The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly

Wei-Hua Chen[1,2,3], Zhong-Jun Qin[1], Jin Wang[1,*] and Guo-Ping Zhao[1,2,4,*]

[1]Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 20032, China, [2]Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai 201203, China, [3]School of Medicine, Shanghai Jiao Tong University, Shanghai 200025, China and [4]State Key Lab of Genetic Engineering and Center for Synthetic Biology; Department of Microbiology and Microbial Engineering, School of Life Sciences, Fudan University, Shanghai 200032, China

ABSTRACT

Techniques for assembly of designed DNA sequences are important for synthetic biology. So far, a few methods have been developed towards high-throughput seamless DNA assembly *in vitro*, including both the homologous sequences-based system and the type IIS-mediated system. Here, we describe a novel method designated 'MASTER Ligation', by which multiple DNA sequences can be seamlessly assembled through a simple and sequence-independent hierarchical procedure. The key restriction endonuclease used, MspJI, shares both type IIM and type IIS properties; thus, it only recognizes the methylation-specific 4-bp sites, ^{m}CNNR (R = A or G), and cuts DNA outside of the recognition sequences. This method was tested via successful assembly of either multiple polymerase chain reaction amplicons or restriction fragments of the actinorhodin biosynthetic cluster of *Streptomyces coelicolor* ($\sim$29 kb), which was further heterologously expressed in a fast-growing and moderately thermophilic strain, *Streptomyces sp.* 4F.

INTRODUCTION

Synthetic biology technology adapts the basic concepts of engineering so that pre-designed and characterized DNA parts are combined to form complex modules for further construction of biological systems with desired new properties (1,2). As enormous DNA sequences with various encoding capacities are being determined in the post-genomic era, the demand for extended ability to analyse these data and to synthesize DNA sequences of any length with any kinds of complexity is increasing (3).

Start with a pool of chemically synthesized oligos, DNA molecules of 1–3 kb in length can usually be constructed through polymerase chain assembly (4). For downstream assembly of larger DNA molecules, multiple homologous sequences-based systems have been developed, including the *in vitro* sequence and ligation-independent cloning (SLIC) strategy (5), the USER fusion method (6), the in-Fusion BioBrick system (7), the Gibson assembly (8) and several *in vivo* assembly strategies (9–13). Among them, Gibson assembly, as a hierarchical method that avoids recurrent polymerase chain reaction (PCR) steps and works through *in vitro* recombination of the long homologous sequences between the ends of the fragments to be assembled, greatly increases the efficiency of synthesis of large DNA molecules, even at the scale of a bacterial genome (e.g. the 1.08-Mb genome of *Mycoplasma mycoides* JCVI-syn1.0) (14,15). However, these homologous recombination-based methods are unsuitable for assembly of DNA molecules with repetitive sequences (16), such as the TALEN DNA-binding modules (17,18), polyketide modules (19), CRISPR gene cluster (20) and any DNA sequence that appears more than once in different building blocks (e.g. a common promoter sequence).

For complex DNA sequences, methods based on type IIS restriction endonucleases (REs) can be adopted (18,21). Unlike the type IIP REs, the type IIS REs

*To whom correspondence should be addressed. Tel: +86 21 54924002; Fax: +86 21 54924015; Email: gpzhao@sibs.ac.cn
Correspondence may also be addressed to Jin Wang. Tel: +86 21 54971126; Fax: +86 21 54924015; Email: cmg.wangjin@gmail.com

cleave DNA sequences outside of their recognition sites, and thus can be used to generate short non-palindromic overhangs (usually 4 bp), making seamless assembly of DNA sequences, including complex sequences, achievable in a high-throughput manner (18,21,22). Notwithstanding, as the constructs become larger, the increasing wastage of unique type IIS restriction sites becomes the key limitation factor in practise. Although these problems can be solved by removal of the internal restriction sites from the primary sequences (23–25), unpredicted phenotypic changes may be caused by DNA base substitution (25). Recently, the pairwise selection assembly (PSA) strategy was developed, and the internal type IIS restriction sites can be sealed by methylation (26). However, four steps must be taken to specifically block the additional sites.

MspJI, a remote homologue of the *Escherichia coli*'s methylation modification-dependent RE (type IIM) Mrr (27), was recently characterized from *Mycobacterium* sp. JLS (28,29) too possess additional type IIS RE properties. It recognizes ^{m}CNNR (R = A or G) sites and cleaves DNA at N12/N16 from the 3′-side of the modified cytosine, producing four-base 5′-overhangs of any sequences (30). In this study, we developed an MspJI-mediated seamless DNA assembly method and tested it via successful assembly of either multiple PCR amplicons or the actinorhodin biosynthesis gene cluster (∼29 kb) from *Streptomyces coelicolor*, which was then heterologously expressed in *Streptomyces sp*. 4F.

MATERIALS AND METHODS

Bacterial strains, media, reagents and primers

Streptomyces strains were grown in mannitol-soy flour (MS) media for preparation of spores and on R2YE media for production of pigmented actinorhodin (31). *E. coli* strains were grown in Luria-Bertani (LB) media (32). *E. coli* strain DH5α was used for DNA cloning and strain ET12567 was used for DNA demethylation *in vivo* and for conjugation with *Streptomyces sp*. 4F (33).

Unless specified, restriction endonucleases were purchased from NEB. GeneRuler^{TM} 1-kb DNA ladder (250–10 000 bp), thermosensitive alkaline phosphatase (FastAP^{TM}) and T4 DNA ligase were purchased from Thermo Fermentas. Plasmid Maxi Kit from Qiagen was used for preparation of plasmids.

All primers used in this study were listed in Supplementary Table S1.

