Published online 29 October 2020 *Nucleic Acids Research, 2020, Vol. 48, No. 22* ***e130***
doi: 10.1093/nar/gkaa956

RedEx: a method for seamless DNA insertion and deletion in large multimodular polyketide synthase gene clusters

Chaoyi Song[1,†], Ji Luan[1,†], Ruijuan Li[1], Chanjuan Jiang[1], Yu Hou[1], Qingwen Cui[1], Tianqi Cui[1], Long Tan[1], Zaichao Ma[1], Ya-Jie Tang[1], A. Francis Stewart[2], Jun Fu[1], Youming Zhang[1] and Hailong Wang[1,*]

[1]State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Helmholtz International Lab for Anti-infectives, Shandong University–Helmholtz Institute of Biotechnology, Shandong University, Qingdao, Shandong, 266237, China and [2]Genomics, Biotechnology Center, Center for Molecular and Cellular Bioengineering, Technische Universität Dresden, Tatzberg 47–51, 01307 Dresden, Germany

Received August 06, 2020; Revised September 08, 2020; Editorial Decision October 06, 2020; Accepted October 08, 2020

ABSTRACT

Biosynthesis reprograming is an important way to diversify chemical structures. The large repetitive DNA sequences existing in polyketide synthase genes make seamless DNA manipulation of the polyketide biosynthetic gene clusters extremely challenging. In this study, to replace the ethyl group attached to the C-21 of the macrolide insecticide spinosad with a butenyl group by refactoring the 79-kb gene cluster, we developed a RedEx method by combining Redαβ mediated linear-circular homologous recombination, *ccdB* counterselection and exonuclease mediated *in vitro* annealing to insert an exogenous extension module in the polyketide synthase gene without any extra sequence. RedEx was also applied for seamless deletion of the rhamnose 3′-*O*-methyltransferase gene in the spinosad gene cluster to produce rhamnosyl-3′-desmethyl derivatives. The advantages of RedEx in seamless mutagenesis will facilitate rational design of complex DNA sequences for diverse purposes.

INTRODUCTION

Polyketides continue to be an invaluable resource for development of medicines and agrochemicals. Besides naturally occurring polyketides, artificial derivatization is an alternative way to obtain new structures for bioactivity assay or to improve their pharmacological properties. Chemical derivatization is usually inaccessible due to their structural complexity. Biosynthetic engineering can be utilized to achieve structural diversification and optimization. Polyketide biosynthetic gene clusters are known to be difficult to modify due to the large repetitive sequences in their biosynthesis gene clusters.

Spinosad, the mixture of spinosyns A and D (Figure 1), is a polyketide-derived macrolide originally isolated from the soil actinomycete *Saccharopolyspora spinosa* (1). It is the active ingredient in several highly effective commercial insect control agents with excellent environmental and mammalian toxicological profile (2). Biosynthesis of spinosad involves 23 genes including 5 polyketide synthase (PKS) genes, 4 polyketide bridging genes, 8 sugar biosynthesis genes, 2 glycosyltransferase genes and 4 sugar methylation genes (3). To improve the efficacy and spectrum of spinosad, diverse approaches including synthetic modification, biotransformation and genetic engineering have been explored. Semisynthetic modification of the spinosad structure generated >1000 analogues and none of them showed significant improvements over the naturally occurring major component spinosyn A (4). Because some regions of the spinosad molecule are recalcitrant to chemical modification, particularly the C-21 position, Sheehan *et al.* replaced the loading module of spinosad PKS with the loading module from the erythromycin PKS by genetic engineering in *Sa. spinosa* (5). The recombinant strain harboring the hybrid PKS pathway was fed with a range of carboxylic acids and produced 16 new spinosad derivatives, among which 21-cyclobutyl-spinosyn A and its semisynthetic 5,6-dihydro derivative showed improved insecticidal activity.

Another example that altering substituents attached to the C-21 of spinosad leads to activity improvement is the naturally occurring butenyl-spinosyns A and D (Figure 1), discovered from *Saccharopolyspora pogona* (6). The

*To whom correspondence should be addressed. Tel: +86 532 67720928; Fax: +86 532 58631501; Email: wanghailong@sdu.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

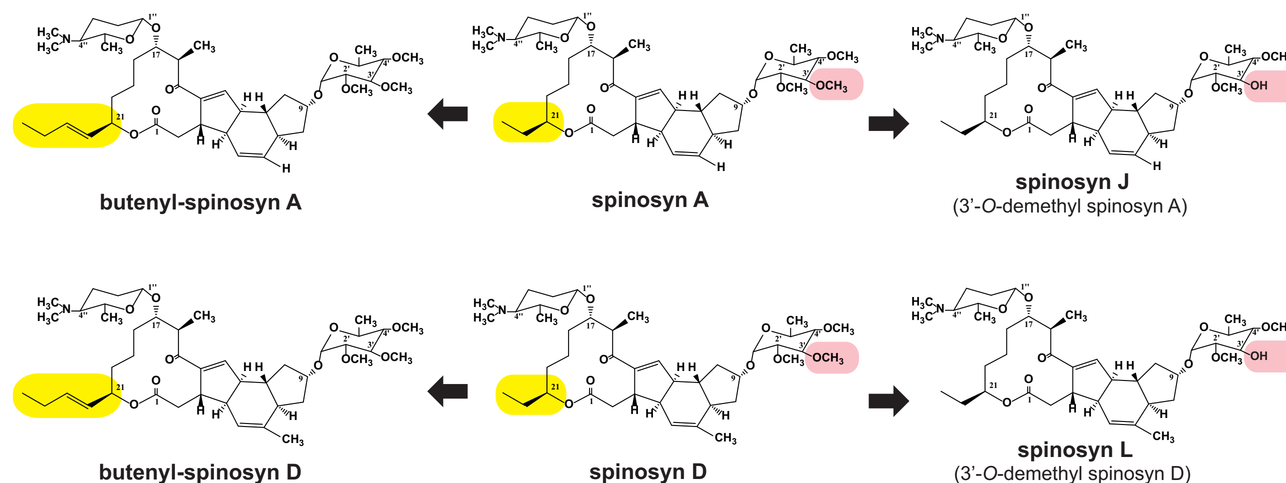


Figure 1. Structures of spinosad and its derivatives generated in this study. The group substitution at C-21 and C-3′ is highlighted in yellow and pink respectively.

butenyl-spinosyn A biosynthetic genes (*bus*) from *Sa. pogona* and the spinosyn A biosynthetic genes (*spn*) from *Sa. spinosa* are organized in the same order and orientation in the gene cluster. The *bus* genes show strong similarity in both the DNA (91–97%) and protein (81–97%) sequence to their *spn* gene counterparts, except that a 5301-bp sequence in the *busA* gene does not have a counterpart in the *spnA* gene. This 5301-bp region codes for the acetyltransferase (AT), dehydratase (DH), ketoreductase (KR) and acyl carrier protein (ACP) domains of module 1b and the ketosynthase (KS) domain of module 1a in the polyketide synthase. The extension module 1b is supposed to be responsible for two additional carbons in the C-21 tail of the butenyl-spinosyn A.

Combinatorial biosynthesis has been proven to be a powerful method to enrich the diversity of natural products (7–9). However, its application in generating spinosad derivatives has been impaired by the extremely low transformation efficiency of *Sa. spinosa*. Recently, the spinosad gene cluster was cloned and heterologously expressed in *Streptomyces albus* (10,11) and *Saccharopolyspora erythraea* (12). These work provided a new platform for combinatorial biosynthesis of spinosad through genetic engineering of the gene cluster in heterologous hosts.

In this study, to facilitate combinatorial biosynthesis of spinosad, we developed a RedEx method by combining Redαβ mediated linear-circular homologous recombination (LCHR) (13), *ccdB* counterselection (14) and exonuclease mediated *in vitro* annealing (15) for efficient seamless DNA insertion and deletion in PKS and nonPKS regions of large multimodular polyketide gene clusters in *Escherichia coli*. Using RedEx, we successfully incorporated AT1b-KS1a domains of the BusA protein into the position between the KS domain and the AT domain in the first extension module of the SpnA protein. Seamless deletion of the rhamnose 3′-*O*-methyltransferase gene, *spnK*, in the spinosad gene cluster was also achieved by RedEx. Refactored gene clusters produced 2.36 mg l^{-1} of butenyl-spinosyn A and 7.34 mg l^{-1} of spinosyn J (Figure 1) in *St. albus* J1074.

