

Site-specific recombination in *Schizosaccharomyces pombe* and systematic assembly of a 400kb transgene array in mammalian cells using the integrase of *Streptomyces* phage ϕ BT1

Zhengyao Xu, Nicholas C. O. Lee, Felix Dafhnis-Calas, Sunir Malla, Margaret C. M. Smith and William R. A. Brown*

Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

Received August 28, 2007; Revised November 10, 2007; Accepted December 3, 2007

ABSTRACT

We have established the integrase of the *Streptomyces* phage ϕ BT1 as a tool for eukaryotic genome manipulation. We show that the ϕ BT1 integrase promotes efficient reciprocal and conservative site-specific recombination in vertebrate cells and in *Schizosaccharomyces pombe*, thus establishing the utility of this protein for genome manipulation in a wide range of eukaryotes. We show that the ϕ BT1 integrase can be used in conjunction with *Cre* recombinase to promote the iterative integration of transgenic DNA. We describe five cycles of iterative integration of a candidate mouse centromeric sequence 80 kb in length into a human mini-chromosome within a human-Chinese hamster hybrid cell line. These results establish the generality of the iterative site-specific integration technique.

INTRODUCTION

Site-specific recombinases and occasionally telomeres have been used as tools for manipulating eukaryotic genome structure. *Cre* (1) and *Flp* (2) were the first site-specific recombinases to be used for this purpose. Both catalyse reversible reactions between identical sites of about 35 bp. These recombinases have been most widely used to promote deletion reactions, where the instability of one product at cell division renders the reaction effectively irreversible. Heterospecific sites (3) have been advocated as a way of rendering such reactions irreversible and have been successfully exploited to allow efficient integration reactions (4). *Cre* and *Flp* belong to the tyrosine recombinase family of site-specific recombinases. Other members of this family of site-specific recombinases such

as the integrase of phage λ promote unidirectional or irreversible reactions between sites differing in sequence, but do so only in the presence of additional proteins, limiting their utility for genome manipulation. A second mechanistically distinct family of site-specific recombinases is the serine recombinase family. This family includes the integrase of the *Streptomyces* phage ϕ C31, which promotes an irreversible site-specific recombination reaction between the related but distinct attachment sites of the phage and bacterial genomes: *attP* and *attB*. The discovery that this reaction can be promoted by the purified integrase protein in the absence of any additional proteins (5) has led to ϕ C31 integrase being used as a genome manipulation reagent (6,7). The ϕ C31 integrase is a member of the 'large serine recombinase sub-family' which includes many different members with a similar pattern of domain organization (8). Several of these proteins have been studied and shown to be similar to the ϕ C31 integrase in promoting unidirectional reactions between non-identical sites in the absence of additional proteins. These other proteins should also be useful as reagents for genome manipulation but only a few have been used as such (9).

Site-specific recombinases have been used to approach a range of different problems but they are most widely employed in the generation of tissue-specific conditional loss of function alleles in metazoan genomes. It is often only necessary to generate and study one such allele in isolation, but for the appropriate modelling of human cancers it is necessary to be able to generate independently loss and gain of function alleles at several different loci in the same animal. This type of complex manipulation requires a large palette of enzymes of different specificities. The 'large serine recombinase sub-family' potentially satisfies this requirement. A second related area in which site-specific recombinases have been used has been in the

*To whom correspondence should be addressed: Tel: +44 (0)115 823 0386; Fax: +44 (0)115 823 0338; Email: William.brown@nottingham.ac.uk
Present address:

Margaret C. M. Smith, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK

introduction of long segments of human genomic DNA into the genome of an experimental organism such as a mouse (10). Such humanization of an experimental organism may be useful for the study of human genetic disease. A major technical challenge confronting this approach is that of integrating long tracts of DNA. We have described a method termed 'Iterative site-specific integration' (ISSI), which allows this challenge to be overcome (11). ISSI combines the activity of a unidirectional site-specific integrase with that of a reversible recombinase to allow the serial integration of transgenic DNA and thus potentially allows the assembly of transgenic arrays of arbitrary size. We have previously used the ϕ C31 integrase in combination with *Cre* to iteratively integrate two inserts of 70 kb and 80 kb in size cloned in bacterial artificial chromosomes into a human mini-chromosome in a chicken human hybrid cell line.

Here, we investigate the potential of an additional member of the serine recombinase family; the integrase of the *Streptomyces* phage ϕ BT1 as a tool for genome manipulation. We show firstly that the ϕ BT1 integrase functions efficiently in vertebrate cells and in cells of the yeast *Schizosaccharomyces pombe*. The results obtained in fission yeast indicate that the ϕ BT1 integrase is likely to function in most eukaryotic cell types. We then show that the ϕ BT1 integrase can be used in conjunction with *Cre* recombinase to promote the iterative integration of transgenic DNA. Thus, we describe five cycles of iterative integration of a candidate mouse centromeric sequence 80 kb in length into a human mini-chromosome within a human-Chinese hamster hybrid cell line. This allowed us to assemble a transgenic tract of DNA 400 kb in length. The demonstration that ISSI can be implemented using different combinations of site-specific recombinases and in different cell lines establishes the versatility of the technique.

MATERIALS AND METHODS

Plasmid construction

Plasmids were constructed by standard techniques. The sequences of the plasmids and of the integrases used in this work can be obtained from <http://www.nottingham.ac.uk/genetics/staff/williambrown/reg.html>. The nuclear localization signal used to tag the ϕ BT1 and ϕ C31 integrases was derived from the large T antigen of SV40 virus and included the residues MPKKKRKV. The sequences of the ϕ BT1 $attP$ and $attB$ sites used were respectively: TTCGGGTGCTGGGT GTTGTCTCTGGACAGTG ATCCATGGGAAACTACTCAGCACCA and GCCG TCCTTGA CCAGGTTTTGACGAAAGTGATCCA GATGATCCAGCTCCACACCC. In each case, these att sites were engineered into plasmids using synthetic

DNA provided by Invitrogen. The plasmids were checked by restriction site mapping and in all cases by sequencing. The blasticidin resistance (BSR) (12) gene was a kind gift from Hiroshi Arakawa of the GSF, Munich. The hygromycin resistance gene used was present in the counterselectable hygromycin-thymidine kinase fusion (HyTk) (13). The CCAG promoter (14) was a kind gift of Ian Chambers (Edinburgh University). All plasmids, sequences and vectors are available from William Brown (WRAB) subject to a materials transfer agreement. Plasmid DNA was purified by alkali lysis and precipitation with polyethylene glycol(PEG). PACs were purified by alkali lysis and PEG precipitation (experiments shown in Figure 2). The CCAG ϕ BT1 integrase IRES zeo was linearized at the unique PstI site at position 3557 and the CCAG *Cre* IRES Neo at the unique NotI site at position 3762.

