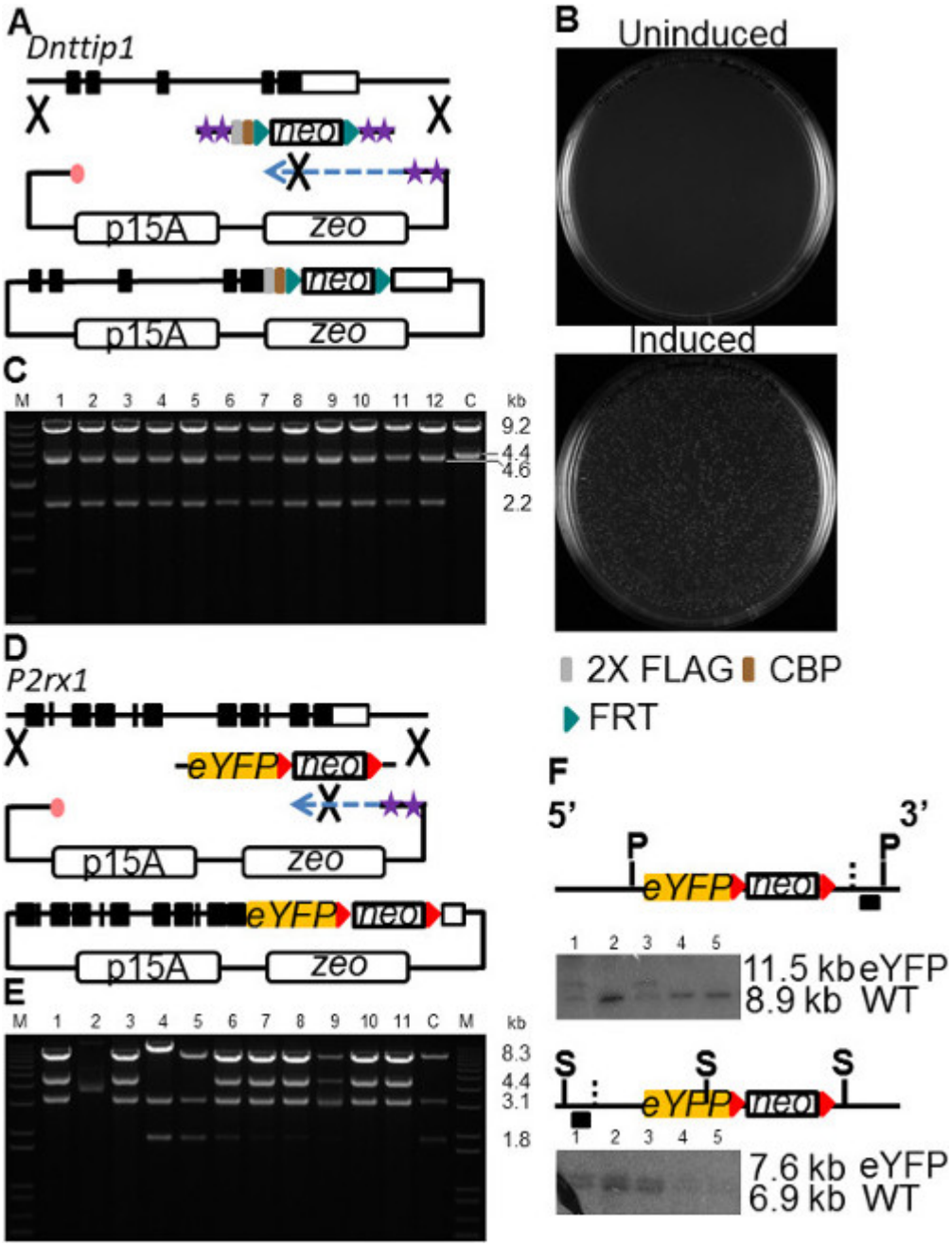

Figure 3. SPI enables construction of difficult knockin vectors. (A) Schematic of the SPI strategy used in the construction of the *Dnttip1* dual tagged vector. Arrow indicates the direction of replication on the BAC clone. **(B)** Plating results of the uninduced and induced samples of the *Dnttip1* SPI experiment. **(C)** EcoRI digest of *Dnttip1* SPI clones. M, 1 kb ladder (NEB); C, p15A *Dnttip1* gap repaired plasmid lacking the dual tag cassette. Fragment sizes are: tagged, 9.2+4.4+2.2 kb; control; 9.2+4.6 kb. **(D)** SPI based *P2rx1-eYFP* knockin vector construction. **(E)** EcoRI digest of *P2rx1-eYFP* SPI clones. M, 1 kb+ ladder (Invitrogen); C, p15A *P2rx1* gap repaired plasmid lacking the *eYFP* cassette. Fragment sizes are: tagged, 8.3+4.4+3.1+0.03 kb; control; 8.3+3.1+1.8+0.03 kb. **(F)** Southern blot analysis of *P2rx1-eYFP* gene targeting in JM8.N4 ES cell cells. Top panel, Southern blot with the 3' end probe using PshAI digest. Shown is the screening result of 5 clones. Bottom panel, Southern blot using the 5' end probe and SpeI digest of the positives identified from the 3' end screening. PshAI RE site; S, SpeI RE site. Dashed line represents the end of the vector HA. Black box denotes southern probe. Expected restriction fragments are, PshAI: WT, 8.9 kb; *eYFP-neo*, 11.5 kb; SpeI: WT, 6.9 kb; *eYFP-neo*, 7.6 kb. Please click here to view a larger version of this figure.

