Biotechnology Advances 37 (2019) 730–745

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Synthetic biology approaches for chromosomal integration of genes and pathways in industrial microbial systems

Lei Li[a], Xiaocao Liu[a,b], Keke Wei[c], Yinhua Lu[d,*], Weihong Jiang[a,*]

[a] Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China
[b] School of Life Sciences, Henan University, Kaifeng 475004, China
[c] Department of Biochemistry, Shanghai Institute of Pharmaceutical Industry, Shanghai 201210, China
[d] College of Life Sciences, Shanghai Normal University, 200232, China

ARTICLE INFO

Keywords:
Industrial biotechnology
Synthetic biology
Chromosomal integration
CRISPR-Cas
Site-specific recombination
Homology-mediated end joining

ABSTRACT

Industrial biotechnology is reliant on native pathway engineering or foreign pathway introduction for efficient biosynthesis of target products. Chromosomal integration, with intrinsic genetic stability, is an indispensable step for reliable expression of homologous or heterologous genes and pathways in large-scale and long-term fermentation. With advances in synthetic biology and CRISPR-based genome editing approaches, a wide variety of novel enabling technologies have been developed for single-step, markerless, multi-locus genomic integration of large biochemical pathways, which significantly facilitate microbial overproduction of chemicals, pharmaceuticals and other value-added biomolecules. Notably, the newly discovered homology-mediated end joining strategy could be widely applicable for high-efficiency genomic integration in a number of homologous recombination-deficient microbes. In this review, we explore the fundamental principles and characteristics of genomic integration, and highlight the development and applications of targeted integration approaches in the three representative industrial microbial systems, including *Escherichia coli*, actinomycetes and yeasts.

1. Introduction

High-yield production of industrially important chemicals and pharmaceuticals is reliant on native pathway engineering or foreign pathway introduction (Lee and Kim, 2015; Nielsen and Keasling, 2016). Microbial production of manifold compounds is a promising and renewable strategy for replacing traditional petrochemical production processes (Cho et al., 2015), achieving rapid industrial-scale production of phytochemicals (Li et al., 2018a) as well as advancing genome mining of novel bioactive natural products (Rutledge and Challis, 2015). The development of synthetic biology (Cameron et al., 2014), combined with continued progress in metabolic engineering and system biology (Lee and Kim, 2015; Nielsen and Keasling, 2016), has led to increased feasibility to produce heterologous compounds not only nature's chemical repertoire in industrial microbial systems (Cho et al., 2015; Kotopka et al., 2018; Zhang et al., 2016b).

Generally, two distinct strategies, namely, plasmid-based systems and chromosomal integration, are employed for expression of metabolic genes/pathways. The use of engineered plasmids is a ubiquitous and indispensable method for expression and analysis of functional genes/pathways due to the ease of manipulation and regulated expression with wide dynamic ranges. However, there are inherent and prominent problems associated with artificial plasmid-based systems, including genetic instability, limited cloning sizes and the need for selection pressure for plasmid maintenance (Carrier et al., 1998; Da Silva and Srikrishnan, 2012; Zhang et al., 1996). In particular, plasmid genetic instability, including segregation instability, structural instability and allele segregation, leads to cell-to-cell variations in protein concentration, thus generating high- and low-performance variants in a culture (Tyo et al., 2009). These drawbacks significantly affect the endeavors to enhance the titer, rate and yield (TRY) of target products, particularly during large-scale and long-term industrial fermentation. Therefore, integration of metabolic genes and pathways into host chromosomes is a preferable approach over plasmid-based expression. On one hand, chromosome-based genetic constructs can be stably maintained in the absence of antibiotics or other selective agents and exhibit consistent gene expression strengths at the single-cell level due to the absence of copy number and segregation variations (Xiao et al., 2016). On the other hand, the metabolic burden of multi-copy plasmid maintenance is also absent, and thus the use of chromosomally integrated genes/

* Corresponding authors.
E-mail addresses: yhlu@shnu.edu.cn (Y. Lu), whjiang@sibs.ac.cn (W. Jiang).

https://doi.org/10.1016/j.biotechadv.2019.04.002
Received 11 January 2019; Received in revised form 1 April 2019; Accepted 1 April 2019
Available online 03 April 2019
0734-9750/

pathways will likely lead to microbial overproduction of target products (Santos et al., 2013; Tyo et al., 2009).