MATERIALS AND METHODS

Bacteria strains, plasmids and culture conditions

Escherichia coli strains were cultivated at 37°C in LB medium supplemented with appropriate antibiotics (15 μg ml^{-1} of chloramphenicol, 15 μg ml^{-1} of kanamycin, 100 μg ml^{-1} of ampicillin, 20 μg ml^{-1} of apramycin). *St. albus* J1074 were grown at 30°C on mannitol soya flour agar plates for sporulation and conjugation. Exconjugants were maintained on brain heart infusion agar plates supplemented with appropriate antibiotics (40 μg ml^{-1} of apramycin, 25 μg ml^{-1} of thiostrepton, 25 μg ml^{-1} of nalidixic acid).

Preparation of linear DNA fragments

The oligos and templates for PCR were listed in Supplementary Table S1. The fragment *spn(4113–4244)-busA(4245–9546)-spnA(4301–5625)* flanked with BstZ17I sites was synthesized by GENEWIZ (Suzhou, China) and maintained on the pUC57-kan vector. Fragments *spnA(4033–4112)-PacI-ampccdB-PacI-spn(4093–4244)-busA(4245–9546)-spnA(4301–5705)*, *spnA(4033–4112)-PacI-ampccdB-PacI-spn(4073–4244)-busA(4245–9546)-spnA(4301–5705)* and *spnA(4033–4112)-PacI-ampccdB-PacI-spn(4033–4244)-busA(4245–9546)-spnA(4301–5705)* were respectively released by BstZ17I digestion of a R6K plasmid which was constructed by assembling the PCR amplified R6K replication origin, the PCR amplified *ampccdB* selection-counterselection marker, and the synthesized fragment *spn(4113–4244)-busA(4245–9546)-spnA(4301–5625)* with LLHR in *E. coli* GBdir-gyrA462-pir116. *E. coli* GBdir-gyrA462-pir116 was constructed by inserting a copy-up *pir116* gene downstream of the *gyrA* gene in the chromosome of *E. coli* GBdir-gyrA462 (13) to maintain R6K plasmids in high copy number.

The fragment *spnA(4063–4244)-busA(4245–9546)-spnA(4301–5675)* was released by BstZ17I digestion of pR6K-amp-busA which was constructed by assembling the PCR amplified R6K replication origin, the PCR am-

plified *amp* selection marker, and the synthesized fragment *spn(4113–4244)*-*busA(4245–9546)*-*spnA(4301–5625)* with LLHR in *E. coli* GBdir-pir116 (13). PCR products and DNA fragments generated by restriction digestion were purified with the Universal DNA purification kit (Tiangen, China).

RedEx seamless insertion of AT1b-KS1a domains from the *busA* gene into the *spnA* gene

pBAC-spnNEW was electroporated into *E. coli* GBred-gyrA462 (14) and expression of Redαβ recombinases was induced by L-arabinose. Five hundred nanograms of the fragment *spnA(4033–4112)*-*PacI*-*ampccdB*-*PacI*-*spnA(4093–4244)*-*busA(4245–9546)*-*spnA(4301–5705)* was electroporated into *E. coli* GBred-gyrA462. The recombinant BAC carrying the *ampccdB* cassette in the *spnA* gene (pBAC-spnNEW-ampccdBinspnA) was identified with ApaLI restriction analysis of colonies selected on LB plates supplemented with ampicillin. We obtained five correctly recombined BAC out of eight colonies (Supplementary Figure S1).

Two hundred nanograms of PacI digested pBAC-spnNEW-ampccdBinspnA was treated with 0.2 U T4pol (New England BioLabs, cat. no. M0203) in a 20 μl reaction at 25°C for 1 h, 75°C for 20 min, 50°C for 2 h and then held at 4°C in a thermocycler. The reaction mixture was dropped on Millipore Membrane Filters (Merck-Millipore, cat. no. VSWP01300) floating on ddH_2O and kept at room temperature for 30 min for desalting by drop dialysis. Five microliters of desalted reaction mixture was electroporated into *E. coli* GB2005 cells. The recombinant BAC carrying AT1b-KS1a domains from the *busA* gene in the *spnA* gene (pBAC-spnNEWbusA) was identified with XmnI restriction analysis of colonies selected on LB plates supplemented with chloramphenicol.

RedEx seamless deletion of the *spnK* gene in the spinosad gene cluster

Five hundred nanograms of the *ampccdB* PCR product was electroporated into *E. coli* GBred-gyrA462 harboring pBAC-spnNEW and expressing Redαβ recombinases. The recombinant BAC carrying the *ampccdB* cassette between the *spnI* gene and the *spnH* gene (pBAC-spnNEW-ampccdBdelK) was identified with PstI restriction analysis of colonies selected on LB plates supplemented with ampicillin. DNA sequence of the *ccdB* gene, PacI sites and homology-arm regions were confirmed by Sanger sequencing. Two hundred nanograms of PacI digested pBAC-spnNEW-ampccdBdelK was treated with T4pol in a thermocycler. Five microliters of desalted reaction mixture was electroporated into *E. coli* GB2005 cells. The recombinant BAC without the *spnK* gene (pBAC-spnNEWJL) was identified with PvuII restriction analysis of colonies selected on LB plates supplemented with chloramphenicol.

Insertion of the AT1b-KS1a region from the *busA* gene into the *spnA* gene on the chromosome of *St. albus* J1074 using the CRISPR-Cas9 system

The 1A and 2A protospacer sequences upstream of the base 3460 in *spnA* and the 1B and 2B protospacer sequences downstream of the base 5626 in *spnA* were selected according to the rules described by Wang *et al.* (16) (Figure 4A). The 1A and 1B single guide RNA (sgRNA) cassettes, and the 2A and 2B sgRNA cassettes were synthesized by GENEWIZ (Suzhou, China) in two separate DNA fragments which contain BbsI sites at both ends. The two pairs of sgRNA cassettes were inserted into pCRISPomyces-2 (17) with the Golden Gate assembly (18). Each sgRNA cassette on pCRISPomyces-2 was driven by the same strong promoter gapdhp(EL). Then, the apramycin resistance gene on pCRISPomyces-2 was replaced with a cassette consisting of the ampicillin resistance gene and the thiostrepton resistance gene using LCHR to get pCRISPomyces-thio1 and pCRISPomyces-thio2. The oligos and templates for PCR were listed in Supplementary Table S1.

The fragment *spnA(1956–2952*-*(2953–2972)**-*2973–3357*-*(3358–3377)**-*3378–4244)*-*busA(4245–9546)*-*spnA(4301–5776*-*(5777–5796)**-*5797–5817*-*(5818–5840)**-*5841–6826)*-*cm* was released by BstZ17I digestion of pR6K-spnbusA which was constructed by assembling the PCR amplified R6K replication origin, *cm* selection marker, the fragment *spnA(4093–4244)*-*busA(4245–9546)*-*spnA(4301–5705)*, the PCR amplified fragment *spnA(1956–2952*-*(2953–2972)**-*2973–3357*-*(3358–3377)**-*3378–4244)* and the PCR amplified fragment *spnA(4301–5776*-*(5777–5796)**-*5797–5817*-*(5818–5840)**-*5841–6826)* via LLHR. Synonymous mutations were introduced to protospacer nucleotides with asterisks to avoid the Cas9–sgRNA complex cutting the targeting plasmid.

Finally, the fragment *spnA(1956–2952*-*(2953–2972)**-*2973–3357*-*(3358–3377)**-*3378–4244)*-*busA(4245–9546)*-*spnA(4301–5776*-*(5777–5796)**-*5797–5817*-*(5818–5840)**-*5841–6826)*-*cm* was inserted into pCRISPomyces-thio1 and pCRISPomyces-thio2 with LCHR respectively. The resulting targeting plasmids pCRISPomyces-spnCasbus1 and pCRISPomyces-spnCasbus2 were transformed into *St. albus* J1074 harboring the artificial spinosad gene cluster on the chromosome respectively. Exconjugants selected on the BHI plate supplemented with thiostrepton and nalidixic acid were identified by colony PCR using sentinel primers listed in Supplementary Table S1.