Cell culture

DT40 were maintained and electroporated (11) as described previously and as summarized in <http://pheasant.gsf.de/DEPARTMENT/dt40.html>. CHO cells were manipulated as described previously (15). The following refinements were made: blasticidin was used at 30 μ g/ml to select for DT40 transformants and at 6 μ g/ml to select for Chinese hamster ovary (CHO) transformants. Hygromycin was used at 2 mg/ml to select for DT40 transformants and 0.6 mg/ml to select for CHO transformants. Zeocin and neomycin were used at 1 mg/ml and 2 mg/ml respectively for DT40 and 0.25 and 0.45 mg/ml, respectively for CHO cells. The site-specific integration experiments in DT40 cells were carried out as follows: cells were electroporated using previously established conditions with the indicated amount of DNA and immediately after transfection were plated out in 96-well dishes. Selection was applied 18 h later. DT40 colonies were counted after 12–14 days. Efficiencies given in Table 1 are the total number of resistant colonies recovered divided by the total number of cells electroporated. We refer to this figure as the absolute efficiency of integration in the table headings and text. During the course of the ISSI experiments, 500 μ g P1 artificial chromosome (PAC) DNA was transfected into 3×10^7 CHO cells by electroporation with settings of 400V, 250 μ F and 25 Ω . Typically, between 40 and 60 colonies were recovered in each experiment.

PCR, FISH and filter hybridization analysis

Pulsed field gels and fluorescent in situ hybridization (FISH) were as described previously (16) except that Alexafluor dyes were used for the FISH. The conditions for the PCR of the $attL$ are as follows for 35 cycles: 94°C for 30 s, 55°C for 30 s and 72°C for 60 s. The $attR$ used

Table 1. Efficiencies of site-specific integration promoted by the ϕ BT1 integrase in DT40 cells

Cell line	Transfected DNA	Amount (μ g)	Number of cells transfected	% plated	Resistant clones	Efficiency
DTCCAGattP''BSD ϕ BT1-5' nls-1	attB''HyTk	100	5.0E+0.7	25	405	$3 \times 10E-5$
DTCCAGattP''BSD ϕ BT1-5' nls-2	attB''HyTk	100	1.5E+0.7	25	476	$12 \times 10E-5$
DTCCAGattP''BSD ϕ BT1-5' nls-2	attB''HyTk	100	5.0E+0.7	25	362	$2.9 \times 10E-5$

