

Assisted large fragment insertion by Red/ET-recombination (ALFIRE)—an alternative and enhanced method for large fragment recombineering

Adolfo Rivero-Müller^{1,*}, Svetlana Lajić² and Ilpo Huhtaniemi^{1,2}

¹Department of Physiology, University of Turku, Kiinamylynkatu 10, Turku 20520, Finland and

²Institute of Reproductive and Developmental Biology (IRDB), Imperial College London, Du Cane Road, W12 0NN London, UK

Received October 30, 2006; Revised March 22, 2007; Accepted April 5, 2007

ABSTRACT

Functional genomics require manipulation and modification of large fragments of the genome. Such manipulation has only recently become more efficient due to the discovery of different techniques based on homologous recombination. However, certain limitations of these strategies still exist since insertion of homology arms (HAs) is often based on amplification of DNA sequences with PCR. Large quantities of PCR products longer than 4–5 kb can be difficult to obtain and the risk of mutations or mismatches increases with the size of the template that is being amplified. This can be overcome by adding HAs by conventional cloning techniques, but with large fragments such as entire genes the procedure becomes time-consuming and tedious. Second, homologous recombination techniques often require addition of antibiotic selection genes, which may not be desired in the final construct. Here, we report a method to overcome the size and selection marker limitations by a two- or three-step procedure. The method can insert any fragment into small or large episomes, without the need of an antibiotic selection gene. We have *humanized* the mouse luteinizing hormone receptor gene (*Lhcgr*) by inserting a ~55 kb fragment from a BAC clone containing the human *Lhcgr* gene into a 170 kb BAC clone comprising the entire mouse orthologue. The methodology is based on the rationale to introduce a counter-selection cassette flanked by unique restriction sites and HAs for the insert, into the vector that is modified. Upon enzymatic digestion, *in vitro* or in *Escherichia coli*, double-strand breaks are generated leading to recombination

between the vector and the insert. The procedure described here is thus an additional powerful tool for manipulating large and complex genomic fragments.

INTRODUCTION

With completion of the human and mouse genome projects, a large volume of data for genes with unknown function will be available. In the present and forthcoming era of functional genomics, sequenced bacterial artificial chromosomes (BACs) (1), P1 vectors (2) and P1 artificial chromosomes (PACs) (3) will be very useful in the analysis of gene function, since they can be used for construction of complex vectors and subsequent creation of transgenic, knockout and knock-in mice. The difficult manipulation of such large episomes with conventional cloning techniques has until now limited their complete usefulness. However, the recent development of techniques based on homologous recombination in *Escherichia coli* (4–8) has improved, simplified and facilitated the manipulation of both large and small vectors, thus becoming a very important tool in the design of targeting vectors and the subsequent functional analysis of newly discovered genes.

Manipulation of plasmid DNA by homologous recombination was first reported in yeast (9) and since then techniques have been developed to include phage-based vectors for convenient and efficient use in *E. coli*. The use of these plasmids to carry out genetic engineering has been called *recombinogenic engineering* or *recombineering* (6). Recombineering exploits the phage derived protein pairs, either RecE/RecT from the *Rac* phage or Red α /Red β from the λ phage, to assist in the cloning or subcloning of fragments of DNA into vectors without the need of restriction enzyme sites or ligases (7,8). A limitation of the original homologous recombination technique was due to

*To whom correspondence should be addressed. Tel: +44 1223 402474; Fax: +44 1223 402070; Email: adoriv@utu.fi

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

the fact that bacterial RecBCD nuclease degrades linear DNA and initially the event had to be studied in RecBCD-deficient strains (7). This was overcome by the discovery that Red α and Red β were assisted by Red γ , which inhibits RecBCD nuclease activity making it possible to use the technique in *E. coli* and other commonly used bacterial strains (10). In addition, the recombination efficiency was increased 10–100 times (11). The combination of these three enzymes (α , β and γ , or E, T and γ) in one vector was named Red/ET recombination and the basic principles of the method are that it requires two homology regions of >42 bp in a linear fragment, double strand breaks (DSBs) in both ends, and another linear or circular plasmid in order for recombination to take place. DSBs are essential so that RecE or Red α can bind and degrade one chain of the DNA (5' to 3') and at the same time load RecT or Red β to the single strand chain that is exposed (7). The single DNA strand loaded with the RecT or Red β recombinase finds a perfect match sequence and joins the two sequences by either chain invasion or annealing.

However, this system requires insertion of homology arms (HAs), which are included in the oligonucleotides that are used for amplification of the PCR products used as linear substrates for the recombination event. The limitations of the PCR reaction with long primers make it difficult to generate large quantities of fragments that are longer than ~4–5 kb, and even in the event of being able to PCR longer templates the rate and risk of mutations/mismatches increases with the size of the template. If longer fragments of DNA are needed for the cloning procedures then the HAs may be inserted with conventional restriction/ligation techniques into primary constructs that are excised and used for recombination. Moreover, the insertion of DNA by Red/ET cloning is highly efficient only if a selection marker gene is included, which may be undesirable in the final construct. This has been addressed by the use of elegant selection/counter-selection systems (12,13) but the limitations of the PCR-size of the insert has not yet been solved. Insertion of large fragments with recombineering may thus be a time consuming multi-step process.