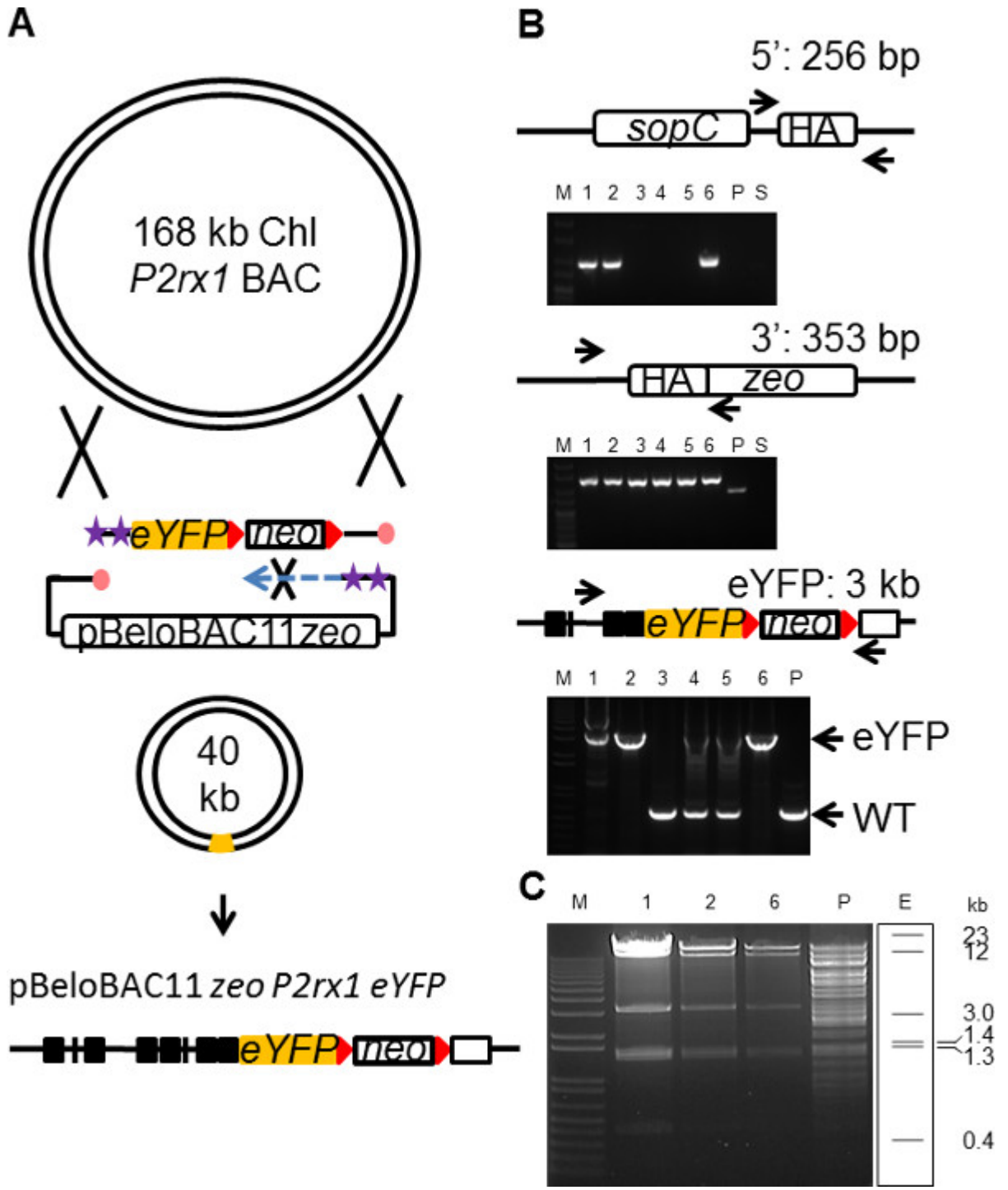

Figure 4. **Simplified BAC trimming using SPI. (A)** Schematic illustrating the concept of BAC trimming using SPI. Key to symbols is described in **Figure 1**. **(B)** PCR amplification across the 5' and 3' ends of the subcloned insert and across the *P2rx1-eYFP* insertion site. The screening strategy is shown for each type of PCR. M (top panel), Hyperladder 25 bp, (bottom panels), 1kb+; P, *P2rx1* BAC; S, pBeloBAC11 *zeo* subcloning plasmid. **(C)** HindIII digest of the three *eYFP* positive SPI BAC clones from **(B)**; E, expected HindIII restriction pattern of the trimmed *eYFP* BAC. Please click here to view a larger version of this figure.