In the last ten years, synthetic biology has provided a large variety of methodologies for low-cost gene synthesis and large pathway assembly (Juhas and Ajioka, 2017; Kosuri and Church, 2014; Ma et al., 2012). Meanwhile, adaptation of the CRISPR-Cas systems for genome editing applications has revolutionized modern industrial biotechnology (Choi and Lee, 2016; Donohoue et al., 2018; Jakociunas et al., 2016). Via rapid and programmed introduction of double-strand breaks (DSBs), which greatly increase the rate of homologous recombination (HR) (Barrangou and Doudna, 2016; Hille et al., 2018), CRISPR-Cas-mediated genetic engineering efficiently eliminates the need for plasmids to introduce foreign DNA into cells and significantly advances multi-copy integration of target genes/pathways (Li et al., 2017a; Liu et al., 2017; Ou et al., 2018).

In this review, we present a comprehensive overview of the basic principles and common strategies for genomic integration of genes and pathways in three important prokaryotic and eukaryotic microbial systems. Several representative applications are also discussed to demonstrate the advantages of gene/pathway integration.

2. Mechanisms and characteristics of chromosomal integration of genes and pathways

2.1. Efficient transformation: the premise of DNA integration

Transformation is a kind of genetic events that a cell directly uptakes and incorporates exogenous DNA materials (e.g., plasmids and linearized DNA molecules) from its surroundings through cell membranes, which holds great importance in synthetic biology and metabolic engineering. For transformation to take place, the recipient bacteria or fungi should be in a state of competence, which might occur in nature or be artificially induced. However, a large number of microbial species do not naturally possess the ability to take up DNA (Domingues et al., 2012; Johnston et al., 2014). Therefore, a wide variety of transformation approaches, including chemical transformation, electroporation, conjugation, spheroplast transformation and LiAc/ssDNA/PEG transformation, have been developed, and the efficiency of each method has been continually improved over the past three decades. Interested readers could refer to the detailed reviews of DNA transformation methods in different bacteria and fungi (Gietz and Woods, 2001; Kawai et al., 2010; Kieser et al., 2000; Yoshida and Sato, 2009). Notably, restriction-modification systems of recipient cells, which can degrade foreign DNA, are the primary genetic barriers to transformation and should be carefully considered (Oliveira et al., 2014; Vasu and Nagaraja, 2013). For instance, proper methylation of exogenous DNA prior to genetic transformation may efficiently overcome the host restriction barrier and thus improve the probability of successful transformation in a number of solvent-producing clostridia (Mermelstein and Papoutsakis, 1993; Yang et al., 2016; Zhang et al., 2012).

2.2. Principles of genomic integration of foreign DNA constructs

2.2.1. Homologous recombination (HR)-mediated DNA integration

HR is an intrinsic biochemical pathway that enables DNA exchange between two regions of identical sequences and contributes to important cellular processes, including DSB repair and horizontal gene transfer (Cromie et al., 2001; Sung and Klein, 2006). Generally, DNA duplexes with 3′-ended single stranded tails are firstly yielded through end resection. Then, DSB repair happens at Holiday junctions after gap-repair DNA synthesis and ligation. HR occurs naturally across all the three domains of life as well as virus, and is very powerful tool in genetic engineering. Genomic integration of foreign DNA (several kb) was first described in *Saccharomyces cerevisiae* by homology-directed repair (HDR) more than three decades ago (Fig. 1A) (Orr-Weaver et al., 1981; Scherer and Davis, 1979). However, following the introduction of homology templates, gene incorporation at desired loci occurs at very low rates of 1 in every 10^3 to 10^9 strains (Smithies et al., 1985; Thomas et al., 1986). Additionally, the efficiency of HDR-mediated integration often decreases rapidly when large metabolic pathways are introduced or multiple DNA constructs are simultaneously inserted (Lee and Da Silva, 1997; Yamada et al., 2010). To promote integration of foreign genes/pathways, researchers have adopted several strategies:

i. Introduction of selection markers. For efficient chromosomal integration, selection markers, including antibiotic resistance and auxotrophic markers, have been widely employed in a wide variety of bacteria and fungi. In particular, increasing antibiotic concentrations have also been used to select for high-productivity engineered strains via repeated transformation (Shi et al., 2014). Therefore, it is very important to develop accessible selection markers for gene integration, particularly in genetically intractable microorganisms. However, the use of selection markers could hamper multi-round integration due to the limited number of dominant markers. Therefore, markerless integration has long been the main direction for future research, which will be discussed in depth in the following section.
ii. Optimization of homology arm (HA) lengths. HA lengths positively influence the recombination efficiency. For example, 50-bp HAs are sufficient for gene integration in *S. cerevisiae*, but > 500-bp HAs might be required in non-conventional yeasts, such as *Kluyveromyces lactis* and *Scheffersomyces stipitis* (Blazeck et al., 2014; Horwitz et al., 2015). Therefore, HA lengths should be carefully considered in different experimental designs and industrial microorganisms.
iii. Sequence-specific introduction of DSBs. In general, two types of endonucleases, namely, meganucleases (or homing endonucleases) and programmable nucleases, are employed to introduce DSBs at target loci of interest (Chevalier and Stoddard, 2001; Kim and Kim, 2014). Traditionally, the recognition site of a meganuclease (e.g., I-*Sce*I) requires prior introduction into the host chromosome. To address this limitation, a number of programmable nucleases have been discovered or developed, such as zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and RNA-guided Cas endonucleases (e.g., the Cas9 protein from *Streptococcus pyogenes* and the Cpf1 or Cas12a protein from *Francisella novicida*) (Gaj et al., 2013). A systematic comparison of these three endonuclease systems was recently reviewed and will not be discussed in detail here (Gaj et al., 2013; Kim and Kim, 2014). Notably, except for the modest limitation associated with protospacer-adjacent motif (PAM) requirements, researchers need to only design a single guide RNA by Watson-Crick base-pairing for cleavage of each genomic locus of interest with Cas endonucleases. In addition, the CRISPR/Cas system exerts a negative screening pressure for efficient genome editing, allowing simple markerless chromosomal integration without the aid of selection markers. In recent years, CRISPR-Cas9- or CRISPR-Cpf1-assisted HDR systems have been broadly used for genomic integration of genes and pathways in a wide variety of industrial microbial systems, including *Escherichia coli*, actinomycetes, conventional or non-conventional yeasts and *Aspergilli* (Donohoue et al., 2018; Jakociunas et al., 2016; Shi et al., 2017).
iv. Introduction of foreign recombination systems. This technology, termed recombineering, only requires short homologies of 35–50 bp and enables the incorporation of DNA constructs in dsDNA or ssDNA donors in cells (Court et al., 2002). Two different recombination systems, including lambda (λ) phage recombinases (Gam, Exo, and Beta) and Rac phage proteins (RedE and RecT), are often heterologously reconstituted for genomic integration in some important industrial microbes, such as *E. coli* (Yu et al., 2000), *Pseudomonas putida* (Choi et al., 2018) and *Corynebacterium glutamicum* (Cho et al., 2017). Integration of up to 50-kb DNA fragments into *E. coli* has been reported using the λ-Red system with the aid of

