Nucleic Acids Research, 2005, Vol. 33, No. 22 ***e189***
doi:10.1093/nar/gni192

Iterative *in vivo* assembly of large and complex transgenes by combining the activities of ϕC31 integrase and Cre recombinase

Felix Dafhnis-Calas, Zhengyao Xu, Steve Haines, Sunir K. 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 September 21, 2005; Revised October 17, 2005; Accepted November 25, 2005

ABSTRACT

We have used the ϕC31 integrase to introduce large DNA sequences into a vertebrate genome and measure the efficiency of integration of intact DNA as a function of insert size. Inserts of 110 kb and 140 kb in length may be integrated with about 25% and 10% efficiency respectively. In order to overcome the problems of constructing transgenes longer than ~150 kb we have established a method that we call; 'Iterative Site Specific Integration' (ISSI). ISSI combines the activities of ϕC31 integrase and Cre recombinase to enable the iterative and serial integration of transgenic DNA sequences. In principle the procedure may be repeated an arbitrary number of times and thereby allow the integration of tracts of DNA many hundreds of kilobase pairs long. In practice it may be limited by the time needed to check the accuracy of integration at each step of the procedure. We describe two ISSI experiments, in one of which we have constructed a complex array of vertebrate centromeric sequences of 150 kb in size. The principle that underlies ISSI is applicable to transgenesis in all organisms. ISSI may thus facilitate the reconstitution of biosynthetic pathways encoded by many different genes in transgenic plants, the assembly of large vertebrate loci as transgenes and the synthesis of complete genomes in bacteria.

INTRODUCTION

Site-specific recombinases are used to manipulate chromosomal DNA sequence organization *in vivo*, *in vitro* (1) and in a range of different experimental systems. The enzymes Cre and Flp (2) are the most widely employed, both are members of the tyrosine recombinase family and catalyse reversible reactions between identical sites of ~35 bp in the absence of any accessory proteins. The reversibility of the reactions catalysed by these enzymes has limited their applicability. They are generally used to promote deletion reactions since integration reactions are kinetically unfavourable. However Cre and Flp have also been used to engineer integrations and translocations in cells in culture where it has been possible to provide the enzyme transiently and trap the potentially unstable product, often but not always, by selection (3,4).

A technology that would allow the construction of large segments of transgenic DNA would be of widespread utility. Thus it would be valuable to be able to assemble large tracts of transgenic DNA at defined chromosomal loci in order to construct animal models of human genetic disease and to study the factors controlling expression of large genes. This goal has become more attractive with the ability to precisely and efficiently manipulate DNA cloned in *Escherichia coli* by homologous recombination (5,6) and with the availability of a complete genome sequence and well-characterized libraries of bacterial artificial chromosome (BAC) clones. Similarly it would be useful to be able to build plants with multiple transgenes at a single locus. Finally it would be interesting to be able to construct synthetic bacterial genomes in order to identify the minimal set of genes consistent with independent life (7). Strategies describing how Cre or Flp might be used to promote irreversible integrations have been presented (8–10) However all of these approaches suffer from the limitation that they only allow one cycle of site-specific recombination and thus cannot be used iteratively. It is difficult therefore to see how they could be used to construct large tracts of transgenic DNA.

*To whom correspondence should be addressed. Tel: +44 115 849 3244; Fax: +44 115 970 9906; Email: William.brown@nottingham.ac.uk
Present address:
Margaret C. M. Smith, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK

Here we investigate the use of an integrase from the *Streptomyces*' phage ϕC31 (11); this integrase belongs to the serine recombinases family of proteins and is thus releated to the resolvase/invertases. ϕC31 integrase is one of an expanding group of so-called 'large serine recombinases' because of their much higher molecular weight compared with the resolvase/invertases. The serine recombinases differ evolutionarily and mechanistically from the tyrosine recombinases such as Cre and Flp. Several phage integrases of the large serine recombinases have been shown to promote unidirectional or irreversible recombination between non-identical sites of ~50 bp *in vitro* in the absence of additional proteins (11,12). They would thus be potentially ideal for engineering precisely those types of re-arrangement that cannot be promoted by Cre and Flp. Several such phage integrases have been shown to be active in vertebrate cells (13). However these reports did not include unambiguous measurements of the efficiency of the reactions promoted by these proteins but emphasized the suggestion that exchange reactions could occur ectopically between one or other of the specific sites and several genomic sites (13). Such ectopic recombination would, of course, compromise the utility of these enzymes for genome engineering but, it was argued, enhance their utility as reagents for gene therapy.

Here we measure the efficiency of the class V serine recombinase: ϕC31 integrase (11), in chicken DT40 cells. We show that this enzyme can be used to promote the efficient integration of plasmids and fragments as large as 100 kb into vertebrate chromosomes. Sequences larger than 100 kb are more frequently deleted than integrated intact. In order to overcome this limitation we have established a method that enables the iterative integration of transgenic DNA sequences by combining the site-specific recombinase activities of the ϕC31 integrase and Cre recombinase. This method which we call 'iterative site-specific integration' (ISSI) potentially enables the reconstitution of transgenes of unlimited length and defined sequence organization. We describe proof of principle experiments in which we have constructed an array of plasmid DNA sequences and a complex array of vertebrate centromeric sequences of 150 kb in size. Although our experiments have been carried out in vertebrate cells the principle that underlies ISSI is simple and applicable to transgenesis in all organisms: plants and bacteria as well as animals. ISSI should thus enable the reconstruction of complete genomes in bacteria and of biosynthetic pathways encoded by many different genes in transgenic organisms. A merit of assembling transgene arrays at a single locus in whole organisms using a technique such as ISSI is that doing so would avoid the difficulties that would arise as a result of the segregation of independent transgenic loci during meiosis.

In our proof of principle experiments we have used site-specific recombination to introduce long, defined tracts of tandemly repeated DNA into hyper-recombinogenic cells derived from the DT40 cell line. Thus we have demonstrated that long tracts of repeated DNA sequences can be introduced into a cell line where such sequences might be expected to be particularly unstable. These experiments therefore suggest that ISSI may be useful in a wide variety of cell types and for a variety of different sequences. However some cell types may be less readily transfected with large DNA than DT40 cells or may process such sequences differently from DT40 cells. Serine recombinases have been shown to work in many cells types but there may be some where they do not and in these implementation of ISSI would be impractical.

MATERIALS AND METHODS

Plasmid construction

Plasmids were constructed by standard techniques including recombineering using the DY380 strain of Court and coworkers (14). The sequences of the plasmids and of the integrases used in this work can be obtained from http://www.nottingham.ac.uk/genetics/brown/genomeengineering.php. The Y chromosome alphoid (DYZ3) DNA BAC used in these experiments was 442H19 from the CITB Human BAC DNA (B and C libraries) Release 4. The original BAC had an insert of 150 kb of human Y alphoid DNA but was unstable in culture and derivatives of varying sizes could be isolated from colonies left on plates for several weeks. The BAC 442H19 was a kind gift of Jonathan Flint and colleagues of the Wellcome Trust Centre for Human Genetics, Oxford. The X chromosome alphoid (DXZ1) DNA was derived from a PAC subclone of BAC RPCI-11 242E23 kindly provided by Hunt Willard and colleagues. The nuclear localization signal used to tag the ϕC31 was derived from the large T antigen of SV40 virus and included the residues MPKKKRKV. The *attP* and *attB* used sites were as follows: gtagtgccccaactggggtaacctttgagttctctcagttgggggcgtag and ccgcggtgcgggtgccagggcgtgcccttgggctccccgggcgcgtactcc for the standard integration. The sequence of the modified *attB* site used in the ISSI constructs was accgcggtgcgggtgccagggtgtgcccttgggctccccagggcaccctccac. 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 across the *att* sites. The gene encoding resistance to apramycin was as described by Kuhstoss *et al.* (15). The blasticidin resistance (BSR) (16) gene was a kind gift from Hiroshi Arakawa of the GSF, Munich. The hygromycin resistance gene used was present in the counter selectable hygromycin–thymidine kinase fusion (HyTk) (17). The CCAG promoter (18) was a kind gift of Ian Chambers (Edinburgh University). Primers used in the PCR are given in Supplementary Table 1. All plasmids and vectors are available from WRAB subject to a materials transfer agreement. Plasmid DNA was purified by alkali lysis and precipitation with polyethylene glycol. BACs and PACs were purified initially by alkali lysis and density gradient centrifugation in mixtures of caesium chloride and ethidium bromide (experiments shown in Figure 4) and subsequently by alkali lysis and PEG precipitation (experiments shown in Figure 8).