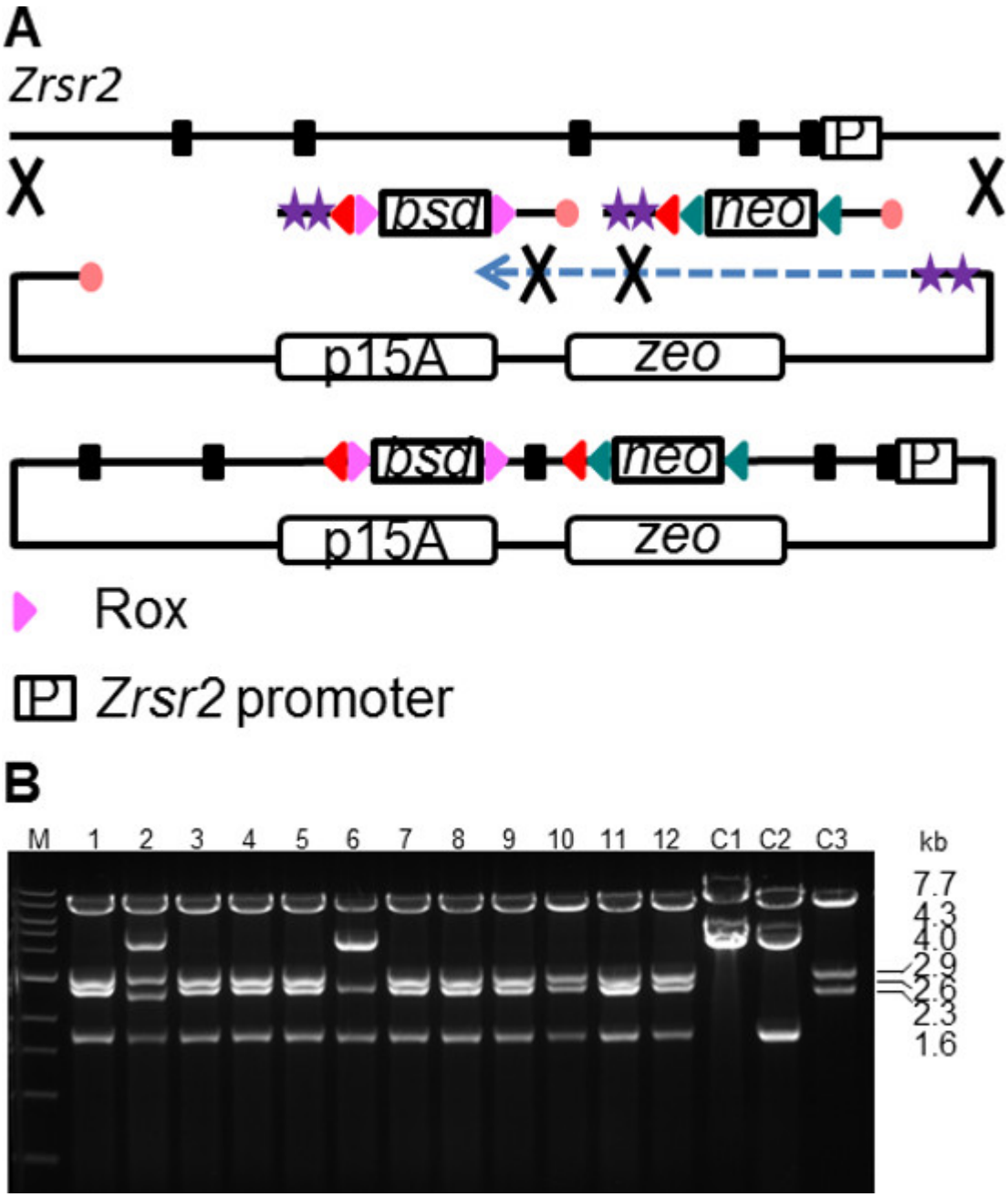

Figure 5. Conditional knockout vector generation using SPI cloning. (A) Schematic of the simultaneous insertion of two different LoxP flanked selection cassettes during subcloning of the *Zrsr2* allele. Key to symbols is described in **Figure 1**. **(B)** EcoRI digests of *Zrsr2* SPI clones. L, 1 kb ladder (NEB); C1, p15A *ZrSr2* gap repaired plasmid, C2, p15A *ZrSr2* gap repaired plasmid containing the *neo* insert; C3, p15A *ZrSr2* gap repaired plasmid containing the *bsd* insert. Fragment sizes are: cko, 7.7+2.9+2.6+1.6 kb; C1, 7.7+4.0 kb; C2, 7.7+4.3+1.6 kb; C3, 7.7+2.9+2.3 kb. Please click here to view a larger version of this figure.

No of steps	Conventional recombineering pipeline[a]	Multiplex recombineering[b]
Step 1	Transformation of recombineering plasmid into BAC host	Transformation of recombineering plasmid into BAC host. Preparation of the targeting cassettes and subcloning vectors.
Step 2	Insertion of R1/R2 Gateway cassette	Multiplex gap repair cloning
Step 3	Insertion of floxed Kan cassette	O/N culture from single colonies
Step 4	Gap repair into R3/R4 plasmid	Plasmid preparation and verification
Step 5	Transformation into Cre+ *E. coli*	
Step 6	Plasmid preparation and verification	
Step 7	O/N three-way Gateway reaction	
Step 8	Transformation of the three-way Gateway reaction into DH10B *E. coli* cells	
Step 9	Overnight culture from single colonies	
Step 10	Plasmid preparation and sequence verification	
[a]Knockout mouse program (KOMP) high throughput vector construction pipeline. Average time of 3 weeks to verified clone[26].		
[b]Average time of 4 days to verified clone		