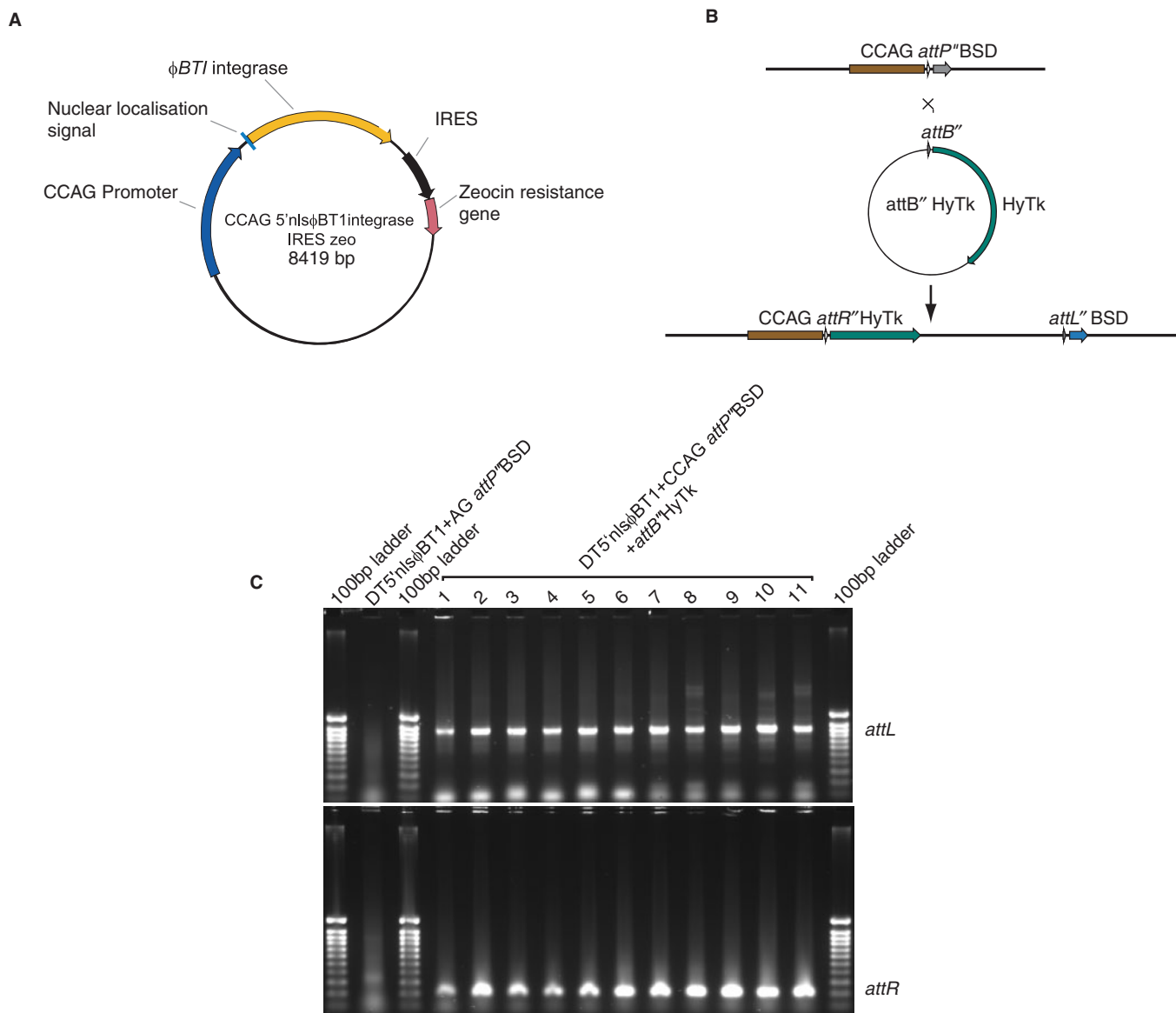


Figure 1. ϕ BT1 integrase-mediated site-specific integration into vertebrate chromosomes. (A) CCAG ϕ BT1 integrase IRES zeo is the ϕ BT1 integrase expression plasmid. (B) Representation of the promoter trapping strategy used to measure ϕ BT1 integrase mediated site-specific integration into chicken chromosomes. The linearized expression plasmid CCAG ϕ BT1 integrase IRES zeo was introduced into DT40 cells by electroporation, the plasmid CCAG *attP''* BSD was then randomly integrated into DT40 cells and site-specific recombination measured by the number of hygromycin resistant cells recovered following transfection with the promoter-less plasmid *attB''* HyTk. (C) ϕ BT1 integrase mediates conservative and reciprocal site-specific integration into chicken chromosomes. PCR across the indicated junctions was used to confirm the site specific nature of the integration reaction described in (B). Details of the primers and predicted product sizes are given in Table 1 of the Supplementary data.

similar cycle times but an annealing temperature of 62°C. The buffer used was a 16 mM (NH₄)₂SO₄, 67 mM Tris-HCl, pH 8.8, 0.1% Tween-20, 1 mM MgCl₂ and the Taq polymerase was from Biorline. The primer sequences and locations are given in Table 1 of the Supplementary data. The primers were used at 200 nM final concentration except for the BSDNotR, which was used at 400 nM. Markers used in the PCR in Figure 1C were the 100 bp ladder from New England Biolabs.

Yeast methods

The genotypes of the key strains are given in the Results section. Media and culture were as described on the

comprehensive site of Professor Susan Forsburg available at <http://www-rcf.usc.edu/~forsburg/index.html> except for the following specific departures. *S. pombe* were transformed using the lithium acetate procedure as described in the paper of Bahler and colleagues (17). DNA was extracted from 10 ml of saturated culture using DNazol. Briefly, cells were spheroplasted using lyticase, concentrated by centrifugation and then lysed with 1 ml of DNazol. After incubation and mixing in 1 ml DNazol, insoluble material was cleared from the extract by centrifugation at 900 g for 10 min. Nucleic acids were precipitated from the supernatant with 1 ml of ethanol and dissolved in 0.5 ml of TE containing RNAase A at

200 µg/ml and SDS at 0.1%. The solution was incubated at 37°C for 45 min and then pronase was added to 50 µg/ml. The sample was then incubated for a further 30 min at 55°C, phenol chloroform extracted two or three times until the interface was free of precipitate and then precipitated with ethanol. The pellet was dissolved in 200 µl TE and stored at 4°C. For western blotting, we used a protocol kindly described to us by Stephen Kearsley of Oxford University. Fifty microlitre of cells were grown to an OD of 0.2 in minimal medium, concentrated by centrifugation, rinsed in 0.9M sorbitol and resuspended in 100 µl of 20% trichloroacetic acid (TCA) and transferred to a 1.5 ml microcentrifuge tube. A quantity of 0.5 mm glass beads was added to produce a total volume of 0.5 ml. The tube was then agitated for five 1 min periods using a Scientific Instruments Disruptor Genie. A total of 900 µl of 5% TCA was added, the beads were allowed to settle and the protein precipitate was collected from the supernatant by centrifugation at 900g for 10 min. The pellet was dissolved in 250 µl of SDS-PAGE loading buffer. Twenty-five microlitre of this material corresponding to about 50 µg of protein was loaded onto a 10% standard SDS polyacrylamide gel and analysed by western blotting using a rabbit anti integrase first antibody and a peroxidase conjugated donkey anti-rabbit IgG second antibody (Amersham), the binding of which was detected using the ECL kit from Amersham Pharmacia.

RESULTS

ϕBT1 integrase functions in vertebrate cells

First of all, we wanted to establish the ϕBT1 integrase functions in vertebrate cells. We established a chicken DT40 cell line expressing the ϕBT1 integrase using the expression plasmid CCAG ϕBT1 integrase IRES Zeo (Figure 1A). We then integrated a linearized plasmid in which a ϕBT1 integrase *attP* (referred to as *attP'*) site is placed between the CCAG promoter and the coding region of a gene conferring resistance to blasticidin (CCAG *attP'* BSD; Figure 1B) and selected for resistance to blasticidin. These cells were then transfected with a circular plasmid in which a ϕBT1 integrase *attB* (*attB'*) site abutted a gene conferring resistance to hygromycin (*attB'* HyTk; Figure 1B), but which lacked a promoter. Site-specific recombination between the stably integrated *attP'* site and the *attB'* would place the hygromycin resistance gene under the control of the CCAG promoter and, thus confer, hygromycin resistance upon the transfected cells (Figure 1B). Following transfection, hygromycin resistant cells were recovered efficiently (Table 1). Eleven hygromycin resistant clones were picked and site-specific recombination was confirmed by PCR (Figure 1C). Sequencing the products as uncloned PCR products from two such cell lines confirmed that the reaction was reciprocal and conservative (data not shown). These results thus demonstrate that the ϕBT1 integrase promotes efficient, reciprocal, conservative site-specific recombination in vertebrate cells. The efficiency of recombination seen with the ϕBT1 integrase in this type of experiment was comparable to what we had detected in

analogous experiments carried out using the ϕC31 integrase (11).