Generating PCR amplified target DNA fragments

Three pairs of primers were used to amplify three continuous fragments (fragments A, B and C) of the actinorhodin biosynthetic gene cluster of *S. coelicolor* (Figure 1A). The PCR amplicons were designed to be ∼4 kb in size and each overlapped 4 bp with the adjacent fragments. A common tag (GGCAGCCGTCTCCCTGGAAG) was attached to the 5′-termini of all primers, and the non-methylated MspJI recognition site (CTGG) was set nine bases before the overlapping four bases (for details refer Supplementary Figure S1). A universal primer kjz.ml2 (GGCAGCCGTCTCC^{m}CTGGAAG) was used

to introduce methylation modification (^{m}CTGG) at both ends of the aforementioned PCR amplicons by a second round of PCR amplification. The cleavage sites of MspJI on designed primers are highlighted in bold. The non-palindromic cohesive ends of the fragments are carefully chosen to ensure that only the adjacent cohesive ends are complementary after MspJI digestion. The pre-designed complementary sticky ends determine the order and direction of the annealed DNA fragments. PCR amplification was carried out in a 25-μl reaction mixture consisting of 0.3 μM primers, 1.5 mM Mg^{2+}, 200 μM dNTP and 0.5 U KOD-401 DNA polymerase (TOYOBO). Thirty cycles were conducted with the PCR procedures of 30 s at 97°C, 30 s at 65°C and 3 min at 72°C, with a final extension of 10 min. DNA normalization was performed with a Nanodrop ND-2000 Spectrophotometer (Thermo). MspJI digestion (30) was carried out at 37°C in 30-μl reaction mixture consisting of 0.2–1 μg of PCR products, 3 μl of 10× NEBuffer 4, 1 μl of 100× bovine serum albumin, 1 μl of 10 μM MspJI activator solution, 2–4 U of MspJI (NEB) and nuclease-free water to 30 μl. The MspJI activator was prepared through self-annealing of the oligo kjz.ml1, which was first incubated at 65°C for 1 h in 1× saline-sodium citrate (SSC) buffer using a PCR machine and then slowly cool down in room temperature. After MspJI digestion, fragments with complementary cohesive ends were purified by agarose gel electrophoresis and were then ligated with T4 DNA ligase at 16°C for 2 h. Ligation reaction was terminated by adding 5 μl of 6× DNA loading buffer, denatured at 65°C for 15 min and then analysed by agarose gel electrophoresis.

Constructing the integrative plasmid for heterologous expression of the actinorhodin gene cluster (*act*)

To test the accuracy of the assembled whole actinorhodin gene cluster, an integrative vector pHI was constructed for heterologous expression of actinorhodin by integration into the chromosome of the moderately thermophilic *Streptomyces sp*. 4F (33). Specifically, a 4648-bp fragment 'H' containing the replication region of Supercos 1 (34) and the thiostrepton resistance gene (*tsr*) were amplified from pHAQ31 (35) with primers kly.HAQ1 and kly.HAQ2. A 3035-bp fragment 'I' containing *oriT* and the phiC31 integrase encoding gene was amplified from pSET152 (36) with primers kly.Int1 and kly.Int2. Both fragments were digested with XbaI, and fragment 'H' was treated with FastAP^{TM} before being ligated with fragment 'I' using T4 DNA ligase. The ligation mixture was then used as a template for PCR amplification with primers lae.fHAQ and lae.rHAQ, and the PCR product was then digested with NheI and allowed for self-ligation to generate the heterologous expression vector pHI.

The protocol for seamless assembly of *act* biosynthetic gene cluster

The process for seamless assembly of endonuclease restriction fragments is illustrated in Figure 1B. The *act* gene cluster was *in silico* divided into four pieces of four bases-overlapping fragments (fragments I–IV), which were amplified by PCR and cloned into the XbaI site of the
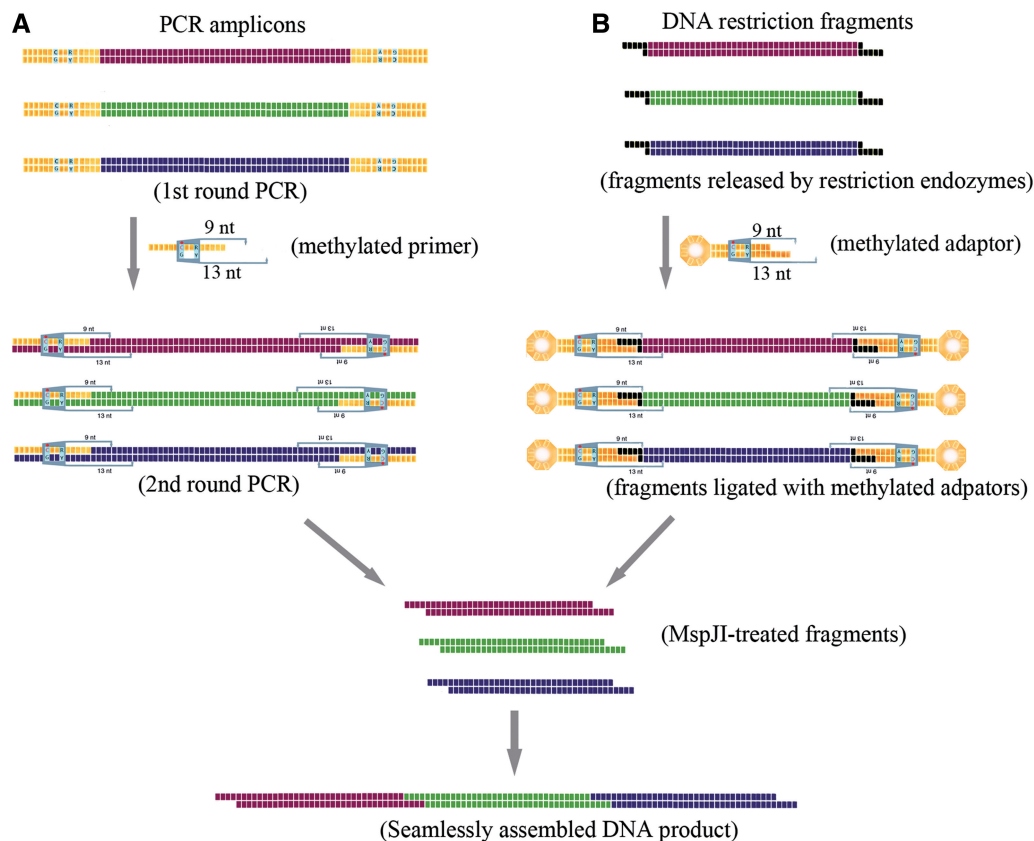
Figure 1. Rationale scheme for *in vitro* seamless assembly of multiple PCR amplicons (**A**) or restriction fragments (**B**). The universal sequence was illustrated by yellow bricks (A), and the restriction site was shown by black bricks (B). The methylation of cytosine was indicated by a red dot above the 'C' letter. Details of PCR amplification, DNA ligation and MspJI digestion can be found in the 'Materials and Methods' section.