Cell culture

DT40 cells and DT40 somatic cell hybrids were as described (19) and were maintained and electroporated as also described previously (19) except that the medium used was RPMI 1640 including 446 mg/l L-alanyl-L-glutamine with 10% foetal bovine serum, 1% chicken serum, 10^{-5} M 2-mercaptoethanol, 10 U/ml penicillin and 10 μg/ml streptomycin. This medium gave cleaner selection after addition of antibiotics than our earlier DMEM based medium. Targeting of the

Table 1. Efficiency of site-specific integration of plasmid DNA promoted by ϕC31 integrase alone and in conjunction with Cre recombinase (ISSI)

Transfection	Cell line	Transfected plasmid	Amount (μg)	Number of cells transfected	% plated out	Number of stably transfected colonies	Absolute efficiency
4.8.02	A9-CCAGattBHyk ϕc31 integrase	*attP*neo	50	5.7×10^7	100	0	$<2 \times 10^{-8}$
2.5.01	A9-CCAGattBHyk ϕc31-5′NLS1	*attP*neo	100	3.2×10^7	100	2912	9.1×10^{-5}
4.5.01	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	100	3.2×10^7	100	3920	12.2×10^{-5}
39	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	25	2.34×10^7	25	90	1.53×10^{-5}
40	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	250	2.34×10^7	25	379	6.47×10^{-5}
41	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	25	2.45×10^8	25	389	0.66×10^{-5}
42	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	25	1.24×107	25	31	0.96×10^{-5}
43	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	250	1.24×10^7	25	398	12.8×10^{-5}
44	A9-CCAGattBHyk ϕc31-5′NLS2	*attP*neo	25	1.24×10^8	25	315	1.01×10^{-5}
89	A9-CCAGattBHyk ϕc31-3′NLS2	*attP*neo	100	2×10^7	25	433	8.66×10^{-5}
98	A9-CCAGattBHyk ϕc31-3′NLS2	*attP*neo	25	3.8×10^7	25	152	1.6×10^{-5}
304	288-ISSI	BBloxPBSR	20	5×10^7	25	226	1.8×10^{-5}

Table 2. Assaying the efficiency of integration of BAC DNA promoted by the ϕC31 integrase alone and during ISSI into a mini-chromosome DT40 cells

Transfection number	Cell line	BAC	Amount μg	Number of cells transfected	% plated out	Resistant clones	Number of clones analysed	No intact
19	5′NLS2	33 kb α *attP*neoBAC	48	2.0×10^8	100	41	20	13
51	5′NLS2	70 kb α *attP*neoBAC	190	2.0×10^8	100	1078	39	20
72	5′NLS2	110 kb α *attP*neoBAC	520	2.75×10^8	100	158	39	7
53	5′NLS2	140 kb α *attP*neoBAC	88	2.0×10^8	100	288	39	4
419	ISSI 293	80 kbX α BBLoxPBSR BAC	500	5.7×10^8	100	>2000	20	4

CCAG HyTk constructs containing attachment or *loxP* sites to the mini-chromosomes was detected by digestion with SacI, hybridization with an alphoid DNA probe and the replacement of a 20 kb SacI fragment by one of 17 kb in size. The site-specific integration experiments were carried out as follows: cells were electroporated in the usual 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 and counts were corrected using the Poisson distribution for the small underestimate arising from the random distribution of transfected cells into a limited number of wells. The efficiencies given in Tables 1 and 2 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.

PCR and filter hybridization analysis

Conventional agarose and pulsed-field gels were as described previously (19). Filter hybridization was as described (19). The sequences of the primers used in the PCR to check the breakpoints in the ϕC31 integration reactions and in the ISSI reaction sequence (Figure 7) are given in the Supplementary Data (Table 1). The markers used in Figure 4C were the mid-range I PFG set from New England Biolabs. The sizes of the informative marker molecules are 15, 33, 48.5, 63.5, 82, 97, 112, 130.5, 145.5 and 160.5 kb. The conditions for the PCR in Figures 3 and 4 and Supplementary Figure 2 (*attL*) are as follows 35 cycles: 92°C for 20 s, 62°C for 20 s and 72°C for 30 s. The buffer used was a 16 mM $(NH_4)_2SO_4$, 67 mM Tris–HCl, pH 8.8, 0.1% Tween-20, 1 mM $MgCl_2$ and the *Taq* polymerase was from Bioline. The primer sequences used are given in Supplementary Table 1. The *attB* primers were used at 400 nM and the *attP* primers at 200 nM final concentrations. Markers used in the PCR in Figure 4 were the 100 bp ladder from NEB.

RESULTS

ϕC31 integrase in DT40 cells

We initiated our experiments by studying the ability of the ϕC31 integrase (11) to promote site-specific integration of a closed circular plasmid into a genomic locus. In order to do this we first of all used sequence targeting to introduce an *attB* site into a defined position on a human mini-chromosome contained in a DT40 hybrid cell line. The *attB* site was placed between the CCAG promoter (18) and the coding region of the hygromycin–thymidine kinase (17) fusion gene. The mini-chromosome used in these experiments was the 49B(A)A9 mini-chromosome (Figure 1A) that was derived from the human Y chromosome and is described in detail elsewhere (19). This mini-chromosome is largely composed of DYZ1,2 repeated sequences but contains two arrays of alphoid DNA at one end of the chromosome. These are 15 and 90 kb in size. The *attB* targeting construct disrupted the G418 resistance gene at the left hand end of the chromosome (Figure 1B) (19). Next we wanted to express the ϕC31 integrase in the cells containing the modified mini-chromosome. The ϕC31 integrase is 68 kDa in molecular weight and is of bacterial origin. It therefore seemed likely that it would need a nuclear localization signal to function in our experiments. Therefore we introduced sequences encoding either the native ϕC31 integrase or the integrase tagged at either the N- or C-termini with the SV40 virus large T antigen nuclear localization signal into an expression vector, CCAG IRES zeo (Figure 2A) that conferred zeocin resistance upon

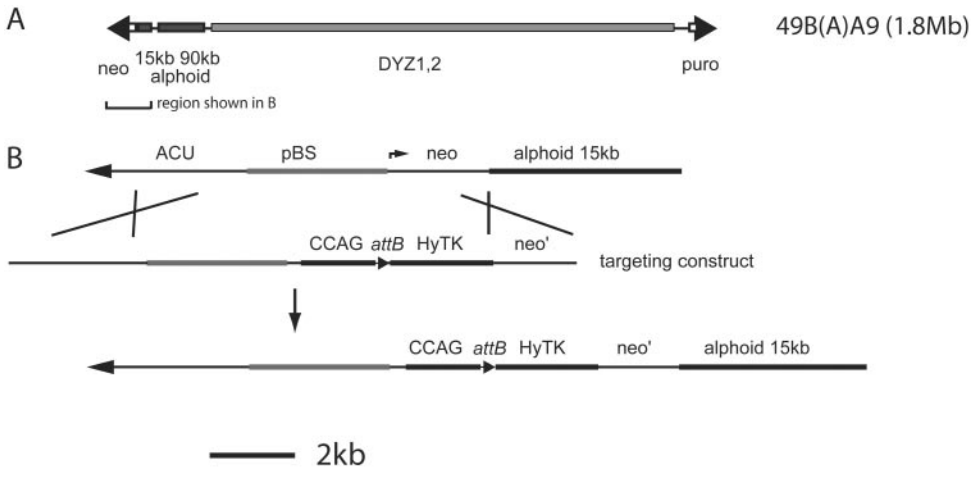

Figure 1. Engineering a mini-chromosome with an *attB* site for the ϕC31 integrase. (**A**) The mini-chromosome 49B(A)A9 is ∼1.8 Mb long and consists mainly of two arrays of repeated sequences. There is a 1.6 Mb array consisting of interspersed DYZ1 and DYZ2 sequences at the end of the mini-chromosome nearest a sub-telomeric puromycin resistance gene. The other end of the mini-chromosome includes sequences necessary for mini-chromosome maintenance in *Saccharomyces cerevisiae* (ACU), plasmid vector sequences (pBS), a gene conferring resistance to the antibiotic G418 (neo), a 15 kb array of alphoid DNA and a 90 kb array of alphoid DNA. The two alphoid DNA arrays are separated by between 30 and 50 kb of conventional euchromatic DNA that has been cloned and mapped in bacterial vectors. Details of the organization of the mini-chromosome and of the cloning and mapping have been described elsewhere (19). (**B**) The *attB* site was targeted into the mini-chromosome with a construct cloned in pBS that included the ACU sequences and homology to the G418 resistance gene. The targeting construct also contained a hygromycin resistance thymidine kinase fusion gene driven by a CCAG promoter with the promoter and coding sequence being interrupted by an *attB* site for the ϕC31 integrase. The arrow above the neo in the upper most diagram indicates the position of the SV40 early region promoter driving the G418 resistance gene in this chromosome.

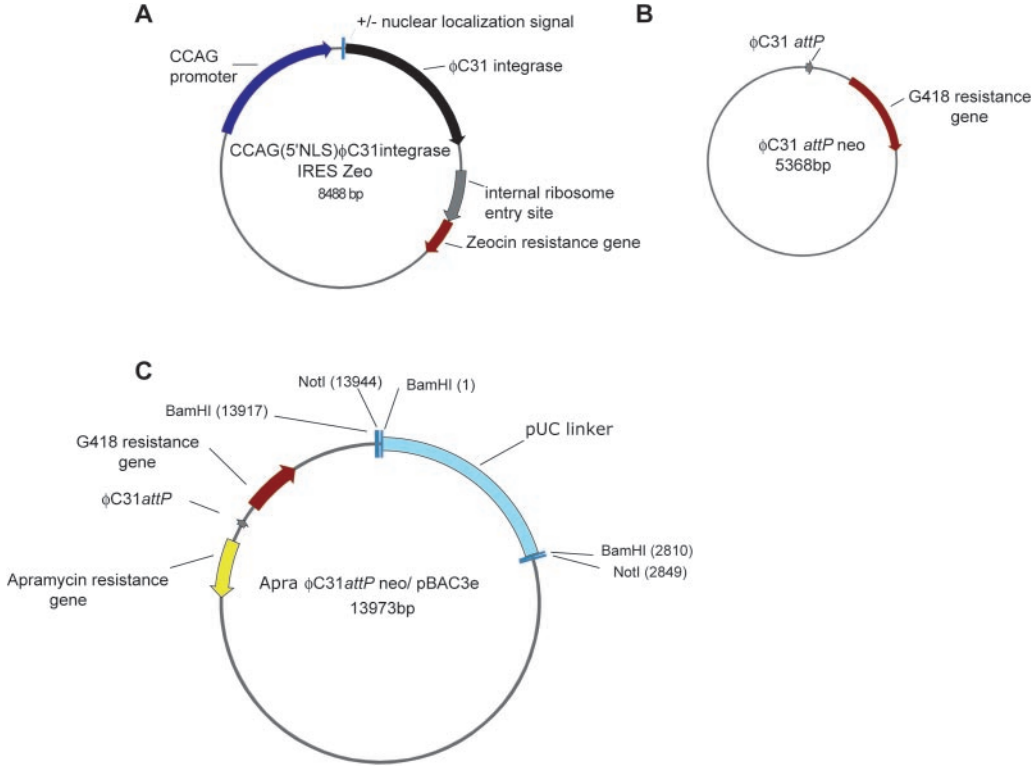

Figure 2. Plasmids used to study the activity of the ϕC31 integrase in DT40 cells. (**A**) CCAG(5′NLS)ϕC31integrase IRES Zeo is the canonical integrase expression plasmid. It contains the ϕC31integrase gene either with or without an N-terminal nuclear localization signal cloned into the internal ribosome entry site containing expression vector CCAG IRES Zeo that confers resistance to the antibiotic zeocin. (**B**) ϕC*attP*neo is the promoterless trapping vector in which a coding region potentially capable of conferring resistance to the antibiotic G418 is abutted by a ϕC31*attP* site. (**C**) Apra ϕC31*attP*neo/pBAC3e is the BAC vector designed for trapping with the ϕC31 integrase. As in pBACe3.6 the cloning sites are interrupted by a pUC plasmid.