Fermentation and high performance liquid chromatography (HPLC)–mass spectrometry (MS) analysis of spinosyns from *St. albus* J1074

Seed cultures were prepared by inoculating *St. albus* J1074 strains into 250 ml flasks containing 30 ml tryptic soy broth and incubating at 30°C with 220 rpm shaking for 72 h. Thirty milliliters of the seed culture was inoculated into a 5 l bioreactor (Bailun, Shanghai, China) containing 3 l fermentation broth (4% glucose, 1% glycerol, 3% soluble starch, 1.5% soytone, 1% beef extract, 0.65% peptone, 0.05% yeast extract,0.1% MgSO4, 0.2% NaCl, 0.24% $CaCO_3$) and cultivated for 10 days. The agitator and temperature were maintained at 500 rpm and 30°C respectively. Thirty six milliliters of 500 g l^{-1} glucose was fed into the bioreactor every day. One milliliter of the culture was withdrawn every two days since the fourth day and mixed with 4 ml of MeOH. The mixture was sonicated for 20 min followed by

standing for 2 h and filtered through a 0.22 μm filter before HPLC-MS analysis. Each fermentation was performed twice in the bioreactor and the average yield of butenyl-spinosyn A and spinosyn J was shown.

The high-resolution mass spectrometry analysis was done on an Ultimate 3000 UHPLC-DAD system (Thermo Fisher Scientific) coupled with an Impact HD micrOTOF-Q III mass spectrometer (Bruker Daltonics, Bremen, Germany) using the standard ESI source equipped with an Acclaim RSLC 120, C18, 2.2 μm, 2.1 × 100 mm (Thermo Scientific) at a flow rate of 0.3 ml min^{-1} and detected UV at 200–600 nm. The mobile phase was H_2O with 0.1% (v/v) formic acid (solvent A) and acetonitrile (ACN) with 0.1% (v/v) formic acid (solvent B). The gradient solvent condition was set as follows: 5% B to 50% B within 5 min, maintain at 50% B for 15 min, 50% B to 95% B within 5 min, followed by 5 min 95% B and 5 min 5% B. Mass spectra were acquired in centroid mode ranging from 100 to 1500 m/z with positive ionization mode and auto MS^2 fragmentation.

Quantification of butenyl-spinosyn A and spinosyn J was performed on the Ultimate3000 UHPLC system (Thermo Fisher Scientific) coupled with the amaZon SL ion trap mass spectrum system (Bruker) under the same condition as the above high-resolution mass spectrometry analysis.

Isolation of butenyl-spinosyn A and spinosyn J for NMR analysis

For butenyl-spinosyn A, 300 ml seed culture was inoculated into a 50 l bioreactor (Bailun, Shanghai, China) containing 30 l fermentation broth and incubated for 8 days. Then, 600 ml of the adsorber resin Amberlite XAD-16 was added and the culture was incubated for another 2 days. Cells and the XAD-16 were collected by centrifugation and lyophilized to dryness. The dry mass was extracted three times with 5 l ethyl acetate (EtOAc) and cell debris was removed by centrifugation. The extraction solution was concentrated in vacuo, applied to a silica gel column and eluted by a CH_2Cl_2–MeOH gradient solvent system. Elution fractions of CH_2Cl_2:MeOH from 70:1 to 20:1 were combined, concentrated in vacuo and resuspended in 5 ml MeOH. The solution was loaded on a Sephadex LH-20 column and eluted with MeOH at a flow rate of 0.8 ml min^{-1}. Elution fractions were collected in one tube per 10 min and analyzed by HPLC-MS. Solutions from tubes containing significant amounts of butenyl-spinosyn A were pooled, evaporated and resuspended in 2 ml MeOH. This solution was subjected to semi-preparative HPLC (ODS-A, C18, 5 μm, 20 × 250 mm, YMC) under gradient solvent condition of 0–60 min 20–100% ACN with the mobile phase H_2O-ACN. Fractions eluted form 42 min to 46 min were concentrated and further purified by HPLC (ODS-A, C18, 5 μm, 10 × 250 mm, YMC) with the mobile phase 10 mM NH_4AC (solvent A) and ACN with 20% (v/v) MeOH (solvent B) at a flow rate of 2.5 ml min^{-1}. The gradient solvent condition was 5% B for 5 min, 5–100% B in 45 min, followed by 100% B for 25 min. Fractions eluted from 59 to 60 min were concentrated and lyophilized to give 8.5 mg butenyl-spinosyn A for NMR analysis and quantification of the fermentation yield in the 5 l bioreactor.

For spinosyn J, the EtOAc extraction solution from 3 l fermentation culture was concentrated by evaporation under reduced pressure, resuspended in 20 ml MeOH and subjected to semi-preparative HPLC (ODS-A, C18, 5 μm, 20 × 250 mm, YMC) under gradient solvent condition of 0–90 min 20–100% ACN with the mobile phase H_2O–ACN. Fractions eluted from 60 to 65 min were concentrated and further purified by HPLC (ODS-A, C18, 5 μm, 10 × 250 mm, YMC) with the mobile phase 10 mM NH_4AC (solvent A) and ACN with 20% (v/v) MeOH (solvent B) at a flow rate of 2.5 ml min^{-1}. The gradient solvent condition was 40% B for 5 min, 40–100% B in 45 min, followed by 100% B for 30 min. Fractions eluted from 53 to 54 min were concentrated and lyophilized to give 9.0 mg spinosyn J for NMR analysis and quantification of the fermentation yield in the 5 l bioreactor.

NMR spectra were obtained at 600 MHz (^1H), and 150 MHz (^{13}C) with a Bruker Avance 600 spectrometer. $^{13}C\{^1H\}$ NMR spectra were obtained using composite pulse decoupling.

RESULTS AND DISCUSSION

The RedEx strategy

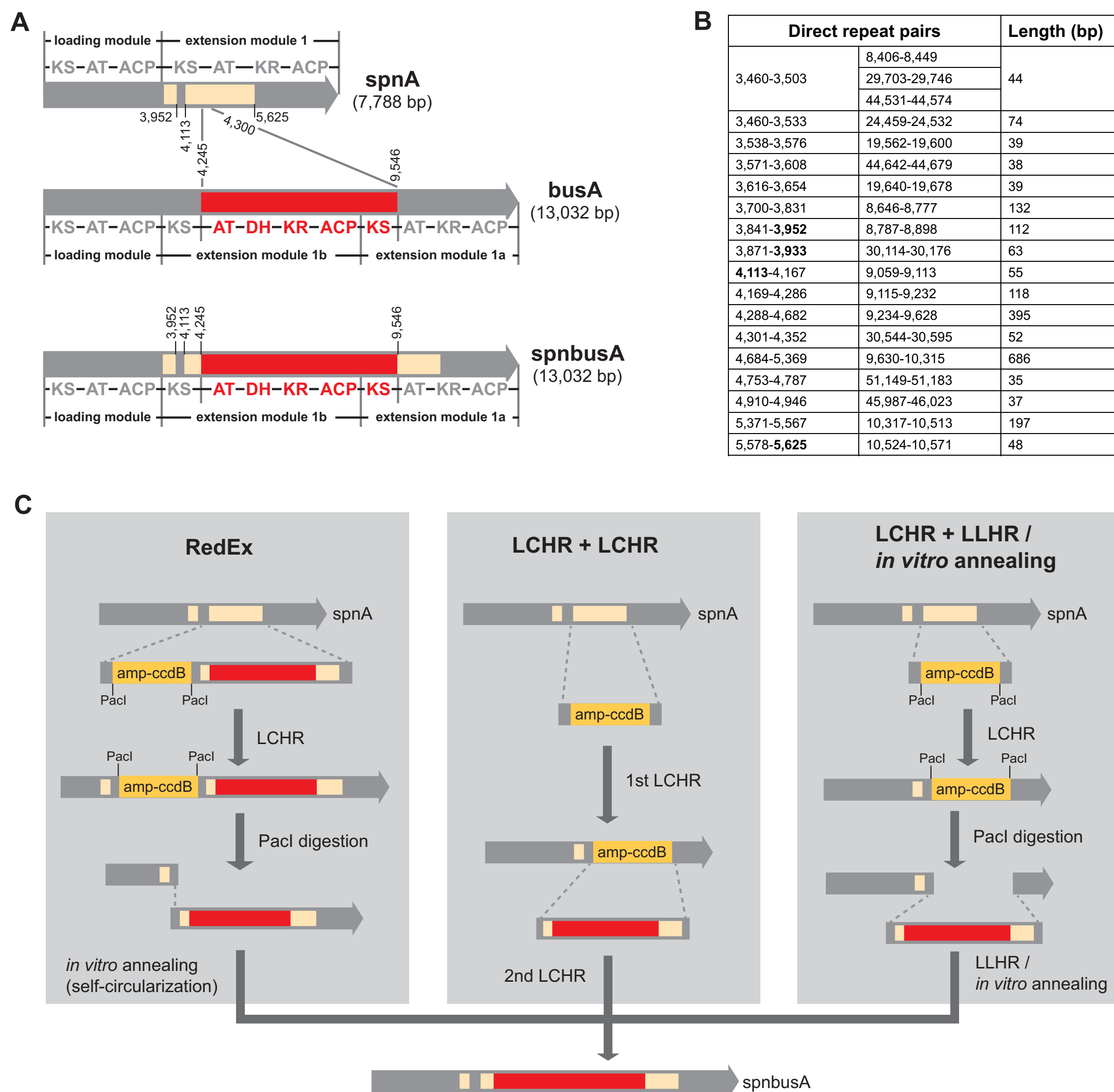
The general RedEx strategy is outlined in Figure 2C. To achieve seamless DNA insertion and deletion on DNA molecules, RedEx firstly uses Redαβ mediated LCHR to insert the designed DNA cassette together with the selection-counterselection marker flanked by unique restriction sites into the target site of the gene cluster. Then, the vector was linearized with restriction digestion to expose terminal 20-bp homology arms at both ends. Finally, the linear DNA molecule with terminal overlaps was circularized by exonuclease mediated *in vitro* DNA annealing. The *ccdB* counterselection was used to avoid the background resistance by ensuring that circular DNA molecules from uncompleted endonuclease digestion cannot propagate in the sensitive *E. coli* host.