pBluescript II KS vector. These subclones were verified via DNA sequencing and transformed into methylation-deficient *E. coli* ET12567 for DNA demethylation. The methylation-free fragments were then released by XbaI digestion. To remove the scars of XbaI restriction sites at the ends of fragments to be assembled, a methylated adaptor klf.ML2 (Supplementary Table S1), which was compatible with the XbaI site, was designed to form a stem–loop structure on annealing, containing the methylated recognition sites (^{m}CTGG) of MspJI. The position of the restriction sites was exactly set to guide MspJI to remove the scars produced by XbaI restriction and generate cohesive ends for the following assembly. Preparation of methylated adaptor klf.ML2 was the same as preparing the MspJI activator kjz.ml1. Fragments with methylated adaptors were digested with MspJI in the presence of its activator kjz.ml1 (Supplementary Table S1).

The MspJI-digested fragment I (7420 bp) and fragment II (8171 bp) were ligated to form the assembly intermediate fragment I–II (15587 bp). With pBluescript II KS as the template, primers Lae.VE1 and Lae.VE4 were used to PCR amplify a 2.4-kb 'vector' product, which excluded the lacZα gene and was digested by BbsI to form compatible cohesive ends with fragment I–II. Notably, both primers contained an XbaI restriction site. Through PCR amplification, the XbaI site was introduced to both

ends of the 'vector'. The BbsI-treated 'vector' product was ligated with fragment I–II to obtain plasmid pSCactI–II, and the newly introduced XbaI site on the 'vector' can be used to release the cloned fragment I–II for the next round of assembly. In the same way, fragments III (6410 bp) and IV (6686 bp) were ligated to form fragment III–IV, which was subsequently ligated with the 'vector' products prepared with primers Lae.VE5 and Lae.VE8, forming pSCactIII–IV.

Plasmids pSCactI–II, pSCactIII–IV and pHI were transformed into methylation-deficient *E. coli* strain ET12567 for demethylation. Fragments I–II and III–IV were released by XbaI digestion, and pHI was digested with NheI. All the fragments were purified, ligated with methylated adaptors and then treated with MspJI digestion before they are ligated together to form pHIW, which contained the whole *act* biosynthetic cluster. Plasmid pHIW, as well as the control plasmid pHI, was transformed into *E. coli* ET12567/pUZ8002 and was then conjugated into the moderately thermophilic strain *Streptomyces sp.* 4F (33). Exconjugants were selected on MS media with nalidixic acid (50 μg/μl) and aparmycin (50 μg/μl) to obtain the recombinant strains. Candidates were then verified by PCR amplification (Supplementary Table S1 and Supplementary Figure S2), and positive recombinant strains were then cultured on R2YE medium at 30°C to allow for heterologous expression of actinorhodin in 4F.

Accessing the accuracy of DNA ligation

For assessing the ligation of the PCR amplicons, plasmid pBluescript II KS (Stratagene) was used as a template for PCR amplification to prepare the 'vector' fragment for ligation with the MspJI-treated fragments. Specifically, primers accu-A and accu-F1 were used to amplify the 2.5-kb 'vector' fragment, which was then purified, digested by BbsI and ligated with MspJI-digested fragment A. Similarly, 'vector' fragments amplified by primer pairs of accu-A/accu-F2 and accu-A/accu-F3 were used to ligate with two (fragments A and B) and three (fragments A, B and C) MspJI-treated fragments, respectively. DNA ligation was performed at 22°C overnight, and the product was transformed into DH5α competent cells. For accessing the accuracy of ligation of restriction fragments, i.e. construction of pSCactI–II, pSCactIII–IV and pHIW, the ligation procedures were the same as described earlier in the text.

The ligation accuracy was first checked by the sizes of the plasmids, and then candidate plasmids of correct sizes were further verified either by DNA sequencing of the ligation sites or by BamHI restriction analysis (Table 1).

RESULTS

A type IIM/IIS RE, MspJI was used to develop a sequence-independent seamless DNA assembly method, where the key step is to generate a set of target DNA fragments flanked by MspJI recognition sites of methylated ^mCNNR. The sites can be designed either in primers and easily introduced to the ends of the DNA sequences through PCR or in the methylated adaptors with sticky ends compatible with those of the target restriction DNA fragments generated by an RE cleavage.

These two procedures are illustrated in Figure 1 and have been tested with the *S. coelicolor act* biosynthetic gene cluster as the target sequence.

Seamless assembly of multiple PCR amplicons

In case that PCR amplicons are desired to be the targets for ligation, the strategy is illustrated in Figure 1A. Part of the *act* gene cluster, a 12-kb DNA sequence, was used to testify this strategy for seamless assembly (Figure 2A).

Table 1. The efficiency of fragments assembly

Fragments	Size (kb)	Plasmid sizes[a] (positive/ tested)	Verification of positive clones[b] (correct/checked)
PCR amplicons			
Vector + A	6.5	30/30 (100%)	DNA sequencing (S) (10/10)
Vector + A + B	10.5	26/30 (87%)	S (10/10)
Vector + A + B + C	14.6	17/30 (57%)	S (10/10)
Restriction fragments			
pSCactI-II	18	22/30 (73%)	S (10/10)
pSCactIII-IV	15.5	15/30 (50%)	S (10/10)
pHIW	36.3	2/30 (6.7%)	BamHI restriction (2/2)