Table 1: **Comparison of conventional recombineering with SPI in the construction of conditional knockout vectors.**

RE buffer	5 µl
DNA	1 µg plasmid or purified PCR products
RE	1 µl (5 units or more)
TE	up to 50 µl
Incubate at 37 °C for at least 1 hr. Heat inacitvate according to the manufacturers instructions	

Table 2: **RE digest.**

PCR materials	Final concentration
PCR buffer	1x
dNTP	200 nM
$MgSO_4$	1.5 mM
Betaine	1.3 M
DMSO	1%
Forward primer	200 nM
Reverse primer	200 nM
DNA polymerase	1 U
Template	10 ng of multicopy plasmids or 2.5 µl of miniprep DNA for genotyping PCRs
Water	up to 50 µl[a]
[a]For a standard 50 µl PCR reaction with multicopy plasmids. Long range genotyping PCRs were set up in 25 µl PCR reactions.	
PCR conditions	
95°C	2 min
92°C	10 sec
55°C	30 sec
72°C	30 sec
30 cycles[b]	
[b]Cycle no may be extended to 35 for BAC PCR genotyping	

Table 3: **PCR Set-up and Conditions.**

Antibiotics	Concentration[a] (μg ml^{-1})
Ampicllin	50
Blasticidin[b]	40
Chloramphenicol	12.5
Gentamicin	2
Hygromycin[c]	30
Kanamycin[b]	15
Tetracycline	4
Trimethoprim[c]	10
Zeocin	5
[a]Recommended for use with BACs and multicopy plasmids when used in combinations in multiplex recombineering	
[b]Blasticidin (35 μg ml^{-1}) and Kanamycin (6 μg ml^{-1}) when used together in combination	
[c]Hygromycin and Trimethoprim are not recommended for selection with single copy BACs.	

Table 4. Recommended antibiotic concentrations for use in SPI experiments.

Discussion

Construction of ES cell lines and mouse models has historically involved gene targeting using plasmid constructs that contained the modified allele[4]. However, the construction of these complex gene targeting vectors has proved to be a significant bottleneck in the timely production of such models. The development of recombineering based vector construction strategies has allowed improved vector designs and more efficient vector assembly. Nonetheless, current recombineering protocols still involve multiple steps, require intermediate plasmid purification and use different bacterial strains. Subcloning plus insertion offers a novel approach to vector construction that can be performed in one electroporation event in the resident BAC host strain. The utility of SPI in gene targeting was tested here in a variety of vector construction applications. In all instances examined here, SPI proved to be efficient and the correct recombinant plasmid was produced. In the majority of the cases, the multiple different cassettes were correctly inserted in the targeting vector. Large DNA cassettes and vectors were easily accommodated in the SPI protocol and demonstrated the flexibility of this system.

SPI relies on the use of long homology sequences and phosphorothioate (PTO) protection of the linear DNA cassettes. PTO modification confers protection against exonucleases to the linear DNA[20,21] and the long HA increases recombination efficiency to permit multiplexing. However, synthesis of longer oligo sequences increases the chances of accumulating errors especially deletions. Mutations in the oligo can be particularly detrimental if they cover protein coding regions. DNA sequencing across the HA and covering the full-length of the inserted cassette is highly recommended to eliminate clones with any sequence alterations. Use of a high-fidelity DNA polymerase system is also suggested to avoid introduction of any PCR errors. The length and composition of the HA of the subcloning plasmid is more critical relative to that of the insertion cassette (data not shown). For particularly sensitive applications like construction of knockin vectors, where any mutations in exon regions are not tolerated, the HA of the insertion cassette can be shortened (50-120 bp) to avoid problems associated with long oligos. The insertion cassette can also be left unmodified or dual phosphorothioated (where knowledge of the direction of replication is not available). But multiplexing in these cases still requires long protected HA subcloning plasmids. A caveat of this particular strategy is the lowering of multiplexing efficiency that can potentially impact SPI cloning at different loci.