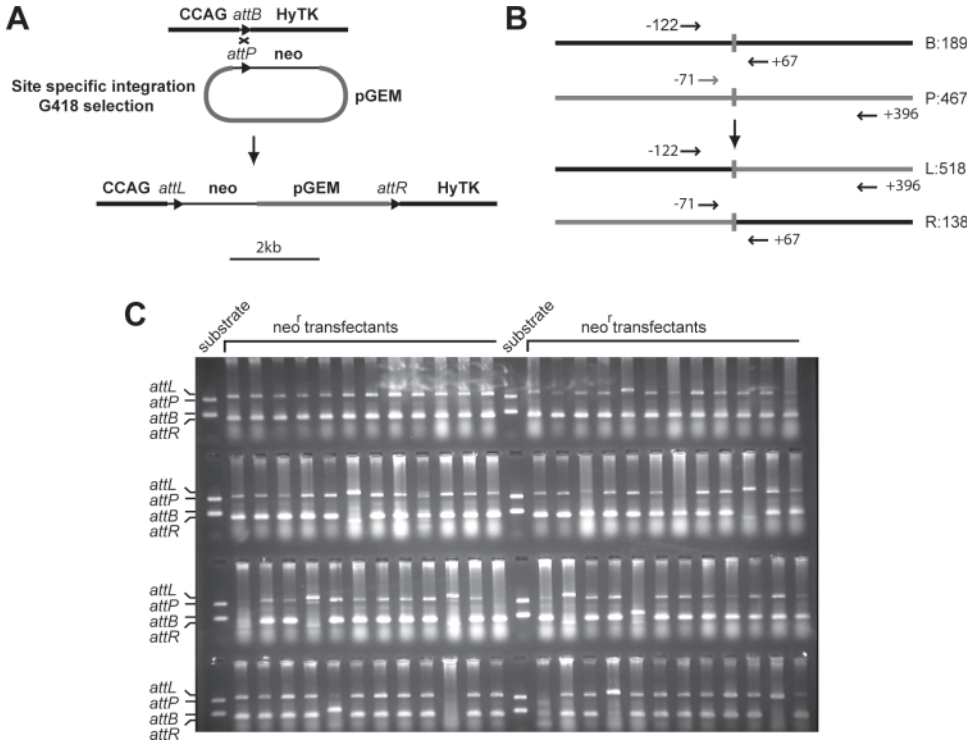

Figure 3. Assaying the accuracy of integration promoted by the ϕC31 integrase in DT40 cells. (**A**) schematic representation of a ϕC31 integrase promoted site-specific integration of the promoterless ϕC31 *attP*neo trapping plasmid into the CCAG *attB* HyTk gene within the mini-chromosome A9 CCAG *attB* HyTk. (**B**) PCR assay for the accuracy of the integration promoted by the ϕC31 integrase was established using four primers which flanked asymmetrically both the *attB* and *attP* sites and the product *attL* and *attR* sites. The diagram indicates the respective positions of the different primer combinations. (**C**) The results of the PCR assay when applied to 96 clones isolated as resistant to G418 following transfection of DT40 cells containing the mini-chromosome A9 CCAG *attB* HyTk and expressing the ϕC31 integrase with ϕC31*attP*neo. The controls used in this experiment were extracts of DT40 cells doped with plasmids CCAG *attB* HyTk and ϕC31 *attP*neo to single copy levels. Details of the PCR are given in Materials and Methods and in Supplementary Data.

vertebrate cells and allowed expression of the integrase gene as a bi-cistronic mRNA. We then introduced the linearized expression vector into DT40 cells containing the *attB*-modified mini-chromosome, selected for stably transfected clones and analysed expression of the integrase gene by western blotting (data not shown). The results indicated that the integrase comprise 0.01% of the total protein in the respective cell extracts. We also used immunocytochemistry to confirm that the protein did not enter the nucleus unless modified appropriately (data not shown).

In order to assay the activity of the integrase in the DT40 cells we assembled a plasmid, ϕC31 *attP*neo (Figure 2B), that contains an *attP* site for the ϕC31 integrase 5′ to the coding region of a promoterless G418 resistance gene. The plasmid ϕC31 *attP*neo was then transfected by electroporation into DT40 cells containing the *attB*-modified mini-chromosome and expressing either the native or nuclear localization signal modified ϕC31 integrase. Transfectants were recovered in 96-well dishes and G418 selection applied. No G418 resistant clones were recovered from the clone (Table 1) expressing the native integrase indicating that the background of integration of the *attP*neo into the host genome is undetectable. The two clones expressing integrase tagged with the nuclear localization signal each gave G418 resistant clones at a frequency of $\sim 10^{-4}$ per electroporated cell (Table 1). In these experiments the proteins tagged at the N- and C-termini with nuclear localization signals were equally efficient in promoting integration reactions (transfections 89 and 98). These results therefore do not confirm those of others (20) who observed that the protein tagged at the C-terminus functions about three times more efficiently than that tagged at the N-terminus. These differences may reflect the nature of the cell lines used in our respective experiments. We analysed the products of the transfections of 2.5.01 and 4.5.01 by PCR (Figure 3). The *attL* and *attR* sites were detected in 81 out of 96 G418 resistant clones. In the other 15 clones one or other or both of the two anticipated products were absent (e.g. clones 43, 49, 52, 82, 85 and 95 as numbered from the top left), the clone contained only an *attB* site (clones 65 and 77) and had presumably escaped selection or the PCR products were ambiguously sized (clones 30, 46, 52, 58, 60, 62 and 88). The clones giving rise to these incorrect products were not analysed further but their existence indicates the need to confirm by molecular techniques that the products of any selection are as intended. We sequenced the *attR* and *attL* sites of two clones (1 and 7) as representative of the majority and found that both were precisely as predicted for accurate site-specific recombination. Overall the results of this analysis indicate that the integrase promotes accurate site-specific recombination in chicken DT40 cells. The fidelity of the plasmid integration reaction into the mini-chromosome was confirmed by gel electrophoresis, filter transfer and hybridization for two such clones. The data on one of these clones are illustrated in Supplementary Figure 1.

We wanted to establish how the efficiency of the integrase reaction varied as a function of cell number and the amount of DNA in the transfection. We therefore took one cell line (A9-CCAGattBHyk ϕC31–5′NLS2) expressing the ϕC31 integrase tagged at the N-terminus and carried out two further series of transfections (numbered 39–41 and 42–44) in which we varied each of these parameters with one of the cell lines termed NLS2. The results of these experiments (Table 1) demonstrated that the number of recovered clones increased in proportion both to the amount of DNA and the number of electroporated cells and therefore extend the observations of others who have used the ϕC31 integrase in a different context (21).

The results in Table 1 also indicate the reproducibility of the reaction; reactions carried out within days of one another are reproducible to within a factor of two but the efficiency varies more over longer intervals. Our earlier reactions (2.5.01 and 4.5.01) were ∼4-fold more efficient than those carried out later. We are not sure why this is so. There may be variations in the quality of the ϕC31 *attP*neo DNA used in the transfection, the cell culture medium or in the epigenetic status of the target DNA. Despite this relatively small variation the results demonstrate that the ϕC31 integrase along with its cognate attachment sites provides a robust and reliable tool for integrating DNA into vertebrate and other genomes.

Site-specific integration of large DNA fragments cloned in a BAC vector using ϕC31 integrase