BACs and large plasmids (e.g. subcloned genomic fragments from BACs into minimal vectors) are important tools for the creation of transgenic organisms since they contain most of the regulatory elements needed for optimal expression. There are many large mammalian genes which are of the same order of magnitude or larger than the average insert of the BAC libraries and it is often difficult to find a single BAC spanning the entire gene including its controlling elements. It may thus be desirable to fuse or exchange fragments from different BACs in order to obtain larger functional regions of DNA or to create *humanized* mouse genes, where the gene of interest in the mouse BAC is exchanged for the human genomic counterpart. Fusion of the regulatory elements of one BAC with another BAC that contains the coding regions of a gene of interest thus becomes a powerful strategy to study tissue or cell specific gene expression. This can be

achieved by using one end of a BAC that overlaps an area of another as homology regions or by adding the HAs into the insert-containing vector by either recombination or conventional cloning (14). However, finding overlapping BACs limits the uses of this strategy. Moreover, the method cannot introduce other non-overlapping fragments such as the same gene from different species or a different gene and it seems to be efficient only if a selection marker (Sm) is present in the inserting fragment.

We report here a method, Assisted Large Fragment Insertion by Red/ET-recombination (ALFIRE), where we have overcome both the size limitations and the restrictions of the cloning/subcloning procedures used in the Red/ET recombineering system. We describe the transfer of a very large DNA fragment (~55 kb) from one BAC into another unrelated BAC by the means of a modified *RpsL*-counter-selection (CS) system (15) combined with *in vitro* or *in vivo* restriction enzyme digestion (16,17).

ALFIRE thus represents an improvement of tools that can be used in order to facilitate construction of large and complex vectors utilized in molecular biology and functional genomics.

MATERIALS AND METHODS

The Red/ET system and the counter-selection cassette (*RpsL*-neo)

The Red/ET recombination system (pSC101BAD $\gamma\beta\alpha$ A[tet]), and the *RpsL*-neo counter-selection cassette (CS) were purchased from GeneBridges, Dresden, Germany.

BAC clones

BACs containing the entire mouse or human luteinizing hormone/chorionic gonadotropin receptor genes (*Lhcgr*) were purchased from Research Genetics (clones RP23-18D7 and RP11-73L19, respectively). Individual BAC clones were propagated in *E. coli* in Luria–Bertani (LB) media containing chloramphenicol (Cm, 20 μ g/ml)/streptomycin (Stp, 75 μ g/ml) and purified using the Large construct kit (Qiagen, Valencia, CA, USA).

PCR

PCR reactions using oligonucleotides carrying HAs were performed using a proofreading polymerase and buffers included in the kit Triple Master Mix (Eppendorf, Hamburg, Germany). The standard reaction conditions were 96°C for 3 min, followed by 30 cycles with 96°C, 45 s, 57°C 45 s, 72°C, 1–3 min, depending on the size of the product, and a final elongation of 5 min at 72°C. An aliquot of each PCR product was analysed by gel electrophoresis. The rest of the PCR product was ethanol-precipitated, resuspended in Tris–HCl (10 mM pH 8.5) and used for transformation into electrocompetent bacteria. PCR was also used to screen for positive recombinants and to verify correct integration.

Primers

The short primers (20–30 nt) were purchased from TAG, Copenhagen, Denmark, whereas the longer primers (60–140 nt) were purchased from Thermo Electron, Hamburg, Germany. Oligonucleotide sequences used for design of the I-SceI-CS-*hmLhcgr*-BAC: CS (I-SceI) to *hmLhcgr*-F (CCAGCATACTGGCCTAGCCACCGGA GCTCACACTCAGGCTGGCGGGCCATGAAGCAGC GGTCTCGGCGCTGCAGCTGCTGAAGCTGCTGC TGCTGCTaggataacaggtaatGGCCTGGTGATGATG GCGGGATCG); CS (I-SceI) to *hmLhcgr*-R (GTGG ACTTTTTGGGGGAACATATTTAGATACAATTC AGTAATGCAGTTAACACTCTGTGTAGCGAGTCT TGTCTAGGAGAGCTGTACCTTGACAGTattaccctgtt atccctaTCAGAAGAACTCGTCAAGAAGGCG).

Recombination

Recombination events were performed in electrocompetent DH10B or HS996 cells harbouring the pSC101BAD $\gamma\beta\alpha$ [hygro] or the pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] plasmids and using standard Red/ET protocols (8). Electrocompetent cells were prepared as described in the general manuals from GeneBridges. Electroporation was performed at 1.35 kV, 25 μ F, 200 Ω using an Eppendorf electroporator (2510). Bacteria were incubated in 1 ml LB media at 37°C, 1100 rpm for 1–2 h before plating on agar plates conditioned with the appropriate antibiotic(s).

Media and antibiotics

Standard LB media and agar plates were used. Antibiotics were used at the following concentrations for selection of high copy vectors on agar plates: ampicillin (Ap) 50 μ g/ml, kanamycin (Km) 25 μ g/ml; and in liquid media: Ap 50 μ g/ml, Km 50 μ g/ml. For BACs and low copy vectors, on plates and in liquid media: chloramphenicol (Cm) 20 μ g/ml, Ap 50 μ g/ml and Km 15 μ g/ml. Tetracycline (Tc) concentrations for selection of pSC101BAD $\gamma\beta\alpha$ [tet] was 5 and 3 μ g/ml on plates and in liquid media, respectively. Hygromycine (Hyg) was used at 30 μ g/ml on plates and 15 μ g/ml in liquid media for selection of the pSC101BAD $\gamma\beta\alpha$ [hygro] vector and streptomycin (Stp) was used at a concentration of 75 μ g/ml.

Heat-treated chlor-tetracycline (cTc) was used to inactivate the Tet repressor. The inducer cTc was suspended in LB at a concentration of 400 μ g/ml and autoclaved in a pressure cooker for 20 min, then stored in dark environment. Induction of I-SceI was achieved by adding cTc to the growing cells at a final concentration of 20 μ g/ml.