Seamless insertion of the *busA* polyketide-synthase domains into the spinosad gene cluster using RedEx

The 13 032-bp *busA* gene encodes one loading module and two extension modules, whereas the 7788-bp *spnA* gene encodes one loading module and one extension module. In the *busA* gene, the first 4245 bp codes for the loading module and the following KS1b domain, and the last 3,486 bp codes for the AT1a, KR1a and ACP1a domains. The above two regions have counterparts in the *spnA* gene. The middle 5301-bp region in the *busA* gene without a counterpart in the *spnA* gene encodes the AT1b, DH1b, KR1b, ACP1b and KS1a domains (6) (Figure 2A).

Recently, we constructed a 79-kb artificial spinosad gene cluster by assembling 23 biosynthetic genes into 7 operons under control of strong constitutive *Streptomyces* promoters (11). The spinosad production was enhanced by 328-fold in *St. albus* J1074 harboring the artificial gene cluster compared with the native gene cluster. To test if it is possible to substitute the ethyl group at C-21 of spinosyn A into a butenyl group by refactoring the spinosad gene cluster, we used RedEx to replace the sequence between the loading

 Nucleic Acids Research, 2020, Vol. 48, No. 22 ***e130***



Direct repeat pairs		Length (bp)
3,460-3,503	8,406-8,449	44
	29,703-29,746	
	44,531-44,574	
3,460-3,533	24,459-24,532	74
3,538-3,576	19,562-19,600	39
3,571-3,608	44,642-44,679	38
3,616-3,654	19,640-19,678	39
3,700-3,831	8,646-8,777	132
3,841-**3,952**	8,787-8,898	112
3,871-**3,933**	30,114-30,176	63
4,113-4,167	9,059-9,113	55
4,169-4,286	9,115-9,232	118
4,288-4,682	9,234-9,628	395
4,301-4,352	30,544-30,595	52
4,684-5,369	9,630-10,315	686
4,753-4,787	51,149-51,183	35
4,910-4,946	45,987-46,023	37
5,371-5,567	10,317-10,513	197
5,578-**5,625**	10,524-10,571	48

Figure 2. Seamless insertion of AT1b-KS1a domains from the *busA* gene into the *spnA* gene. (**A**) Sequence comparison of the *spnA* gene, the *busA* gene and the hybrid *spnbusA* gene. (**B**) Direct repeats longer than 35 bp identified in the *spnA* gene. (**C**) Three methods tried to insert AT1b-KS1a domains from the *busA* gene into the *spnA* gene.

module and the first extension module (bases 4245–4300) of the *spnA* gene with the middle 5301-bp sequence coding for AT1b-KS1a domains from the *busA* gene.

Because the multi-modular PKS genes usually contain repeated sequences, cautions must be taken to design homology arms for Redαβ mediated LCHR to reduce unwanted recombination. We analyzed direct repeats (DRs) longer than 35 bp in the *spn* PKS region using the Unipro UGENE software (19), because homology sequences longer than 35 bp can mediate efficient recombination by lambda Red proteins. Sixty-eight pairs of DRs longer than 35 bp were identified in the whole *spn* PKS region and 19 pairs of DRs have counterparts in the *spnA* gene (Figure 2B). DRs flanking the bases 4245–4300 should be avoided to be homology arms for LCHR. Therefore, we selected the 80-bp sequence upstream of the base 4113 and downstream of the base 5625 located in regions without DRs as homology arms for Redαβ mediated LCHR (Figure 2A). The fragment *spnA(4033–4112)-PacI-ampccdB-PacI-spn(4093–4244)-busA(4245–9546)-spnA(4301–5705)* in which two PacI sites, not existing on pBAC-spnNEW, are flanked by 20-bp DRs (spn(4093–4112)) was prepared (Figure 2A and C). This fragment was used to replace bases 4113–5625 in the *spnA* gene on pBAC-spnNEW in *E. coli* GBred-gyrA462 (14) which expresses Redαβ proteins for LCHR and carries $gyrA_{\text{Arg462Cyc}}$ mutation conferring CcdB resistance. Terminal 20-bp homologies (spn(4093–4112)) at both ends were exposed after the *ampccdB* cassette was removed by PacI digestion of the recombinant BAC. The linear BAC was restored to be circular by *in vitro* annealing between terminal 20-bp DRs mediated by T4 DNA polymerase (T4pol) (15). *E. coli* GB2005 (14), sensitive to CcdB toxin, was used to recover the final recombinant BAC (pBAC-spnNEWbusA). The

Table 1. Colony numbers per ml of recovery cultures after electroporation and accuracies in brackets (correct/number examined by restriction digestion) of T4pol *in vitro* annealing mediated self-circularization of linear 108-kb BACs

	Length of terminal homology arms		
DNA amount	20 bp	40 bp	80 bp
50 ng	0	0	0
100 ng	26 ± 12 (5/12)	12 ± 3 (9/12)	12 ± 6 (9/12)
200 ng	244 ± 41 (10/12)	240 ± 54 (10/12)	265 ± 12 (8/12)
400 ng	88 ± 40 (9/12)	61 ± 17 (11/12)	95 ± 55 (7/12)

Each experiment was done in triplicate. DNA analyses are shown in Supplementary Figure S2.

leftover undigested BAC carrying the *ccdB* gene was lethal to *E. coli* GB2005.

In the course of RedEx optimization, we found that the DNA amount used in the T4pol *in vitro* annealing step significantly affected the efficiency of linear BAC self-circularization (Table 1). The highest efficiency was obtained when 200 ng of the PacI linearized 108-kb BAC was used. The efficiency of using 200 ng linear BAC DNA was 4 times of using 400 ng DNA and 20 times of using 100 ng DNA when 40-bp terminal homology was used. Self-circularization of the linear BAC failed completely in the 50-ng experiment. We also evaluated the impact of terminal homology length on linear BAC self-circularization mediated by T4pol *in vitro* annealing (Table 1). Results were largely indifferent when 20-, 40- and 80-bp terminal DRs were used. Therefore, the optimal length of terminal homology is 20 bp and the optimal amount of the linear 108-kb BAC is 200 ng for RedEx.

Heterologous expression of butenyl-spinosyns A and D

The pBAC-spnNEWbusA, which carries the 5,301-bp region encoding the AT1b, DH1b, KR1b, ACP1b and KS1a domains from the *busA* gene in the 79-kb artificial spinosad gene cluster, was transformed into *St. albus* J1074 and integrated onto the phiC31 attB sites on the chromosome. The butenyl-spinosyns A/D was successfully detected as the most abundant spinosyns in the fermentation broth (Figure 3B and C). The structure of butenyl-spinosyn A was confirmed by NMR analysis (Supplementary Table S2 and Supplementary Figures S3 and S4). When 3 l medium in the 5 l bioreactor was used for a 12-day fermentation, the highest yield, 2.36 mg l^{-1}, of butenyl-spinosyn A in *St. albus* J1074 harboring pBAC-spnNEWbusA was obtained at the 10th day (Figure 3D). Production of butenyl-spinosyn A did not increase after 10 days. The yield of butenyl-spinosyn A was 16 times lower than the yield of spinosyn A in *St. albus* J1074 under the same fermentation condition at the 10th day. This result suggested that insertion of the AT1b-KS1a domains from the BusA protein into the SpnA protein generated a functional hybrid polyketide assembly line. Enzymes in the spinosyn assembly line can accept and process both the spinosyn polyketide substrates and butenyl-spinosyn polyketide substrates.