We wanted to use the ϕC31 integrase for integration of large fragments of DNA into defined loci so we next built a BAC vector, Apra ϕC31*attP*neo/pBAC3e (Figure 2C) containing the *attP*neo trapping cassette (Figure 2B). We confirmed by PCR and by gel electrophoresis and filter hybridization that this vector integrated in site specifically (Supplementary Figure 1). In order to investigate the ability of the ϕC31 integrase to integrate large fragments of DNA we needed a suitable test sequence. Alphoid DNA is a tandemly repeated sequence found at the centromere of human chromosomes and has been shown to be mediate centromere function. We chose this as our test sequence for two reasons. First, it is of low sequence complexity and so it has a sequence organization that is easy to analyse; this is a particularly important when one wishes to establish whether a sequence has been integrated intact. Second, tandemly repeated sequences are particularly prone to re-arrangement and thus if we were able to integrate this sequence intact then we could argue confidently that one should be able to use the system to integrate a sequence more typical of vertebrate genomic DNA. We cloned four differently sized segments of Y chromosome alphoid DNA (Figure 4A) of 33, 70, 110 and 140 kb into a BAC vector Apra ϕC31*attP*neo/pBAC3e and transfected them into the DT40 cell line A9-CCAGattBHyTK ϕC31–5′NLS2. The efficiency varied between experiments as would be expected given variations in the quality of DNA cloned in low copy number vectors (Table 2) and were lower than with the *attP*-neo plasmid. We analysed the G418^r clones recovered after transfection for the presence of the cloned insert by pulsed-field gel electrophoresis (PFGE), blotting and PCR for the recombinant *attL* site (Figure 4C and D and Supplementary Figure 2). In interpreting the data in Figure 4 it is important to know that the 49B(A)A9 mini-chromosome (19), derived as it is, from the human Y chromosome contains alphoid DNA similar in sequence to that cloned in the BAC vector, Apra ϕC31*attP*neo/pBAC3e and which consequently hybridizes to the probe used to detect the sequence we are integrating into the chromosome. This sequence that we are introducing into the chromosome by site-specific recombination therefore appears in addition to pre-existing chromosomal fragments of 15 and 90 kb that are present on the native 49B(A)A9 mini-chromosome and are detected on the blots shown in Figure 4B. These blots of the transfected clones demonstrated that we could recover clones in which the DNA had integrated intact (lanes denoted as N in Figure 4C). These blots also show a change in size of the 15 kb alphoid fragment present on the native chromosome in those clones in which the BAC has integrated. This change in fragment size arises because the 15 kb alphoid fragment is present on the same SacI fragment as the integration site (for the positions of these two sequences see Figure 1). Site-specific integration introduces both the alphoid DNA insert and the BAC vector and thereby introduces a new SacI site immediately adjacent to the 15 kb alphoid DNA. This leads to the change in size of the smaller fragment. Not surprisingly the efficiency with which we recovered clones with intact BAC insert declined with the size of the insert so that with the 140 kb BAC we were only able to recover intact BAC in 10% of the G418^r clones (Figure 4C and Table 2). The G418^r clones which did not contain the intact insert were found to be of two types (indicated above the lanes in Figure 4C). Type 1 re-arrangements were often mixtures of inserts of two different sizes. These clones contained both recombinant *attL* and *attR* sites (Figure 4D). This type of clone was predominant in the case of the 70 kb BAC. We suggest that these re-arrangements arose as a result of the incoming BAC being nicked or otherwise damaged on one strand and that this damage was repaired by gene conversion at DNA replication from the undamaged sister chromatid. We suggest that the deletions in the products of the repair arise as a result of the gene conversion events occurring between tandemly repeated arrays that were aligned out of register. In some cases the BAC was detectably repaired from the alphoid DNA already present on the A9 mini-chromosome as judged by a deletion in this DNA; these events are referred to as 1* and one of these is shown in the panel illustrating the data obtained with the 140 kb BAC. The type II events contained no detectable insert but are G418^r. PCR and filter hybridization analysis showed that these clones were deleted for sequences flanking the expected *attL* and *attR* sites (Figure 4D). These clones may have arisen as a result of a double-strand break in the CCAG-ϕC*attB*-HyTk gene being resected and repaired by ligation to the neo coding sequence native to the 49B(A)A9 mini-chromosome. Alternatively they may have arisen by site-specific integration of a BAC with a double-strand break being resected prior to a similar non-homologous repair process. Since we see few such clones in the control transfections where we do not transfect with BAC DNA we consider the second explanation more likely but have no way of proving this interpretation. The goal of our project and of others is to be able to integrate DNA intact rather than re-arranged and so we did not characterize these re-arrangements in any more detail.

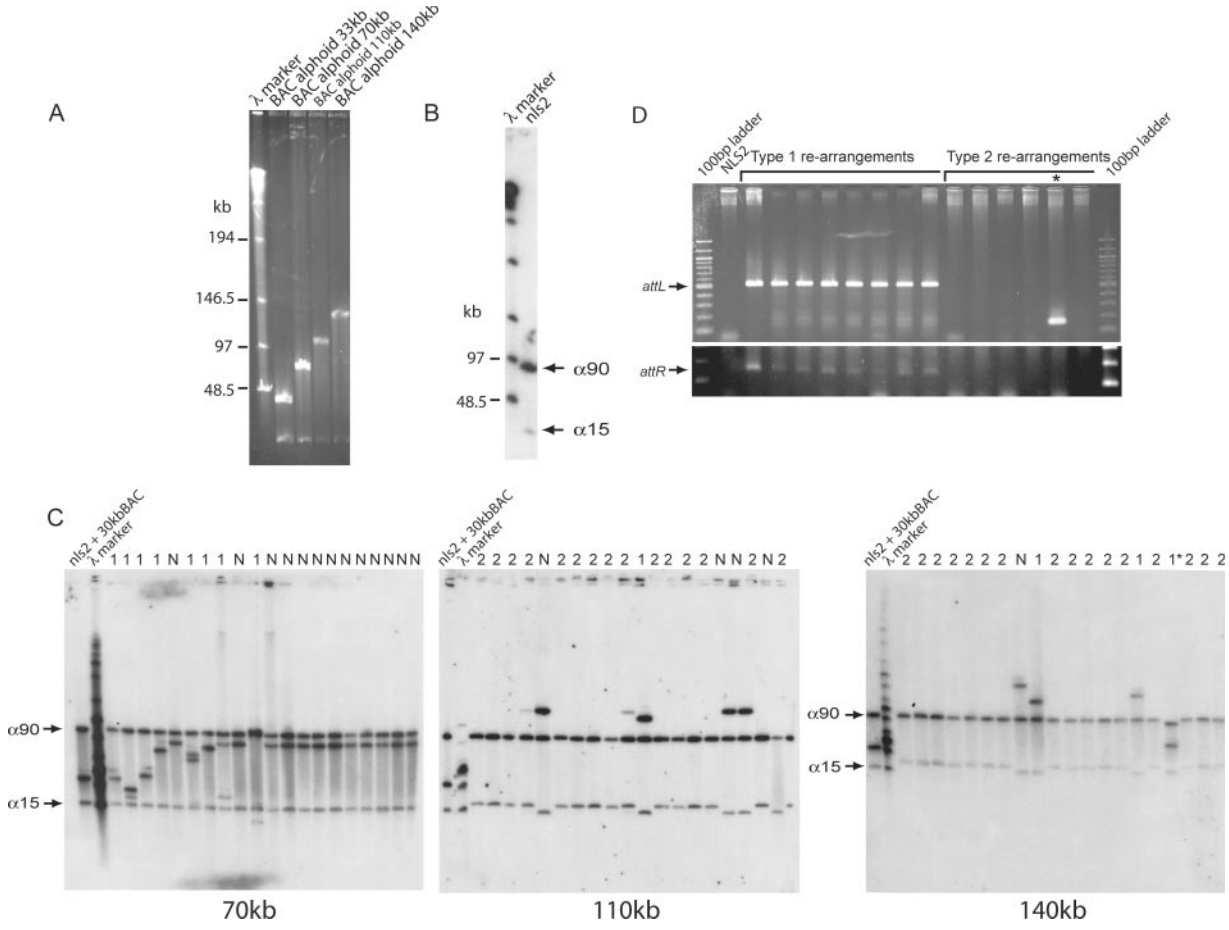

Figure 4. ϕC31 integrase mediated integration of alphoid DNA inserts cloned in a BAC trapping vector into mini-chromosome A9 CCAG *attB* HyTk in DT40 cells. (**A**) Pulsed-field gel analysis of Y alphoid DNA sequences of 30, 70, 110 and 140 kb cloned in BAC vector AprαϕC31*attP*neo/pBAC3e. (**B**) Restriction and filter hybridization analysis of the mini-chromosome A9 CCAG *attB* HyTk. Following digestion with SacI the DNA was size fractionated by PFGE. Hybridization analysis was carried out using a probe specific for alphoid DNA. The arrows indicate the 90 and 15 kb alphoid DNA fragments endogenous to this starting mini-chromosome. The marker used in this gel was a collection of multimers of the phage λ cI857 genome (48.5 kb). (**C**) Restriction and filter hybridization analysis of the mini-chromosome A9 CCAG *attB* HyTk following site-specific integration of the BACs containing inserts of Y chromosome alphoid DNA of the indicated sizes. Following digestion with SacI the DNA was size fractionated by PFGE. Hybridization analysis was carried out using a probe specific for alphoid DNA. The arrows indicate the 90 and 15 kb alphoid DNA fragments endogenous to the starting mini-chromosome. The left most track in each panel was of a SacI digest of DNA extracted from a site-specific recombinant between a BAC containing a 33 kb insert and the mini-chromosome A9 CCAG *attB* HyTk. The penultimate track in the 110 kb panel includes degraded DNA but was shown to derive from an undeleted integrant in a separate gel. The marker used in this gel was a collection of concatamers of the intact and XhoI digested phage λ cI857 genome (48.5 kb). Sizes in Materials and Methods. (**D**) PCR analysis of the recombinant *attL* and *attR* sites generated by site-specific recombination between the incoming BAC DNA and the resident mini-chromosome A9 CCAG *attB* HyTk in a selection of clones containing re-arranged DNA. The aberrantly sized *attL* site in the type 2 re-arrangement indicated with an asterisk is consistent with deletion of sequences between the two primer binding sites.

Iterative site-specific integration of transgenic DNA fragments

The results described in Table 2 and Figure 4 demonstrate that while it is possible to integrate large DNA fragments clones containing the intact integrated molecule are a small percentage of the stably transfected clones if the incoming molecule is much bigger than 100 kb. It would be useful however to be able to assemble larger stretches of transgenic DNA than are implied by this figure. For example it would be valuable to be able to build transgenic mice containing human gene families of medical importance. This goal could be achieved by combining the activities of a unidirectional and reversible recombinase as outlined in Figure 5. In this strategy a modified selectable marker gene is first introduced into a chromosome and then used as a seed from which the transgenic sequences are grown. The modifications of the selectable marker are envisaged as follows; a site for a reversible recombinase; e.g. a *loxP* site for Cre recombinase is placed between the promoter and coding region of the gene and an *att* site for a unidirectional site-specific recombinase such as ϕC31 integrase is placed at the 3′ end of the gene. In this example the site is chosen, arbitrarily to be an *attP* site. Plasmids that direct expression of the Cre recombinase and a unidirectional site-specific recombinase are then introduced into the cells containing the modified selectable marker gene. This modified selectable marker gene can then be used as the site from which inserts in circular DNA molecules; typically cloned in plasmid, cosmid or BAC vectors are iteratively introduced into the chromosome as described in the following.