The length of the insertion cassette and the subcloning plasmid is another important parameter in SPI cloning. Larger DNA molecules electroporate less efficiently and the effect is cumulative, given the requirement to introduce all the cassettes in the same cell (data not shown). Multiplexing is most efficient with smaller insertion cassettes. DNA fragments larger than 3 kb also place a limit on Red mediated ssDNA processing, which is most efficient up to 3 kb[30]. Insertion cassettes exceeding 3 kb are dual resected and recombine less efficiently via a beta independent pathway[20]. Therefore, screening of sufficient colonies to identify the correct clone becomes important in these cases. A longer duration of recombination post electroporation also increases the chances of recovery of the correct clone in difficult SPI exercises. Up to four small cassettes (< 1.5 kb) can be inserted simultaneously with the SPI process, though the multiplexing efficiency decreases with each additional cassette (data not shown). However, this reflects the outcome of an optimal SPI experiment and it is advisable to consider the limits of multiplexing when using many large cassettes.

The development of genome editing tools like clustered regularly interspaced short palindromic repeats (CRISPR)-cas9 endonuclease system has enabled the creation of novel genome modifications and has led to more efficient genome engineering[36]. However, these newer technologies have complemented gene targeting vectors rather than supplanted them. It is envisaged that the CRISPR-cas system could replace antibiotic selection and further refine the multiplex recombineering protocol.

Disclosures

T.R.R is the inventor of a British patent application on multiplex recombineering technology.

Acknowledgements

T.R.R conceived the project, designed and performed experiments and prepared the manuscript. E.J.K and S.E.M performed experiments. All authors contributed intellectually to the final manuscript. This work was funded by a University of Leicester Innovation Fellowship to SEM and by a UK Medical Research Council (MRC) Senior Research Fellowship to S.M.C (MR/J009202/1). DNA sequencing was performed by Protein and Nucleic Acid Laboratories (PNACL), Centre for Core Biotechnology Services, University of Leicester. We wish to thank A. Francis Stewart for providing the Rox and Dre plasmids. pL451 and pL452 plasmids were obtained from National Cancer Institute (NCI), Frederick, USA.