Saccharopolyspora erythraea is another heterologous host which was successfully used to express spinosyns (12). We failed to get exconjugants when the 108-kb pBAC-spnNEWbusA was used to transform *Sa. erythraea.* Huang *et al.* integrated the entire spinosyn gene cluster onto the *Sa. erythraea* chromosome via four rounds of double crossover (12). We truncated the pBAC-spnNEWbusA into different length and found that the vectors larger than 60 kb cannot be transformed into *Sa. erythraea.* Cloning the whole *bus* cluster from *Sa. pogona* is another way to generate a butenyl–spinosyn producing strain. However, the butenyl-spinosyn production from the native gene cluster would be very low in the heterologous *Streptomyces* hosts. To improve its production in heterologous host, the gene cluster needs to be refactored to over-express biosynthetic genes, such as what we did for the spinosad gene cluster (11). Engineering of the artificial over-expression spinosad gene cluster is a good choice to achieve considerable heterologous butenyl-spinosyn production.

Comparison of RedEX with alternative seamless mutagenesis methods

Modular type I PKSs assemble coenzyme A-derived substrates into the core scaffold of final polyketide compounds. Ideally, module swapping in PKS genes should be seamless to avoid any extraneous sequence introduced during mutagenesis interfering with protein expression. Type I PKS gene clusters are usually larger than 50 kb, and the traditional restriction digestion and ligation method is limited by restriction sites and the size of gene clusters (20,21). Type IIS restriction endonucleases cleave DNA outside of their recognition sequences and are particularly useful in seamless assembly of DNA fragments even for those with repetitive sequences (18,22). However, sequence analysis revealed that all of 49 available commercial type IIS restriction enzymes have multiple recognition sites on the pBAC-spnNEW.

The counterselection-based recombineering strategy has been optimized to achieve seamless point mutations in DNA with repeated sequences, such as the bacterial 16-kb *plu3263* nonribosomal peptide synthetase gene on a plasmid and the 170-kb human *BRD4* gene on a BAC (14). In this study, the sequential selection/counterselection based recombineering strategy (14) was also tried to insert the 5301-bp AT1b-KS1a region from the *busA* gene into the spinosad gene cluster on the pBAC-spnNEW. In the first round of LCHR, bases 4113–5625 in the *spnA* gene on pBAC-spnNEW (11) were replaced with the *ampccdB* cassette. In the second round of LCHR, the *ampccdB* cassette was replaced with the fragment *spnA(4063–4244)-busA(4245–9546)-spnA(4301–5,675)* (the *busA* fragment) (Figure 2A and C). However, recovery of correct recombinants failed completely after counterselection, even though we used Redα omission strategy in the second round of LCHR (14). Incorrect recombinants all resulted from unwanted intramolecular recombination between different DRs (Supplementary Figure S5). We then tried another strategy by placing PacI restriction sites, which do not exist on pBAC-spnNEW, to flank the *ampccdB* cassette used for the first round of LCHR (Figure 2C). Then, the recombinant BAC was linearized with PacI digestion to expose terminal 50-bp homology arms to both ends of the *busA* fragment. RecET mediated *in vivo* linear-linear homologous recombination (LLHR) (23) or T4 DNA polymerase (T4pol)

 Nucleic Acids Research, 2020, Vol. 48, No. 22 ***e130***

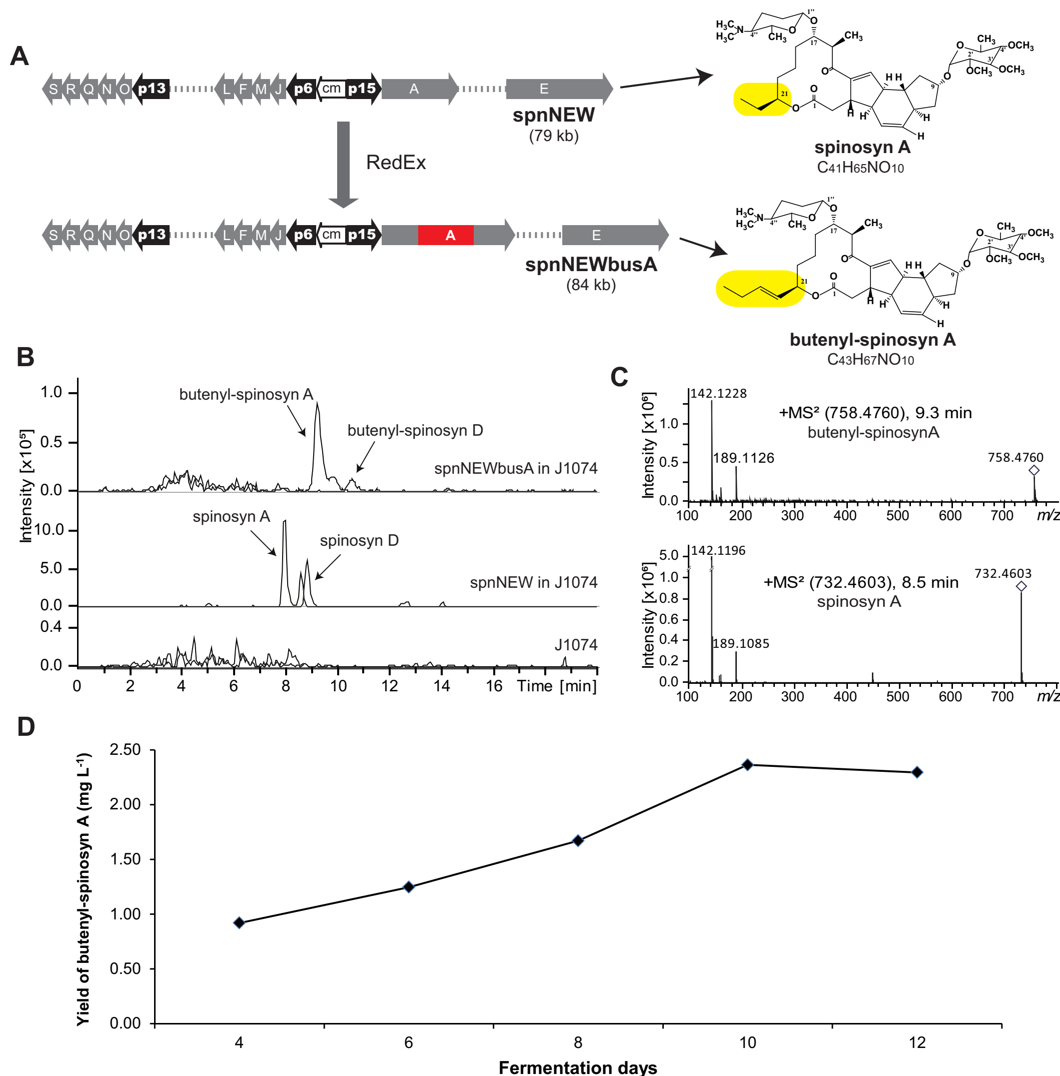


Figure 3. Heterologous expression of the hybrid spnNEWbusA gene cluster. (**A**) Generation of the hybrid spnNEWbusA gene cluster with RedEx. The group change at C-21 of spinosad was highlighted. (**B**) HPLC–MS analysis of butenyl-spinosyns A/D production in *St. albus* J1074 harboring the hybrid spnNEWbusA gene cluster. (**C**) Comparison of MS^2 fragmentation patterns of butenyl-spinosyn A and spinosyn A produced in *St. albus* J1074. Peaks at m/z 142.1 and 189.1 are the forosamine and trimethylrhamnose fragments respectively (34). Peaks at m/z 758.5 and 732.5 are the molecular ion of butenyl-spinosyn A and spinosyn A respectively. (**D**) Time course of the butenyl-spinosyn A yield resulted from *St. albus* J1074 harboring the hybrid spnNEWbusA gene cluster in the 5 l bioreactor containing 3 l medium. Each fermentation was performed twice and average data were shown.

mediated *in vitro* annealing (15) were tried to recombine the linear BAC and the *busA* fragment. However, no correct recombinant was obtained. RecET mediated *in vivo* LLHR promoted self-circularization of the linear BAC using DRs and generated a large number of colonies harboring smaller BACs with deletions (Supplementary Figure S6). We failed to obtain any colony in the T4pol *in vitro* annealing experiment. This suggested T4pol *in vitro* annealing can neither recombine large linear DNA molecules (102 and 6.6 kb in this case), nor promote self-circularization of linear DNA molecules using internal DRs as homology arms.