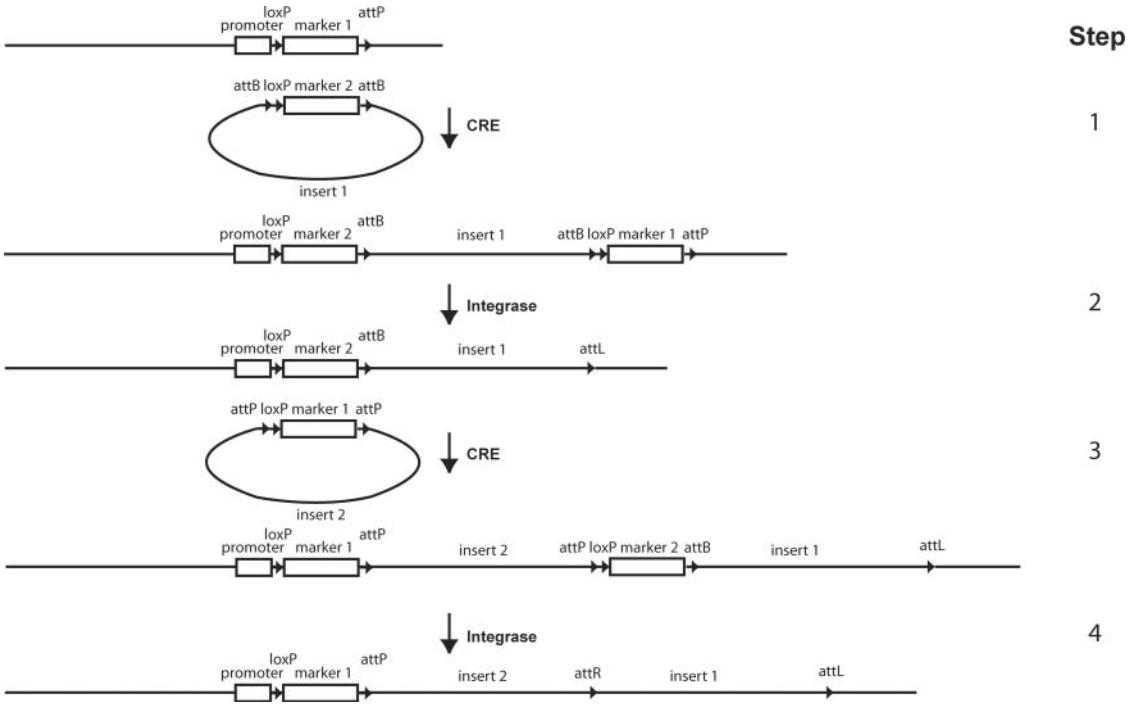

Figure 5. A schematic illustration of how ISSI might occur. Details are discussed in the text.

Step 1: A desired segment of DNA is cloned in what is here termed vector 1. Vector 1 consists of a circular piece of recombinant DNA containing a second selectable marker gene and is abutted at the 5′ end of the coding region by a *loxP* site. Two copies of an *att* site reciprocal to that already integrated into the genome flank the insert cloning site in vector 1. In this example the *att* site would be an *attB* site. The recombinant DNA molecule in vector 1 is introduced into the cells containing the chromosome modified as described above and expressing both the Cre recombinase and unidirectional site-specific recombinase. Selection is applied such that cells expressing the marker gene in vector 1 survive. These cells would be expected to be enriched in those in which the Cre recombinase has catalysed recombination between the resident chromosomal loxP site and the loxP site cloned in vector 1.

Step 2: As a result of the action of the Cre recombinase the insert contained within vector 1 is introduced into the chromosomal site and the two *attB* sites originally present in vector 1 are separated by the insert such that one is closer to the previously resident *attP* site. One of the two *loxP* products of the Cre mediated reaction is between the promoter and the marker gene contained within vector 1 while the second is between *attP* and *attB* sites. The next reaction to consider is that promoted by the unidirectional site-specific integrase which leads to the excision of the second *loxP*. The excision reaction may occur between the *attP* site and either of the two *attB* sites present in the incoming vector 1. If, as illustrated, the *attB* site nearest to the *attP* site participates in the unidirectional recombination reaction then the original marker gene and the *loxP* site adjacent to marker 1 is excised on a small circular fragment which would be expected to be lost from the nucleus at cell divison. The insert and *attB* site from vector 1 remain but loss of one of the two *loxP* sites renders the Cre catalysed integration irreversible. Alternatively if the *attB* site nearest to the second marker gene participates in the unidirectional recombination reaction then the first marker, a *loxP* site and the insert in the first vector will be excised. The first of these two alternative unidirectional recombination reactions is necessary for the next step in the sequence because the integrated product can function as a substrate for another sequence of site-specific recombination reactions that can lead to the integration of a second insert.

Step 3: A desired segment of DNA is cloned in what is here termed vector 2. Vector 2 consists of a circular piece of recombinant DNA containing the original selectable marker gene and is abutted at the 5′ end of the coding region by a *loxP* site. Two copies of an *att* site that can function as a substrate for the integrase and reciprocal to the site now integrated into the genome flank the insert cloning site in vector 2. In this example the *att* site would be an *attP* site. The recombinant DNA molecule in vector 2 is introduced into the cells containing the chromosome modified as described above as a products of steps 1 and 2 and expressing the Cre recombinase and unidirectional site-specific recombinase. Selection is applied such that cells expressing the marker gene in vector 2 survive. These cells would be expected to be enriched in those in which the Cre recombinase has catalysed recombination between the resident chromosomal *loxP* site and the *loxP* site cloned in vector 2.

Step 4: As a result of the action of the Cre recombinase the insert contained within vector 2 is introduced into the chromosomal site and the two *attP* sites originally present in vector 2 are separated by the insert two such that one is closer to the *attB* site arising from the preferred reaction occurring in step 2 of the process. As above one of the two *loxP* products of this

Cre mediated reaction is between the promoter and the marker gene contained within vector 1 while the second is between the *attB* site and *attP* site. Once again we now consider the reaction promoted by the unidirectional site-specific integrase which as in step 2 leads to the excision of the second *loxP* site. The excision reaction may occur between the *attB* site and either of the two *attP* sites originally present in the incoming vector 2. If, as illustrated, the *attP* site nearest to the *attB* site participates in the unidirectional recombination reaction then the marker gene 2 and the adjacent *loxP* site are excised to generate an unstable product. The insert and *attP* site from vector 2 remain but the loss from the cell of the small circular molecule generated by the integrase would render the Cre catalysed reaction irreversible. Were the *attP* site nearest to the second marker gene to participate in the unidirectional recombination reaction then marker 2, a *loxP* site and the second insert would be excised. The first of these two alternative unidirectional recombination reactions is, of course, preferred because as before the integrated product can function as a substrate for another sequence of site-specific recombination reactions that can lead to the integration of a third and subsequent inserts because it creates a target identical to that with which the process was started. The reaction series of steps 1–4 may therefore be repeated an arbitrary number of times. We refer to this strategy as ISSI.

We have described the ISSI strategy as though the Cre promoted reaction occurred first because this description is particularly easy to follow mechanistically. In fact the order of the two site-specific recombination reaction makes no difference to the outcome and indeed they may occur concurrently. The only feature of the reaction that is critical is which of the two attachment (*att*) sites in the incoming vector recombines with the *attB* or *attP* site adjacent to the resident selectable marker gene.

Implementation of iterative site-specific integration

We have implemented ISSI using a combination of the Cre recombinase and ϕC31 integrase. First of all we built the necessary plasmids, BAC and PAC vectors (Figure 6). In