References

1. Dovey, O. M., Foster, C. T., & Cowley, S. M. Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proc. Natl. Acad. Sci. USA.* **107**, 8242-8247, doi:10.1073/pnas.1000478107 (2010).
2. Harvey, M. *et al.* Spontaneous and carcinogen-induced tumorigenesis in p53-deficient mice. *Nat. Genet.* **5**, 225-229, doi:10.1038/ng1193-225 (1993).
3. Waterhouse, P. *et al.* Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. *Science.* **270**, 985-988, doi:10.1126/science.270.5238.985 (1995).
4. Thomas, K. R., & Capecchi, M. R. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell.* **51**, 503-512 (1987).
5. Hasty, P., Rivera-Pérez, J., & Bradley, A. The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell. Biol.* **11**, 5586-5591, doi:10.1128/mcb.11.11.5586 (1991).
6. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., & Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature.* **317**, 230-234 (1985).
7. Hirata, R., Chamberlain, J., Dong, R., & Russell, D. W. Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat. Biotechnol.* **20**, 735-738 (2002).
8. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., & Stahl, F. W. The double-strand-break repair model for recombination. *Cell.* **33**, 25-35 (1983).
9. Bradley, A., Evans, M., Kaufman, M. H., & Robertson, E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature.* **309**, 255-256 (1984).
10. Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W., & Stewart, A. F. A recombineering pipeline to make conditional targeting constructs. in: *Methods in Enzymology.* Volume **477**, Wassarman, P.M.., & Soriano P. M., eds., Academic Press, 125-144, (2010).
11. Zhang, Y., Buchholz, F., Muyrers, J., & Stewart, A. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123 - 128 (1998).
12. Copeland, N., Jenkins, N., & Court, D. Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* **2**, 769 - 779 (2001).
13. Muyrers, J., Zhang, Y., Testa, G., & Stewart, A. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27**, 1555 - 1557 (1999).
14. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* **97**, 5978 - 5983 (2000).
15. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G., & Court, D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protoc.* **4**, 206-223, doi:10.1038/nprot.2008.227 (2009).
16. Wang, J. *et al.* An improved recombineering approach by adding RecA to λ red recombination. *Mol. Biotechnol.* **32**, 43-53 (2006).
17. Datta, S., Costantino, N., & Court, D. L. A set of recombineering plasmids for gram-negative bacteria. *Gene.* **379**, 109-115, doi:10.1016/j.gene.2006.04.018 (2006).
18. Fu, J. *et al.* Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **30**, 440-446 (2012).
19. Muyrers, J., Zhang, Y., Buchholz, F., & Stewart, A. RecE/RecT and Redalpha/Redbeta initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Develop.* **14**, 1971 - 1982 (2000).
20. Maresca, M. *et al.* Single-stranded heteroduplex intermediates in lambda Red homologous recombination. *BMC Mol. Biol.* **11**, 54 (2010).
21. Mosberg, J. A., Lajoie, M. J., & Church, G. M. Lambda Red recombineering in *Escherichia coli* occurs through a fully single-stranded intermediate. *Genetics.* **186**, 791-799, doi:10.1534/genetics.110.120782 (2010).
22. Liu, P., Jenkins, N. A., & Copeland, N. G. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**, 476-484, doi:10.1101/gr.749203 (2003).
23. Chan, W. *et al.* A recombineering based approach for high-throughput conditional knockout targeting vector construction. *Nucleic Acids Res.* **35**, e64, doi:10.1093/nar/gkm163 (2007).
24. Zhang, Y., Muyrers, J. P. P., Testa, G., & Stewart, A. F. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314-1317 (2000).
25. Lee, E. C. *et al.* A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics.* **73**, 56-65, doi:10.1006/geno.2000.6451 (2001).
26. Skarnes, W. C. *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature.* **474**, 337-342 (2011).
27. Valenzuela, D. M. *et al.* High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* **21**, 652-659 (2003).
28. Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat. Biotechnol.* **29**, 255-265, doi:10.1038/nbt.1759 (2011).
29. Mulryan, K. *et al.* Reduced vas deferens contraction and male infertility in mice lacking P2X1 receptors. *Nature.* **403**, 86-89, doi:10.1038/47495 (2000).
30. Subramanian, K., Rutvisuttinunt, W., Scott, W., & Myers, R. The enzymatic basis of processivity in lambda exonuclease. *Nucleic Acids Res.* **31**, 1585 - 1596 (2003).

31. Antoch, M. P. *et al.* Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell.* **89**, 655-667 (1997).
32. Rostovskaya, M. *et al.* Transposon-mediated BAC transgenesis in human ES cells. *Nucleic Acids Res.* **40**, e150, doi:10.1093/nar/gks643 (2012).
33. Giraldo, P., & Montoliu, L. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res.* **10**, 83-103 (2001).
34. Hill, F. *et al.* BAC trimming: minimizing clone overlaps. *Genomics.* **64**, 111-113, doi:10.1006/geno.1999.6106 (2000).
35. Warming, S., Costantino, N., Court, D. L., Jenkins, N. A., & Copeland, N. G. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36, doi:10.1093/nar/gni035 (2005).
36. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science.* **339**, 819-823, doi:10.1126/science.1231143 (2013).