Construction of the pSC101BAD $\gamma\beta\alpha$ [hygro] vector

The pSC101BAD $\gamma\beta\alpha$ [hygro] plasmid was based on the pSC101BAD $\gamma\beta\alpha$ [tet] vector (GeneBridges) and was constructed by standard Red/ET recombination procedures (8) using the primers Hygro-F (5' AATGCGGTAGTTTATCACAGTTAAATTGCTAAC GCAGTCAGGCACCGTGTATGGATAGATCCGGA AAGCCTGAA 3') and Hygro-R (5' TCCAATTCTTGG

AGTGGTGAATCCGTTAGCGAGGTGCCGCCGGC TTCCATCTATTCCTTTGCCCTCGGACGAGT 3') in order to amplify the *hygromycin phosphotransferase* gene from the pIRESHYG3 plasmid (BD Biosciences, Palo Alto, CA, USA). Positive recombinants were selected on LB-agar plates supplemented with Hyg 30 μ g/ml.

Construction of the pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector

The pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector was assembled by PCR amplification of the I-SceI gene together with the Tet-Repressor, Tet-Promoter and *ampicillin resistance* gene sequences from the PsiI-digested pST98AS plasmid (AF170483) (16) [oligo sequences I-SceI to ET Hyg-F (AATGCGGTAGTTTATCACAGTTAAATTGCTAAC GCAGTCAGGCACCGTGTGACCAATTCGGGTGCA CTTAT)/I-SceI to ET Hyg-R (TCCAATTCTTGG AGTGGTGAATCCGTTAGCGAGGTGCCGCCGGCT TCCATTGGTCATGAGATTATCAAAAAGGA)]. The PCR product which contained HAs for pSC101BAD $\gamma\beta\alpha$ [hygro] was co-transformed with the pSC101BAD $\gamma\beta\alpha$ [hygro] vector into HS996 electrocompetent *E. coli* cells harbouring the pSC101BAD $\gamma\beta\alpha$ [tet] vector and recombination was achieved using standard Red/ET protocols (8) and selection with Ap 50 μ g/ml.

Verification of positive recombinants

PCR was used to check for correct insertion into plasmids or BACs by using one primer annealing to the vector and the other to the insert. In some cases the full insert was amplified by flanking primers. PCR reactions were performed using Biotoools polymerase, Madrid, Spain. The general conditions included an initial denaturation at 96°C for 3 min, followed by 30 cycles with 95°C for 45 s, 57°C for 45 s and 72°C for 1–3 min, depending on the length of the amplified product. Products were analysed by gel electrophoresis and when needed DNA was recovered or minipreps were purified and used for sequencing. Sequencing was performed in both directions when the product was longer than 1 kb or when both integration points were analysed. Correct recombination was also verified by restriction enzyme digestion of positive clones. A complete list of the oligonucleotide sequences used for screening of recombinants can be obtained from the authors upon request.

Single BAC clones were grown in LB medium conditioned with the Cm/Stp and screened by PCR over the inserted areas (the *mLhcgr* promoter to the *hLhcgr* intron1 area; and the *hLhcgr* exon 11 to the *mLhcgr* 3' area) using primers *mLhcgr*P-F (CCAGCATACTGGCCTAGCC AC)/*hLhcgr*1-R (AGTACACAGCGCTCCCGTC) and *hLhcgr*end-F (CGAAACCCAGAATTAATGGCTA)/*mLhcgr*3-R (CAATTCACCTGAAGTGCTTAAAGA), respectively. Correctly integrated clones were further screened by PCR of areas covering exons 5 and 6 of the inserted *hLhcgr* using primers Exon 5-6-F (GCATG AGGGACTTCTAAATTGC)/Exon 5-6-R (TGCTCTTT TTAAGCCAGGAAAG), and by lack of PCR amplification of either the *hLhcgr*-BAC [primers *hLhcgr*end-F/*hLhcgr*3-R (GTTTTTAGTGTGGCAGTGGTCA)] or the *mLhcgr*-BAC [primers *mLhcgr*end-F (TACCC TTACAGTCATCACTCTGGA)/*mLhcgr*3-R].

RESULTS

Designing the tools for assisted large fragment insertion

The pSC101BAD $\gamma\beta\alpha$ [hygro] vector. The pSC101BAD $\gamma\beta\alpha$ [hygro] plasmid was assembled by replacing the *tet* gene of pSC101BAD $\gamma\beta\alpha$ [tet] with the *hygro* gene via standard Red/ET recombination protocols (8). Plasmid DNA was isolated from positive (hygromycin selected) recombinants and digested with NcoI to further verify correct recombination (Figure 1a) and three clones were additionally sequenced over the integration sites. The rationale why the pSC101BAD $\gamma\beta\alpha$ [hygro] vector was constructed was to facilitate manipulation of targeting vectors that are subcloned into the commonly used tetracycline resistant minimal vector pACYC184 (New England Biolabs, Beverly, MA, USA). ALFIRE can be used for manipulation of any vector and targeting vectors for construction of knockout or knock-in animals are often assembled by subcloning DNA fragments from BAC clones with the help of Red/ET recombination. In addition, the pSC101BAD $\gamma\beta\alpha$ [hygro] plasmid was used as the basic vector for the construction of the pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector.

The pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector. The pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector (Supplementary Figure 1) was constructed to promote *in vivo* enzyme digestion followed by recombination and thereby facilitating modification of BAC clones using ALFIRE. It was assembled by introducing the tetracycline (cTc)-inducible I-SceI homing endonuclease fragment from the pST98AS plasmid (16) to the pSC101BAD $\gamma\beta\alpha$ [hygro] plasmid with standard Red/ET protocols (8). The entire *hygro* gene was thus exchanged with the I-SceI-Tet-Rep/Tet-Prom-amp unit generating pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp]. Positive (Ap selected) clones were verified by SspI restriction enzyme digestion (Figure 1b) and three clones were sequenced over the modified areas. The vector data were submitted to EMBL (AM403094).