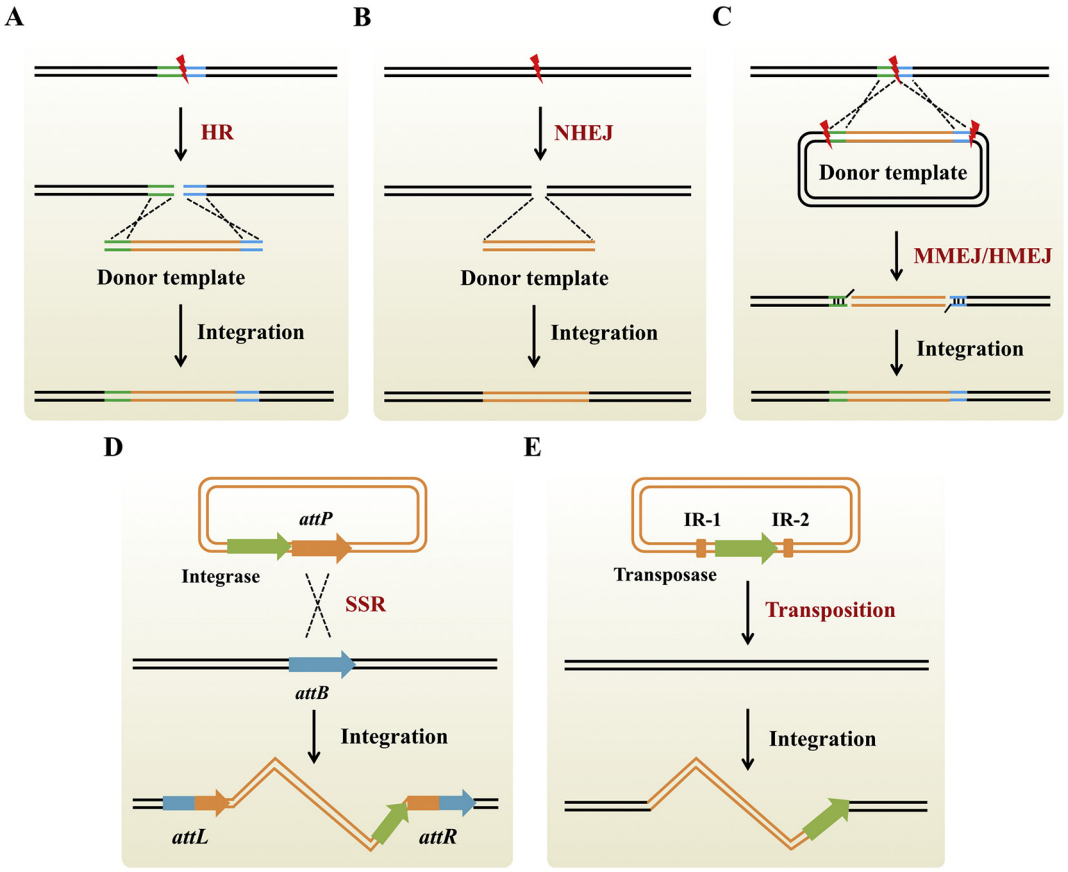

Fig. 1. Principles of genomic integration of foreign DNA constructs. (A) Homologous recombination (HR) -mediated DNA integration. (B) Non-homologous end joining (NHEJ)-mediated random insertion of foreign DNA fragments. (C) (Micro)homology-mediated end joining (MMEJ/HMEJ)-assisted gene integration. (D) Site-specific recombination (SSR). (E) Transposition-mediated chromosomal random integration of DNA constructs. IR: inverted repeated sequences.

selection markers (Juhas and Ajioka, 2016). In recent years, the meganuclease (e.g., I-*Sce*I) or CRISPR-Cas9 system has been coupled with recombineering to increase the scarless integration efficiency of large DNA constructs (Ou et al., 2018).

