


RESEARCH ARTICLE

High fidelity one-pot DNA assembly using orthogonal serine integrases

Jumai Abioye¹ | Makeba Lawson-Williams² | Alicia Lecanda¹ | Brecken Calhoon¹ |
Arlene L. McQue¹ | Sean D. Colloms¹ | W. Marshall Stark¹ | Femi J. Olorunniji^{1,2} ¹School of Molecular Biosciences, University of Glasgow, Glasgow, UK²School of Pharmacy and Biomolecular Sciences, Faculty of Science, Liverpool John Moores University, Liverpool, UK

Correspondence

Femi J. Olorunniji, School of Pharmacy and Biomolecular Sciences, Faculty of Science, Liverpool John Moores University, James Parsons Building, Byrom Street, Liverpool L3 3AF, UK.

Email: F.J.Olorunniji@ljmu.ac.uk

Abstract

Background: Large serine integrases (LSIs, derived from temperate phages) have been adapted for use in a multipart DNA assembly process in vitro, called serine integrase recombinational assembly (SIRA). The versatility, efficiency, and fidelity of SIRA is limited by lack of a sufficient number of LSIs whose activities have been characterized in vitro.**Methods and Major Results:** In this report, we compared the activities in vitro of 10 orthogonal LSIs to explore their suitability for multiplex SIRA reactions. We found that Bxb1, ϕ R4, and TG1 integrases were the most active among the set we studied, but several others were also usable. As proof of principle, we demonstrated high-efficiency one-pot assembly of six DNA fragments (made by PCR) into a 7.5 kb plasmid that expresses the enzymes of the β -carotenoid pathway in *Escherichia coli*, using six different LSIs. We further showed that a combined approach using a few highly active LSIs, each acting on multiple pairs of *att* sites with distinct central dinucleotides, can be used to scale up “poly-part” gene assembly and editing.**Conclusions and Implications:** We conclude that use of multiple orthogonal integrases may be the most predictable, efficient, and programmable approach for SIRA and other in vitro applications.

KEYWORDS

genome editing, large serine integrases (LSIs), SIRA, site-specific recombination, synthetic biology

1 | INTRODUCTION

Serine integrases are site-specific recombinases used by lysogenic phages for the integration of their genomic DNA into the bacterial host genome.^[1] The phages also encode a second protein called the recombination directionality factor (RDF) that modifies the properties of the

integrases and activates the excision of the prophage from the host during the lytic phase (Figure 1A).^[1,2]

Serine integrases catalyze recombination between a specific site in the phage genome (*attP*) and a corresponding site in the bacterial genome (*attB*). The *att* sites are relatively short sequences of about 40–50 bp; each integrase has its own cognate *attP* and *attB* sites. Recombination of *attP* and *attB* DNA results in the formation of new sites called *attR* and *attL*, both of which are inert to recombination by the integrase alone. The reverse reaction (*attL* x *attR*) requires the interaction of the RDF with the integrase (Figure 1A). Like all

Abbreviations: LSI, large serine integrase; RDF, recombination directionality factor; SIRA, serine integrase recombinational assembly.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. Biotechnology Journal published by Wiley-VCH GmbH.

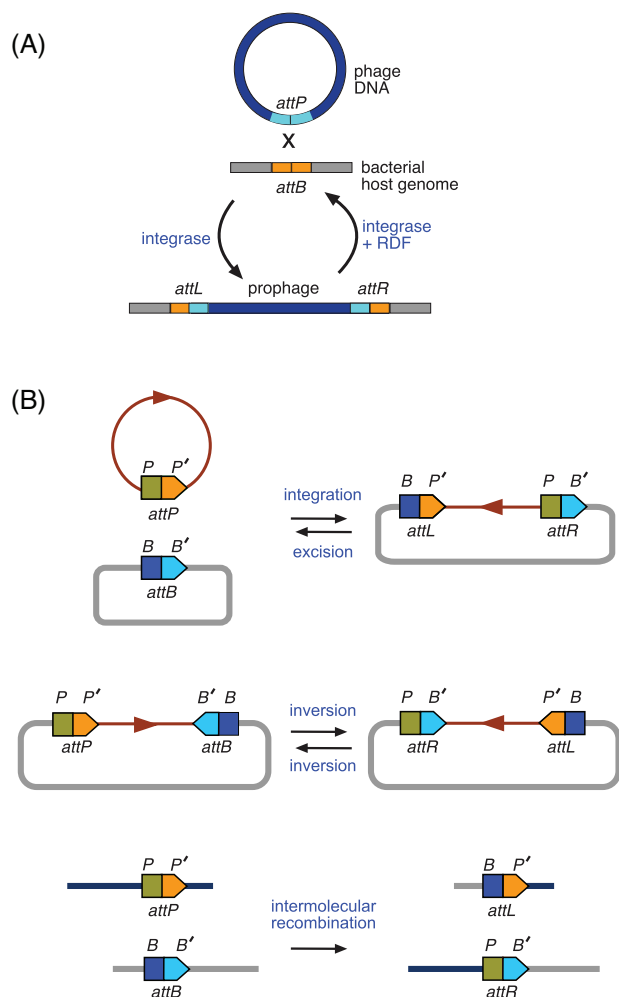


FIGURE 1 Site-specific recombination by LSI systems, and outcomes of recombination events. (A) Phage integration into bacterial host genome. In their natural context, LSI systems recombine an *attP* site in the phage DNA and an *attB* site in the genome of the bacterial host, resulting in integration of the phage DNA as a prophage. This unidirectional reaction generates two new recombination sites: *attL* (formed from the left end of *attB* and right end of *attP*), and *attR* (formed from the left end of *attP* and right end of *attB*). In the reverse reaction, a second host protein termed the RDF binds to the LSI and changes its specificity to recombine *attR* and *attL*, thereby reforming *attB* and *attP*. Adapted from Olorunniji *et al.*^[26] (B) Outcomes of site-specific recombination by LSI systems. Intermolecular recombination between *att* sites results in integration of a circular DNA molecule into a corresponding genomic locus. In the reverse reaction, intramolecular recombination of compatible *att* sites arranged in “head to tail” (or “direct repeat”) orientation results in an excision (resolution) outcome. In contrast, when sites are in “head to head” (or “inverted repeat”) orientation, the outcome of the recombination event is inversion, that is, “flipping” of the DNA segment flanked by the recombining *att* sites. Finally, intermolecular recombination involving linear DNA molecules generates two different linear recombinant molecules. The functional annotation of *attP* (*P*-*P'*) and *attB* (*B*-*B'*) sites are shown to illustrate the site orientations in the recombinant *attR* (*P*-*B'*) and *attL* (*B*-*P'*) sites, respectively. Adapted from Olorunniji *et al.*^[3] LSI, large serine integrase; RDF, recombination directionality factor.

site-specific recombination systems, the outcome of large serine integrase (LSI)-catalyzed reactions can be determined by the location and orientation of the recombining *att* sites (Figure 1B). Specific arrangements of sites can be used to achieve transgene integration, gene deletion, inversion of DNA regions flanked by the *att* sites, and intermolecular joining of linear DNA fragments.^[2,3]

Sets of orthogonal LSIs (where each integrase will act only on its own *att* sites) have found increasing use for applications in biotechnology, both in vivo and in vitro.^[4–7] Although they have a wide range of potential applications in biotechnology, few LSIs have been characterized extensively to identify the most active ones and to establish the optimal conditions for their in vitro reactions. Thorough characterization of the in vitro reaction properties of LSIs is necessary to standardize their use in routine DNA assembly protocols and other applications for which they are used. In this study, we have used a simple intramolecular resolution reaction (Figure 2A,B) and intermolecular joining reaction (Figure 2C,D) to characterize the activities of purified selected LSIs.

In particular, the sequence specificity of LSIs and the unidirectional nature of their reactions make them useful tools in DNA assembly applications.^[3,8–10] In one method, serine integrase recombinational assembly (SIRA) (Figure 3A),^[11,12] the DNA fragments are made by amplifying gene cassettes from their plasmid or genomic sources using PCR. Appropriate *att* sites are appended to each fragment by including their sequences in the PCR primers (Figure 3B). Equimolar amounts of the fragments are mixed together with a plasmid “vector,” also containing *att* sites. The appropriate LSIs are then added to promote recombination, and the correctly assembled construct can be recovered following transformation of *Escherichia coli*. This strategy has been applied to achieve rapid metabolic pathway optimization.^[9,11,13]

Earlier SIRA studies used a single LSI, and recombination was restricted to particular pairs of *att* sites by varying the 2 bp sequence at the center of the *att* sites where DNA strands from the two parental sites overlap in the recombinant products (the so-called “central dinucleotide”). Recombination can be completed only if the two recombining sites are identical in this sequence, and thus can make Watson–Crick basepairs in the products. By using DNA fragments with *att* sites differing in the central dinucleotide sequence, a specific assembly could be specified (Figure 3C).^[11,13] While the use of central dinucleotide sequence as the determinant in arranging DNA fragments in a particular order has been successful, this method of distinguishing the *att* sites has some shortcomings. Only six variants of the central dinucleotide sequence are useful in practice^[11] (Figure 3C) and the variants can have widely differing reactivity; furthermore, sites with different central dinucleotides can potentially recombine to some extent, giving undesired products. A superior alternative may be to use *att* site pairs which are targeted by different, orthogonally acting LSIs, but to date too few LSIs have been characterized in sufficient detail for this alternative approach to be practical.