[a]Transformants were first checked by the plasmid sizes and the efficiency was shown in the bracket.
[b]Except plasmid pHIW, 10 positive clones of each construction were chosen for DNA sequencing verification, among which five clones were sequenced with primer Mseq-1, whereas the other five with Mseq-2 (Supplementary Table S1). Based on the sequencing results, the DNA sequences of the ligation sites were correct in all tested clones. Plasmid pHIW was verified by the BamHI restriction (Figure 3D). Two obtained clones with the desired sizes were proven correct based on the BamHI restriction analysis.
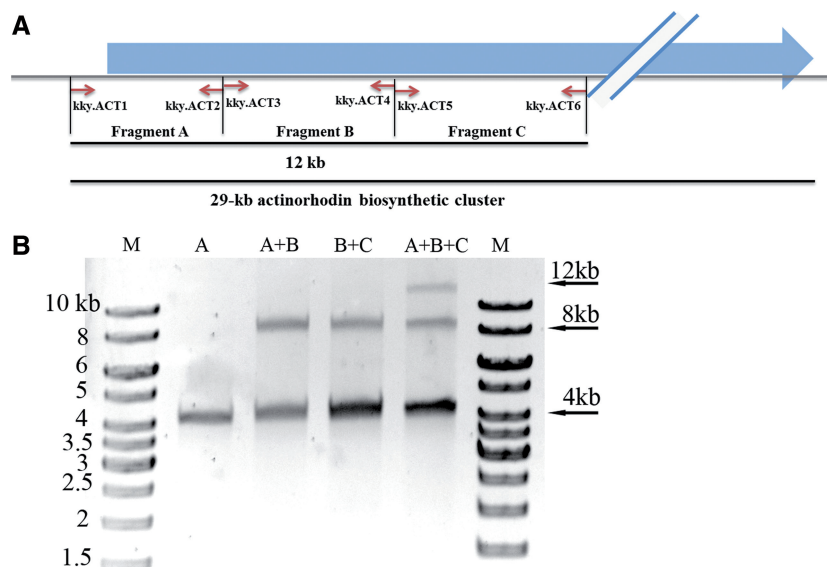

Figure 2. *In vitro* seamless assembly of 3 PCR amplicons. (**A**) The schematic diagram of the PCR amplicons. The fragments A (4092 bp), B (4000 bp) and C (4084 bp) were amplified from the actinorhodin biosynthetic cluster with designed primers (Supplementary Table S1). (**B**) The PCR amplicons were digested with MspJI and used for ligation at 16°C for 2 h. Fragments A, B and C were indicated on each lane. 'M' stood for 1-kb DNA ladder, and the sizes of the ladder were labelled.

The DNA sequence was first *in silico* divided into three fragments (A, B and C) of ∼4 kb each in length, which were then amplified with three pairs of PCR primers individually. The MspJI recognition sites were introduced at the ends of the amplicons through a second round of PCR amplification ('Materials and Methods' section). After MspJI cleavage, combinations of fragments A, B and C were ligated with T4 DNA ligase (Figure 2B). Specifically, when fragment A was incubated with fragment B, an 8-kb ligated product of fragment A–B could be clearly observed, and so as for the ligation between fragments B and C. When fragments A, B and C were ligated, both 8- and 12-kb fragments were obtained, which were the products of two and three fragments, respectively, and the final product of fragment A–B–C was ∼10% of the total products.

The MspJI-digested fragments were ligated with pBluescript II KS to check the accuracy of DNA ligation ('Materials and Methods' section). The ligation products were used to transform *E. coli* competent cells, and the plasmids from the transformants were verified. Based on their sizes, the accuracy was 100% for ligation of one fragment and was 87 and 57% for two and three fragments, respectively. Plasmids of correct sizes were then randomly selected for DNA sequencing, and the ligation sites were proved to be correct in all tested clones (Table 1).

Seamless assembly of multiple restriction fragments using methylated adaptors

Although PCR is well known for its convenience in efficient amplification of any desired regions of DNA fragments, it has not achieved the fidelity as high as that in the process of *in vivo* replication, e.g. 4.4×10^{-7} error rate of Phusion high-fidelity DNA polymerase (Thermo scientific) versus 1×10^{-9} in *E. coli* chromosomal replication (37). Therefore, with the increase of the sizes of DNA to be assembled, it is more and more difficult to get successful and truthful PCR amplification, and usually repeated site-directed mutagenesis is required to correct the mutations introduced by PCR, bringing more complicated and time-consuming work burden. To guarantee the fidelity, a large DNA sequence can be divided into several small sub-fragments, each of which can be generated either by direct PCR amplification or by *de novo* DNA synthesis and be individually cloned and verified by DNA sequencing before they are used for the assembly of the large DNA construct.

The MspJI-mediated sequence-independent seamless assembly strategy for multiple DNA fragments excised from the aforementioned sequencing-verified plasmids via RE cleavage is illustrated in Figure 1B. With this strategy, we successfully assembled the whole actinorhodin biosynthetic gene cluster of 29 kb, which product is an antibiotic produced by the model *S. coelicolor* and is blue when the medium pH is >8.5 (31). The cluster was first *in silico* divided into four DNA fragments (I–IV) of 6–8 kb (Figure 3A and 'Materials and Methods' section), all of which were PCR amplified and cloned into the XbaI site of the plasmid pBlueScript II KS individually for DNA sequencing. Plasmids containing errorless fragments were

demethylated via amplification in the methylation-deficient *E. coli* strain ET12567 host. Fragments (I–IV) were then released from their respective vectors with XbaI digestion and were ligated with the methylated adaptor klf.ML2 with the MspJI recognition sites (^{m}CTTG). Because the distance between the recognition site and the four-base overlapping region is specially designed to be nine bases, the MspJI digestion may produce the cohesive ends exactly at the preset overlapping region (Supplementary Figure S1). MspJI-digested fragments I and II as well as fragments III and IV were ligated to form fragments I–II and III–IV (Figure 3B), respectively, which were then purified and ligated to obtain the full-length actinorhodin gene cluster (fragment I–II–III–IV) and finally cloned in the heterologous expression vector pHI to obtain pHIW (Figure 3C).

Plasmid pHIW was first verified by BamHI restriction map (Figure 3D and Table 1) and then introduced by conjugation from *E. coli* ET12567 into *Streptomyces sp.* 4F and integrated into the attB site of the chromosome catalysed by the phiC31 integrase (36). Exconjugants were selected by apramycin and nalidixic acid resistance, and the integration between the pHIW-attP site and the 4F-attB site was confirmed by PCR verification (Supplementary Figure S2). The recombinant 4F strains were then cultured on R2YE agar at 30°C, and the successful production of the blue pigment of actinorhodin demonstrated the practicability of this novel methodology (Figure 3E).