Seamless genome editing methods based on endonucleases, such as the homing endonuclease I-SceI (24) and the clustered regularly interspaced short palindromic repeats (CRISPR)-associated nuclease Cas9 (25,26), have been established in diverse organisms including actinomycetes (17,27–31). I-SceI can cleave double stranded DNA at the 18-bp recognition sequence (24) which does not exist in the genome of *Streptomyces* determined so far. In *Streptomyces*, I-SceI was used to create DNA double-strand breaks on the genome to promote homologous recombination and facilitate selection of desired mutants possessing the correctly edited genotype (27–29). However, the I-SceI recognition sequence has to be integrated into the genome firstly. The Cas9-guide RNA complex cleaves target DNA at the site three nucleotides upstream of a protospacer-adjacent motif (PAM) and produces the blunt-ended double-strand DNA break (25,26). Its simplicity and customizability have markedly promoted the genome engineering of *Streptomyces* (17,30,31). In this study, we

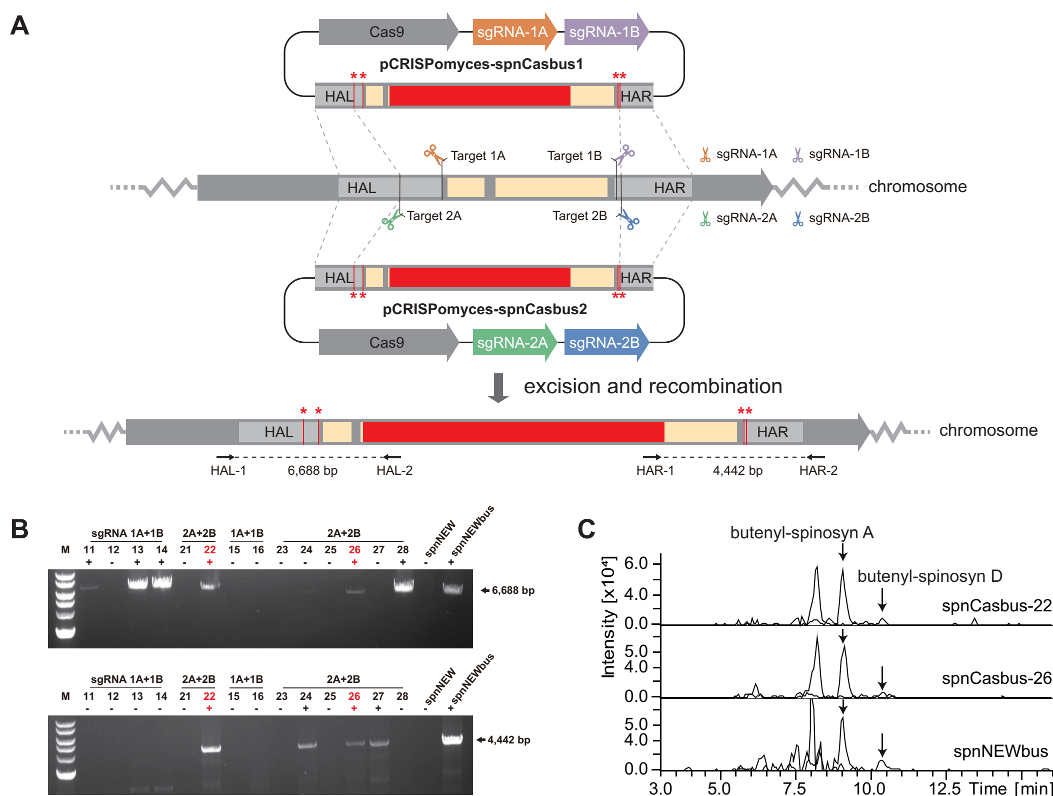


Figure 4. Generation of the butenyl-spinosyn gene cluster in *St. albus* J1074 using the CRISPR-Cas9 system. (**A**) The CRISPR-Cas9 strategy for inserting the AT1b-KS1a region from the *busA* gene into the *spnA* gene on the J1074 chromosome. Mutated protospacer sequences with synonymous mutations were indicated with asterisks. (**B**) Genotyping of exconjugants with PCR using primers indicated in (A). The upper panel is the agarose electrophoresis of PCR products with primers HAL-1 and HAL-2. The bottom panel is the agarose electrophoresis of PCR products with primers HAR-1 and HAR-2. Correct recombinants were highlighted in red. (**C**) HPLC analysis of butenyl-spinosyns A/D production in *St. albus* J1074 harboring the hybrid gene cluster generated with the CRISPR-Cas9 method (spnCasbus) and RedEx (spnNEWbus).

also tried inserting the 5301-bp AT1b-KS1a region from the *busA* gene into the spinosad gene cluster on the chromosome of *St. albus* J1074 using the CRISPR-Cas9 system. To achieve this, we constructed two targeting plasmids based on the pCRISPomyces-2 (17), pCRISPomyces-spnCasbus1 and pCRISPomyces-spnCasbus2 (Figure 4A). Each contains two single guide RNA (sgRNA) cassettes targeting two different positions outside of the region containing repetitive sequences in the *spnA* gene. Two 1-kb homology arms for recombination-driven repair were included in the *spnA* sequences flanking the 5301-bp *busA* region on the targeting plasmids. Both the targeting plasmids and the J1074 chromosome contain target sites of the Cas9–sgRNA complex. To avoid the Cas9–sgRNA complex cutting the targeting plasmid, five to six synonymous mutations were introduced to nucleotides in protospacers and their adjacent PAMs in the *spnA* sequences on targeting plasmids. When pCRISPomyces-spnCasbus1 and pCRISPomyces-spnCasbus2 were transferred into *St. albus* J1074 harboring the spinosad gene cluster, 3 ± 1 and 4 ± 2 exconjugants displaying thiostrepton resistance were obtained. For the pCRISPomyces-spnCasbus1, genotyping results suggested that no exconjugants possessed the desired recombination (Figure 4B). For the pCRISPomyces-spnCasbus2, two out of eight exconjugants displayed the desired recombination (Figure 4B). HPLC analysis of fermentation extracts from above two correctly edited *St. albus* J1074 exconjugants suggested that butenyl-spinosyns A/D were successfully produced and their yields are almost the same as the *St. albus* J1074 harboring the hybrid spnNEW-busA gene cluster generated by RedEX (Figure 4C).

In *Streptomyces*, homologous recombination is the major pathway to repair DNA double-strand DNA breaks created by CRISPR-Cas9. Currently, insertion of promoter cassettes, usually around 1 kb, into biosynthetic gene clusters has been greatly facilitated by the CRISPR-Cas9 system in *Streptomyces* (31). To the best of our knowledge, the 8-kb DNA fragment containing the *busA* and *spnA* sequences is the largest among those that have been inserted into the *Streptomyces* chromosome using the CRISPR–Cas9 system. In this experiment, we only got two correctly edited recombinants out of a few exconjugants from one of two targeting plasmids. The relative low endogenous homologous recombination activity of *Streptomyces* make insertion of large DNA fragments into the chromosome is challenging. The insertion efficiency would be improved if the homologous recombination is optimized in *Streptomyces* in future.

Gene synthesis is another choice to refactor pathways. However, *de novo* synthesis of large polyketide synthase genes is extremely challenging due to its highly repetitive nature. Furthermore, when the synthesized gene cluster needs