In this study, we have used a simple intramolecular resolution reaction (Figure 2A,B) and intermolecular joining reaction (Figure 2C,D) to characterize the activities of 10 purified selected LSIs, thus informing the selection of active LSIs for in vitro applications. We then demonstrate a significant improvement on the original SIRA method by using

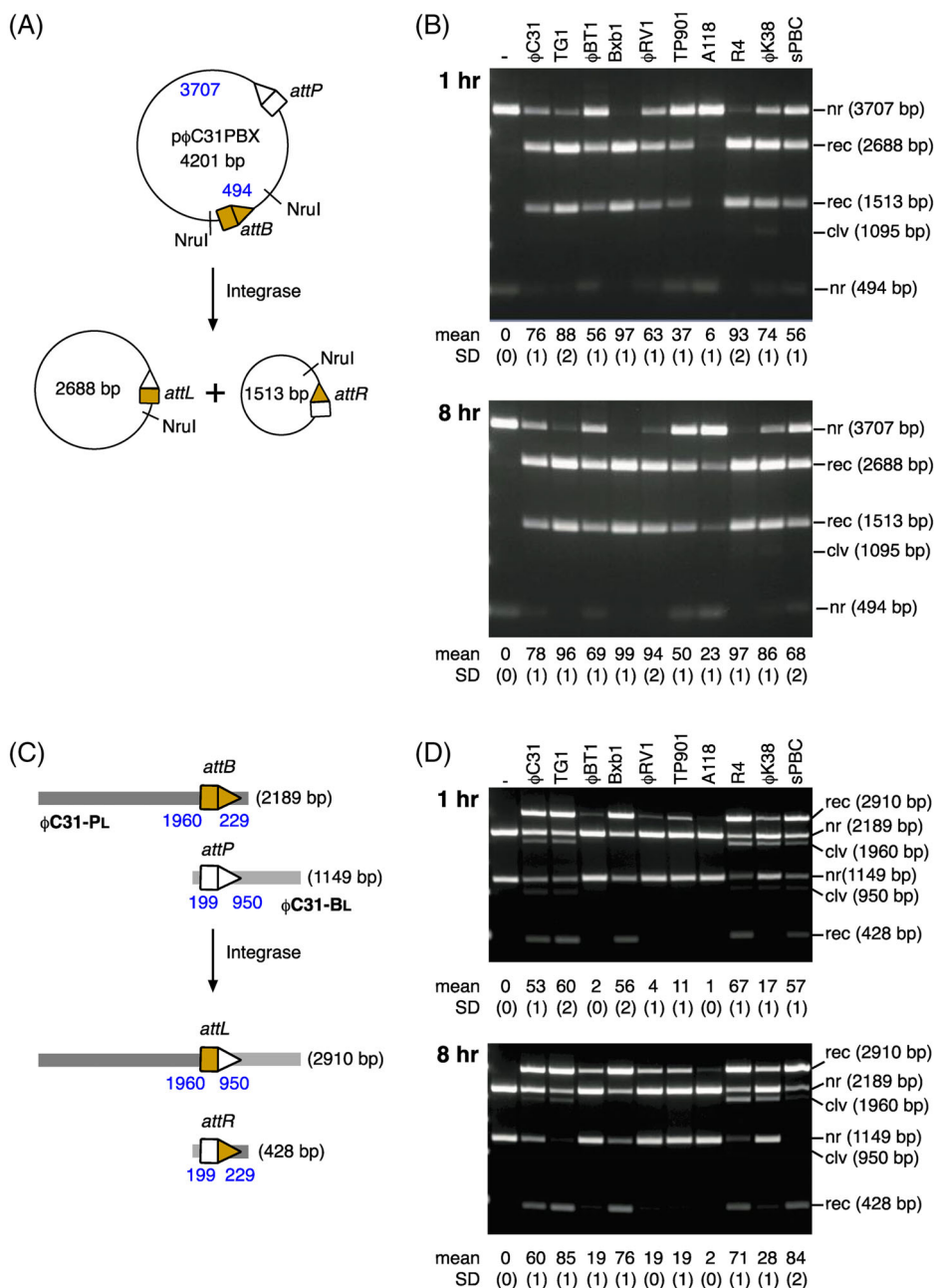


FIGURE 2 in vitro activities of *attP* \times *attB* recombination activities of serine integrases. (A) Scheme illustrating the in vitro intramolecular recombination assay (substrate plasmid p ϕ C31PBX for ϕ C31 integrase is shown; other integrase substrates are similar). The plasmid substrates are named after the large serine integrase (LSI) (ϕ C31) and the *att* sites that will recombine (PBX; *attP* \times *attB* resolution reaction). (B) Intramolecular recombination of *attP* \times *attB* plasmid substrates by selected serine integrases (ϕ C31, TG1, ϕ BT1, Bxb1, ϕ RV1, TP901, A118, ϕ R4, ϕ K38, and SPBC). Reactions were incubated for 1 or 8 h in the reaction buffer described in Section 2. Reaction products were digested with the restriction endonuclease NruI prior to 1.2% agarose gel electrophoresis. The final integrase concentrations were 800 nM in all cases. Top panel: 1-h reactions; Bottom panel: 8-h reactions. The bands on the gel are labeled nr (non-recombinant, i.e., substrate), rec (recombination product), clv (cleavage product). The sizes of the products of the recombination reactions are shown next to each band on the gel image. The mean extent of recombination and standard deviation (%) from quantitation of triplicate experiments are given below each lane. (C) Scheme illustrating the in vitro intermolecular recombination assay using linear DNA substrates (ϕ C31-PL and ϕ C31-BL for ϕ C31 integrase is shown; other integrase substrates are similar). The sizes of the linear DNA substrates and the position of the *att* sites on them are designed to allow clear separation and identification of the reaction substrates and products using agarose gel electrophoresis. The sizes of the products of the recombination reactions are shown next to each band on the gel image. (D) Intermolecular recombination of *attP* \times *attB* linear substrates by selected serine integrases. Reactions were incubated for 1 or 8 h in the reaction buffer described in Section 2. Reaction products were separated by 1.2% agarose gel electrophoresis. The final integrase concentration was 800 nM in all cases. Top panel: 1-h reactions; Bottom panel: 8-h reactions. The bands on the gel are labeled nr (non-recombinant, i.e., substrate), rec (recombination product), clv (cleavage product). The mean extent of recombination and standard deviation (%) from quantitation of triplicate experiments are given below each lane.

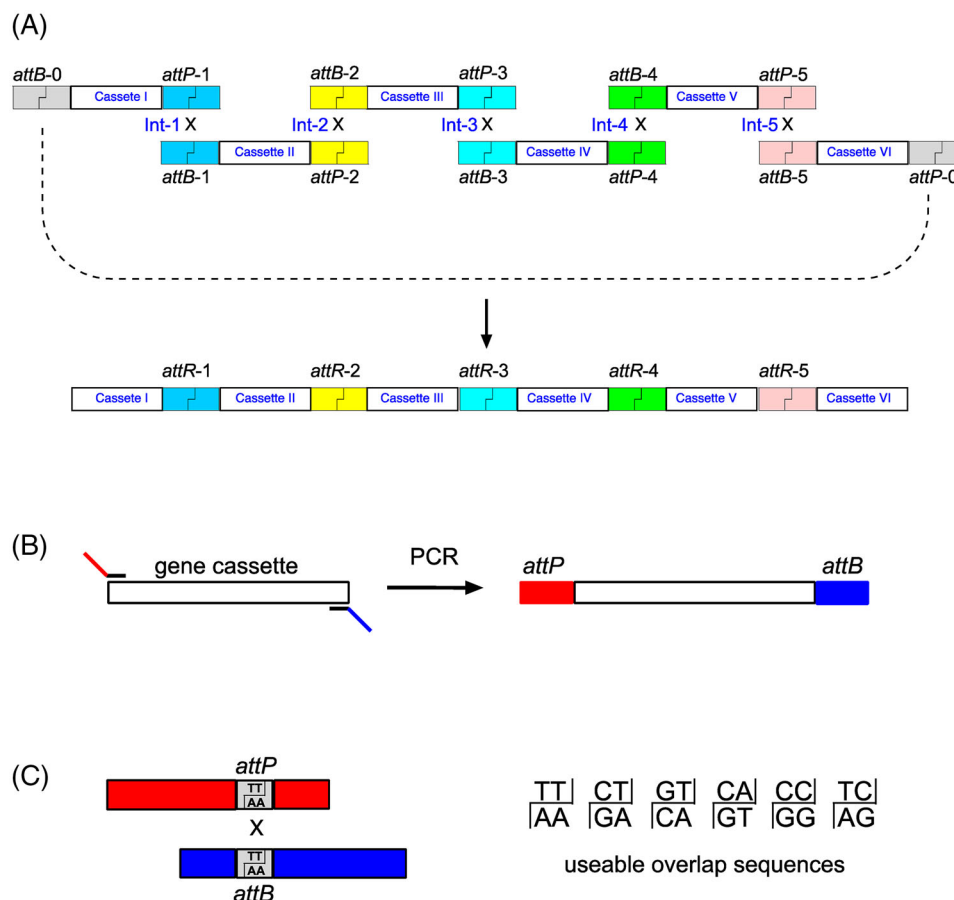


FIGURE 3 One-pot orthogonal serine integrase recombinational assembly. (A) Orthogonal serine integrase recombinational assembly. In this strategy, each serine integrase recombines its cognate pair of *attP* and *attB* sites, resulting in highly specific and defined assembly of gene cassettes. Orthogonal pairs of *attP* and *attB* sites are shown using the same color. Recombination of *attB-0* and *attP-0* will result in formation of a circular product. (B) PCR amplification of gene cassettes used in SIRA reactions. Primers are designed with either *attP* (red) or *attB* (blue) sites added 5' upstream of homologous sequence on the target template. (C) Use of the *att* site central dinucleotide as the determinant of SIRA orthogonality. An LSI recombines its cognate *attB* and *attP* sites only if they have identical central dinucleotide sequences. During recombination, the integrase cuts both strands of the *att* site near the centers to generate a 2-bp overlap. These two basepairs are referred to as the central dinucleotide. The central dinucleotides must match in the *attP/attB* pair for recombination to take place. On the right, the six possible central dinucleotides are shown. LSI, large serine integrase; SIRA, serine integrase recombinational assembly.

multiple integrases to assemble a plasmid containing the genes for the β -carotenoid biosynthetic pathway in a one-pot reaction.^[11,14–18]

2 | EXPERIMENTAL SECTION

2.1 | in vitro plasmid substrates for intramolecular recombination

Substrate plasmids for analysis of intramolecular integrase activity in vitro were constructed by cloning synthetic double-stranded oligonucleotides containing the appropriate *att* sites into pFM122, as described previously.^[19] The sequences of the DNA oligonucleotides used for cloning each *att* site are shown in Supplementary Table S1. In this report, each plasmid was named according to the LSI *att* sites inserted in it, as exemplified in Figure 1A. In this example, pC31PB contained ϕ C31 *attP* and *attB* sites. The *att* sites were in “head to tail” ori-

entation, an arrangement that results in resolution of the substrate plasmid into two smaller circles upon recombination (Figure 2A). The *att* sites in these substrates were ϕ C31 (FM16), TG1 (FM138), ϕ BT1 (FM158), Bxb1 (FM97), ϕ RV1 (FM16), TP901 (FM134), A118 (FM122), R4 (FM157), ϕ K38 (FM159), and SPBc (FM213). The plasmids were available upon request. Supercoiled plasmid DNA was prepared from transformed *E. coli* DH5 cells using a Qiagen miniprep kit according to the manufacturer's instructions. DNA concentrations were estimated by measuring absorbance at 260 nm, and the quality was verified by electrophoresis on 1.2% agarose gels.