2.2.2. *Non-homologous end joining (NHEJ)-mediated random insertion of foreign DNA constructs*

In addition to HR, NHEJ is often used to repair DSBs without the need for homologous templates in all eukaryotes and many prokaryotes, which results in stochastic deletion and insertion of nucleotides and even translocation at the desired loci (Bowater and Doherty, 2006; Pitcher et al., 2007; Shuman and Glickman, 2007). A typical NHEJ system consists of a specialized DNA ligase (LigD in bacteria and Lig4 in eukaryotes) and a single-stranded DNA-end-binding protein Ku (Shuman and Glickman, 2007). NHEJ is typically more efficient than HR in some important industrial microbes, including *Bacillus*, filamentous fungi and non-conventional yeasts (Ayora et al., 2011; Jiang et al., 2013; Wagner and Alper, 2016). Therefore, NHEJ system is also applied to the multi-locus integration of foreign genes/pathways in many HR-deficient bacteria or fungi (Fig. 1B) (Cui et al., 2019; Shi et al., 2017; Yao et al., 2018). Due to random genomic insertion of DNA segments, NHEJ-based targeted integration can easily enable multivariate modular library construction for enhanced biochemical biosynthesis of metabolic pathways (Cui et al., 2019). Noteworthy, suppression of the dominant NHEJ system, such as deletion of the dedicated *ku* or ligase gene, is used as a common strategy to enhance HR efficiency and accuracy (Qin et al., 2017; Verbeke et al., 2013).

2.2.3. *(Micro)homology-mediated end joining-assisted gene integration*

Microhomology-mediated end joining (MMEJ) is an error-prone repair of DSBs with the aid of microhomologous sequences (5–25 bp) internal to the broken ends, resulting in random DNA mutations (Sfeir and Symington, 2015). However, in the presence of short homologous sequences (10–40 bp) flanking the genomic target locus, MMEJ could be used to mediate precise, in-frame integration of exogenous DNA segments (Fig. 1C) (Nakade et al., 2014). MMEJ-mediated integration has been widely used in mammalian cells, providing a useful complementary strategy to the HR-based strategy (Nakade et al., 2014). However, among industrial microorganisms, the only reported MMEJ-based application is the precise insertion of an exogenous GFP gene in *Aspergillus fumigatus*, in which integration of target genes is very challenging due to the inefficient HR ability (Zhang et al., 2016a). Recently, a homology-mediated end joining (HMEJ)-based strategy with long HAs (e.g., ~800 bp) was reported to achieve transgene integration in mammalian cells with an efficiency much greater than that of HR- or MMEJ-based methods (Yao et al., 2017). Similarly, HMEJ-based scarless integration was twice as efficient as HR-based integration in *Yarrowia lipolytica* (Fig. 1C) (Gao et al., 2018). Although the precise DSB repair mechanism of HMEJ requires further investigation, this newly discovered pathway harbors longer and more stable HAs, and can mediate genomic integration with higher accuracy and efficiency than MMEJ. Therefore, combined with the CRISPR-Cas9 system, HMEJ-mediated gene/pathway integration is expected to be widely used in many HR-deficient industrial microorganisms in the future.

2.2.4. *Site-specific integration*

Site-specific recombination (SSR), a process involving DNA breakage and reunion, results in integration, excision or inversion of defined DNA segments (Grindley et al., 2006). There are two evolutionarily distinct site-specific recombinases with different

recombination mechanisms, including tyrosine recombinases (e.g., Cre recombinase) and serine integrases (e.g., ΦC31 integrase) (Fogg et al., 2014). Generally, foreign DNA constructs can be efficiently integrated into the bacterial chromosomal attachment site (*attB*). Interestingly, unlike tyrosine recombinase, which requires an accessory host protein (IHF) for DNA integration, serine integrase itself enables the mediation of *attP* × *attB* unidirectional recombination (Fig. 1D) (Stark, 2017). Additionally, serine recombinases are capable of promoting efficient genomic integration of large DNA constructs (> 100 kb), such as natural product biosynthetic gene clusters (BGCs) from actinomycetes (Myronovskyi and Luzhetskyy, 2013). Tyrosine recombinases have been used for chromosomal insertion of genes with multiple copies in *E. coli* and *Halomonas* sp. (Gu et al., 2015; Yin et al., 2015). Moreover, site-specific serine integration systems have been used mainly to produce stable recombinants of actinomycetes (Baltz, 2012). Notably, a selection marker is usually required to identify the occurrence of site-specific integration after transformation, which possibly hampers multi-round integration of foreign genes/pathways. In addition, the non-programmable feature of DNA target specificity further hampers widespread application of site-specific integrases. To increase the number of sites amenable to targeted integration, the catalytic domains of some dedicated serine recombinases have been fused to widely used programmable DNA-binding proteins, such as zinc finger, transcription activated-like effector (TALE) and nuclease-deactivated Cas9 (dCas9) (Chaikind et al., 2016; Mercer et al., 2012). Although the fusion of dCas9-serine recombinase has already been employed for scarless, discrete integration of the *eGFP* reporter gene with three copies in mammalian cells (Chaikind et al., 2016), application of this construct in industrial microorganisms requires further investigation.