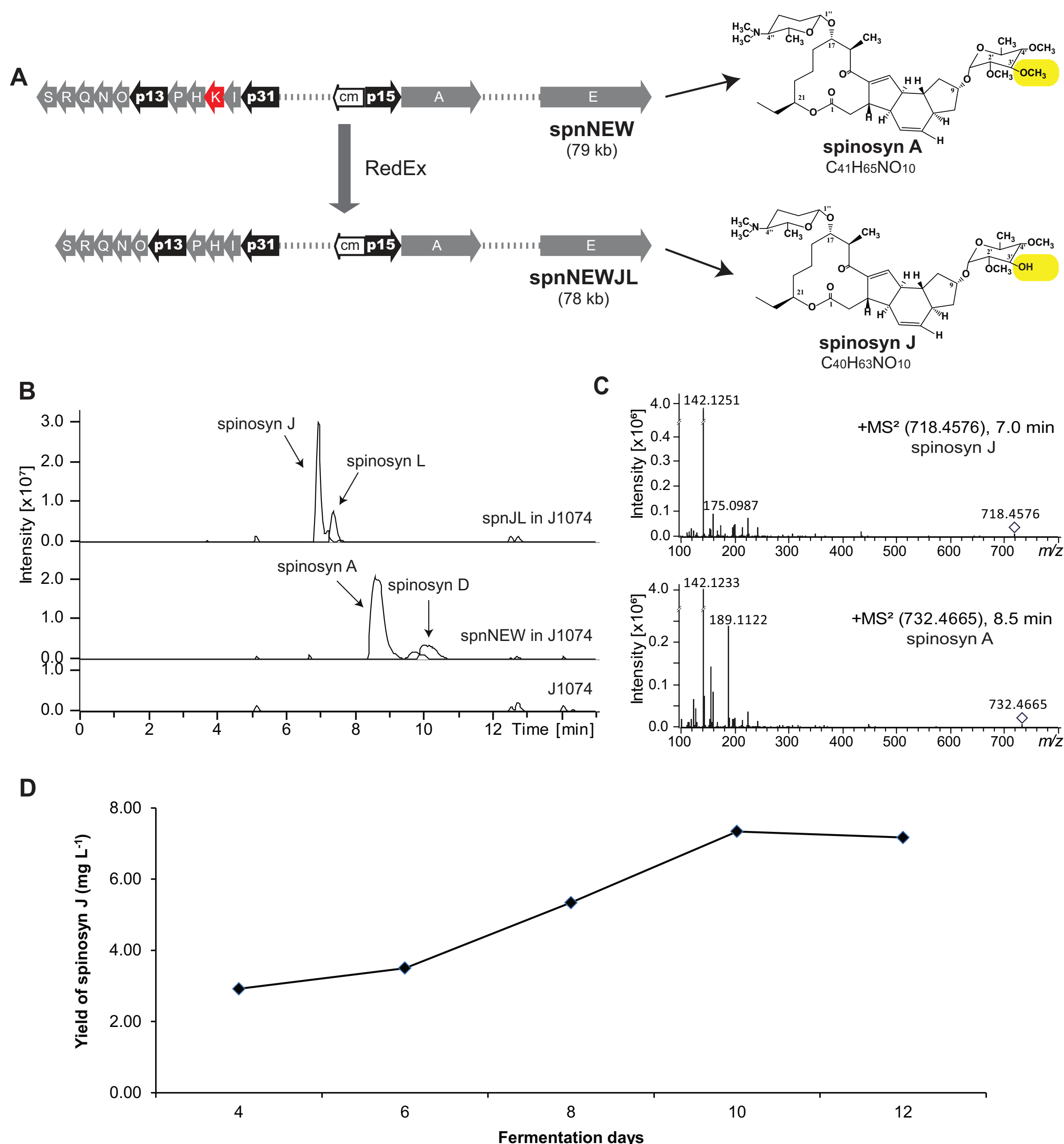


Figure 5. Heterologous expression of the spinosyns J/L biosynthetic pathway. (**A**) Seamless deletion of the *spnK* gene with RedEx. The group change at C-3′ of spinosad was highlighted. (**B**) HPLC–MS analysis of spinosyns J/L production in *St. albus* J1074 harboring the spnNEWJL gene cluster. (**C**) Comparison of MS^2 fragmentation patterns of spinosyn J and spinosyn A produced in *St. albus* J1074. The peak at m/z 175.1 in the upper panel is the 2,4-dimethyl-rhamnose fragment from spinosyn J (34). The peak at m/z 189.1 in the lower panel is the trimethylrhamnose fragment from spinosyn A (34). Peaks at m/z 718.5 and 732.5 are the molecular ion of spinosyn J and spinosyn A respectively. (**D**) Time course of the spinosyn J yield resulted from *St. albus* J1074 harboring the spnNEWJL gene cluster in the 5 l bioreactor containing 3 l medium. Each fermentation was performed twice and average data were shown.

to be modified, reliable methods such as the RedEx developed in this study are favorable.

Seamless deletion of the *spnK* gene in the artificial spinosad gene cluster using RedEx and heterologous expression of spinosyns J and L

Spinetoram is the next-generation spinosyn-based insecticide which received the Presidential Green Chemistry Challenge Award in 2008 because it is more active than spinosad and has a broader spectrum while maintaining the exceptional environmental and mammalian safety profile (4,32). Spinetoram is the mixture of 3′-*O*-ethyl-5,6-dihydro-spinosyn J and 3′-*O*-ethyl-spinosyn L produced by chemically modifying spinosyns J (3′-*O*-demethyl spinosyn A) and L (3′-*O*-demethyl spinosyn D). Spinosyns J and L were identified in a *Sa. spinosa* strain with point mutations in the *spnK* gene resulting in elimination of the rhamnose 3′-*O*-methyltransferase activity of SpnK (4).

To construct a gene cluster producing spinosyns J and L in heterologous host *St. albus* J1074, we used RedEx to seamlessly delete the *spnK* gene in the spnIKHP operon. The pBAC-spnNEWJL was generated from pBAC-spnNEW (11) which carries the artificial spinosad gene cluster.(Figure 5A). Briefly, the *spnK* gene on pBAC-spnNEW was replaced with the cassette *PacI-ampccdB-PacI* using Redαβ mediated LCHR in *E. coli* GBred-gyrA462 to obtain pBAC-spnNEW-ampccdBdelK. Then, pBAC-spnNEW-ampccdBdelK was digested with PacI to remove the *ampccdB* cassette and expose terminal 20-bp homology arms at both ends. The linear BAC was circularized using T4pol mediated *in vitro* annealing. When we evaluated the colonies from the selectable agar plate by restriction analysis, 100% accuracy of the recombinant pBAC-spnNEWJL BAC was observed (six out of six were correct, Supplementary Figure S7). This indicated that RedEx is highly efficient in seamless engineering of non-PKS regions in polyketide gene clusters.

Spinosyns J and L were successfully detected as the most abundant spinosyns in the fermentation broth of *St. albus* J1074 harboring pBAC-spnJL (Figure 5B and C). The structure of spinosyn J was confirmed by NMR analysis (Supplementary Table S3 and Supplementary Figures S8 and S9). When 3 l medium in the 5 l bioreactor was used for a 12-day fermentation, the highest yield, 7.34 mg l^{-1}, of spinosyn J was obtained at the 10th day (Figure 5D). Production of spinosyn J did not increase after 10 days. The yield of spinosyn J was five times lower than the yield of spinosyn A in *St. albus* under the same fermentation condition at the 10th day.

General design guides for RedEx

(1) The DNA sequence should be analysed using the Unipro UGENE software (19) to find the region without repetitive sequences flanking the target site. Direct repeats longer than 35 bp should be recognized as repetitive sequences, because they can mediate efficient LCHR.
(2) For seamless deletion, the *PacI-ampccdB-PacI* cassette can be prepared with PCR. For seamless insertion, the *PacI-ampccdB-PacI* cassette can be stitched with the insertion sequence using overlap extension PCR or DNA assembly methods. Two PacI sites should be flanked by 20-bp direct repeats.
(3) The *PacI-ampccdB-PacI* cassette is inserted into the target site with LCHR in the *ccdB* resistant *E. coli* strain GBred-gyrA462 (14).
(4) Correct recombinants can be selected with restriction analysis.
(5) PacI digestion is used to remove the *ampccdB* cassette and expose the terminal 20-bp overlaps.
(6) The circular finally recombined DNA is recovered by transforming the T4pol *in vitro* reacted products into the *ccdB* sensitive *E. coli* strain with neither RecET nor Redαβ recombinase expression, such as GB2005 (14), to minimize unexpected recombination.
(7) Finally, correct recombinants can be selected with restriction analysis. Usually, one restriction analysis is sufficient to ensure the integrity of BAC. More than one restriction analysis using different enzymes can be performed to ensure that there is no unwanted fragment loss caused by unexpected recombination.

Concluding remarks

Biosynthetic pathways involving modular enzymes such as PKSs are still the focus of synthetic biology inspired engineering. Recently, Eng *et al.* released a web-based toolkit, ClusterCAD, to streamline and simplify the design of chimeric PKSs (33). Our increased knowledge and information on the structural blueprint of PKS megaenzymes and the interplay of individual catalytic domains will facilitate bring us closer to the promise of rational engineering of PKSs. Therefore, it is useful to enrich the toolbox for genetic engineering of PKS genes. To exploit the potential of combinatorial biosynthesis in producing spinosyn analogues with altered polyketide carbon skeletons or methylation of sugar moiety, we developed the RedEx method which is highly efficient in seamless DNA insertion and deletion in large multimodular polyketide gene clusters. The advantages of RedEx in seamless genetic engineering will facilitate reprograming of complex biosynthetic pathways including both polyketide assembly lines and others, particularly for those in the host in which the gene editing method has not been established.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Jingyao Qu, Jing Zhu and Zhifeng Li from State Key Laboratory of Microbial Technology of Shandong University for help and guidance in LC–MS.