The vector pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] encodes the Red γ , Red β , Red α and RecA proteins as well as the homing endonuclease I-SceI. Upon activation with *L*-arabinose and chloro-tetracycline (cTc), respectively, these enzymes will be expressed and cells harbouring this plasmid will become capable of performing digestion and recombination of the resulting ends containing the homologous sequences. Plasmids or BACs containing the 18 bp recognition site of the homing endonuclease I-SceI can be linearized *in vivo* and the double-strand breaks thus formed promote recombination to identical homology sequences. Efficient recombination is especially important when large DNA fragments are being recombined such as in BAC fusion experiments.

To test the functionality of the I-SceI meganuclease, 10 bacterial colonies harbouring the I-SceI-CS-*hmLhcgr*-BAC and the pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] were grown overnight at 30°C in LB media with Km/Cm/Ap. Next day, three parallel tubes with 1 ml LB media were set up and inoculated with 30 μ l of each culture and incubated at 1100 rpm with the following conditions (1) cTc/Cm/Ap, 37°C for 1 h; (2) cTc/Cm/Ap, 30°C for 5 h; (3) Cm/Ap, 30°C for 6 h. None of the ten clones propagated in the cTc-induced tubes incubated at 37°C, showing that I-SceI was expressed and that the I-SceI-CS-*hmLhcgr*-BAC was linearized at the I-SceI sites. All of the ten clones grew in the non-induced tubes incubated at 30°C indicating that the I-SceI endonuclease was not expressed and that the I-SceI-CS-*hmLhcgr*-BAC clone and the pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] vector could continue to propagate. Another control experiment was also set up where a clone harbouring only the I-SceI-CS-*hmLhcgr*-BAC was incubated at 37°C in LB media (cTc/Cm) for 6 h, 1100 rpm. This clone also propagated indicating that cTc is not toxic to *E. coli* at the appointed concentration.

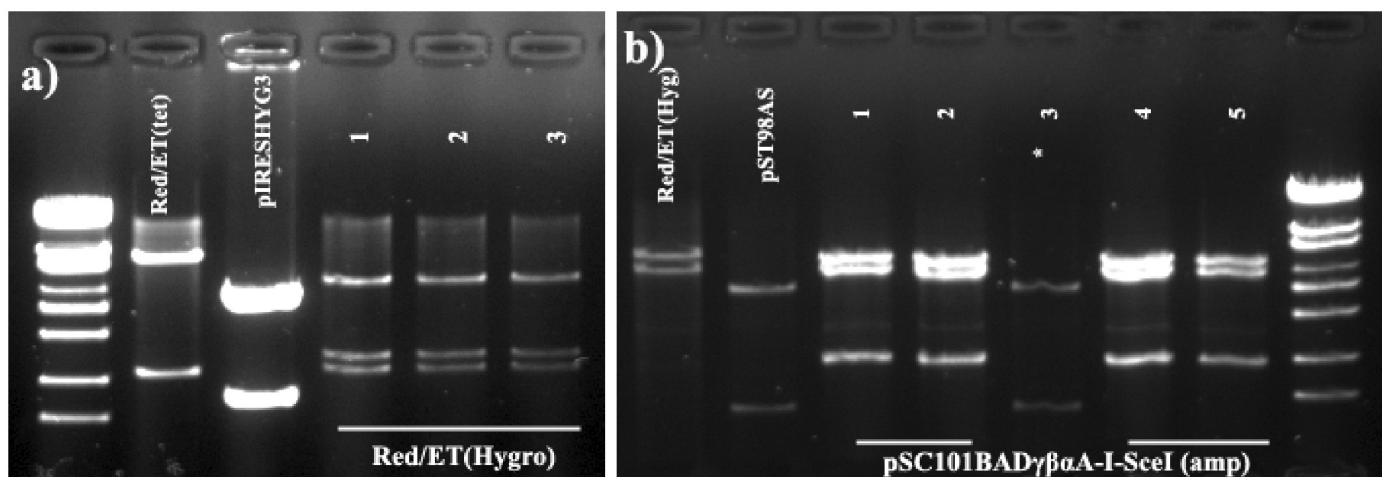


Figure 1. Verification of correct pSC101BAD $\gamma\beta\alpha$ [hygro] (Red/ET (hygro) and pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] recombinants. (a) Correct pSC101BAD $\gamma\beta\alpha$ [hygro] recombinants were identified by digestion with NcoI (fragment sizes 4.8, 2.3 and 2.0 kb). (b) Correct pSC101BAD $\gamma\beta\alpha$ -I-SceI[amp] clones were identified by digestion with SspI (fragment sizes 4.6, 4.1 and 2.0 kb). *Negative clone.

Assisted large fragment insertion with Red/ET-recombination (ALFIRE)

ALFIRE is based on the rationale that HAs can be easily inserted directly into the accepting vector at the same time when the integration of an insert/selection cassette is performed. We used long oligonucleotides containing ~50 bases homologous to the acceptor plasmid insertion site, ~55 bases homologous to the insert, and in order to generate the DSBs needed for recombination a unique restriction site (RS) sequence was included prior to the primer binding sequence. The counter-selection cassette *RpsL*-neo (CS) was used for primary selection in order to reduce the background of the undigested accepting vector. Accepting and donor vectors were linearized *in vitro* and transformed into Red/ET competent bacteria for recombination to take place (Figure 2). With this strategy ALFIRE can be used efficiently irrespective of the vector or insert size and we show here the exchange of a very large fragment (~55kb) from one BAC (*hLhcgr*) to another unrelated BAC (*mLhcgr*).