2.2 | Linear substrates for in vitro intermolecular recombination

Linear substrates used for intermolecular recombination were prepared by amplifying a 2169 bp (*attB*) and a 1149 bp (*attP*) fragment

from the in vitro plasmid substrate of each LSI (Figure 2C). PCR was performed using high-fidelity Phusion DNA polymerase (New England Biolabs). Primers for PCR were obtained from Eurofins-MWG and Integrated DNA Technologies (IDT). Typical reactions (100 μ l) contained 1 \times high fidelity Phusion polymerase buffer supplemented with additional 0.5 mM MgCl₂, 200 μ M of each deoxynucleotide triphosphate, 3% dimethyl sulfoxide, 0.5 μ M of each primer, 2 ng of plasmid DNA template, and 2 μ l Phusion polymerase. The polymerase chain reactions were started with an initial denaturation step of 98°C for 1 min, followed by 30 cycles of amplification (98°C, 20 s; 55°C, 30 s; and 72°C; 60 s). The process was completed with a final step of 72°C for 10 min. PCR products were purified using a commercial kit (QIAGEN) according to the manufacturer's instructions, with an additional 500 μ l wash with buffer PE. To determine the purity and concentration of the PCR products, 5 μ l of the purified sample were mixed with 15 μ l TE (1 \times) and 5 μ l loading buffer (25 mM Tris-HCl [pH 8.2], 20% w/v Ficoll, 0.5% sodium dodecyl sulfate [SDS], 1 mg/ml protease K, and 0.25 mg/ml bromophenol blue). Reaction products were separated by electrophoresis on 1% agarose gels in Tris-Acetate buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA) at 125 V for 120 min. The gels were stained with ethidium bromide or SYBR Safe and gel images were taken with a Bio-Rad GelDoc scanner. Standard 1 kb ladder was used to estimate fragment sizes. The sequences of primers used for amplifying *att* sites from in vitro plasmid substrates are shown in Supplementary Table S2.

2.3 | Design and preparation of DNA cassettes for SIRA reactions

Amplification of DNA fragments for SIRA was performed by PCR using high-fidelity Phusion DNA polymerase (New England Biolabs). Primers for PCR were obtained from Eurofins-MWG and IDT. PCR reactions and purification/quantification of the products were essentially as described above for preparation of linear recombination substrates.

2.4 | Serine integrase expression vectors

Plasmids for inducible overexpression of serine integrases in *E. coli* were constructed by inserting codon-optimized protein-coding DNA fragments into pET-28a(+) (Novagen), between the NdeI and XhoI sites. The source and accession numbers of the LSIs used in this study are shown in Supplementary Table S3. Synthesized codon-optimized LSI genes were obtained from Thermo Fisher. The proteins expressed from these plasmids have an N-terminal hexahistidine tag (MGSSHH-HHHHSSGLVPRGSHM followed by integrase amino acid 2) to allow purification via nickel affinity chromatography. For Bxb1 integrase overexpression, the coding sequence with a C-terminal polyhistidine tag (TSHHHHHH) was inserted into pSA1101^[20] between the NdeI and Acc65I sites. Each of these plasmids carried a kanamycin resistance gene to allow selection in protein expression cultures.

2.5 | Overexpression and purification of serine integrases

All the proteins used in this work were purified using essentially the same protocol as described in Olorunniji *et al.*^[13] The strain BL21(DE3)pLysS was transformed with the relevant overexpression plasmid (Table S3). The expression strain for each LSI was grown in 2 \times YT-broth at 37°C to mid-log phase and induced with 0.5 mM IPTG, after which the cultures were grown for a further 16 h at 20°C. The proteins were purified by nickel affinity chromatography using a His-Trap FF pre-packed column (GE Healthcare). Fractions containing the protein of interest were dialyzed against Protein Dilution Buffer (PDB; 25 mM Tris-HCl [pH 7.5], 1 mM DTT, 1 M NaCl, and 50% v/v glycerol), and stored at –20°C.

2.6 | in vitro recombination of supercoiled plasmid substrates, and product analysis

Purified integrases were stored at –20°C and diluted to the appropriate concentrations in a buffer that contains 25 mM Tris-HCl (pH 7.5), 1 mM DTT, 1 M NaCl, and 50% v/v glycerol. For recombination reactions using supercoiled plasmid substrates, integrase (4 μ M, 5 μ l) was added to a 40 μ l solution containing the appropriate plasmid substrate (25 μ g/ml), 50 mM Tris-HCl (pH 7.5), 100 μ g/ml BSA, 5 mM spermidine, and 0.1 mM EDTA. Samples were incubated at 30°C for 1 or 8 h, after which the reactions were stopped by heating at 80°C for 10 min. The samples were then cooled and treated with restriction enzymes to facilitate separation and analysis of the reaction products by agarose gel electrophoresis. This was done by mixing a 30 μ l aliquot of the reaction mixture with 28 μ l of B103 buffer (90 mM Tris-HCl pH 7.5, 20 mM MgCl₂) prior to addition of 20 units (2 μ l) NruI (New England Biolabs). The restriction digests were carried out at 37°C for 2 h, then 7.5 μ l loading buffer (25 mM Tris-HCl pH 8.2, 20% w/v Ficoll, 0.5% SDS, 1 mg/ml protease K, and 0.25 mg/ml bromophenol blue) was added. The reaction products were separated by agarose gel electrophoresis, then stained with SYBR safe and visualized as previously described, using a BioRad GelDoc apparatus.^[21,22] The intensities of DNA bands on the gel were quantitated with the volume analysis tool of Image Lab software (BioRad) using automatic band detection. These were manually corrected to exclude artifact bands. The molar proportions of recombinant products were determined by factoring in the fragment size of the DNA in each band.

2.7 | in vitro recombination of linear substrates, and product analysis

For intermolecular recombination reactions, the protocols described above for plasmid substrate recombination were followed with the following differences. Equimolar concentrations of the *attP* and *attB* linear DNA substrates were mixed in the same buffer and reaction conditions

as described above. Recombination reactions were neither heated nor treated with restriction enzymes prior to the addition of the loading buffer. Electrophoresis and visualization were carried out in the same way as described for supercoiled plasmid reactions.

2.8 | Multipart SIRA assembly using serine integrases

In this report, a nomenclature that identified an SIRA plasmid product based on the number of fragments assembled in the SIRA reaction (F) and the number of LSIs (I) used in the assembly reaction was used. For example, two fragments-one integrase (2F1I), three fragments-two integrases (3F2I), and so on. In this project, six different integrases were selected for assembly of the complete carotenoid pathway using six DNA fragments in a one-pot reaction (6F6I; Figure 4B). Supplementary Table S4 shows the DNA templates used for PCR, the sizes of DNA fragments used for 6F6I assembly, and the *att* sites appended to the ends of each fragment. Supplementary Table S5 shows the sequences of the primers used for the PCR steps.

For assembly of a nine-fragment plasmid with three different integrases (9F3I; Figure 5A), Supplementary Table S6 shows the DNA templates used for PCR, the sizes of DNA fragments used, and the *att* sites appended to the ends of each fragment. The sequences of the primers used for PCR are shown in Supplementary Table S7.

Typical SIRA reactions were carried out in a total volume of 50 μ l and contained each DNA fragment (10 nM), 25 mM Tris-HCl (pH 7.5), 0.1 mM EDTA (pH 8), 2 mM DTT, 5 mM spermidine, and 0.1 mg/ml bovine serum albumin (BSA). Recombination reactions were initiated by the addition of 200 nM (5.5 μ l) of each integrase in PDB and incubated at 30°C for 16 h. The assembled DNA was recovered by ethanol precipitation (0.22 volume of 5 M NH_4Ac and 2.5 volume of 100% ethanol). The precipitated DNA was re-dissolved in 20 μ l of double-distilled H_2O , and 5 μ l of the purified DNA was used to transform 50 μ l of electrocompetent *E. coli* DS941 cells.

2.9 | Restriction endonuclease digests

Restriction endonuclease digests were used to analyze SIRA reaction products (Figure 2A,B). Master mixes containing 2 μ l restriction enzyme buffer, 13 μ l distilled H_2O , and 3 μ l enzyme were made, allowing for 2 μ l (250 $\mu\text{g}/\text{ml}$) of plasmid DNA to be added to 18 μ l of Master Mix in individual microcentrifuge tubes. Restriction enzymes PciI, MluI, and NsiI were added to NEBuffer 3.1 (100 mM NaCl, 50 mM Tris-HCl [pH 7.9], 10 mM MgCl_2 , 100 $\mu\text{g}/\text{ml}$ BSA); AgeI to NEBuffer 1.1 (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl_2 , 100 $\mu\text{g}/\text{ml}$ BSA, at pH 7.0); BsrGI to NEBuffer 2.1 (50 mM NaCl, 10 mM Tris-HCl [pH 7.9], 10 mM MgCl_2 , 100 $\mu\text{g}/\text{ml}$ BSA); and BamHI-HF, SpeI, Sall-HF, and RsrII to NEBuffer CutSmart (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 $\mu\text{g}/\text{ml}$ BSA, at pH 7.9). Restriction digests were carried out for 2 h at 37°C, after which the reactions were stopped by the addition of SDS loading buffer. Products of restriction

enzyme digests were analyzed on 1.2% agarose gel electrophoresis as described above.