FUNDING

National Key Research & Development Program of China [2018YFA0900400 to H.W., J.F.]; National Natural Science Foundation of China [31700073 to H.W., 31670097 to Y.Z.]; 111 Project [B16030 to Y.Z.]; Recruitment Program of Global Experts in Shandong University (to Y.Z.); Qilu Youth Scholar Startup Funding of SDU (to H.W.); Taishan Scholar Program of Shandong Province (to H.W. and J.F.). Funding for open access charge: Taishan Scholar Program of Shandong Province.
Conflict of interest statement. A.F.S. and Y.Z. are shareholders in Gene Bridges GmbH, which holds exclusive commercial rights to recombineering.

REFERENCES

1. Mertz,F.P. and Yao,R.C. (1990) *Saccharopolyspora spinosa* sp. nov. isolated from soil collected in a sugar mill rum still. *Int. J. Syst. Evol. Microbiol.*, **40**, 34–39.
2. Kirst,H.A. (2010) The spinosyn family of insecticides: realizing the potential of natural products research. *J. Antibiot. (Tokyo)*, **63**, 101–111.
3. Madduri,K., Waldron,C., Matsushima,P., Broughton,M.C., Crawford,K., Merlo,D.J. and Baltz,R.H. (2001) Genes for the biosynthesis of spinosyns: applications for yield improvement in *Saccharopolyspora spinosa*. *J. Ind. Microbiol. Biotechnol.*, **27**, 399–402.

4. Galm,U. and Sparks,T.C. (2016) Natural product derived insecticides: discovery and development of spinetoram. *J. Ind. Microbiol. Biotechnol.*, **43**, 185–193.
5. Sheehan,L.S., Lill,R.E., Wilkinson,B., Sheridan,R.M., Vousden,W.A., Kaja,A.L., Crouse,G.D., Gifford,J., Graupner,P.R., Karr,L. *et al.* (2006) Engineering of the spinosyn PKS: directing starter unit incorporation. *J. Nat. Prod.*, **69**, 1702–1710.
6. Hahn,D.R., Gustafson,G., Waldron,C., Bullard,B., Jackson,J.D. and Mitchell,J. (2006) Butenyl-spinosyns, a natural example of genetic engineering of antibiotic biosynthetic genes. *J. Ind. Microbiol. Biotechnol.*, **33**, 94–104.
7. Weissman,K.J. (2016) Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.*, **33**, 203–230.
8. Hertweck,C. (2015) Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci.*, **40**, 189–199.
9. Smanski,M.J., Zhou,H., Claesen,J., Shen,B., Fischbach,M.A. and Voigt,C.A. (2016) Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol.*, **14**, 135–149.
10. Tan,G.Y., Deng,K., Liu,X., Tao,H., Chang,Y., Chen,J., Chen,K., Sheng,Z., Deng,Z. and Liu,T. (2017) Heterologous biosynthesis of spinosad: an omics-guided large polyketide synthase gene cluster reconstitution in *Streptomyces*. *ACS Synth. Biol.*, **6**, 995–1005.
11. Song,C., Luan,J., Cui,Q., Duan,Q., Li,Z., Gao,Y., Li,R., Li,A., Shen,Y., Li,Y. *et al.* (2019) Enhanced heterologous spinosad production from a 79-kb synthetic multi-operon assembly. *ACS Synth. Biol.*, **8**, 137–147.
12. Huang,J., Yu,Z., Li,M.H., Wang,J.D., Bai,H., Zhou,J. and Zheng,Y.G. (2016) High level of spinosad production in the heterologous host *Saccharopolyspora erythraea*. *Appl. Environ. Microbiol.*, **82**, 5603–5611.
13. Wang,H., Li,Z., Jia,R., Hou,Y., Yin,J., Bian,X., Li,A., Müller,R., Stewart,A.F., Fu,J. *et al.* (2016) RecET direct cloning and Redαβ recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression. *Nat. Protoc.*, **11**, 1175–1190.
14. Wang,H., Bian,X., Xia,L., Ding,X., Müller,R., Zhang,Y., Fu,J. and Stewart,A.F. (2014) Improved seamless mutagenesis by recombineering using *ccdB* for counterselection. *Nucleic Acids Res.*, **42**, e37.
15. Wang,H., Li,Z., Jia,R., Yin,J., Li,A., Xia,L., Yin,Y., Muller,R., Fu,J., Stewart,A.F. *et al.* (2018) ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Res.*, **46**, e28.
16. Wang,Y., Cobb,R.E. and Zhao,H. (2016) High-efficiency genome editing of *Streptomyces* species by an engineered CRISPR/Cas system. *Methods Enzymol.*, **575**, 271–284.
17. Cobb,R.E., Wang,Y. and Zhao,H. (2015) High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.*, **4**, 723–728.
18. Engler,C., Kandzia,R. and Marillonnet,S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647.
19. Okonechnikov,K., Golosova,O., Fursov,M. and team,U. (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, **28**, 1166–1167.
20. Donadio,S., McAlpine,J.B., Sheldon,P.J., Jackson,M. and Katz,L. (1993) An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 7119–7123.
21. McDaniel,R., Thamchaipenet,A., Gustafsson,C., Fu,H., Betlach,M. and Ashley,G. (1999) Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 1846–1851.
22. Scior,A., Preissler,S., Koch,M. and Deuerling,E. (2011) Directed PCR-free engineering of highly repetitive DNA sequences. *BMC Biotechnol.*, **11**, 87.
23. Fu,J., Bian,X., Hu,S., Wang,H., Huang,F., Seibert,P.M., Plaza,A., Xia,L., Müller,R., Stewart,A.F. *et al.* (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.*, **30**, 440–446.
24. Monteilhet,C., Perrin,A., Thierry,A., Colleaux,L. and Dujon,B. (1990) Purification and characterization of the *in vitro* activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron. *Nucleic Acids Res.*, **18**, 1407–1413.
25. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
26. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
27. Siegl,T., Petzke,L., Welle,E. and Luzhetskyy,A. (2010) I-SceI endonuclease: a new tool for DNA repair studies and genetic manipulations in *Streptomyces*. *Appl. Microbiol. Biotechnol.*, **87**, 1525–1532.
28. Lu,Z., Xie,P. and Qin,Z. (2010) Promotion of markerless deletion of the actinorhodin biosynthetic gene cluster in *Streptomyces coelicolor*. *Acta Biochim. Biophys. Sin. (Shanghai)*, **42**, 717–721.
29. Fernandez-Martinez,L.T. and Bibb,M.J. (2014) Use of the meganuclease I-SceI of *Saccharomyces cerevisiae* to select for gene deletions in actinomycetes. *Sci. Rep.*, **4**, 7100.
30. Zeng,H., Wen,S., Xu,W., He,Z., Zhai,G., Liu,Y., Deng,Z. and Sun,Y. (2015) Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl. Microbiol. Biotechnol.*, **99**, 10575–10585.
31. Zhang,M.M., Wong,F.T., Wang,Y., Luo,S., Lim,Y.H., Heng,E., Yeo,W.L., Cobb,R.E., Enghiad,B., Ang,E.L. *et al.* (2017) CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.*, **13**, 607–609.
32. Sparks,T.C., Crouse,G.D., Dripps,J.E., Anzeveno,P., Martynow,J., Deamicis,C.V. and Gifford,J. (2008) Neural network-based QSAR and insecticide discovery: spinetoram. *J. Comput. Aided Mol. Des.*, **22**, 393–401.
33. Eng,C.H., Backman,T.W.H., Bailey,C.B., Magnan,C., Martin,G., Katz,L., Baldi,P. and Keasling,J.D. (2018) ClusterCAD: a computational platform for type I modular polyketide synthase design. *Nucleic Acids Res.*, **46**, D509–D515.
34. Lewer,P., Hahn,D.R., Karr,L.L., Duebelbeis,D.O., Gilbert,J.R., Crouse,G.D., Worden,T., Sparks,T.C., Edwards,P.M. and Graupner,P.R. (2009) Discovery of the butenyl-spinosyn insecticides: novel macrolides from the new bacterial strain *Saccharopolyspora pogona*. *Bioorg. Med. Chem.*, **17**, 4185–4196.