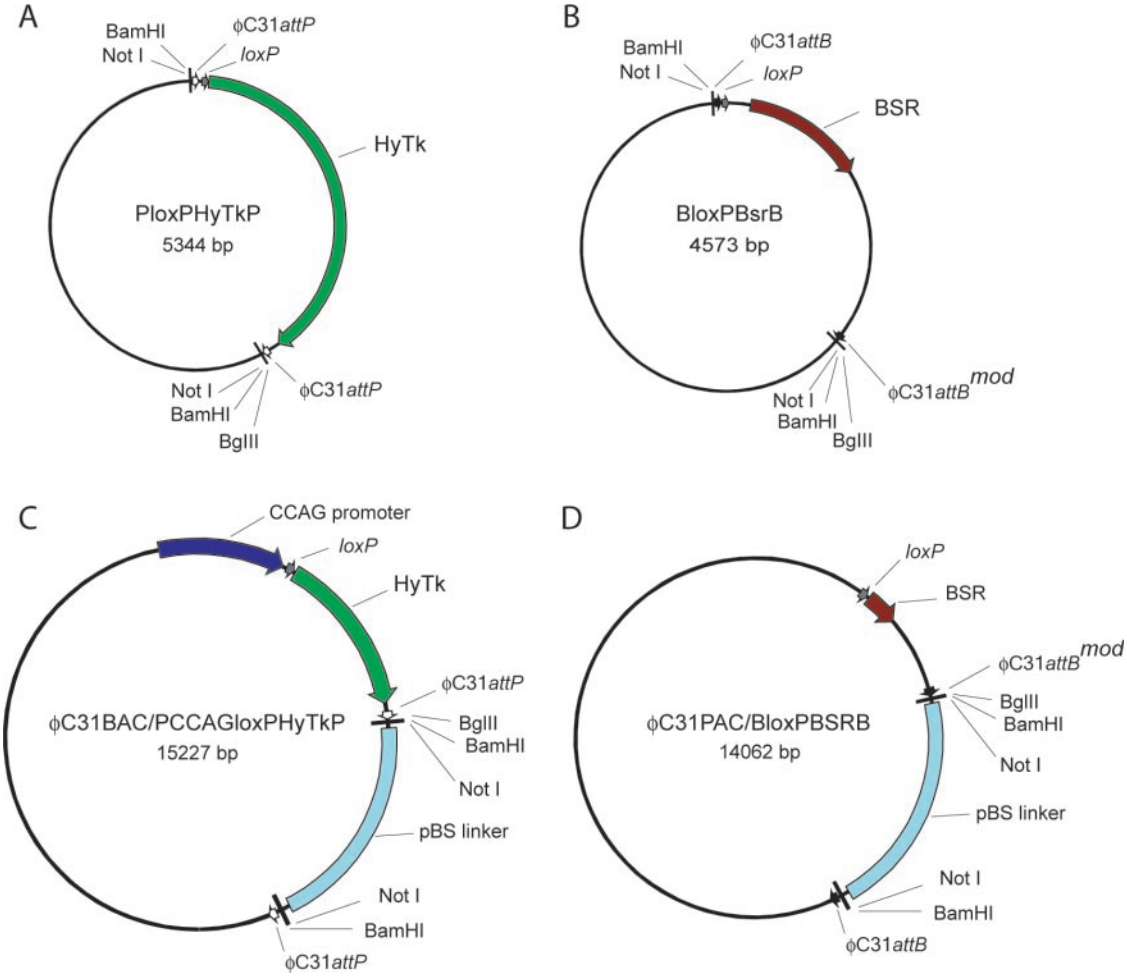

Figure 6. Plasmid and large insert vectors used in ISSI. (**A**) Plasmid *attPloxP*Hytk*attP* used in the ISSI sequence shown in Figure 7. (**B**) Plasmid *attBloxP*BSR*attB*^{mod} used in the ISSI sequence shown in Figure 7. (**C**) BAC vector ϕC31BAC *attP*CCAG*loxP*HyTk*attP*. This vector differs from the others in that it includes a promoter upstream of the selectable marker. We originally expected that this modification would improve the efficiency of the integration compared with a promoterless vector however comparisons revealed no difference in the respective efficiencies. This vector was used in the experiments shown in Figure 8. Promoterless BAC and PAC equivalents have been engineered (Supplementary Data). (**D**) PAC vector ϕC31PAC *attBloxP*BSR*attB*^{mod}. This vector was used as the vector for the human X chromosome alphoid DNA in the experiment shown in Figure 8. A BAC equivalent has been engineered (Supplementary Data).

order to do this we noted the results of others which demonstrated that ϕC31 *attB* sites are less recombinationally active when integrated into vertebrate genomic DNA than when present in cloned DNA (21). The ϕC31 *attB* site includes seven CpG sites that are potentially methylatable in vertebrate genomic DNA. This suggested to us that cytosine methylation at the 5′ position within these CpG dinucleotides was inhibiting the ability of a genomic *attB* site to participate in site-specific recombination reactions. We therefore investigated the effect of CpG methylation of the *attB* site upon its ability to participate in site-specific recombination reactions *in vitro*. These results (S. K. Malla unpublished data) demonstrated that CpG methylation does indeed inhibit the ability of the *attB* site to participate in ϕC31 mediated site-specific recombination *in vitro*. We therefore devised an unmethylatable analogue of the *attB* site termed $attB^{mod}$ and have shown that it functions as efficiently as the unmethylated native *attB in vivo*. We have therefore used $attB^{mod}$ in the ISSI plasmids shown in Figure 6. Although the details of the work describing the effects of CpG methylation of the *attB* site upon ϕC31 site-specific recombination reaction are unpublished the sequence of $attB^{mod}$ is provided in the Materials and Methods. We also used bisulphite sequencing to investigate whether the *attB* site in the CCAG *attB* HyTk construct used in the previous experiments was methylated and find that it was not (S.K. Malla unpublished data). The initial conclusions and the numbers in Table 1 are therefore not complicated by the consequences of *attB* site methylation *in vivo*. The *attB* site in the CCAG *attB* HyTk construct is located adjacent to the *CpG* island defined by the chicken β-actin promoter within the CCAG promoter and thus would not be expected to be readily methylated *in vivo*. However in the work described by Belteki *et al*. (21) the *attB* site lies outside the transcribed region of the gene and not within a *CpG* island and it may therefore be readily methylated. It thus remains to be established whether an inappropriately placed *attB* site can be methylated *in vivo* and when it is so methylated to what extent this inhibits site-specific recombination. Experiments to establish this point are in progress (S. K. Malla and W. R. A. Brown).

The genes that we used in the integrating plasmids were those conferring resistance to the antibiotics blasticidin (16) (BSR) and hygromycin (17) (HyTk).

In order to implement ISSI we targeted the seed construct CCAG *loxP* HyTk *attP* (Figure 5A) to the 49B(A)A9 mini-chromosome in the same way as we had targeted the CCAG *attB* HyTk construct (Figure 1B). We then introduced sequentially the CCAG 5′nls ϕC31 integrase IRES zeo (Figure 2A) and the CCAG Cre IRES ecogpt (Supplementary Data) plasmids that direct expression of the ϕC31 integrase tagged at the N-terminus with a nuclear localization sequence and the Cre recombinase, respectively. This generated several cells lines which should be capable of mediating the ISSI sequence. We used one of these, termed ISSI 2, for our proof of principle experiments.

In the first of our proof of principle experiments we used plasmids to execute two complete cycles of this strategy (Figure 7A). This led to the integration of a concatamer of four pBlueScript and accompanying polylinker sequences into the target. The presence of four other plasmid vector DNA molecules integrated into background of this cell line precluded direct analysis of the integrated DNA and so we used restriction enzyme digestion, filter hybridization and indirect end labelling to confirm the predicted structures of the DNA integrated during the ISSI sequence. DNA extracted

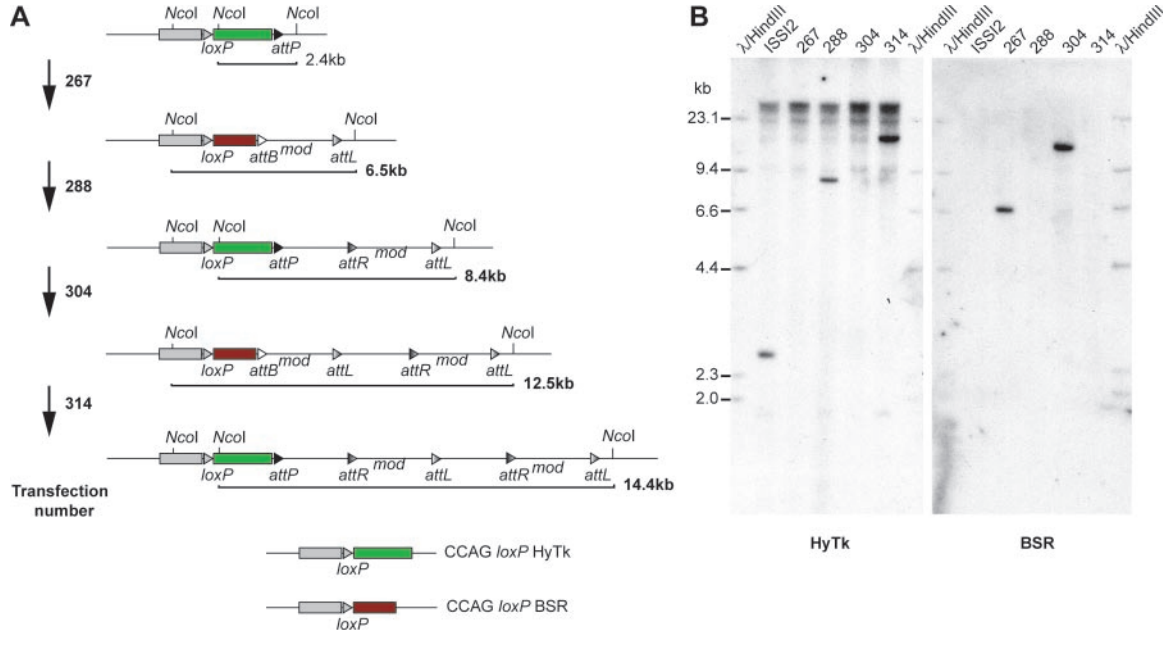

Figure 7. ISSI of plasmid vector molecules into mini-chromosome A9 CCAG *loxP* HyTk *attP* in DT40 cells. (**A**) Illustrates the iterations of plasmid integration reactions into the CCAG *loxP* HyTK *attP* site present in the starting cell line ISSI 2 and analysed in (B). The positions of the substrate and product attachment (*att*) sites and of the restriction enzyme sites used in the mapping are indicated. The cassettes used in these transfections were alternately *attB loxP* BSR $attB^{mod}$ and *attP loxP* HyTk *attP*. (**B**) Filter hybridization analysis of genomic DNA extracted from the cell lines generated in the sequence indicated in A using probes recognizing the selectable marker genes used in the ISSI sequence. The cross hybridization to chicken DNA seen with the HyTk probe at ∼25 kb serves as a loading control in this analysis.

from the cells was analysed by digestion with an enzyme that does not cut in pBluescript; NcoI, and then analysed (Figure 7B) either with a probe that recognizes the gene conferring resistance to the antibiotic Blasticidin (BSR) or with a probe from the hygromycin resistance gene present in the original target and in the products of the second and fourth plasmid integration reactions. These cell lines are referred to in Figure 7B by the number of the transfection from which they were derived which is indicated to the left of Figure 7A. The sizes and distribution of the cognate restriction fragments in these cells lines were precisely as predicted.