ϕBT1 integrase functions in fission yeast cells

We wished to establish the general utility of the ϕBT1 integrase in a variety of eukaryotic cells. We therefore tested the function in the model organism *S. pombe*. Initially we tried a promoter trapping approach using single recombining sites that was analogous to that used in the experiments described above but, despite many attempts, we could detect no site-specific recombination activity. We therefore used the target replacement approach described in Figure 2A and which had been previously used by Thomason and colleagues (18) in their work on the ϕC31 integrase. We needed a positive control and so we chose to compare the activities of the ϕBT1 integrase and the ϕC31 integrase. We targeted a counter selectable marker gene; URA4; flanked by *attB* sites for either of the two integrases to the *leu1*⁺ locus into strain NL16 (*h*⁺, *ade6-M216*, *ura4-D18*, *leu1*⁺, *arg3-D4*). This sequence served as a target for the site-specific recombinase-mediated integration reaction (Figure 2A). We then introduced a plasmid in which the integrase expression was driven either by the strong REP1 or weak REP81 promoter (Figure 2B) into the cells harbouring the target site. We (11) and others (19) have shown that integrase activity in eukaryotic cells is enhanced by the presence of a nuclear localization signal and thus in both cases we used derivatives that had been tagged at the amino termini with the nuclear localization signals of SV40 large T antigen. Cells containing integrase were grown under leucine selection for the expression plasmid, transformed with a circular plasmid (p *attP* Arg3 *attP*) (Figure 2A) containing an Arg 3 gene flanked by the appropriate *attP* sites and selection applied for arginine prototrophy. Selection on the Leu marker for the expression plasmid was relaxed at the time of transformation and the ability of extant integrase to promote integration of the Arg3 gene into the genome was measured. We recovered the arginine prototrophs and tested whether they had lost the URA4 gene on the basis of resistance to fluororotic acid (FOA). The efficiency of the integration reaction was estimated by transformation with a plasmid pARSV40, a plasmid that contains an *S. pombe* autonomously replicating sequence (ARS) sequence as well as the ARG3 gene and is capable of extra-chromosomal replication. The data in Figure 2C and D are normalized with respect to the number of colonies obtained in the same experiment upon transformation with this plasmid. In order to estimate the extent of non-specific integration of the *pattP* Arg3 *attP* plasmid, we also established derivatives of each *leu1::attB-ura4-attB* strain containing an empty REP1 vector and this was also transformed with the appropriate *pattP* Arg3 *attP* plasmid.

The results in Figure 2C and D are presented as the mean of two independent experiments carried out for each integrase. The proportion of the transformants that are ARG⁺URA⁺ or ARG⁺URA⁻ are presented in dark tone and light tones, respectively in Figure 2B and C. The numbers of transformants recovered in the individual experiments are presented in Supplementary data Table 2.