DISCUSSION

In this study, an MspJI mediated-seamless assembly method was developed to facilitate *in vitro* seamless assembly of any length of DNA using the same enzymatic reagents and reaction conditions. This methodology is designated MASTER Ligation, standing for methylation-assisted tailorable ends rational ligation. Comparing with other *in vitro* seamless assembly methods (Table 2), three technical features of this novel technology shall be illustrated later in the text in detail.

The 'methylation-assisted' feature enables MASTER Ligation to assemble DNA in a sequence-independent manner

MspJI belongs to type IIM REs and thus is modification-dependent (30). As the un-methylated internal sequences are not recognized by MspJI, assisted by methylated primers or adaptors, unique methylated restriction site (^{m}CNNR) of MspJI can be easily introduced into the ends of the fragments to be assembled; therefore, the 'methylation-assisted' property is emphasized in the nomenclature. Theoretically, no extra step is needed for the assembly of PCR amplicons, and only one extra step of ligation of adaptors is required for the assembly of restriction fragments. Although a unique restriction site is required to release the restriction fragments to be assembled, the large numbers of available commercial type IIP enzymes and the fact that different enzymes can be used for each fragment make the choice of a unique site
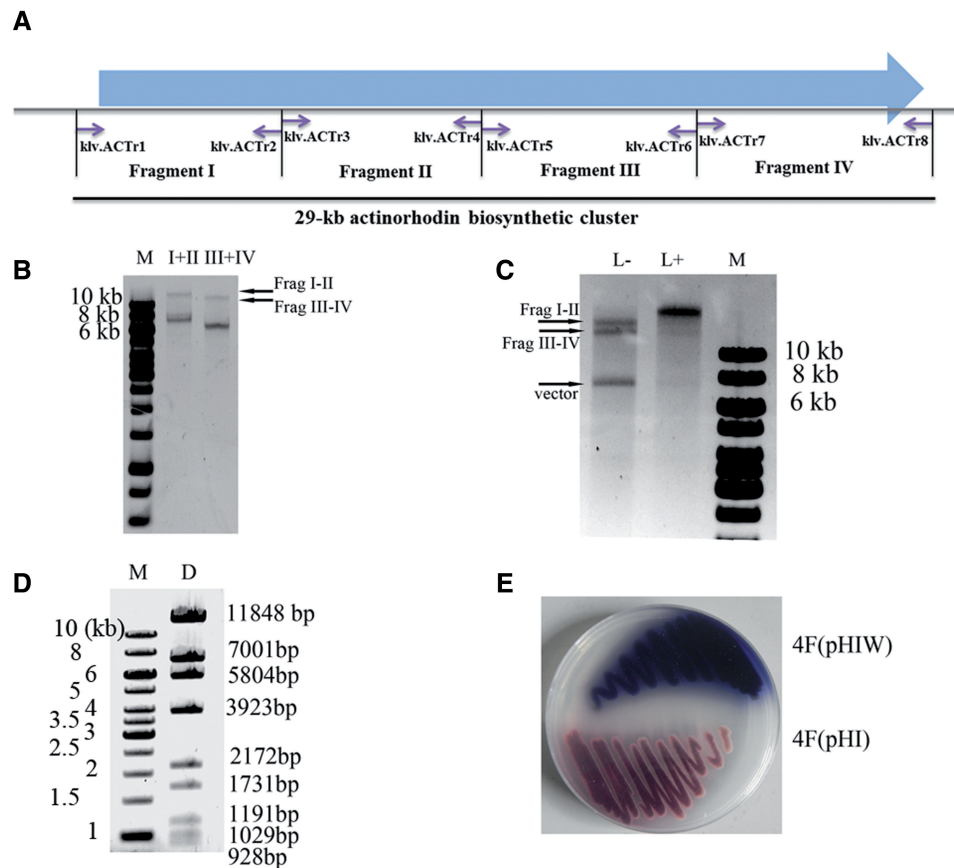
Figure 3. *In vitro* seamless assembly of the whole actinorhodin biosynthetic cluster from multiple restriction fragments. (**A**) The schematic diagram of the PCR amplicons. The 29-kb actinorhodin biosynthetic cluster was divided into four fragments, which were then PCR amplified with primers specified (Supplementary Table S1). (**B**) Errorless demethylated fragments I (7420 bp), II (8171 bp), III (6410 bp) and IV (6686 bp) were released from pBluescript II KS by XbaI digestion, which were then ligated with a designed adaptor klf.ML2 (Figure 1B and Supplementary Table S1) and digested with MspJI as described in the 'Materials and Methods' section. The MspJI-treated fragments were used for ligation with T4 DNA ligase at 16°C for 2 h. (**C**) Fragments I–II and III–IV were ligated with pHI with ('L+') or without ('L−') the addition of T4 DNA ligase. The synthesized DNA could be viewed in the 'L+' lane together with the disappearance of the substrate fragments. The ligation reaction was performed at temperature 22°C for 8 h. (**D**) The BamHI restriction map of the synthesized plasmid pHIW. The theoretical restriction fragments of pHIW contains fragments of 11 848, 7001, 5804, 3923, 2172, 1731, 1191, 1029, 928 and 743 bp in size. The 743-bp fragment was run out of the gel and was not shown. (**E**) Heterologous expression of the assembled actinorhodin biosynthetic cluster in *Streptomyces* strain 4F (4F/pHIW), using 4F integrated plasmid pHI (4F/pHI) as a negative control. Strains were cultured on R2YE medium at 30°C for 2 days.