3 | RESULTS

3.1 | in vitro recombination activities of serine integrases

Several studies have reported the in vitro activities of some LSIs^[3,8,9] and their applications in gene assembly have been documented in other studies.^[11,23–25] These reports demonstrate that some LSIs work very well in vitro, but a definitive comparison of LSI activities under the same recombination conditions is required as a guide to using them for multiplex applications such as SIRA. We therefore compared the activities of 10 LSIs that have been reported in the literature, which were all expressed and purified by the same methods (Section 2). We also standardized the recombination reaction conditions, using a simple buffer which we have used previously and shown to work very well for ϕ C31, Bxb1, and TG1 integrases.^[11,19,26]

Reactions were carried out for 1 h ("short" reactions) and for 8 h ("long" reactions) to assess the relative activities and the nature of reaction products from each LSI on *attP* x *attB* supercoiled plasmid substrates (Figure 2A,B) and linear DNA fragments carrying the *attP* and *attB* sites (Figure 2C,D).

First, we assayed the activities of the integrases using circular supercoiled plasmid substrates in which the *attP* and *attB* sites are arranged in head to tail orientation. This arrangement leads to the deletion of the DNA segment flanked by the *att* sites. The reaction products were treated with restriction enzymes and analyzed by agarose gel electrophoresis, followed by quantification of the bands (Figure 2B). Next, we assayed the recombination activities of the integrases in intermolecular reactions in which the *attP* and *attB* sites are located on separate linear DNA fragments (Figure 2C). These assays give a better indication of the usefulness of the LSIs in SIRA, or any other application involving intermolecular DNA reactions. The reaction products were loaded onto agarose gels without further treatment and separated by electrophoresis (Figure 2D). The mean recombination activities and standard deviation (%) from quantitation of triplicate experiments are shown below each lane.

Unsurprisingly, the different LSIs displayed different levels of activities in the two types of reactions. Among the 10 LSIs we studied, Bxb1, ϕ R4, and TG1 integrases proved to be the most active after 1 h reactions (both intra- and intermolecular). These three enzymes achieved close to complete recombination of the plasmid substrate after 8 h (Figure 2).

The other LSIs display a range of activities, with ϕ C31 integrase approaching the level of intermolecular activities seen with Bxb1, R4, and TG1 after 1-h reaction. ϕ BT1 integrase was efficient in supercoiled plasmid recombination but showed very limited intermolecular recombination activity even after 8 h. In contrast, SPBc integrase performed less efficiently in plasmid recombination, but converted most of the linear substrates into products in the intermolecular reactions.

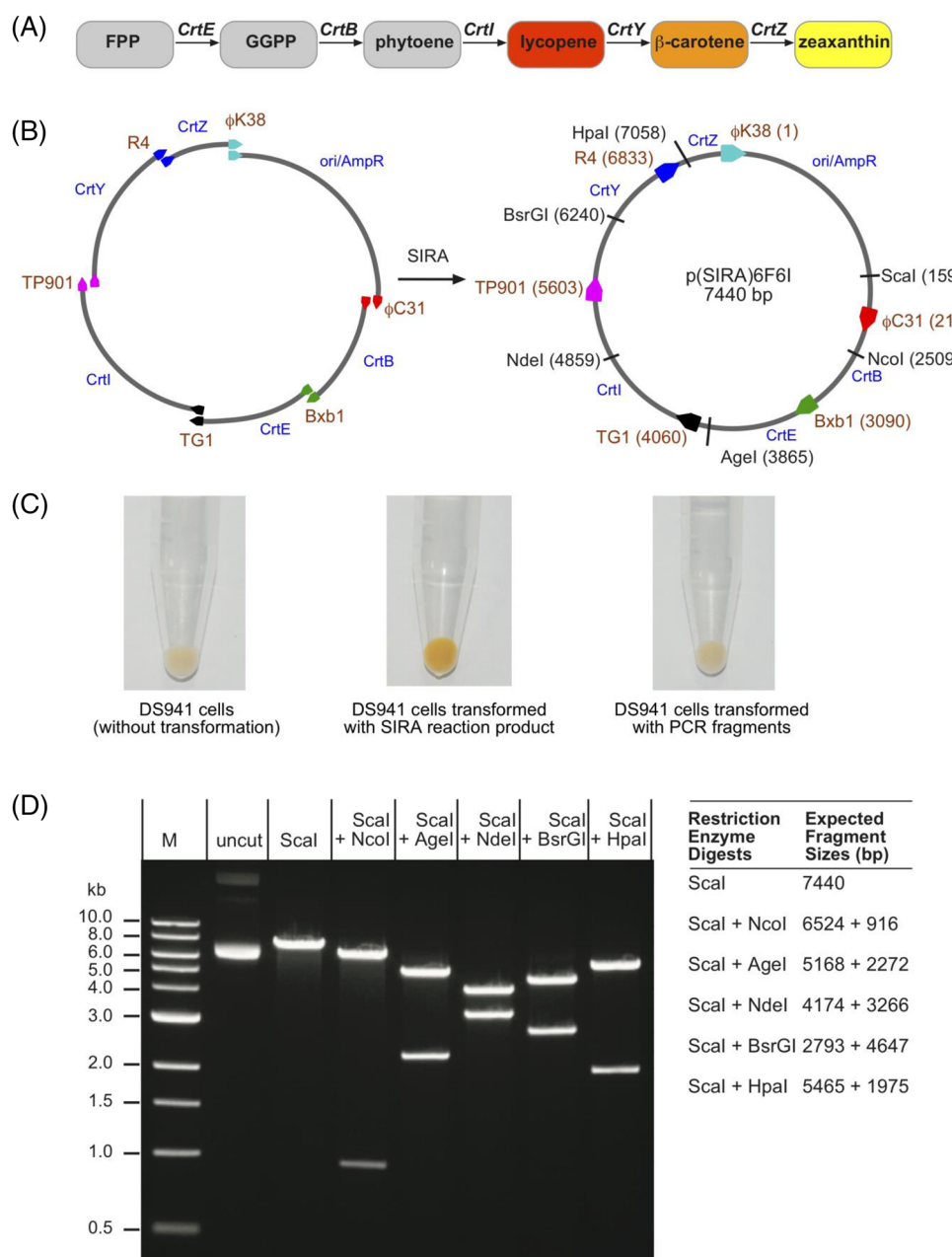


FIGURE 4 in vitro high fidelity one-pot serine integrase recombinational assembly of the carotenoid pathway. (A) Metabolic pathway for the expression of β -carotenoids in *E. coli*. Host genes in *E. coli* produce farnesyl pyrophosphate which is converted into carotenoids by heterologous genes *CrtE*, *CrtB*, *CrtI*, *CrtY*, and *CrtZ*, derived from *P. ananas*. As shown, expression of *CrtE*, *CrtB*, and *CrtI* are required for the expression of lycopene, giving red-colored *E. coli* cells on agar plates and in liquid culture. The production of β -carotene (orange color) and zeaxanthin (yellow color) requires the expression of *CrtY* and *CrtY* + *CrtZ*, respectively. (B) One-pot assembly of a 7.2 kb plasmid p(SIRA)6F6I for the expression of the carotenoid pathway in *E. coli*. Six DNA fragments (all PCR products) were added to the buffered reaction mixture in the presence of ϕ C31, Bxb1, TG1, TP901, ϕ R4, and ϕ K38 integrases. The reactions were incubated for 8 h at 30°C. The DNA cassettes are drawn to scale to reflect the sizes of each gene in the DNA fragment and in the final serine integrase recombinational assembly (SIRA) plasmid product. The unique restriction endonuclease sites used for the analyses in (D) are shown. (C) Cell pellets from cultures of transformants from the SIRA reactions showing expression of the carotenoid pathway products. The first panel shows pellets from cells not transformed with SIRA reaction products; the second shows cells transformed with SIRA reaction products as described above. The pellet shown in the third panel is from cells transformed with a mixture of all six PCR fragments without the integrases. Further details on growth and selection conditions are provided in the text. (D) Restriction endonuclease digest analyses of the SIRA reaction product. The plasmid, p(SIRA)6F6I, was treated with different combinations of restriction endonucleases as described in Section 2. (1) Uncut plasmid, (2) Scal, cuts in pSBA13 vector and to determine plasmid size, (3) Scal + NcoI, cuts in *CrtB*, [4] Scal + AgeI, cuts in *CrtE*, (5) Scal + NdeI, cuts in *CrtI*, (6) Scal + BsrGI, cuts in *CrtY*, (7) Scal + HpaI, cuts in *CrtZ*. Samples were analyzed on a 1.2% agarose gel. The expected sizes of the products of each restriction endonuclease digest are shown in the accompanying table.