In the initial step, a CS cassette flanked by recognition sites of I-SceI and HAs to the insert and the acceptor vector was designed (see Figure 2 and Supplementary Figure 2). The primers were ~140 nt long and this was the limit set by the supplier to guarantee good quality primers. The I-SceI sequence was chosen because the 18 bp long recognition site is rare even in genomic DNA. In fact, the *E. coli* chromosome does not contain a single I-SceI site and the meganuclease has been used to generate DSBs into foreign DNA amplified in *E. coli* (16, 17). First, the CS flanked by two I-SceI restriction sites (I-SceI-CS) was inserted to the mouse *Lhcgr* (*mLhcgr*)-BAC by Red/ET recombination with the following procedure: the I-SceI-CS cassette was amplified by PCR with primers [CS (I-SceI) to hmLhcgr-F/CS (I-SceI) to hmLhcgr-R] containing (a) the HAs for insertion (~50 nt) to the *mLhcgr* promoter and 3' regions; (b) HAs (~55 nt) to subclone the human *Lhcgr* (*hLhcgr*); (c) the 18 nt recognition site of the homing endonuclease I-SceI (TAGGGATAACAGG GTAAT); and (d) the primer sites (24 nt) for the *RpsL*-neo (Figure 2 and Supplementary Figure 2).

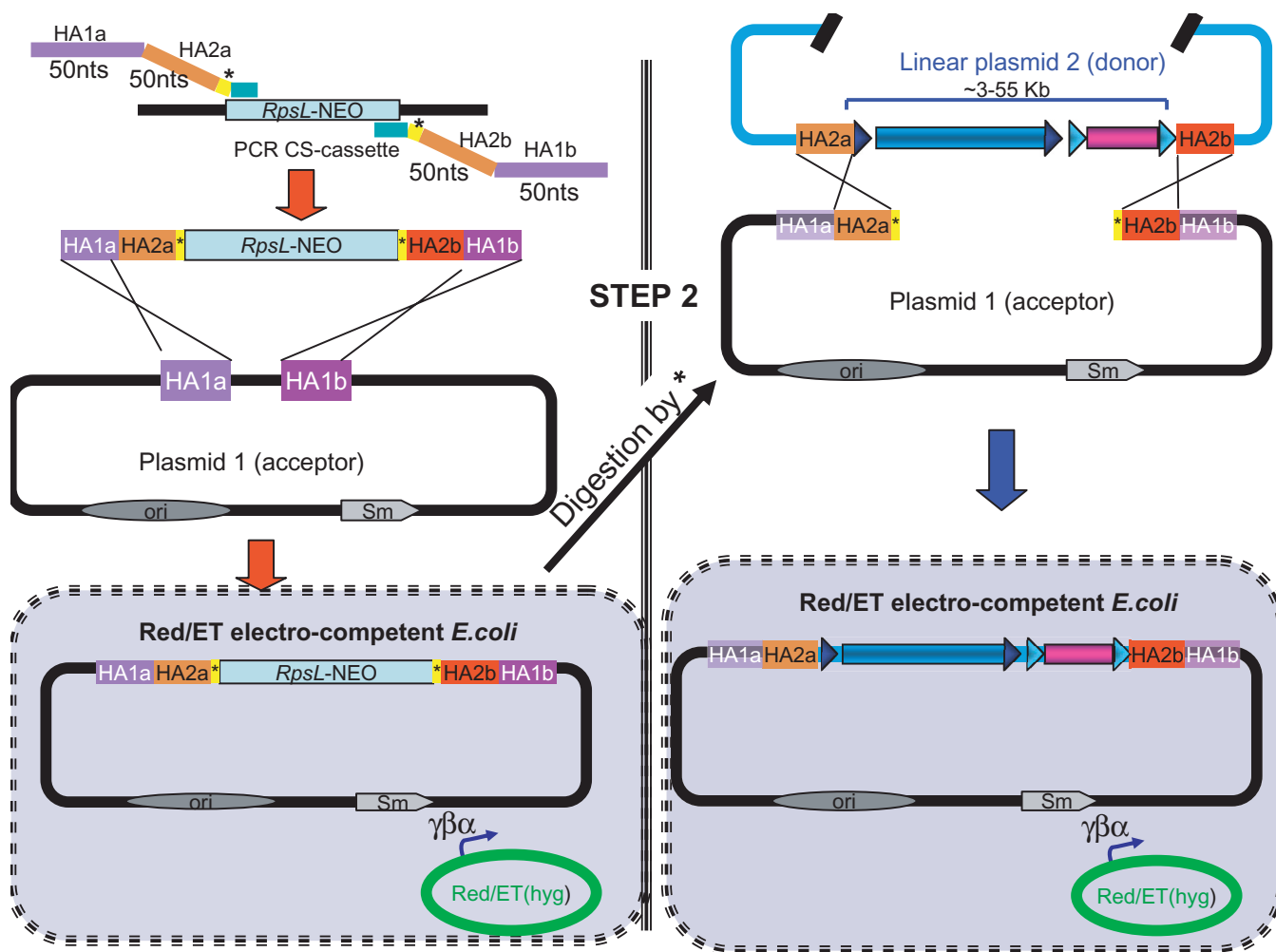


Figure 2. Outline of the ALFIRE procedure. The accepting vector is modified with a counter-selection/selection cassette (*RpsL*-neo) flanked by two unique restriction sites and containing HAs to the fragment to be inserted. The resulting vector is linearized and co-transformed with a linear fragment containing the insert into bacteria expressing Red/ET recombinases. Cells are plated in the appropriate antibiotics. Most of the surviving cells will have undergone recombination between the acceptor vector and the insert.