Table 2. Summary of current *in vitro* seamless assembly methods

Mechanism	Homology				Type IIS			
	Overlap		Recombination		Traditional		Modified	
Method	USER[a]	OE-PCR[b]	SLIC[c]	Gibson[d]	Golden gate (38)	Golden Braid[e]	PSA (26)	MASTER[f]
Sequence-independence	+[g]	+	+	+	−	−	+	+
Hierarchical manner	−[h]	−	−	+	+	+	+	+
Complicated sequence	−	−	−	−	+	+	+	+
Multiple fragments	+	+	+	+	+	−	−	+
Material sources	PCR	PCR	PCR	Both	Both	Plasmid	Plasmid	Both
Sizes (39)	Parts-Genes-Pathways	Parts-Genes-Pathways	Genes-Pathways	Genes-Pathways-Genome	Parts-Genes-Pathways	Parts-Genes-Pathways	Genes-Pathways-Genome	Parts-Genes-Pathways

[a]USER is a representative of USER (6), PLICing (40), Ribocloning (41) and so forth.
[b]OE-PCR is a representative of OE-PCR (42), CPEC (43), SHA (44) and so forth.
[c]SLIC is a representative of SLIC (5), In-Fusion (45), Cold Fusion (System Biosciences), Fast Seamless Cloning (DoGene), CloneEZ (GenScript) and so forth.
[d]Gibson is a representative of Gibson Assembly (8) and SLiCE (46).
[e]Golden Braid is a representative of Golden Braid (47), MoClo (48), NOMAD (49) and so forth.
[f]As the efficiency drops remarkably when large DNA fragments are used for ligation, the application of MASTER Ligation in genome assembly remains to be established.
[g]Plus represents techniques applicable in relevant context.
[h]Minus represents techniques unsuitable in relevant context.

for each fragment easy. Therefore, even for the assembly of restriction fragments, MASTER Ligation can be considered as a 'sequence-independent' method.

Although traditional type IIS REs-based methods, such as Golden Gate (38), can be conveniently used to assemble multiple DNA sequences in a seamless manner, it is difficult to find suitable unique REs when assembling long pieces of DNA because of their relative short recognition sequences (4–8 bp). However, with the usage of MspJI, MASTER Ligation and Golden Gate may be combined for a wider application. Recently, an improved method named pairwise selection assembly (PSA) was developed, aiming to overcome the limitation of type IIS RE and to assemble DNA in a sequence-independent manner. PSA uses four steps to specifically methylate the internal restriction sites of type IIS REs, whereas leaving the two terminal sites protected from methylation by the oligo-directed RecA-mediated structure (26). Although Blake and co-workers (50) successfully assembled the entire 91-kb synIXR fragment using this method, DNA fragments to be assembled via the PSA procedure must be first cloned into specifically designed vectors before they can be hierarchically assembled into a larger DNA construct. Except that PSA is a truly sequence-independent method, the MASTER ligation is simpler in design and easier for manipulation.

The 'tailorable ends' feature enables MASTER Ligation to assemble DNA in a hierarchical manner

In addition to its type IIM property, MspJI also belongs to the type IIS REs, which cleaves DNA sequences outside of their recognition sites. In this particular case, MspJI recognizes ^{m}CNNR sites and cleaves DNA at N12/N16 from the 3′-side of the modified cytosine. The base shift between the recognition site and the restriction site of MspJI can be harnessed to remove restriction scars and produce complementary cohesive ends for the following assembly at the same time, i.e. the so-called 'tailorable ends'. To enable the MASTER Ligation to assemble DNA in a hierarchical manner, a new restriction site is introduced to the termini of the sub-clone vector, which can be used to release the intermediate assembly product for the next round of assembly ('Materials and Methods' section and Supplementary Table S1). Therefore, sequencing-verified sub-clones can, therefore, be used for large DNA assembly, and the accuracy of the assembled sequences can be guaranteed.

With the use of methylated adaptors, MASTER Ligation can easily remove the scars of restriction sites, which, however, hindered many other assembly methods, such as USER (6), In-Fusion (51) and SLIC (5), from assembling DNA in a hierarchical manner. Although Gibson assembly can resect the restriction scars and assemble DNA in a hierarchical way, the scars compromise the efficiency. In principle, other type IIS REs can also be used to tailor the restriction scars. Actually, Golden Gate can assemble DNA in a hierarchical way, whereas MspJI provides an extra crucial advantage because of its sequence-independent property.

Recently, two synthetic type IIS REs, CdnDI and CdnDII, were created by fusing a non-cleaving mutant of the homing endonuclease I-SceI with the catalytic domain of type IIS FokI (52). With a long recognition sequence, these enzymes can provide enough specificity for large DNA assembly. However, the base shift between recognition site and restriction site of these enzymes is only two base pairs in length, which is much less than that of MspJI described in this study. Therefore, these two enzymes are unable to remove restriction scars, which are usually 5 bp in length and at least 3 bp for those frequently used 6-bp blunt-cutting type IIP enzymes.

The 'rational ligation' feature enables MASTER Ligation to assemble complicated DNA sequences and/or multiple fragments

The homologous recombination-based systems, including LIC-PCR (53), SLIC (5), Gibson assembly (8) and several commercially available recombination systems and so forth, are widely used because of their convenience and high efficiency. As the systems are based on homologous ends [e.g. >25 bp for Gibson assembly and >15 bp for In-Fusion^{TM} PCR Cloning (Clontech)], these methods may have difficulties in assembling DNA with highly repeated sequences. Besides, the termini of the DNA sequences to be assembled should avoid to have highly stable secondary structures (e.g. hairpins and stem–loops), which unfortunately frequently exist in promoter and terminator regions (16).

MspJI harbours the properties of both type IIM and type IIS REs, producing four-base 5′-overhangs of any sequences. By choosing non-palindromic overhangs, multiple fragments can be assembled in desired order and orientation, i.e. 'rational ligation'. Like other type IIS-based methods, such as Golden Gate (38,54), MoClo (48) and GoldenBraid (47), MASTER Ligation has the advantages in assembling complicated DNAs because of the short sticky ends produced.

FUTURE PERSPECTIVES

Considering the sequence independent and hierarchical ligation characteristics, MASTER Ligation is more suitable for assembly of long complicated DNA constructs. A foreseeable application of MspJI is to assemble large clusters from different vectors. Usually, large polyketide synthases (PKS) biosynthetic clusters [e.g. the rifamycin biosynthetic gene cluster of ∼90 kb in *Amycolatopsis mediterranei* (55)] contain highly repeated regions in the PKS genes. Besides, because of their large sizes, it is difficult to directly clone such clusters into one vector. Usually, a fosmid library covering the whole-genomic sequences will be constructed, and the large cluster will be split into several overlapping fragments and scattered in several fosmids. As the gap-repair technology is widely used to precisely trim the borders of each fragment and add specific restriction sites to the ends (56), the combination of MASTER Ligation and the gap-repair technology may shed light on the seamless assembly of extremely large clusters with complicated structures.