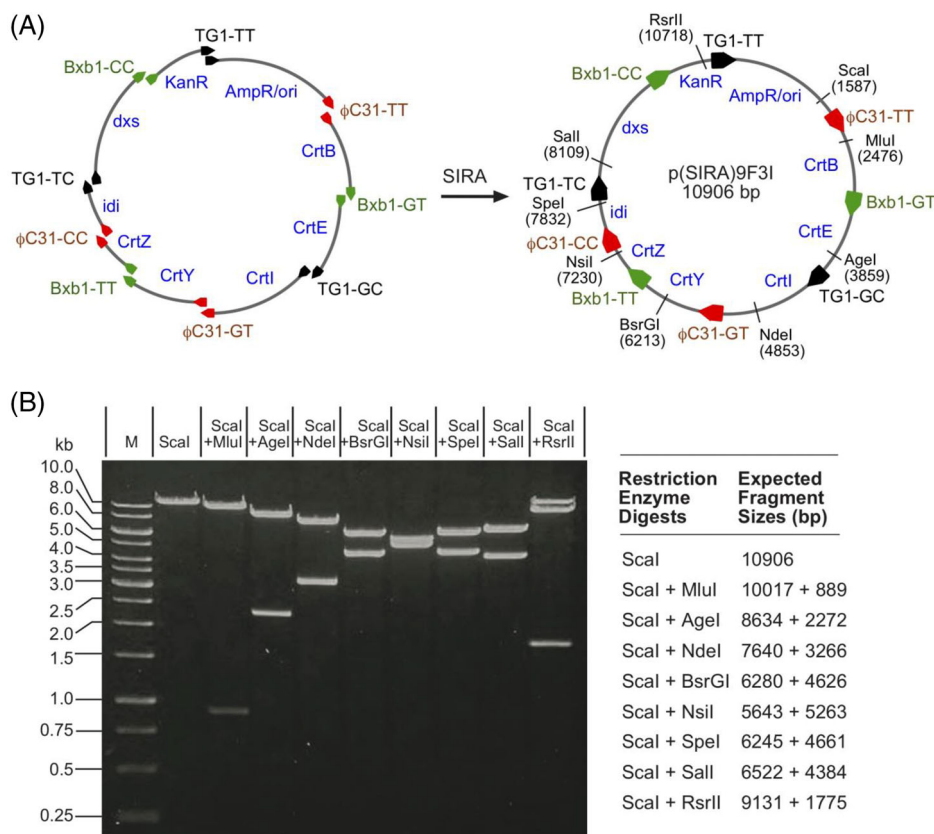


FIGURE 5 One-pot assembly of nine DNA fragments using three orthogonal integrases. (A) One-pot assembly of a 10,906 bp plasmid p(SIRA)9F3I for the expression of the carotenoid pathway in *E. coli*. Reactions were designed and carried out as described in Figure 4. Nine DNA fragments (all PCR products) were added to the buffered reaction mixture in the presence of ϕ C31, Bxb1, and TG1 integrases. The reactions were incubated for 8 h at 30°C. The DNA cassettes are drawn to scale to reflect the sizes of each gene in the DNA fragment and in the final SIRA plasmid product. The unique restriction endonuclease sites used for the analyses in (B) are shown on the right-hand diagram. For each LSI, the three pairs of *attP* and *attB* sites were used. The difference between each pair is the base sequence of the central dinucleotide at the center of the *att* site (see also Figure 1C). (B) Restriction endonuclease digest analyses of the SIRA reaction product. The plasmid, p(SIRA)9F3I, was treated with different combinations of restriction endonucleases as described in Section 2. (1) Scal, SBA13 vector and to determine plasmid size, (2) Scal + MluI, cuts in *CrtB*, (3) Scal + AgeI, cuts in *CrtE*, (4) Scal + NdeI, cuts in *CrtI*, (5) Scal + BsrGI, cuts in *CrtY*, (6) Scal + NsiI, cuts in *CrtZ*, (7) Scal + SpeI, cuts in *idi*, (8) Sall + NsiI, cuts in *dxs*, (9) Scal + RsrII, cuts in *KanR*. Samples were analyzed on a 1.2% agarose gel. The expected sizes of the products of each restriction endonuclease digest are shown in the accompanying table. Incomplete digestion seen in the Scal + RsrII reaction is most likely due to RsrII not working optimally. The full-length linear DNA bands seen are likely due to cleavage by Scal alone since Scal cuts to completion on its own. The same pattern is seen in Supplementary Figures S1–S3.

Other LSIs including ϕ RV1, ϕ K38, and TP901 promoted recombination in both types of assays, with longer (8 h) reactions giving higher levels of recombination. For example, ϕ RV1 integrase achieved 63% formation of recombinant products from the plasmid substrate in the 1-h reaction, but the products increased to 94% after 8 h. The activity of ϕ RV1 integrase in the intermolecular reaction was inefficient, an indication that it may not be very suitable for SIRA reactions. A118 integrase had generally low recombination activity, and very limited intermolecular recombination was seen even after 8 h.

Interestingly, cleaved DNA intermediates were detected in the intermolecular reactions of some LSIs, and these became more pronounced in the 8-h reaction. The sizes of the cleaved DNA intermediates correspond to cleavage at the respective *attP* and *attB* sites (Figure 2). While this may be due to some mechanistic features of the LSIs involved, it could also be because of disparity between the con-

centrations of functionally available *attP* and *attB* sites present in the reactions.

3.2 | One-pot assembly of the carotenoid pathway genes from six DNA fragments using six LSIs

Based on our analyses of the in vitro activities of the LSIs on linear DNA substrates, and *att* site sequence analysis to maximize orthogonal interactions, we chose ϕ C31, Bxb1, TG1, TP901, ϕ R4, and ϕ K38 integrases to synthesize a six-fragment plasmid that carries the genes required for the expression of the carotenoid pathway in *E. coli*, in a one-pot assembly reaction (Figure 4A,B). Previously, we used ϕ C31 integrase to assemble the carotenoid pathway, using the central dinucleotides of the *att* sites as the basis for arranging the order of gene

cassettes.^[11] In this study, we aimed to improve on the technology in two ways: (1) to build the plasmid vector entirely from DNA fragments derived from PCR, and (2) to order the assembly using multiple LSIs and their cognate *att* sites, rather than variation of the central dinucleotides, and thereby to suppress the formation of products with missing or wrongly ordered fragments.

DNA fragments were prepared by PCR as described in Section 2 and purified using a QIAGEN PCR clean-up kit. The PCR products were quantified by spectrophotometry and stored at -20°C in 10 mM Tris-HCl buffer until they were used in the SIRA reactions. The DNA fragments were added together in equimolar concentrations and the SIRA reaction was initiated by adding optimal concentrations of the six LSIs. The reaction was incubated at 30°C for 16 h. The reaction products were then used to transform *E. coli* DS941 cells.^[20,27] Cells were cultured to allow recovery for 2 h at 37°C and spread on L-agar plates with ampicillin selection. The plates were cultured for further 16 h at 37°C . The expected plasmid product is referred to as “p(SIRA)6F6I” indicating six fragments and six integrases are used for its assembly (Figure 4B).

In one representative experiment, ~2000 colonies were recovered with all showing the full yellow color expected when the cells express the complete carotenoid pathway. In contrast, DS941 cells grown without transformation did not show any yellow color. In the second control experiment, a mixture of the SIRA reaction components including the buffer and DNA fragments but without the integrases was used to transform DS941 cells. The transformed cells were cultured for 2 h and spread on plates with and without ampicillin. The cells spread on ampicillin plates did not grow, an indication that no self-replicating plasmid was formed. As expected, growth was observed on the plate without antibiotics but no colored colonies were observed. Individual yellow colonies were picked from the plates and used to inoculate media with ampicillin selection. The cultures were spun down, and the pellets (Figure 4C) showed the expected yellow color of cells expressing the carotenoid pathway genes.

Next, we randomly selected 20 colonies from the SIRA reaction plate and grew them in L-broth containing ampicillin. Plasmid DNA was prepared from the 20 cultures and analyzed by restriction digest to confirm that all the plasmids were of the right size (data not shown). Next, 10 of the plasmids were analyzed further to confirm the exact position of each gene fragment in the plasmid. Figure 4D shows an example of the restriction site mapping used to ascertain that all six DNA cassettes are arranged in the plasmid as intended by the positioning of the *att* sites. The results show that all six cassettes were assembled in the correct order in all 10 of the chosen plasmids.

3.3 | One-pot assembly of an enhanced carotenoid pathway from nine-fragments using three “orthogonal” integrases

To assess the expandability of our multi-LSI strategy by combining it with differentiation of the *att* sites by variation of their central dinucleotides, we attempted to assemble nine fragments using three highly

active LSIs (ϕC31 , Bxb1, and TG1 integrases). In this approach, we aimed to assemble more fragments with fewer LSIs. Each integrase needs to join three pairs of *att* sites. Hence, *att* sites with different (orthogonal) central dinucleotides were used to impose specificity on this joining (Figure 5A).^[11] We selected ϕC31 , Bxb1, and TG1 integrases for this assembly reaction since we know that the identity of the central dinucleotide does not affect their recombination activities. Some LSIs have reduced activity when the native central dinucleotide sequence of the *att* sites is replaced with certain non-cognate dinucleotides (data not shown).

The assembly included fragments encoding the five genes for carotenoid biosynthesis derived from *Pantoea ananatis* (*crtE*, *crtB*, *crtI*, *crtY*, and *crtZ*; see above), and three *E. coli* endogenous genes (*dxs* and *idi*) that are known to increase the production of precursor metabolites needed for the β -carotenoid pathway,^[11,28] along with a plasmid vector fragment. Hence, the nine DNA fragments consist of the seven genes that comprise the zeaxanthin biosynthetic pathway,^[11,29] a kanamycin resistance gene (to provide a more stringent selection), and the pSB1A3 biobrick vector carrying the origin of replication and β -lactamase gene (ampicillin resistance). All nine DNA fragments were prepared as PCR products with the desired integrase *att* sites appended to their ends (Figure 3B) as described earlier for the construction of the six-fragment plasmid (Figure 4). The expected plasmid product is referred to as “p(SIRA)9F3I”; its architecture is shown in Figure 5A.

The DNA fragments were added together in equimolar concentrations and the SIRA reaction was initiated with the addition of the three integrases. The reaction was incubated at 30°C for 16 h, and the reaction product was used to transform *E. coli* DS941 cells. Transformants were selected on plates containing ampicillin and kanamycin. From one transformation experiment, a total of 183 colonies were recovered from one plate, of which 45 (25%) were brightly yellow-colored (an indication of a fully functional carotenoid biosynthesis pathway). The remaining 138 were a mixture of colorless colonies and some colonies showing red and orange colors. These color variations indicate that some of the gene cassettes required to produce the bright yellow color for zeaxanthin production were absent from the assembled plasmid in the transformed *E. coli* colonies. Ten of the bright yellow colonies were picked and grown in liquid media with ampicillin and kanamycin selection. All the cultures retained their bright yellow color. Plasmid DNA samples were prepared from the cultures and restriction digest analyses were used to determine the order in which the DNA cassettes are arranged in the assembled plasmids. The results show that all 10 plasmid clones analyzed were assembled in the correct order, giving the exact plasmid map predicted (see Figure 5A). Figure 5B shows the restriction site mapping for one of the 10 plasmid clones analyzed.