At each step in the sequence of transfections illustrated in Figure 7 we characterized ~10 individual clones in order to identify those in which the ϕC31 *att* site 5′ to the incoming marker had recombined with the resident substrate site and which therefore had a structure that made them potentially available for a subsequent round of integration. At each step we found that all the recombination events were with one or other of the two sites and that each site had recombined with similar efficiency. It seemed to us possible that since ISSI requires the coordinated action of two site-specific recombinases for a stable integration of the desired structure it would be less efficient than the integration reaction promoted by the ϕC31 integrase alone. This was not so and ISSI appeared to be at least as efficient as the single ϕC31 reaction alone for plasmid (Table 1) integrations.

In the second proof of principle experiment we started with the product of transfection 267 described in the plasmid ISSI experiment and successively introduced a 70 kb alphoid DNA array derived from the human Y chromosome cloned in a CCAG *loxP* HyTk *attP attP* BAC and a 80 kb array of alphoid DNA derived from the human X chromosome cloned in a *loxP* BSR *attB attB* BAC (Figure 8A and B). This led to the isolation of cell line 293 containing the Y alphoid DNA and then to cell line 419 containing both the Y and X alphoid DNA (Figure 8B). We characterized these cell lines at successively higher resolution by three methods. Firstly we demonstrated that we could detect the acquisition of the alphoid DNA arrays by PFGE of the intact mini-chromosomes (Figure 8C). Integration of the X chromosome alphoid DNA was indicated by specific hybridization and of both sequences by a slight increase in the size of the mini-chromosome. Secondly we used restriction enzyme digestion, PFGE and filter hybridization to establish physical linkage of the sequences. This analysis made use of the sequence of the junctions and the low sequence complexity of the two sequences. Neither BglII nor BclI cut either alphoid DNA array but BglII cuts the polylinker junction between the two arrays. Consequently BglII releases each different array separately but BclI releases them together on a single large fragment of ~170 kb in size and thus reveal their linkage (Figure 8D). The BclI fragments are 20 kb longer than the integrated DNA because the integration site itself lies on a 20 kb BclI fragment. The 15 kb alphoid DNA cognate fragment is increased in size in the mini-chromosomes in cell lines 293 and 419 as compared with the mini-chromosome in clone NLS2. This increase arises because of the introduction of the pBluescript molecule during transfection 267 immediately prior to the generation of cell line 293 (Figure 8A). Finally we established the presence of the predicted products and substrates of the site-specific recombination reactions in the starting line ISSI 2 and in the three cell lines used in the sequence. There are nine such junctions (Figure 8B). Their presence or absence in the cell lines is as predicted (Figure 8E); sequence analysis (data not shown) confirms the identity of the product in each case. These results thus prove that the ISSI functions as envisaged. The efficiency of the ISSI BAC transformation reactions is greater (Table 2) than seen with the ϕC31 integrase alone which is consistent with the results seen with plasmids.

DISCUSSION

The experiments described here demonstrate that the ϕC31 integrase specifically and efficiently promotes site-specific integration in chicken DT40 cells. The observation that these proteins function in vertebrate cells is consistent with earlier work showing that the purified ϕC31 integrase promotes unidirectional exchange reactions between *attB* and *attP* sites *in vitro* (11) and with some aspects of the results of others who have studied either the ϕC31 and R4 (21,22) integrases in mammalian cells.

We have shown that integration of intact 140 kb BACs into vertebrate cells is very inefficient because they usually break before they integrate. The breakage does not appear to reflect the integrity prior to electroporation because the DNA was purified as closed circular by equilibrium centrifugation on a caesium chloride gradient in the presence of ethidium bromide and was assessed as intact by PFGE. Although there have been some reports of the successful electroporation of BACs into mammalian cells (23) others have reported problems (24) and have developed episomal vector systems in order to partially overcome them. We therefore conclude that the problem of integrating intact, large BACs into genomic DNA might be a general one.

We have addressed the problem of integrating large segments of DNA by establishing a strategy that combines the activities of the ϕC31 integrase and Cre recombinase to permit iterative integration of transgenic sequences. The strategy that we have implemented is termed ISSI is just one of several that one may envisage in which the activities of one or more unidirectional enzymes are combined with either the activity of a reversible recombinase or homologous recombination to permit serial or iterative integration. In one of these the activity of Cre is combined with that of two homing nucleases or meganucleases, oligonucleotides and T4 ligase and has been used for multiple rounds of integration into cloned DNA *in vitro* (25). The use of the two meganucleases rather than a single unidirectional enzyme allowed a single unambiguous resolution product to be generated at each half of the cycle but this is not necessary when one wishes to assemble an array *in vivo* because one would wish to check the integrants for their integrity at each stage. Qualitatively ISSI works exactly as anticipated but quantitatively ISSI works better than one might have supposed. The reaction sequence that leads to the integration of one marker and excision of another is not known but the efficiency with which these steps occur relative to that of a single ϕC31 integrase integration reaction suggests that the ϕC31 integrase and the Cre are cooperating at the rate-limiting step. Such cooperativity could arise if the rate-limiting step was the association of the incoming DNA with the genomic target and both the Cre recombinase and the ϕC31 integrase bound to each of their respective target sites. This would allow

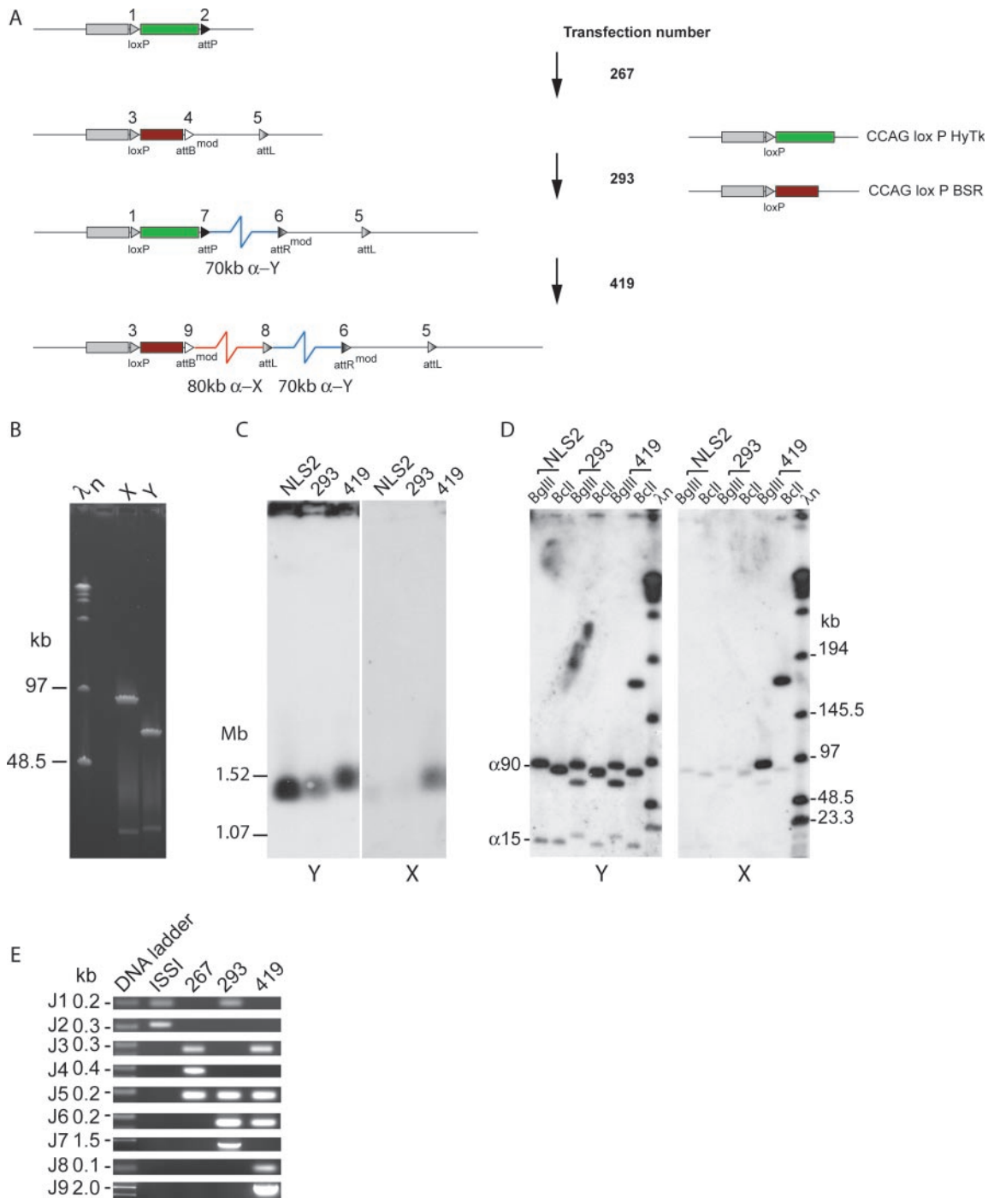

Figure 8. ISSI of large fragments molecules into mini-chromosome A9 CCAG *loxP* HyTk *attP* in DT40 cells. (**A**) Illustrates the iterations of plasmid and large fragment integration reactions into the CCAG *loxP* HyTK *attP* site present in the starting cell line ISSI 2 and analysed in (C–E). The cell line ISSI 2 was transfected with the plasmid *attBloxP*BSR*attB* to yield cell line 267 as described in Figure 7. This cell line was then transfected sequentially with a BAC containing 70 kb of Y alphoid DNA cloned in ϕC31BAC/PCCAGloxPHyTkP to yield cell line 293 and then with a PAC containing 80 kb of X alphoid DNA cloned in ϕC31PAC/BloxPBSRB to yield cell line 419. The recombination sites analysed by PCR in section E below are indicated. (**B**) Pulsed-field gel analysis of 80 kb X and 70 kbY alphoid DNA cloned in the BAC and PAC vectors used to transfect the cells as described in A. The DNA was digested with NotI, which excises the insert, prior to gel electrophoresis. The image is of an ethidium bromide stained gel. (**C**) PFGE and filter hybridization analysis of the intact modified mini-chromosomes in cell lines 293 and 419. The control is the mini-chromosome isolated from the cell line A9-CCAGattBHyk ϕC31-5′NLS2 (abbreviated NLS2) modified with CCAG *attB* HyTk and expressing the ϕC31 integrase. This filter transfers were probed with Y and X alphoid DNA sequences and illustrate the incremental increase in size of the mini-chromosome that follows from the introduction of the 70 and 80 kb alphoid Y and X alphoid DNA sequences. (**D**) Restriction enzyme mapping of the integrated alphoid DNA sequences. DNA extracted from the indicated cell lines was digested with either BglII or BclI, size fractionated by PFGE, filter transferred and the transfer probed with either X or Y alphoid DNA sequences. BglII but not BclI cuts the integrated DNA sequences at the junctions between the insert and vector. The BglII digests therefore identify the DNA sequences that have been introduced into the chromosome while the co-hybridization of the 170 kb fragment in the BclI digest of the 419 cell line establishes linkage between the Y and X alphoid DNA sequences as predicted in (A). (**E**) Identification by PCR of the predicted junctions indicated in (A).