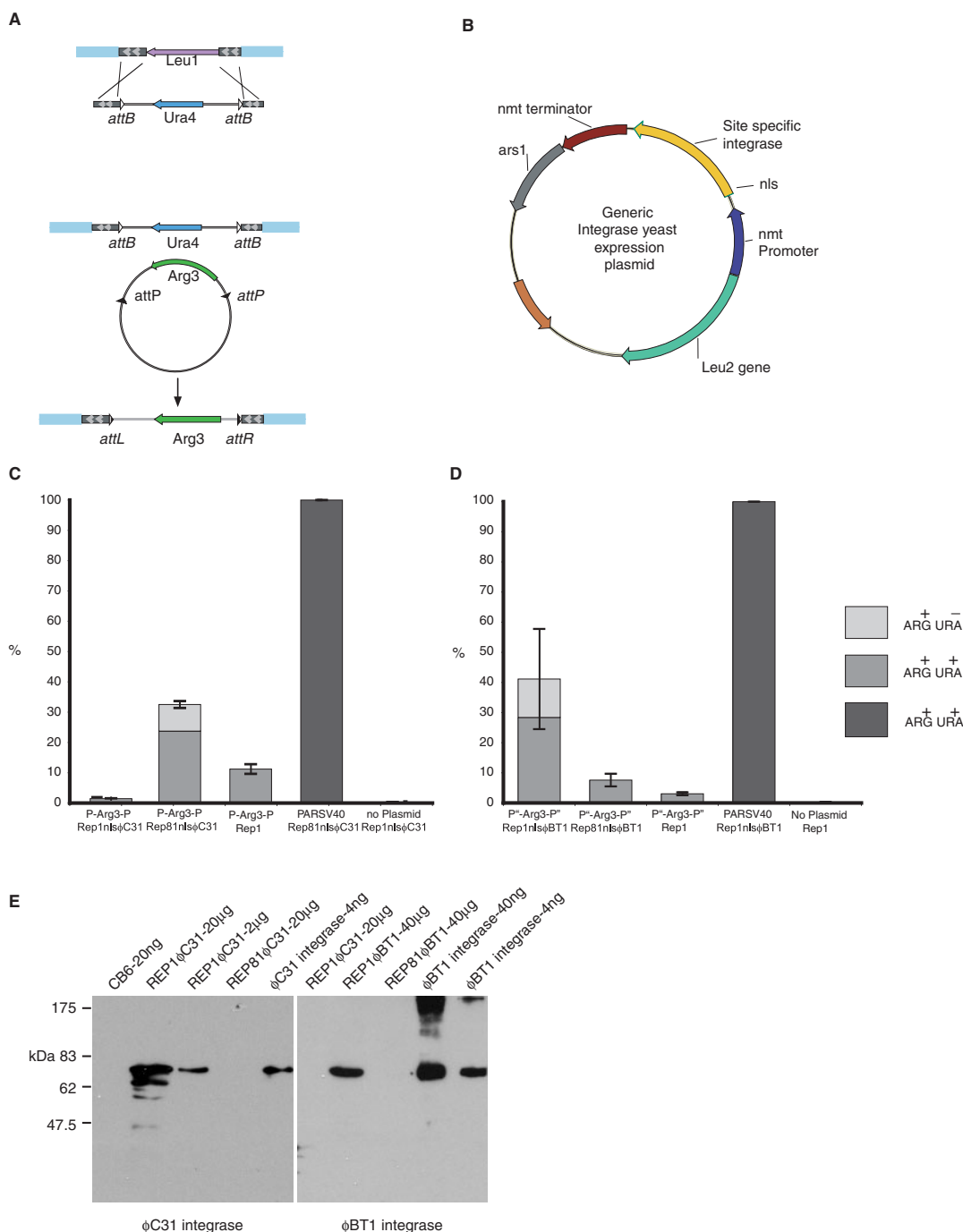


Figure 2. ϕ BT1 and ϕ C31 integrase activities in *S. pombe*. (A) Schematic representation of the strategy used to detect site-specific integration promoted by the serine recombinases in *S. pombe*. The *leu1* gene was replaced with a cassette encoding the counter selectable *ura4* gene flanked by *attB* sites. Site-specific recombination in the phase indicated between both of these *attB* sites and the *attP* sites flanking the *Arg3* gene on an incoming plasmid leads to the replacement of the *ura4* gene and the insertion of the *Arg3* gene. Several other outcomes arising as a result of a single recombination event on one sister chromatid followed by site-specific recombination with the other sister are discussed in the text and described in the supplementary data. (B) The generic integrase expression plasmid used in the experiments are carried out in *S. pombe*. The gene encoding the integrase tagged with a nuclear localization signal was cloned into either the REP1 or REP81 expression plasmids to produce the indicated construct. (C) Site-specific recombination mediated by ϕ C31 integrase at *attB ura4 attB*. The indicated integrase expression plasmid was transformed into the *attB ura4 attB* containing strain and then a plasmid pBS *attP Arg3 attP* introduced, arginine prototrophs recovered and tested for loss of the *ura4* gene by growth on FOA. The ARS containing plasmid PARSV40 was used as a control for the transformation efficiency in the two reactions. (D) Site-specific recombination mediated by ϕ BT1 integrase at *attB' ura4 attB'*. The indicated integrase expression plasmid was transformed into the *attB' ura4 attB'* containing strain and then a plasmid pBS *attP' Arg3 attP'* introduced, arginine prototrophs recovered and tested for loss of the *ura4* gene by growth on FOA. The ARS containing plasmid PARSV40 was used as a control for the transformation efficiency in the two reactions. (E) Western blotting of ϕ C31 and ϕ BT1 integrase expression in *S. pombe*. Purified integrase proteins or proteins from the indicated strains that had been extracted by disruption with glass beads in TCA, size fractionated, transferred and analysed by western blotting. CB6 is a control strain lacking the integrase expression plasmid.

As mentioned earlier, we identify the ϕ BT1 integrase attachment sites by the use of the superscript " in both Figure 2 and Supplementary data Table 2. The results in Figure 2 suggest that both integrases promote site-specific integration of the respective *pattP* Arg3 *attP* into the target locus. Several aspects of the data bear comment. Firstly, the ϕ C31 integrase activity was undetectable when expressed from the REP1 promoter but activity could be detected when expression was driven by the weaker REP81 promoter. We noted that the REP1 ϕ C31 integrase containing colonies were very small and it was often difficult to transform the yeast with this plasmid. We confirmed the identity of the plasmid by sequencing and thus excluded trivial explanations for the result. We therefore examined the transformants for the presence of the integrase by western blotting (Figure 2E). This experiment confirmed that the REP1 expression plasmid was functional. The inability to detect the integrase in the REP81 transformants by western blotting was consistent with what the relative activities of the REP series of expression plasmids <http://www-rcf.usc.edu/~forsburg/index.html>. The fact that the REP1 ϕ C31 integrase containing cells contain several orders of magnitude, more integrase yet do not permit efficient plasmid integration indicates that the integrase is inactive when present at high concentrations in *S. pombe*. In contrast, integration activity of the ϕ BT1 integrase could only be detected using a REP1 expression vector. Comparison of the relative amounts of the two integrases in the respective yeast strains indicated that the likely cause of this difference was that the ϕ C31 integrase was more efficiently expressed than the ϕ BT1 integrase by a factor of about 10-fold (Figure 2E).

Although ARG⁺URA⁻ transformants were detected following transformation with either the REP81 ϕ C31 integrase or the REP1 ϕ BT1 integrase expression constructs, about 70% in each case were ARG⁺URA⁺. Given the background of ARG⁺URA⁺ clones seen with the empty vector plasmid, it was important to determine the proportion of the ARG⁺URA⁺ clones that arose by site-specific recombination and to understand their origin. We checked the identity and determined the structure of the transformants by using PCR, using primers adjacent to the ends of either the incoming or target attachment sites (Supplementary data, Tables 3 and 4 illustrates the relationship between the structure of the integrant and the predicted content of attachment sites). We confirmed the accuracy of the PCR analysis by restriction enzyme digestion with an enzyme that cuts in neither the ARG3 or URA4 loci, gel electrophoresis, filter transfer and hybridization with a probe corresponding to the deleted segment of the URA4 (Δ 18) locus in the original NL16 strain (Supplementary data; Figure 1). With the exception of the single most complicated of the URA⁺ARG⁺ clones the sizes of the cognate fragments were consistent with the structure predicted on the basis of the PCR results. In the case of both the enzymes, > 90% of the ARG⁺URA⁺ clones arose by site-specific integration at the target *leu1::attB-ura4-attB* locus. In 21 cases out of 29, it was necessary to invoke recombination reactions between sister chromatids at G2 during the cell cycle.

Two clones had structures that could be explained if homologous recombination was also involved. A final clone was likely to be random integration events. Hypothetical explanations for the origin of these structures are set out in Figure 2 of the Supplementary data.