We also analyzed some of the colonies which grew on plates containing ampicillin and kanamycin, but which did not express the complete carotenoid pathway (as judged by lack of the bright yellow color due to zeaxanthin production). We hypothesized that fragments could have been omitted from the assembly due to recombination of *att* sites recognized by the same LSI but with non-matching central dinucleotides, as was previously observed.^[11] We selected 20 such

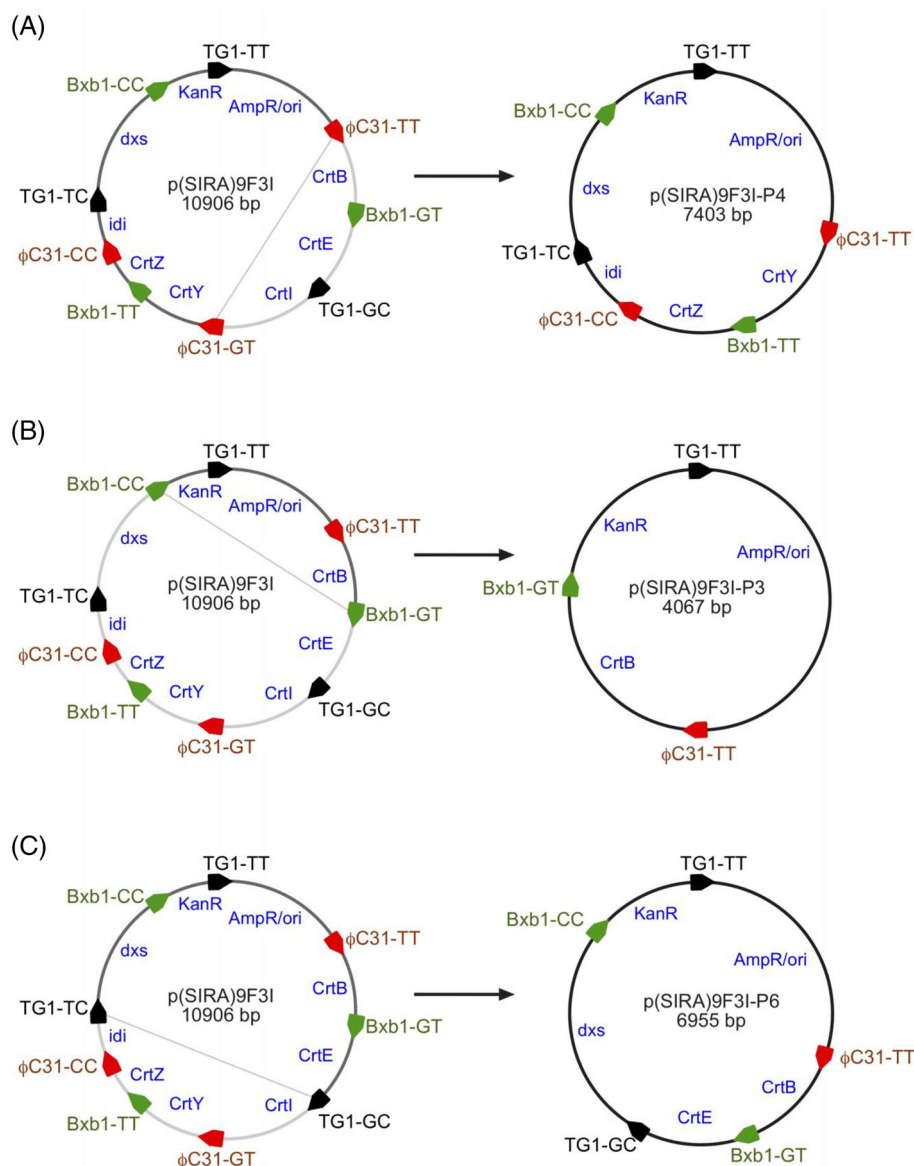


FIGURE 6 Serine integrase recombinational assembly (SIRA) reactions between *att* sites with “incompatible” central dinucleotides resulting in plasmid products with missing gene fragments. Plasmid DNA samples prepared from colonies that did not show the color characteristic of the full carotenoid pathway were analyzed as described in Figure 5. The results are shown in Supplementary Figures S1–S3 and are summarized diagrammatically here. (A) Product of ϕ C31 integrase recombination between *attP*-TT and *attB*-GT, leading to the exclusion of *CrtB*, *CrtE*, and *CrtI* cassettes from the SIRA product containing the plasmid origin of replication and antibiotic resistance marker. (B) Product of Bxb1 integrase recombination between *attP*-GT and *attB*-CT, leading to the loss of all components of the carotenoid pathway except *CrtB*. (C) Product of TG1 integrase recombination between *attP*-CC and *attB*-TC, resulting in the omission of *CrtI*, *CrtY*, *CrtZ*, and *idi* cassettes from the plasmid product. The product diagrams are not drawn to scale.

colonies, and grew them in L-broth with ampicillin and kanamycin selection. Plasmid DNA samples were prepared from all 20 cultures and subjected to restriction site mapping to determine the plasmid sizes and the fragments that were present in each construct. In all cases, some fragments were missing. These omissions could be explained by joining of *attP* and *attB* sites of the same integrase but with different central dinucleotides, resulting in exclusion of the DNA cassettes expected to be located between the two recombining sites. Figure 6 illustrates examples of such non-orthogonal recombination involving each of the three integrases used. The restriction mapping

data for the three examples shown in Figure 6 are in Supplementary Figures S1–S3.

In these assembly experiments, no evidence was found of recombination between *att* sites of two different integrases. Note however that any product that does not contain the *AmpR/ori* and *KanR* cassettes would not be detected due to the ampicillin and kanamycin in the plates used to select transformants. We have in vitro evidence that TG1 integrase can recombine its own *attB*-TT and ϕ C31 *attP*-TT (data not shown), but such products would not be observed because the *KanR* gene cassette would be missing. This shows how more stringent

selection conditions built into SIRA constructs can minimize or prevent selection of “unwanted” or “partial products.”

4 | DISCUSSION

4.1 | LSIs display a range of recombination activities in vitro

Our comparison of LSI activities in vitro provides a guide to which enzymes are optimal for in vitro applications. Several studies have identified Bxb1 integrase as a highly active LSI when compared with a few other ones that have been studied and have shown that it is highly effective in mediating recombination in mammalian systems.^[30–32] Chao *et al.*^[33] show that there is a wide variability in the activities of LSIs in vivo. The authors attempted a formal characterization of LSI activities in HEK293 cells using a combination of experimental data and mathematical models. Jusiak *et al.*^[34] compared the efficiency of Bxb1, Wβ, and ϕC31 integrases for transgenic integration into CHO cells, and showed that Bxb1 integrase was the most efficient of these three systems. In addition, they showed that the Bxb1 *att* site GA central dinucleotide was more efficient than the wild type GC. These studies show the potential of using LSIs for in vitro applications, but no comparative analysis of many LSI in vitro activities has been done. The work reported here presents a comprehensive comparison of the activities in vitro of the most popularly used LSIs.

In most cases, the efficiency of intramolecular activity (observed using plasmid substrates) and intermolecular recombination (between two linear DNA molecules) was correlated, but there were a few exceptions. For example, ϕRV1 integrase recombined 94% of the plasmid substrate after 8 h, but only 19% conversion of the linear substrates was achieved over the same reaction time. The differences might be partially due to the fact that the plasmids are supercoiled. DNA supercoiling might assist synapsis of sites, thereby increasing the rate of recombination.^[35] The LSIs use a highly conserved catalytic domain; hence, it is less likely that large differences in activities could be due to inherent variation in the rate of phosphoryl transfer reactions involved in recombination. The observed differences might be due to the characteristics of the *att* site sequences and the different affinities of the corresponding DNA-recognition domains in each LSI.^[36–40] A better understanding of the determinants of LSI-*att* site interactions may help in optimizing LSI activities, either by engineering them for better interaction with *att* sites or by redesigning *attP/attB* sites for low-activity integrases. It is still unclear to what extent natural *attP* and *attB* sequences are optimal for LSI activity,^[39–45] and it is likely that phages integrate into non-optimal *attB* sites in some organisms. Recent studies have suggested ways through which the *att* sites of LSIs could be redesigned for improved activities by changing the spacing between DNA sections of the *att* site recognized by specific integrase domains,^[41,46,47] but the universality of these ideas is yet to be tested across a wide range of LSIs. Also, optimization of reaction buffer might be needed for each LSI.

Some LSI reactions accumulated non-recombined cleavage products, most noticeably with ϕR4, ϕK38, ϕC31, and TG1 integrases. While this seems to be a characteristic of more active LSIs, no such cleaved reaction intermediates were seen in the Bxb1 integrase reaction. It is also more noticeable with the linear substrates, which suggests that intermolecular recombination could promote accumulation of cleaved intermediates, and/or DNA supercoiling leads to better coordination of the recombination steps. One possible explanation for the accumulation of cleaved intermediates could be differences in the affinity of the LSI for the *attP* and *attB* sites^[40] such that depletion of one substrate may lead to a build-up of cleaved intermediates of the other.

Taken together, we suggest that Bxb1 integrase is the best LSI for typical in vitro applications, because of its high activity and efficiency of recombination product formation. Nevertheless, some other LSIs (TG1 and ϕR4 integrases) also displayed high recombination activity, achieving maximum conversion to products after 1-h reaction. Their rapid and complete conversion of the substrate make these LSIs excellent candidates for SIRA and other in vitro applications. While several reports have identified Bxb1 integrase as an efficient LSI, both in vivo and in vitro,^[24,30,31,48] less attention has been drawn to the comparable activities and efficiency of R4 and TG1 integrases as seen in this study. In contrast, recombination by ϕC31, ϕBT1, ϕK38, and SPBc integrases was slower, with more product after 8 h of reaction than after 1 h.

4.2 | Orthogonal LSIs deliver high fidelity SIRA in a one-pot reaction

We used a 6-LSI SIRA system to assemble a 7.5-kb plasmid from six PCR products. This demonstrates the powerful potential of using LSIs for specific and predictable assembly of DNA fragments in vitro. We have demonstrated that use of multiple orthogonal LSIs, rather than variation of *att* site central dinucleotides, provides the most reliable approach to SIRA, and by projection other in vitro applications where orthogonality and precision activities are required. Although fewer colonies were recovered in our experiments compared to systems in which a single LSI (ϕC31 integrase) was used,^[11] analysis of the constructs show up to 100% accuracy of the samples analyzed, in contrast to the previously published results.^[11] As highlighted above, the process could be optimized by developing a more universal recombination reaction condition for the LSIs and the most optimal ratio of *attB/attP* for each LSI.