Red/ET competent bacteria carrying the *mLhcgr*-BAC were transformed with the I-SceI-CS PCR product and plated on LB agar plates supplemented with Km/Cm. Positive clones were screened by PCR (primers *mLhcgr*P-F/*mLhcgr*3-R) and further selected by their inability to grow (negative selection) on LB agar plates supplemented with Stp/Cm. In addition, the resulting PCR product was digested *in vitro* with I-SceI (NEB) in order to check for functionality of the inserted I-SceI restriction sites. One of three clones contained a mismatch in the I-SceI sequence leading to reduced or abolished cleavage by this meganuclease. Five positive clones were further selected for sequencing of the entire HA area, which needed to be faultless to allow efficient Red/ET recombination. Four out of five clones were without mismatches and one was selected for further experiments. The I-SceI-CS-*hmLhcgr*-BAC construct can thus be used as a general PCR template (I-SceI-CS) and the I-SceI recognition sites would not need to be added each time in the oligonucleotide sequences in future experiments, making the oligos shorter and eliminating the need of checking the I-SceI sites every time.

***In vitro* generation of DSBs.** Linearization of the *hLhcgr*-BAC was performed by digestion with NotI (Promega, Mannheim, Germany), and I-SceI-CS-*hmLhcgr*-BAC was digested with I-SceI (NEB) following the manufacturer's instructions. Both BACs were purified by phenol: chloroform, precipitated with ethanol and resuspended in Tris-HCl buffer (10 mM, pH 8.5). Approximately 3 µg of each linear BAC was co-transformed to Red/ET[hygro] electrocompetent bacteria and selection of positive clones was initially performed on Stp/Cm agar plates and further in liquid LB media (Stp/Cm). Only 2/32 PCR screened colonies (6.3%) showed correct integration of the ~55 kb *hLhcgr* insert (Table 1). This might be due to low-transformation efficiency of long linear fragments. Since BACs are difficult to manipulate, separate and purify by gel electrophoresis, and the *in vitro* digestion resulted in poor recombination efficiency, an *in vivo* digestion system was also created.

ALFIRE in *E. coli*. *In vivo* recombination between the *hLhcgr*-BAC and the I-SceI-CS-*hmLhcgr*-BAC was achieved by activation of the Red/ET recombinases and the I-SceI homing endonuclease expressed by the pSC101BADγβαA-I-SceI[amp] vector in *E. coli*. DSBs on the I-SceI-CS-*hmLhcgr*-BAC were induced as

follows: Bacterial cultures (1.4 ml) in LB conditioned with Cm/Km/Ap for *E. coli* containing both the I-SceI-CS-*hmLhcgr*-BAC and the pSC101BADγβαA-I-SceI[amp] plasmid (*L*-arabinose and cTc inducible) were grown overnight in a shaker (Eppendorf) at 30°C, 1100 rpm. Next day electrocompetent cells were made from 30 ml cultures following the general manual of GeneBridges. At a cell density of OD 0.2, *L*-arabinose was added to the medium (final concentration 0.2%) and the cultures were further grown for 45 min at 37°C, 300 rpm, with the difference that at different time points, inactivated chlorotetracycline (cTc) was added (final concentration 20 µg/ml) and the cells were made electrocompetent. After transformation with 3 µg of linear (NotI digested) *hLhcgr*-BAC cells were plated in LB-agar plates (diameter 25 cm) containing Stp/Cm and cTc, 15 µg/ml, and incubated overnight at 37°C. The number of colonies on the plate was low (<100) but the recombination efficiency was high even with such a long fragment as ~55 kb. Maximum recombination efficiency (69%) was achieved after 15 min induction with cTc (Table 1). Thirty-two of the resulting colonies were screened by PCR amplification of the integration points (5' and 3') between the *mLhcgr* BAC and the inserted *hLhcgr* gene. Twenty-two of 22 clones showed positive results in both ends (Figure 3a). Randomly, five clones were picked and grown further to check their integrity. All five clones contained exons 5 and 6 (Figure 3b), which are in the middle of the inserted region, and had no contamination of either the *hLhcgr*-BAC or the unmodified *mLhcgr*-BAC (Figure 3c). Minipreps (5 ml) from the same clones were made and the DNA was subjected to restriction enzyme analysis. Digestion with XhoI/PacI showed the expected pattern equally in all clones (Figure 3d).

The linear BAC (*hLhcgr*) could not recombine with the circular accepting BAC vector (I-SceI-CS-*hmLhcgr*) when the I-SceI meganuclease was not induced. The lack of recombination efficiency (Table 2) by this latter experiment suggests that the recombination event is dependent on both the DSBs generated by the I-SceI and the activity of the Red/ET recombinases in order to be fully effective. For schematic presentation of the use of ALFIRE in *E. coli* see Supplementary Figure 2.