In previous experiments using the ϕ C31 integrase in vertebrate cells we had identified damage to participating sites as a cause of incomplete reaction (20) and so we checked the sequences of both the substrate (*attB* and *attP*) and (*attL* and *attR*) product sites (10 clones) for both the ϕ C31 and ϕ BT1 integrase. In all cases, these sites were intact and undamaged. Similarly, we checked 30 ARG⁺URA⁻ clones by PCR and sequenced six, which confirmed that they were site-specific replacements of the *ura4*⁺ gene with the *arg3*⁺ gene. In summary therefore our data show that in all of the transformants the reactions had gone to completion and we found no evidence of incomplete reaction as a result of damage to the substrate sites.

The results of these experiments demonstrate that the ϕ BT1 integrase promotes site-specific integration in fission yeast cells with efficiencies that in practical terms are similar to the ϕ C31 integrase and suggest that the ϕ BT1 integrase will be generally useful. The fact that the ϕ BT1 integrase is expressed less efficiently than the ϕ C31 integrase will be discussed.

Iterative site specific recombination (ISSI) in Chinese hamster cells using a combination of the ϕ BT1 integrase and *Cre* recombinase

ISSI (11) is a method that allows the serial assembly of long transgene arrays as a result of the combined action of a unidirectional recombinase and a reversible recombinase such as *Cre*. ISSI has previously been implemented only in chicken DT40 cells using *Cre* and the ϕ C31 integrase. If ISSI is to be widely adapted as a method then it is important that it be shown to be practical in other cell types and to know how effectively it works when used with different combinations of recombinase. Therefore, we set out to establish the ISSI methodology using *Cre* and the ϕ BT1 integrase in Chinese hamster cells. In order to do this, we used a mini-chromosome derivative of the human Y chromosome termed XP4 that had previously been engineered by sequence targeting and telomere directed chromosome breakage in DT40 cells (20). This mini-chromosome includes a CCAG *loxP* HyTK *attP'* cassette integrated by homologous recombination into the DYZ5 array 700 kb distant from the centromeric array of alphoid DNA (Figure 3A). This cassette acts as a substrate for the subsequent integration reactions that constitute the ISSI process.

We used microcell transfer to move this chromosome into the CHO cells, checked the structure of the chromosome by pulsed field gel electrophoresis and confirmed its linear integrity by fluorescent *in situ* hybridization (data not shown). We then introduced the CCAG ϕ BT1 integrase IRES Zeo (Figure 1A) and CCAG *Cre* IRES Neo (Figure 3B) expression plasmids into the hybrid cell line, in order to establish a cell line expressing