Recently, several groups have developed different ways of using multiple LSIs for DNA manipulation in vivo,^[49] for example, using TG1, SV1, and ϕBT1 integrases to mediate multiplex conjugation and integration of plasmids from *E. coli* into *Streptomyces*. Fayed *et al.*^[50] used multiple integrases for in vivo assembly of the polyketide synthase genes to synthesize erythromycin in a streptomyces strain. More recently, Gao *et al.*^[51] used a combination of serine integrase assembly and In-Fusion cloning to assemble complex pathways expressing secondary metabolic products in *Streptomyces*. A key advancement in the work reported here is that all the starting materials for the plasmid

assembly are derived from PCR and no additional cloning technology was used in the one-pot SIRA reaction.

Tomimatsu *et al.*^[52] used Bxb1, ϕ R4, and TP901 integrases to mediate multiple cassette exchange on a mouse chromosome. Some in vivo applications have used central dinucleotides as the basis for specific DNA assembly. Blanco-Redondo and Langenhan^[53] used two different central dinucleotides to orthogonally direct transgenes into separate *att* sites in *Drosophila*. In another system termed serine integrase recombinational engineering (SIRE), Snoeck *et al.*^[54] combined the specificity of ϕ C31 integrase with recombinase-mediated cassette exchange (RMCE) to develop a tool for genome editing in bacteria. Inniss *et al.*^[55] used CRISPR/Cas9 to insert a landing pad for Bxb1 integrase into a CHO cells for the purpose of protein expression and production. These studies show an increasing appreciation of the specific advantages of using LSIs for unidirectional transgene insertion.

SIRA involving the use of LSIs is currently limited by the availability of highly efficient and orthogonal LSIs, so further characterization of LSIs in vitro is needed to enhance the toolkit. As we gain better understanding of how LSIs interact with *att* sites, we might be able to re-design the sites for some of the already characterized but less active LSIs to achieve more efficient recombination.

It is noteworthy that there are other well-known methodologies for achieving ordered DNA assembly reactions. For example, recombineering exploits homologous recombination systems for in vivo DNA assembly and has been used for effectively inserting genes into bacterial genomes,^[56–58] and Gibson Assembly utilizes three complementary enzymatic properties in a one-pot reaction (5' exonuclease activity, DNA polymerase 3' extension activity, and DNA ligase activity) to achieve DNA assembly in vitro.^[59,60] The main advantage of SIRA over these alternative methodologies is the ease with which individual DNA cassettes can be removed and replaced easily without the need to rebuild the entire construct. SIRA would be the DNA assembly of choice in systems where flexibility in editing modular constructs is desirable. The DNA assemblies made from SIRA can be modified by the same LSIs using their corresponding RDFs.^[2,11,13] Another advantage of using LSIs is that they deal with assembly of very large DNA fragments in contrast to cloning methods such as Gibson assembly.^[48,53]

We also show that an SIRA system using 3 LSIs together with *att* sites varying in their central dinucleotides leads to successful in vitro assembly of a 10-kb plasmid from nine PCR product parts. Careful selection of highly active and “reaction-compatible” LSIs together with the use of 2 or 3 *att* site pairs with orthogonal central dinucleotides could provide a solution to building large constructs using just a handful of LSIs.

4.3 | Central dinucleotide specificity is not a strong determinant of orthogonality

Analysis of plasmid DNA from clones in our p(SIRA) 9F3I assembly experiment that did not express the full carotenoid biosynthesis pathway shows that the LSIs likely recombined *attP* and *attB* sites with

different central dinucleotides to give incompletely assembled plasmids (Figure 6 and Supplementary Figures S1–S3). In contrast, no example of “illegitimate” joining of *att* sites for two different integrase sites was observed, suggesting that orthogonality among the chosen LSIs with their cognate *att* sites is more complete. These findings suggest that more work needs to be done to identify central dinucleotide pairs that are most suitable for use as determinants of orthogonality. Such an approach would characterize the central dinucleotide performance for each LSI, and the tendency toward non-orthogonality. The information obtained could inform design when *att* site central dinucleotides are to be used as the basis of orthogonality in multipart SIRA reactions.

The findings from this work emphasize the need for a larger pool of orthogonal LSIs that can be used for in vitro applications. This would avoid the loss of fidelity and reduced yield that we observed when central dinucleotides are used as basis for orthogonal assembly. The availability of a large palette of orthogonal LSIs is becoming more realistic as more of these enzymes are being isolated and characterized using computational genome mining tools.^[6,61,62]

5 | CONCLUSION

LSIs will continue to gain increasing attention as tools for synthetic biology, and specifically genome editing. In this study, we have expanded our toolkits of purified LSIs and shown how they can be used for in vitro applications. We have also demonstrated how multiple LSIs that are completely orthogonal to each other allow for streamlined and highly predictable gene assembly using SIRA, and that combining LSI orthogonality and central dinucleotide orthogonality can significantly scale up multi-part gene assembly. The LSIs described here can further be developed to achieve better activities in vitro, either by re-designing their *att* sites or engineering their DNA recognition domains. Identification and characterization of more LSI RDFs will also enhance the flexibility of these systems for use in precision and programmable DNA rearrangements.^[8,9,22,26,61,62]

AUTHOR CONTRIBUTIONS

Investigation: Jumai Abioye, Makeba Lawson-Williams, Alicia Lecanda, Brecken Calhoon, Arlene McQue. *Conceptualization:* Sean Colloms, Marshall Stark, Femi Olorunniji. *Formal analysis:* Femi Olorunniji. *Project administration:* Marshall Stark. *Supervision:* Marshall Stark. *Methodology:* Femi Olorunniji. *Writing—original draft:* Sean Colloms, Marshall Stark, Femi Olorunniji. *Writing—review & editing:* Sean Colloms, Marshall Stark, Femi Olorunniji.

ACKNOWLEDGMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council (grant number BB/003356/1) and Medical Research Council Proximity (MC PC 16077).

CONFLICT OF INTEREST

The authors declare no commercial or financial conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. In addition, data that supports the findings of this study are available in the supplementary material of this article. The amino acid sequences of the integrases reported in this work are publicly available at <https://www.ncbi.nlm.nih.gov/protein/>. The accession numbers are shown in the supplementary information.

ORCID

Femi J. Olorunniji  <https://orcid.org/0000-0001-9389-2981>

REFERENCES

- Smith, M. C. M. (2015). Phage-encoded serine integrases and other large serine recombinases. In N. L. Craig, M. Chandler, M. Gellert, A. Lambowitz, P. A. Rice, & S. B. Sandmeyer (Eds.), *Mobile DNA III* (pp. 253–272). ASM Press.
- Stark, W. M. (2017). Making serine integrases work for us. *Current Opinion in Microbiology*, 38, 130–136.
- Olorunniji, F. J., Rosser, S. J., & Stark, W. M. (2016). Site-specific recombinases: Molecular machines for the Genetic Revolution. *Biochemical Journal*, 473, 673–684.
- Bonnet, J., Subsoontorn, P., & Endy, D. (2012). Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proceedings of the National Academy of Science USA*, 109, 8884–8889.
- Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P., & Endy, D. (2013). Amplifying genetic logic gates. *Science*, 340, 599–603.
- Yang, L., Nielsen, A. A. K., Fernandez-Rodriguez, J., McClune, C. J., Laub, M. T., Lu, T. K., & Voigt, C. A. (2014). Permanent genetic memory with >1-byte capacity. *Nature Methods*, 11, 1261–1266.
- Roquet, N., Soleimany, A. P., Ferris, A. C., Aaronson, S., & Lu, T. K. (2016). Synthetic recombinase-based state machines in living cells. *Science*, 353, aad8559.
- Fogg, P. C. M., Colloms, S. D., Rosser, S. J., Stark, W. M., & Smith, M. C. M. (2014). New applications for phage integrases. *Journal of Molecular Biology*, 426, 2703–2716.
- Merrick, C. A., Wardrope, C., Paget, J. E., Colloms, S. D., & Rosser, S. J. (2016). Rapid optimization of engineered metabolic pathways with serine integrase recombinational assembly (SIRA). *Methods in Enzymology*, 575, 285–317.
- Haimovich, A. D., Muir, P., & Isaacs, F. J. (2015). Genomes by design. *Nature Reviews Genetics*, 16, 501–516.
- Colloms, S. D., Merrick, C. A., Olorunniji, F. J., Stark, W. M., Smith, M., Osbourn, A., Keasling, J. D., & Rosser, S. J. (2014). Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. *Nucleic Acids Research*, 42, e23–e23.
- Olorunniji, F. J. (2021). Site-specific recombination: Biological functions, reaction mechanisms, and applications. In J. Joseph (ed.), *Encyclopedia of biological chemistry* (3rd ed., vol. 4, pp. 170–180).
- Olorunniji, F. J., Merrick, C., Rosser, S. J., Smith, M. C. M., Stark, W. M., & Colloms, S. D. (2017). Multipart DNA assembly using site-specific recombinases from the large serine integrase family. In N. Eroshenko (ed.), *Site-specific recombinases*, Methods in molecular biology (pp. 303–323). Springer Science & Business Media.
- Nishizaki, T., Tsuge, K., Itaya, M., Doi, N., & Yanagawa, H. (2007). Metabolic engineering of carotenoid biosynthesis in *Escherichia coli* by ordered gene assembly in *Bacillus subtilis*. *Applied and Environmental Microbiology*, 73, 1355–1361.
- Larroude, M., & Nicaud, J.-M. (2022). Golden gate multigene assembly method for *Yarrowia lipolytica*. *Methods in Molecular Biology*, 2513, 205–220.
- Qi, M., Zhang, B., Jiang, L., Xu, S., Dong, C., Du, Y.-L., Zhou, Z., Huang, L., Xu, Z., & Lian, J. (2021). PCR & Go: A pre-installed expression chassis for facile integration of multi-gene biosynthetic pathways. *Frontiers in Bioengineering and Biotechnology*, 8, 613771.
- Xu, X., Tian, L., Tang, S., Xie, C., Xu, J., & Jiang, L. (2020). Design and tailoring of an artificial DNA scaffolding system for efficient lycopene synthesis using zinc-finger-guided assembly. *Journal of Industrial Microbiology and Biotechnology*, 47, 209–222.
- Naseri, G., Behrend, J., Rieper, L., & Mueller-Roeber, B. (2019). COM-PASS for rapid combinatorial optimization of biochemical pathways based on artificial transcription factors. *Nature Communications*, 10, 2615.
- Olorunniji, F. J., Buck, D. E., Colloms, S. D., McEwan, A. R., Smith, M. C., Stark, W. M., & Rosser, S. J. (2012). Gated rotation mechanism of site-specific recombination by ϕ C31 integrase. *Proceedings of the National Academy of Science USA*, 109, 19661–19666.
- Arnold, P. H., Blake, D. G., Grindley, N. D. F., Boocock, M. R., & Stark, W. M. (1999). Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *The EMBO Journal*, 18, 1407–1414.
- Olorunniji, F. J., He, J., Wenwieser, S. V. C. T., Boocock, M. R., & Stark, W. M. (2008). Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Research*, 36, 7181–7191.
- Olorunniji, F. J., Lawson-Williams, M., McPherson, A. L., Paget, J. E., Stark, W. M., & Rosser, S. J. (2019). Control of ϕ C31 integrase-mediated site-specific recombination by protein trans-splicing. *Nucleic Acids Research*, 47, 1142–11460.
- Zhang, L., Zhao, G., & Ding, X. (2011). Tandem assembly of the epothilone biosynthetic gene cluster by in vitro site-specific recombination. *Scientific Reports*, 1, e141.
- Wang, X., Tang, B., Ye, Y., Mao, Y., Lei, X., Zhao, G., & Ding, X. (2016). Bxb1 integrase serves as a highly efficient DNA recombinase in rapid metabolite pathway Assembly. *Acta Biochimica et Biophysica Sinica*, 49, 44–50.
- Ba, F., Liu, Y., Liu, W.-Q., Tian, X., & Li, J. (2022). SYMBIOSIS: Synthetic manipulable biobricks via orthogonal serine integrase systems. *Nucleic Acids Research*, 50, 2973–2985.
- Olorunniji, F. J., McPherson, A. L., Rosser, S. J., Smith, M. C. M., Colloms, S. D., & Stark, W. M. (2017). Control of serine integrase recombination directionality by fusion with the directionality factor. *Nucleic Acids Research*, 45, 8635–8645.
- Anthony, R. (2003). Transformation of competent bacterial cells: Electroporation. In eLS. John Wiley & Sons Ltd. <http://www.els.net> [<https://doi.org/10.1038/npg.els.0003749>]
- Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K., & Harashima, K. (1990). Elucidation of the *Erwinia ure-dovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. *Journal of Bacteriology*, 172, 6704–6712.
- Kim, S. W., & Keasling, J. D. (2001). Metabolic engineering of the non-mevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnology and Bioengineering*, 72, 408–415.
- Russell, J., Chang, D., Tretiakova, A., & Padidam, M. (2006). Phage Bxb1 integrase mediates highly efficient site-specific recombination in mammalian cells. *Biotechniques*, 40, 460–464.
- Xu, Z., Thomas, L., Davies, B., Chalmers, R., Smith, M., & Brown, W. (2013). Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome. *BMC Biotechnology*, 13, e87.
- Duportet, X., Wroblewska, L., Guye, P., Li, Y., Eyquem, J., Rieders, J., Rimchala, T., Batt, G., & Weiss, R. (2014). A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Research*, 42, 13440–13451.
- Chao, G., Travis, C., & Church, G. (2021). Measurement of large serine integrase enzymatic characteristics in HEK293 cells reveals variability and influence on downstream reporter expression. *The FEBS Journal*, 288, 6410–6427.