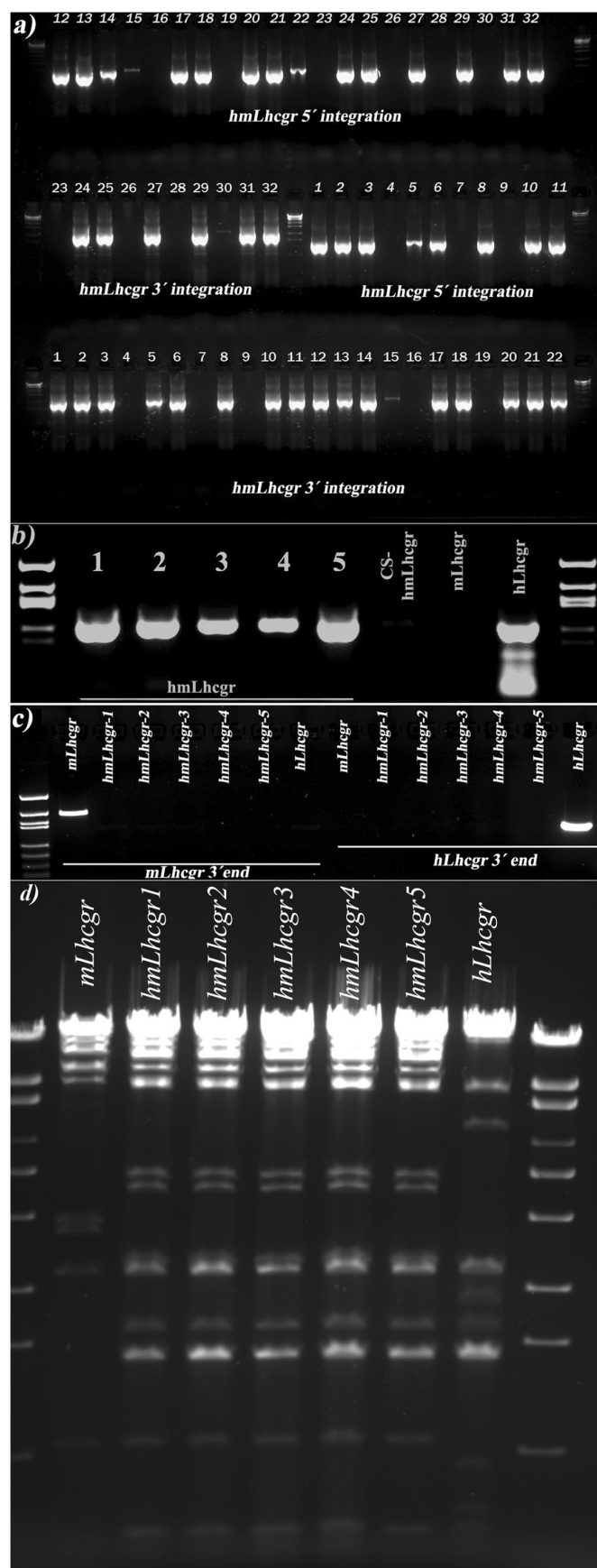
DISCUSSION

The use of homologous recombination to generate and manipulate large fragments of DNA has opened up new avenues in the study of functional genomics. A limitation of traditional recombinogenic engineering is the difficulty of inserting large fragments into large episomes in a single step. It is necessary to overcome these limitations in order to facilitate the manipulation of large genomic sequences needed for the production of transgenic, knockout and knock-in organisms.

We have overcome the major difficulty of inserting HAs into large fragments by incorporating the homologous sequences in the acceptor vector together with unique restriction sites thereby promoting generation of DSBs and facilitating recombination. It is important to note,

Table 1. Efficiency for *in vitro* and *in vivo* BAC fusion (insertion of long sequences) with ALFIRE

	Incubation with cTc (min)	Tested colonies (n)	Positive colonies (n)	Efficiency (%)
<i>In vitro</i>	–	32	2	6.3
<i>In E. coli</i>	0	32	0	0
	15	32	22	69
	20	32	16	50
	30	32	6	19



that often such restriction sites are palindromes and therefore the choice of the endonuclease should be carefully addressed to make sure that the resulting ends have little homology in order to reduce vector re-circularization. It is known that as little as ≥ 6 nt homology near each end might result in 'end joining' (8). Homing endonucleases have long recognition sites that are not palindromes and the presence of these sequences is rare even in genomic size DNA. We chose the homing meganuclease I-SceI recognition sequences as unique sites to generate DSBs in the acceptor BAC carrying the HAS for the insert. However, when the entire *hLhcgr* (~ 55 kb) was subcloned into the region between the promoter and the polyA of a BAC containing the mouse homologue (*mLhcgr*), the efficiency after *in vitro* digestion (I-SceI) and isolation was relatively low (6.3%). The poorer recombination efficiency is most probably due to the limitations of transforming the large BACs, in particular if they are linear, into bacteria and the risk of shredding of the BAC during the DNA purification steps. To overcome these problems we constructed an all-in-one universal vector (pSC101BAD $\gamma\beta\alpha$ A-I-SceI [amp]) to be used in *E. coli*. This system expresses the recombinases Red γ , Red β and Red α as well as the I-SceI homing endonuclease under the tight control of *L*-arabinose and chloro-tetracycline, respectively, and makes it possible to linearize the acceptor vector intracellularly followed by recombination between homologous sequences. In this way only the *hLhcgr* BAC was linear resulting in improved transformation efficiency and recombination (69% positive clones).

Table 2. Efficient recombination of large fragments is dependent on double strand cleavage (I-SceI) of the accepting vector to unveil homology arms to Red/ET recombinases and induce recombination by gap repair

Red/ET	I-SceI	Tested colonies	Positive clones	Efficiency (%)
-	-	11 ^b	0 ^a	0
+	-	6 ^b	0	0
-	+	0 ^a	0	0
+	+	32	22	69

The accepting BAC (I-SceI-CS-*hmLhcgr*) was co-transformed with the linear *hLhcgr* BAC to competent cells either expressing (+) or not expressing (-) the Red/ET and/or I-SceI enzymes.

^aNo colonies on the plate, or no growth in liquid media.

^bAll colonies on the plate were picked.