During the construction of pHIW, we found that when multiple large DNA fragments were used for ligation, the efficiency and accuracy dropped remarkably (Table 1), which suggests that the present system, including the conditions for both the MspJI digestion and DNA ligation, still needs further optimization. Besides, MspJI requires a longer incubation time than usual REs, i.e. 3 h at the presence of activators indicated in this study and in a previous report (30). Although MspJI is stable within at least 16 h and the extended digestion time is not vital to the success of the experiments, to speed-up the assembly, the enzymatic activities of MspJI as well as the digestion conditions may need further improvement. However, with more and more MspJI-family restriction enzymes discovered [e.g. SgeI (Thermo, USA) and FspEI (30)], which have the same restriction properties and the same potential usage as MspJI in seamless DNA assembly, new options can be tested to find an enzyme with higher digestion efficiency.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–2 and Supplementary References [57,58].

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Lu,T.K., Khalil,A.S. and Collins,J.J. (2009) Next-generation synthetic gene networks. *Nat. Biotechnol.*, **27**, 1139–1150.
2. Endy,D. (2005) Foundations for engineering biology. *Nature*, **438**, 449–453.
3. Carlson,R. (2009) The changing economics of DNA synthesis. *Nat. Biotechnol.*, **27**, 1091–1094.
4. Stemmer,W.P., Crameri,A., Ha,K.D., Brennan,T.M. and Heyneker,H.L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*, **164**, 49–53.
5. Li,M.Z. and Elledge,S.J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods*, **4**, 251–256.
6. Geu-Flores,F., Nour-Eldin,H.H., Nielsen,M.T. and Halkier,B.A. (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.*, **35**, e55.
7. Berrow,N.S., Alderton,D., Sainsbury,S., Nettleship,J., Assenberg,R., Rahman,N., Stuart,D.I. and Owens,R.J. (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.*, **35**, e45.
8. Gibson,D.G., Young,L., Chuang,R.Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
9. Kong,Y.P., Yang,T.Z. and Geller,A.I. (1999) An efficient in vivo recombination cloning procedure for modifying and combining HSV-1 cosmids. *J. Virol. Methods*, **80**, 129–136.
10. Shao,Z.Y., Zhao,H. and Zhao,H.M. (2009) DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.*, **37**, e16.
11. Gibson,D.G., Benders,G.A., Axelrod,K.C., Zaveri,J., Algire,M.A., Moodie,M., Montague,M.G., Venter,J.C., Smith,H.O. and Hutchison,C.A. (2008) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl Acad. Sci. USA*, **105**, 20404–20409.
12. Itaya,M., Fujita,K., Kuroki,A. and Tsuge,K. (2008) Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat. Methods*, **5**, 41–43.
13. Wingler,L.M. and Cornish,V.W. (2011) Reiterative Recombination for the in vivo assembly of libraries of multigene pathways. *Proc. Natl Acad. Sci. USA*, **108**, 15135–15140.
14. Gibson,D.G., Benders,G.A., Andrews-Pfannkoch,C., Denisova,E.A., Baden-Tillson,H., Zaveri,J., Stockwell,T.B., Brownley,A., Thomas,D.W., Algire,M.A. *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, **319**, 1215–1220.
15. 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. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52–56.
16. Hillson,N.J., Rosengarten,R.D. and Keasling,J.D. (2012) j5 DNA assembly design automation software. *ACS Synth. Biol.*, **1**, 14–21.
17. Cermak,T., Doyle,E.L., Christian,M., Wang,L., Zhang,Y., Schmidt,C., Baller,J.A., Somia,N.V., Bogdanove,A.J. and Voytas,D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.*, **39**, e82.
18. Zhang,F., Cong,L., Lodato,S., Kosuri,S., Church,G.M. and Arlotta,P. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.*, **29**, 149–153.
19. Ridley,C.P., Lee,H.Y. and Khosla,C. (2008) Evolution of polyketide synthases in bacteria. *Proc. Natl Acad. Sci. USA*, **105**, 4595–4600.
20. Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S., Romero,D.A. and Horvath,P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
21. Scior,A., Preissler,S., Koch,M. and Deuerling,E. (2011) Directed PCR-free engineering of highly repetitive DNA sequences. *BMC Biotechnol.*, **11**, 87.
22. Mandecki,W. and Bolling,T.J. (1988) Foki method of gene synthesis. *Gene*, **68**, 101–107.
23. Yount,B., Denison,M.R., Weiss,S.R. and Baric,R.S. (2002) Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J. Virol.*, **76**, 11065–11078.
24. Kodumal,S.J., Patel,K.G., Reid,R., Menzella,H.G., Welch,M. and Santi,D.V. (2004) Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl Acad. Sci. USA*, **101**, 15573–15578.
25. Menzella,H.G., Reisinger,S.J., Welch,M., Kealey,J.T., Kennedy,J., Reid,R., Tran,C.Q. and Santi,D.V. (2006) Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *J. Ind. Microbiol. Biotechnol.*, **33**, 22–28.
26. Blake,W.J., Chapman,B.A., Zindal,A., Lee,M.E., Lippow,S.M. and Baynes,B.M. (2010) Pairwise selection assembly for sequence-independent construction of long-length DNA. *Nucleic Acids Res.*, **38**, 2594–2602.
27. Waite-Rees,P.A., Keating,C.J., Moran,L.S., Slatko,B.E., Hornstra,L.J. and Benner,J.S. (1991) Characterization and expression of the *Escherichia coli* Mrr restriction system. *J. Bacteriol.*, **173**, 5207–5219.

28. Zheng,Y., Cohen-Karni,D., Xu,D., Chin,H.G., Wilson,G., Pradhan,S. and Roberts,R.J. (2010) A unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Res.*, **38**, 5527–5534.