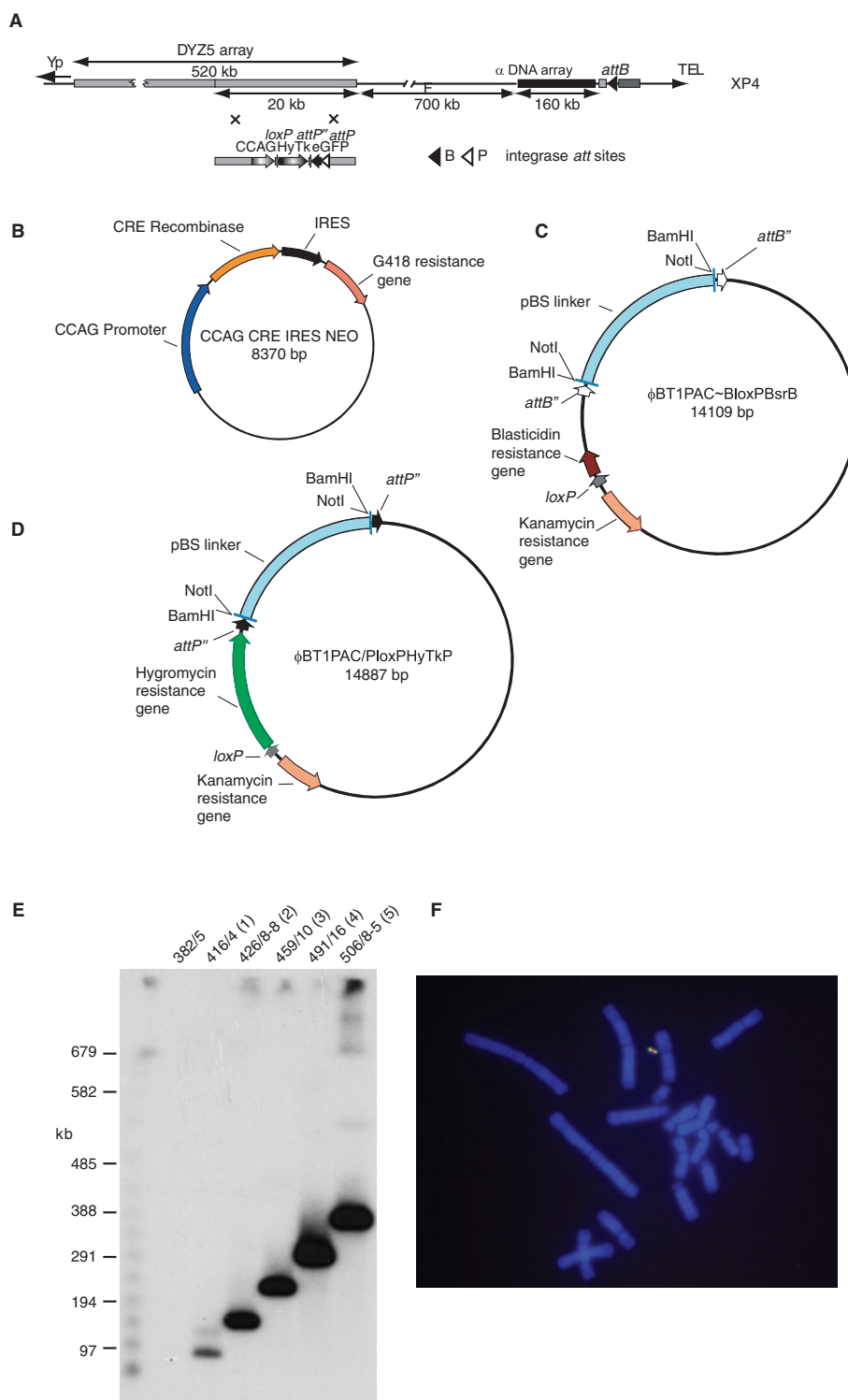


Figure 3. ϕ BT1 integrase and *Cre* recombinase-mediated ISSI into vertebrate chromosomes. (A) A diagram of the centromere adjacent DNA of the mini-chromosome XP4 used in the ISSI reaction sequence is shown in (D). The mini-chromosome was engineered by sequence targeting in the chicken cell line DT40 and then moved into Chinese hamster ovary cells by microcell transfer. (B) CCAG *CRE* IRES Neo is the *Cre* expression plasmid used to establish the cell line used in the ISSI experiments shown in this figure. (C) ϕ BT1PAC~*attB* BloxPBsr *att B* is one of the two large fragment cloning vector used in the ISSI experiments shown in this figure. (D) ϕ BT1PAC~*attP* loxPHYTk *attP* is one of the two the large fragment cloning vector used in the ISSI experiments shown in this figure. (E) ISSI mediated by the ϕ BT1 integrase and *Cre* recombinase in CHO cells. Cells containing the mini-chromosome XP4 were transfected with the ϕ BT1 integrase and *Cre* expression plasmids shown in Figure 1 and then serially transfected with PAC vectors; ϕ BT1PAC~*attB* BloxPBsr *att B* and ϕ BT1PAC~*attP* loxPHYTk *attP* containing 80 kb inserts of a tandemly repeated DNA sequence designed to resemble the centromeric mouse-minor satellite DNA. (F) Cytogenetic characterization of the mini-chromosome 459/10 from the ISSI sequence shown in (D). The mouse-minor satellite DNA sequence and the human Y aliphoid DNA have been detected by fluorescence *in situ* hybridization and are shown in red and green, respectively.

both site-specific recombinases necessary for the ISSI reaction.

We wished to use ISSI to assemble a candidate centromeric sequence on the DYZ5 array of XP4, which we could subsequently test for functionality by excising the pre-existing centromere using the ϕ C31 integrase as described earlier. We therefore, designed a sequence that was based upon the mouse-minor satellite DNA. A tandem array of this sequence was assembled by standard methods of concatamerization using ligation of BamHI and BglII cleaved DNA. The final product of 80 kb in size was cloned into each of two PAC vectors that can participate in the ISSI reaction (Figure 3C and D). One PAC vector; ϕ BT1 PAC *attB'* *attB'* *loxP* BSR is used for the first, third and subsequent odd numbered steps of the ISSI reaction, the other ϕ BT1 PAC *attP'* *attP'* *loxP* HyTk is used for the even numbered steps. We carried out five steps of the ISSI reaction resulting in the assembly of a 400 kb array of the test DNA on the mini-chromosome. At each step, we typically recovered many tens of transformants. We characterized these using both molecular (Figure 3E) and cytogenetic approaches (Figure 3F) in order to select as substrates for the next step in the ISSI sequence cell lines containing a single un-rearranged chromosome that includes an integrated array of the predicted size. Typically, each step gave between three and five clones with such a structure. The integrated DNA in the remainder was either deleted or empty as has been observed and explained previously (11). The success of this sequence thus establishes ISSI as a generally useful approach to the assembly of transgenic DNA arrays and also demonstrates the utility of the ϕ BT1 integrase.

DISCUSSION

The results contained in this article establish a new serine recombinase; ϕ BT1 integrase, as a tool for genome manipulation. We have shown that this novel recombinase functions efficiently in three eukaryotic cell types: *S. pombe*, chicken DT40 cells and Chinese hamster ovary cells. We have now shown that ISSI, an iterative technique for the assembly of long tracts of transgenic DNA, works with two large serine integrases in two different vertebrate cell lines; DT40 cells and Chinese hamster ovary cells. This indicates that ISSI is a robust technique effective in cells other than the hyper-recombinogenic DT 40 chicken cell line where it was originally established. One practical merit of the characterization of a second unidirectional recombinase is that when combined with the ϕ C31 integrase, it should allow one to confidently attempt the substitution of long tracts of genomic DNA with new sequences. Thus, one could use one unidirectional recombinase in combination with *Cre* in a series of ISSI reactions to assemble a sequence at a locus of choice and then use the second enzyme to delete the corresponding stretch of genomic DNA. These results in conjunction with those of others, therefore, suggest that the large serine recombinases will turn out to be useful tools for manipulating precisely the long range structure of eukaryotic genomes. It will, however, be important to test the ISSI technique in mouse embryonal stem cells (ES)

cells and to determine whether the manipulated cells will colonize the germ line. *Cre* recombinase (21) and possibly ϕ C31 integrase (22) have been shown to damage DNA when expressed in vertebrate somatic cells. Therefore, it will be important to establish conditions in which *Cre* recombinase and a serine integrase can be used to drive an ISSI reaction and retain germ line competence in ES cells.

In general, the range of tools available in yeast is such that it may be thought that the ability to use site-specific recombinases in these cells is of no practical worth. The *S. pombe* centromeres (23) are large and complex and currently impossible to manipulate *in vivo*. Thus, despite many years of work the sequence requirements for the efficient function of the CENH3 (CENP-A) binding central core are not defined. Our results raise the possibility that we will be able to manipulate the central core sequence organization using these site-specific recombinases and thereby establish tools to analyse the relationship between kinetochores and heterochromatin assembly. The demonstration that the ϕ BT1 integrase is less efficiently expressed than the ϕ C31 integrase in fission yeast may be a consequence of the differences in codon usage between *Streptomyces* and *S. pombe*. Recent experiments which studied the expression of ϕ C31 integrase in mouse embryos (24) have suggested the importance of using codon optimized ϕ C31 integrase expression constructs. It would seem on this basis desirable for work in whole animals and probably yeast to use codon optimized ϕ BT1 integrase integrase. In preliminary work, one of us (N.C.O.L.) have compared the activities in *S. pombe* of the native and ϕ C31 integrase proteins codon optimized for expression in mammalian cells. These experiments have shown that the proteins function with similar activity, but the yeast containing the codon optimized protein proliferate more quickly than those containing the native protein and can be transformed much more efficiently with DNA. We therefore conclude that it will be preferable to use the ϕ C31 integrase protein that is codon optimized for expression in mammalian cells than the native protein even when working in fission yeast.