for a bivalent association which would be expected to increase the affinity of participating DNA molecules for one another and their overall rates of reaction.

The observation that the two attachment sites in the incoming circular DNA recombine equally efficiently with the reciprocal genomic attachment site is at first surprising. A linear representation of a coordinate reaction process might lead one to suppose that the 3′ site in the incoming DNA would be stereo-chemically favoured. Our observation may reflect a two-step process as envisaged in Figure 5 or the fact that both participating DNA molecules are folded in three dimensions and that recombination between the *loxP* sites favour neither the 3′ nor the 5′ attachment sites.

Vertebrate centromeres are typically composed of several hundred kilobasepairs of tandemly repeated sequences. We have demonstrated that ISSI will enable us to build such structures. We anticipate that ISSI may also be applied to other problems of experimental transgenesis where it is necessary to combine and assemble multiple transgenes. In some of these it will be simpler to 'stack' (25) the transgenes *in vitro* prior to their introduction into the genome but in others, where one wishes, for example, to introduce longer stretches of DNA, this will be impractical and *in vivo* assembly will be the only option. Examples include the assembly of large genomic regions for the purpose of investigating distally located control elements, reconstruction of complex loci such as the major histocompatibility complex as transgenes and the synthesis of bacterial genomes (7). Although we have carried out two proofs of principle experiments it remains to be proven how well ISSI works in cells other than those of the chicken DT40 line. We have argued in the introduction that our experimental approach is robust in so far as we have shown that we can integrate unstable sequences into a hyper-recombinogenic cell line. One limitation may be how well the unidirectional integrase works in different cell types. ϕC31 integrase itself has been shown to function in mouse, human, fission yeast and *Drosophila* cells [reviewed in (26)]. Some uncertainty also surrounds the ease with which large DNA fragments may be introduced into both eukaryotic and prokaryotic cells. Nevertheless BACs are routinely introduced into *E.coli* cells (27) and have been transferred into mouse and human fibroblasts (24). Similarly large DNA fragments may be readily introduced into both budding (28) and fission yeast (29). This body of work suggests that size of the transfected DNA will not fundamentally limit the implementation of ISSI in a wide variety of cell types. Consistently in unpublished data we have used three rounds of ISSI to construct a 210 kb transgene in Chinese hamster ovary cells. In these experiments the ϕBT1 integrase (30) rather than the ϕC31 integrase was used as the unidirectional component in the reaction. Despite these arguments and encouraging preliminary results only sustained application in a variety of systems will establish the utility and limitations of our technique.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

We thank David Brook for discussion and Alistair Chambers for comments on the manuscript. The work was supported by the EU (the DT40 cell line as a genetic model. Contract QLRT-1999-00923), the Leverhulme Trust, the BBSRC and the University of Nottingham through a studentship to S.K.M. Funding to pay the Open Access publication charges for this article was provided by the JISC.

Conflict of interest statement. None declared.

REFERENCES

1. Sauer,B. (1994) Site-specific recombination: developments and applications. *Curr. Opin. Biotechnol.*, **5**, 521–527.
2. Rodriguez,C.I., Buchholz,F., Galloway,J., Sequerra,R., Kasper,J., Ayala,R., Stewart,A.F. and Dymecki,S.M. (2000) High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nature Genet.*, **25**, 139–140.
3. Smith,A.J., De Sousa,M.A., Kwabi-Addo,B., Heppell-Parton,A., Impey,H. and Rabbitts,P. (1995) A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. *Nature Genet.*, **9**, 376–385.
4. Mills,A.A. and Bradley,A. (2001) From mouse to man: generating megabase chromosome rearrangements. *Trends Genet.*, **17**, 331–339.
5. 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.
6. Zheng,B., Mills,A.A. and Bradley,A. (2001) Introducing defined chromosomal rearrangements into the mouse genome. *Methods*, **24**, 81–94.
7. Smith,H.O., Hutchison,C.A.,III, Pfannkoch,C. and Venter,J.C. (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl Acad. Sci. USA*, **100**, 15440–15445.
8. Albert,H., Dale,E.C., Lee,E. and Ow,D.W. (1995) Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J.*, **7**, 649–659.
9. Langer,S.J., Ghafoori,A.P., Byrd,M. and Leinwand,L. (2002) A genetic screen identifies novel non-compatible loxP sites. *Nucleic Acids Res.*, **30**, 3067–3077.
10. Lauth,M., Spreafico,F., Dethleffsen,K. and Meyer,M. (2002) Stable and efficient cassette exchange under non-selectable conditions by combined use of two site-specific recombinases. *Nucleic Acids Res.*, **30**, e115.
11. Thorpe,H.M. and Smith,M.C. (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.
12. Bibb,L.A., Hancox,M.I. and Hatfull,G.F. (2005) Integration and excision by the large serine recombinase phiRv1 integrase. *Mol. Microbiol.*, **55**, 1896–1910.
13. Thyagarajan,B., Guimaraes,M.J., Groth,A.C. and Calos,M.P. (2000) Mammalian genomes contain active recombinase recognition sites. *Gene*, **244**, 47–54.
14. Yu,D., Ellis,H.M., Lee,E.C., Jenkins,N.A., Copeland,N.G. and Court,D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **97**, 5978–5983.
15. Kuhstoss,S., Richardson,M.A. and Rao,R.N. (1991) Plasmid cloning vectors that integrate site-specifically in *Streptomyces* spp. *Gene*, **97**, 143–146.
16. Izumi,M., Miyazawa,H., Kamakura,T., Yamaguchi,I., Endo,T. and Hanaoka,F. (1991) Blasticidin S-resistance gene (bsr): a novel selectable marker for mammalian cells. *Exp. Cell Res.*, **197**, 229–233.
17. Lupton,S.D., Brunton,L.L., Kalberg,V.A. and Overell,R.W. (1991) Dominant positive and negative selection using a hygromycin phosphotransferase–thymidine kinase fusion gene. *Mol. Cell. Biol.*, **11**, 3374–3378.
18. Niwa,H., Yamammura,K. and Miyazaki,J. (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, **108**, 193–201.
19. Yang,J.W., Pendon,C., Yang,J., Haywood,N., Chand,A. and Brown,W.R. (2000) Human mini-chromosomes with minimal centromeres. *Hum. Mol. Genet.*, **9**, 1891–1902.
20. Andreas,S., Schwenk,F., Kuter-Luks,B., Faust,N. and Kuhn,R. (2002) Enhanced efficiency through nuclear localization signal fusion on phage

PhiC31-integrase: activity comparison with Cre and FLPe recombinase in mammalian cells. *Nucleic Acids Res*., **30**, 2299–2306.
21. Belteki,G., Gertsenstein,M., Ow,D.W. and Nagy,A. (2003) Site-specific cassette exchange and germline transmission with mouse ES cells expressing phiC31 integrase. *Nat. Biotechnol*., **21**, 321–324.
22. Olivares,E.C., Hollis,R.P. and Calos,M.P. (2001) Phage R4 integrase mediates site-specific integration in human cells. *Gene*, **278**, 167–176.
23. Hejna,J.A., Johnstone,P.L., Kohler,S.L., Bruun,D.A., Reifsteck,C.A., Olson,S.B. and Moses,R.E. (1998) Functional complementation by electroporation of human BACs into mammalian fibroblast cells. *Nucleic Acids Res*., **26**, 1124–1125.
24. Wade-Martins,R., Smith,E.R., Tyminski,E., Chiocca,E.A. and Saeki,Y. (2001) An infectious transfer and expression system for genomic DNA loci in human and mouse cells. *Nat. Biotechnol*., **19**, 1067–1070.
25. Lin,L., Liu,Y.G., Xu,X. and Li,B. (2003) Efficient linking and transfer of multiple genes by a multigene assembly and transformation vector system. *Proc. Natl Acad. Sci. USA*, **100**, 5962–5967.
26. Groth,A.C. and Calos,M.P. (2004) Phage integrases: biology and applications. *J. Mol. Biol*., **335**, 667–678.
27. 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.
28. Burke,D.T., Carle,G.F. and Olson,M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, **236**, 806–812.
29. Young,D.J., Nimmo,E.R. and Allshire,R.C. (1998) A *Schizosaccharomyces pombe* artificial chromosome large DNA cloning system. *Nucleic Acids Res*., **26**, 5052–5060.
30. Gregory,M.A., Till,R. and Smith,M.C. (2003) Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *J. Bacteriol*., **185**, 5320–5323.