29. Horton,J.R., Mabuchi,M.Y., Cohen-Karni,D., Zhang,X., Griggs,R.M., Samaranayake,M., Roberts,R.J., Zheng,Y. and Cheng,X. (2012) Structure and cleavage activity of the tetrameric MspJI DNA modification-dependent restriction endonuclease. *Nucleic Acids Res.*, **40**, 9763–9773.

30. Cohen-Karni,D., Xu,D., Apone,L., Fomenkov,A., Sun,Z., Davis,P.J., Kinney,S.R., Yamada-Mabuchi,M., Xu,S.Y., Davis,T. *et al.* (2011) The MspJI family of modification-dependent restriction endonucleases for epigenetic studies. *Proc. Natl Acad. Sci. USA*, **108**, 11040–11045.

31. Kieser,T., Bibb,M.J., Buttner,M.J., Chater,K.F. and Hopwood,D.A. (2000) *Practical Streptomyces Genetics*. The John Innes Foundation, Norwich, UK.

32. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

33. Chen,W.H. and Qin,Z.J. (2011) Development of a gene cloning system in a fast-growing and moderately thermophilic *Streptomyces* species and heterologous expression of *Streptomyces* antibiotic biosynthetic gene clusters. *BMC Microbiol.*, **11**, 243.

34. Evans,G.A., Lewis,K. and Rothenberg,B.E. (1989) High-efficiency vectors for cosmid microcloning and genomic analysis. *Gene*, **79**, 9–20.

35. Hai-yang,X., Jun,H., Min-jie,H., Mei-juan,S., Peng-fei,X., Liang,Z., Hai-bin,W. and Zhong-jun,Q. (2009) Construction of an ordered cosmid library of *S. avermitilis* for genetic modification of the industrial strains. *Chin. J. Antibiot.*, **7**, 403–405.

36. Bierman,M., Logan,R., O'Brien,K., Seno,E.T., Rao,R.N. and Schoner,B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*, **116**, 43–49.

37. Kunkel,T.A. (2004) DNA replication fidelity. *J. Biol. Chem.*, **279**, 16895–16898.

38. Engler,C., Kandzia,R. and Marillonnet,S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647.

39. Ellis,T., Adie,T. and Baldwin,G.S. (2011) DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol. (Camb.)*, **3**, 109–118.

40. Blanusa,M., Schenk,A., Sadeghi,H., Marienhagen,J. and Schwaneberg,U. (2010) Phosphorothioate-based ligase-independent gene cloning (PLICing): an enzyme-free and sequence-independent cloning method. *Anal. Biochem.*, **406**, 141–146.

41. Donahue,W.F., Turczyk,B.M. and Jarrell,K.A. (2002) Rapid gene cloning using terminator primers and modular vectors. *Nucleic Acids Res.*, **30**, e95.

42. Horton,R.M., Hunt,H.D., Ho,S.N., Pullen,J.K. and Pease,L.R. (1989) Engineering hybrid genes without the use of restriction enzymes-gene-splicing by overlap extension. *Gene*, **77**, 61–68.

43. Quan,J.Y. and Tian,J.D. (2009) Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One*, **4**, e6441.

44. Jiang,X., Yang,J., Zhang,H., Zou,H., Wang,C. and Xian,M. (2012) In vitro assembly of multiple DNA fragments using successive hybridization. *PLoS One*, **7**, e30267.

45. Sleight,S.C., Bartley,B.A., Lieviant,J.A. and Sauro,H.M. (2010) In-Fusion BioBrick assembly and re-engineering. *Nucleic Acids Res.*, **38**, 2624–2636.

46. Zhang,Y., Werling,U. and Edelmann,W. (2012) SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.*, **40**, e55.

47. Sarrion-Perdigones,A., Falconi,E.E., Zandalinas,S.I., Juarez,P., Fernandez-del-Carmen,A., Granell,A. and Orzaez,D. (2011) Goldenbraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One*, **6**, e21622.

48. Weber,E., Engler,C., Gruetzner,R., Werner,S. and Marillonnet,S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, **6**, e16765.

49. Rebatchouk,D., Daraselia,N. and Narita,J.O. (1996) NOMAD: a versatile strategy for in vitro DNA manipulation applied to promoter analysis and vector design. *Proc. Natl Acad. Sci. USA*, **93**, 10891–10896.

50. Dymond,J.S., Richardson,S.M., Coombes,C.E., Babatz,T., Muller,H., Annaluru,N., Blake,W.J., Schwerzmann,J.W., Dai,J.B., Lindstrom,D.L. *et al.* (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature*, **477**, 471–476.

51. Zhu,B.G., Cai,G.F., Hall,E.O. and Freeman,G.J. (2007) In-Fusion (TM) assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *Biotechniques*, **43**, 356–359.

52. Lippow,S.M., Aha,P.M., Parker,M.H., Blake,W.J., Baynes,B.M. and Lipovsek,D. (2009) Creation of a type IIS restriction endonuclease with a long recognition sequence. *Nucleic Acids Res.*, **37**, 3061–3073.

53. Aslanidis,C. and Dejong,P.J. (1990) Ligation-independent cloning of PCR products (Lic-PCR). *Nucleic Acids Res.*, **18**, 6069–6074.

54. Engler,C., Gruetzner,R., Kandzia,R. and Marillonnet,S. (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One*, **4**, e5553.

55. Zhao,W., Zhong,Y., Yuan,H., Wang,J., Zheng,H.J., Wang,Y., Cen,X.F., Xu,F., Bai,J., Han,X.B. *et al.* (2010) Complete genome sequence of the rifamycin SV-producing *Amycolatopsis mediterranei* U32 revealed its genetic characteristics in phylogeny and metabolism. *Cell Res.*, **20**, 1096–1108.

56. Zhang,Y.M., Muyrers,J.P., Testa,G. and Stewart,A.F. (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18**, 1314–1317.

57. Combes,P., Till,R., Bee,S. and Smith,M.C. (2002) The *Streptomyces* genome contains multiple pseudo-attB sites for the (phi)C31-encoded site-specific recombination system. *J. Bacteriol.*, **184**, 5746–5752.

58. Hall,T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, **41**, 95–98.