In previous work (20) we demonstrated that the ϕ C31 integrase reaction may be interrupted in vertebrate cells by damage to the substrate sites, probably as a result of the activities of host DNA repair activities. With this observation in mind, we examined the products of integration reactions promoted by both the ϕ C31 and ϕ BT1 integrases in *S. pombe* and have failed to detect any evidence that reaction intermediates are liable to such damage in these cells. It would seem unwarranted, however, to draw the conclusions that site damage may not limit the efficiencies of these enzymes in *S. pombe* or that the ϕ BT1 integrase reaction may not be liable to being interrupted by site damage in vertebrate cells. In the experiments of Malla and colleagues (20) the reaction that was being analysed was occurring between ϕ C31 sites that were \sim 1Mb apart and thus it may have been particularly slow due to the difficulties of maintaining the sites in alignment. A slow reaction may increase the likelihood that the double strand breaks in the reaction intermediates are recognized by the host cell DNA repair activities.

Thus, further work will be required to investigate the efficiency of the ϕ BT1 integrase and other serine recombinases for promoting long range re-arrangements and to determine how liable the reaction intermediates are to damage arising from the activities of host DNA repair activity in different cell types.

In summary, the practical conclusion from our work is that we have established a new serine recombinase as a tool for genome manipulation. The availability of two well characterized unidirectional recombinases creates the possibility of creating long range substitutions in eukaryotic genomes using genome manipulation techniques.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Funding to pay the Open Access publication charges for this article was provided by the BBSRC under the Exploiting Genomics Initiative and the EU (Mechanisms of Gene Integration).

Conflict of interest statement. None declared.

REFERENCES

- Sauer, B. (1987) Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **7**, 2087–2096.
- Golic, K.G. and Lindquist, S. (1989) The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell*, **59**, 499–509.
- Hoess, R.H., Wierzbicki, A. and Abremski, K. (1986) The role of the loxP spacer region in P1 site-specific recombination. *Nucleic Acids Res.*, **14**, 2287–2300.
- Soukharev, S., Miller, J.L. and Sauer, B. (1999) Segmental genomic replacement in embryonic stem cells by double lox targeting. *Nucleic Acids Res.*, **27**, e21.
- Thorpe, H.M. and Smith, M.C.M. (1998) In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc. Natl Acad. Sci. USA*, **95**, 5505–5510.
- Thyagarajan, B. *et al.* (2001) Site-specific genomic integration in mammalian cells mediated by phage fC31 integrase. *Mol. Cell Biol.*, **21**, 3926–3934.
- Belteki, G. *et al.* (2003) Site-specific cassette exchange and germline transmission with mouse ES cells expressing fC31 integrase. *Nat. Biotechnol.*, **21**, 321–324.
- Smith, M.C. and Thorpe, H.M. (2002) Diversity in the serine recombinases. *Mol. Microbiol.*, **44**, 299–307.
- Nkrumah, L.J. *et al.* (2006) Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods.*, **3**, 615–621.
- Wallace, H.A. *et al.* (2007) Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell*, **128**, 197–209.
- Dafhnis-Calas, F. *et al.* (2005) Iterative in vivo assembly of large and complex transgenes by combining the activities of fC31 integrase and Cre recombinase. *Nucleic Acids Res.*, **33**, e189.
- Izumi, M. *et al.* (1991) Blasticidin S-resistance gene (bsr): a novel selectable marker for mammalian cells. *Exp. Cell Res.*, **197**, 229–233.
- Lupton, S.D. *et al.* (1991) Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. *Mol. Cell Biol.*, **11**, 3374–3378.
- Niwa, H., Yamamura, K. and Miyazaki, J. (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, **108**, 193–201.
- Brown, K.E. *et al.* (1994) Dissecting the centromere of the human Y chromosome with cloned telomeric DNA. *Hum. Mol. Genet.*, **3**, 1227–1237.
- Yang, J.W. *et al.* (2000) Human mini-chromosomes with minimal centromeres. *Hum. Mol. Genet.*, **9**, 1891–1902.
- Bahler, J. *et al.* (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast*, **14**, 943–951.
- Thomason, L.C., Calendar, R. and Ow, D.W. (2001) Gene insertion and replacement in *Schizosaccharomyces pombe* mediated by the *Streptomyces* bacteriophage fC31 site-specific recombination system. *Mol. Genet. Genomics*, **265**, 1031–1038.
- Andreas, S. *et al.* (2002) Enhanced efficiency through nuclear localization signal fusion on phage fC31-integrase: activity comparison with Cre and FLP recombinase in mammalian cells. *Nucleic Acids Res.*, **30**, 2299–2306.
- Malla, S. *et al.* (2005) Rearranging the centromere of the human Y chromosome with fC31 integrase. *Nucleic Acids Res.*, **33**, 6101–6113.
- Loonstra, A. *et al.* (2001) Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc. Natl Acad. Sci. USA*, **98**, 9209–9214.
- Ehrhardt, A. *et al.* (2006) Molecular analysis of chromosomal rearrangements in mammalian cells after fC31-mediated integration. *Hum. Gene Ther.*, **17**, 1077–1094.
- Wood, V. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature*, **415**, 871–880.
- Raymond, C.S. and Soriano, P. (2007) High-efficiency FLP and fC31 site-specific recombination in mammalian cells. *PLoS ONE*, **2**, e162.