Figure 3. Integration of the *hLhcgr* gene into the *mLhcgr* BAC by ALFIRE. Resulting recombinants were screened over the 5' and 3' integration sites. The integrity of the BAC (*hmLhcgr*) was first checked by PCR of a fragment including exons 5 and 6 (in the middle of the inserted region) and by restriction enzyme digestion pattern. (a) *hLhcgr* gene integration into the area between the mouse BAC promoter and 3' end; (b) five clones were further screened by PCR over the middle region (exons 5 and 6) of the 55 kb insert; (c) the same clones were checked for contamination by either the *hLhcgr*-BAC or the *mLhcgr*-BAC by PCR amplification of the 3' regions and (d) enzymatic digestion of *mLhcgr*, *hLhcgr* and *hmLhcgr* BAC clones with XhoI/Pacl.

Thus, a 10-fold higher efficiency was achieved by *in vivo* compared to *in vitro* DSBs generation by I-SceI (Table 1) and the background was also reduced significantly with *in vivo* enzyme digestion.

We show that by using appropriate unique restriction enzymes sites, and subcloning-HAs, large fragments can be copied in 2 or 3 steps (see Figure 2 and Supplementary Figure 2 for general overview of the ALFIRE method). By adding the HAs for subcloning into the oligos used for amplification of the I-SceI-CS cassette we eliminated the time-consuming use of transfer constructs. The presence of these short exact sequences (~55 nt) is thus sufficient to result in a highly efficient and accurate recombination. We also show that long oligos, up to ~140 nt, can generate satisfactory HAs for insertion and for subcloning. Alternatively, the double HAs could be generated by PCR using four shorter oligos (>80 nt) instead of two in the same PCR reaction.

The methodology described is highly efficient and an easy protocol to follow in order to insert, fuse, exchange and manipulate virtually any size of DNA or plasmid without long template PCR and undesired selection markers. It is thereby an extra tool in addition to existing methods (8,16) for the generation and modification of large constructs, and it has unlimited uses from insertion of point mutations to incorporation of large fragments of DNA into BAC, PAC or cosmid vectors. The fact that a 55 kb fragment from genomic DNA can be subcloned into a BAC suggests that high throughput manipulation of BAC libraries and cloning of large fragments of genomic DNA is possible with this straightforward methodology, as well as generation of complex transgenic, knockout and knock-in constructs.

AUTHOR'S CONTRIBUTIONS

A.R.-M. and S.L. created the idea, performed the experiments and drafted the manuscript. I.H. supervised the research. All authors read and approved the final manuscript.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors wish to thank Dr György Posfai, Institute of Biochemistry, Biological Research Center, Szeged, Hungary, for the kind gift of the pST98AS plasmid. This work was supported by the Academy of Finland; the Wellcome Trust; the Wenner-Gren Foundation; the European Society for Paediatric Endocrinology and

the European Commission. Funding to pay the Open Access publication charges for this article was provided by Academy of Finland.

Conflict of interest statement. None declared.

REFERENCES

- Shizuya,H., Birren,B., Kim,U.J., Mancino,V., Slepak,T., Tachiiri,Y. and Simon,M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA*, **89**, 8794–8797.
- Sternberg,N.L. (1992) Cloning high molecular weight DNA fragments by the bacteriophage P1 system. *Trends Genet.*, **8**, 11–16.
- Ioannou,P.A., Amemiya,C.T., Garnes,J., Kroisel,P.M., Shizuya,H., Chen,C., Batzer,M.A. and de Jong,P.J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.*, **6**, 84–89.
- Copeland,N.G., Jenkins,N.A. and Court,D.L. (2001) Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.*, **2**, 769–779.
- Muyrers,J.P., Zhang,Y. and Stewart,A.F. (2000) ET-cloning: think recombination first. *Genet. Eng. (NY)*, **22**, 77–98.
- Muyrers,J.P., Zhang,Y. and Stewart,A.F. (2001) Techniques: recombinogenic engineering – new options for cloning and manipulating DNA. *Trends Biochem. Sci.*, **26**, 325–331.
- Zhang,Y., Buchholz,F., Muyrers,J.P. and Stewart,A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.*, **20**, 123–128.
- Zhang,Y., Muyrers,J.P., Testa,G. and Stewart,A.F. (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18**, 1314–1317.
- Baudin,A., Ozier-Kalogeropoulos,O., Denouel,A., Lacroute,F. and Cullin,C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.
- Murphy,K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.*, **180**, 2063–2071.
- Murphy,K.C., Campellone,K.G. and Poteete,A.R. (2000) PCR-mediated gene replacement in *Escherichia coli*. *Gene*, **246**, 321–330.
- 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.
- Wong,Q.N., Ng,V.C., Lin,M.C., Kung,H.F., Chan,D. and Huang,J.D. (2005) Efficient and seamless DNA recombineering using a thymidylate synthase A selection system in *Escherichia coli*. *Nucleic Acids Res.*, **33**, e59.
- Zhang,X.M. and Huang,J.D. (2003) Combination of overlapping bacterial artificial chromosomes by a two-step recombinogenic engineering method. *Nucleic Acids Res.*, **31**, e81.
- Zhang,Y., Muyrers,J.P., Rientjes,J. and Stewart,A.F. (2003) Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol. Biol.*, **4**, 1.
- Posfai,G., Kolisnychenko,V., Bereczki,Z. and Blattner,F.R. (1999) Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res.*, **27**, 4409–4415.
- Jamsai,D., Orford,M., Nefedov,M., Fucharoen,S., Williamson,R. and Ioannou,P.A. (2003) Targeted modification of a human beta-globin locus BAC clone using GET Recombination and an I-SceI counterselection cassette. *Genomics*, **82**, 68–77.