34. Jusiak, B., Jagtap, K., Gaidukov, L., Duportet, X., Bandara, K., Chu, J., Zhang, L., Weiss, R., & Liu, T. K. (2019). Comparison of integrases identifies Bxb1-ga mutant as the most efficient site-specific integrase system in mammalian cells. *ACS Synthetic Biology*, 8, 16–24.
35. Watson, M. A., Boocock, M. R., & Stark, W. M. (1996). Rate and selectivity of synapsis of res recombination sites by Tn3 resolvase. *Journal of Molecular Biology*, 257, 317–326.
36. Thorpe, H. M., Wilson, S. E., & Smith, M. C. (2000). Control of directionality in the site-specific recombination system of the *Streptomyces* phage phic31. *Molecular Microbiology*, 38, 232–241.
37. Adams, V., Lucet, I. S., Tynan, F. E., Chiarezza, M., Howarth, P. M., Kim, J., Rossjohn, J., Lyras, D., & Rood, J. I. (2006). Two distinct regions of the large serine recombinase TnpX are required for DNA binding and biological function. *Molecular Microbiology*, 60, 591–601.
38. Ghosh, P., Pannunzio, N. R., & Hatfull, G. F. (2005). Synapsis in phage bxb1 integration: Selection mechanism for the correct pair of recombination sites. *Journal of Molecular Biology*, 349, 331–348.
39. Mandal, S., Dhar, G., Avliyakov, N. K., Haykinson, M. J., & Johnson, R. C. (2013). The site-specific integration reaction of Listeria phage A118 integrase, a serine recombinase. *Mobile DNA*, 4, 2.
40. Li, H., Sharp, R., Rutherford, K., Gupta, K., & Van Duyne, G. (2018). Serine integrase attP binding and specificity. *Journal of Molecular Biology*, 430, 4401–4418.
41. Van Duyne, G. D., & Rutherford, K. (2013). Large serine recombinase domain structure and attachment site binding. *Critical Reviews in Biochemistry and Molecular Biology*, 48, 476–491.
42. Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S., & Calos, M. P. (2001). Site-specific genomic integration in mammalian cells mediated by phage phic31 integrase. *Molecular and Cellular Biology*, 21, 3926–3934.
43. Singh, S., Ghosh, P., & Hatfull, G. F. (2013). Attachment site selection and identity in bxb1 serine integrase-mediated site-specific recombination. *PLoS Genetics*, 9, e1003490.
44. Keravala, A., Groth, A. C., Jarrahan, S., Thyagarajan, B., Hoyt, J. J., Kirby, P. J., & Calos, M. P. (2006). A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Molecular Genetics and Genomics*, 276, 135–146.
45. Gupta, M., Till, R., & Smith, M. C. M. (2007). Sequences in attB that affect the ability of phic31 integrase to synapse and to activate DNA cleavage. *Nucleic Acids Research*, 35, 3407–3419.
46. Rutherford, K., Yuan, P., Perry, K., Sharp, R., & Van Duyne, G. D. (2013). Attachment site recognition and regulation of directionality by the serine integrases. *Nucleic Acids Research*, 41, 8341–8356.
47. Mandal, S., & Johnson, R. (2021). Control of the serine integrase reaction: Roles of the coiled-coil and helix E regions in DNA site synapsis and recombination. *Journal of Bacteriology*, 203, e0070320.
48. Low, B. E., Hosur, V., Lesbirel, S., & Wiles, M. V. (2022). Efficient targeted transgenesis of large donor DNA into multiple mouse genetic backgrounds using bacteriophage Bxb1 integrase. *Scientific Reports*, 12, 5424.
49. Gao, H., & Smith, M. C. M. (2021). Use of orthogonal serine integrases to multiplex plasmid conjugation and integration from *E. coli* into *Streptomyces*. *Access Microbiology*, 3(3), 000291.
50. Fayed, B., Ashford, D. A., Hashem, A. M., Amin, M. A., El Gazayerly, O. N., Gregory, M. A., & Smith, M. C. M. (2015). Multiplexed integrating plasmids for engineering of the erythromycin gene cluster for expression in *Streptomyces* spp. and combinatorial biosynthesis. *Applied and Environmental Microbiology*, 81, 8402–8413.
51. Gao, H., Taylor, G., Evans, S. K., Fogg, P. F. M., & Smith, M. C. M. (2020). Application of serine integrases for secondary metabolite pathway assembly in *Streptomyces*. *Synthetic and Systems Biotechnology*, 5, 111–119.
52. Tomimatsu, K., Kokura, K., Nishida, T., Yoshimura, Y., Kazuki, Y., Narita, M., Oshimura, M., & Ohbayashi, T. (2017). Multiple expression cassette exchange via TP901-1, R4, and Bxb1 integrase systems on a mouse artificial chromosome. *FEBS Open Bio*, 7, 306–317.
53. Blanco-Redondo, B., & Langenhan, T. (2018). Parallel genomic engineering of two *Drosophila* genes using orthogonal attB/attP sites. *G3: Genes, Genomes, Genetics*, 8, 3109–3118.
54. Snoeck, N., De Mol, M. L., Van Herpe, D., Goormans, A., Maryns, I., Coussement, P., Peters, G., Beauprez, J., De Maseseneire, S. L., & Soetart, W. (2018). Serine integrase recombinational engineering (SIRE): A versatile toolbox for genome editing. *Biotechnology and Bioengineering*, 116, 364–374.
55. Inniss, M. C., Bandara, K., Jusiak, B., Lu, T. K., Weiss, R., Wrobleksa, L., & Zhang, L. (2017). A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO cells. *Biotechnology and Bioengineering*, 114, 1837–1846.
56. Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., & Court, D. L. (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, 97, 5978–5983.
57. Ellis, H. M., Yu, D., DiTizio, T., & Court, D. L. (2001). High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences USA*, 98, 6742–6746.
58. Court, D. L., Sawitzke, J. A., & Thomason, L. C. (2002). Genetic engineering using homologous recombination. *Annual Review of Genetics*, 36, 361–388.
59. Gibson, G. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6, 343–345.
60. Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., ... Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329, 52–56.
61. Brown, W. R. A., Lee, N. C. O., Xu, Z., & Smith, M. C. M. (2011). Serine recombinases as tools for genome engineering. *Methods (San Diego, Calif.)*, 53, 372–379.
62. Durrant, M. G., Fanton, A., Tycko, J., Hinks, M., Chandrasekaran, S. S., Perry, N. T., Schaepe, J., Du, P. P., Lotfy, P., Bassik, M. C., Bintu, L., Bhatt, A. S., & Hsu, P. D. (2022). Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nature Biotechnology*, <https://doi.org/10.1038/s41587-022-01494-w>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Abioye, J., Lawson-Williams, M., Lecanda, A., Calhoon, B., McQue, A. L., Colloms, S. D., Stark, W. M., & Olorunniji, F. J. (2023). High fidelity one-pot DNA assembly using orthogonal serine integrases. *Biotechnology Journal*, 18, e2200411.
<https://doi.org/10.1002/biot.202200411>