2.2.5. Transposition-mediated random insertion of large DNA constructs

Transposons, which are widely distributed in prokaryotic and eukaryotic microbes, were originally described as “jumping genes” and consist of a transposase gene flanked by inverted terminal repeats (Hayes, 2003). Generally, transposases randomly bind to any target site in the genome and promote the insertion of transposable elements (Fig. 1E). Therefore, compared to HDR- or SSR-mediated integration, transposition-assisted DNA insertion exhibits random and unstable features. There are at least two classes of transposons, including retrotransposons and DNA transposons. Currently, the mariner DNA transposon family has been extensively used due to the following reasons: these transposons have a broad host range, do not require any host factors and have no sequence preference for insertion (Aziz et al., 2010; Lampe et al., 1996). Transposition is used not only to non-specifically knockout single genes for functional genomic studies but also to mediate the integration of foreign DNA constructs. So far, this strategy has been widely used for heterologous expression of large secondary metabolite BGCs in *Myxococcus xanthus*, *P. putida* and *Burkholderiales* (Bian et al., 2017; Fu et al., 2008). Intriguingly, the integration efficiency achieved by transposition is sometimes higher than that of HR-mediated integration in *M. xanthus* (Fu et al., 2008). Here, a brief description related to mechanisms of chromosomal integration of foreign DNA constructs was summaried in Table 1.

2.3. Chromosome position effects

Chromosome position effects, i.e., the influence of gene location on expression, are fundamental features of bacterial or fungal genomes and are involved in important physiological processes, such as sporulation, competence and biofilm formation (Narula et al., 2015; Slager and Veening, 2016). The main factors that affect gene expression in different chromosomal loci are (i) level of DNA compaction; (ii) distance to DNA replication initiation site; (iii) availability of regulatory factors with effects on gene expression (Bryant et al., 2014; Chen and Zhang, 2016). By introducing foreign reporter genes into different loci of chromosomes, distinct position effects have been demonstrated in a

Table 1
| A brief summary related to principles of chromosomal integration of foreign DNA constructs.

Principles	Abbreviations	Definitions	Distributions	Applications
Homologous recombination	HR	An intrinsic biochemical pathway enabling DNA exchange between two identical sequences	All the three domains of life as well as virus	Integration of foreign DNA constructs at desired loci with the aid of homology templates
Non-homologous end joining	NHEJ	An error-prone DSB repair pathway without the need for homologous templates	All eukaryotes and many prokaryotes	Random chromosomal insertion of foreign DNA fragments
Microhomology-mediated end joining	MMEJ	An error-prone DSB repair process with the aid of microhomologous sequences (5–25 bp)	Many eukaryotes, such as human cell and fungi	Precise integration of foreign DNA constructs at desired loci
Homology-mediated end joining	HMEJ	A precise DSB repair process with the aid of long homology templates (such as 800–1000 bp)	Many eukaryotes, such as human cell and yeast	Integration of foreign DNA segments at desired loci
Site-specific recombination	SSR	An intrinsic process involving DNA breakage and reunion mediated by recombinase	Many prokaryotes, such as *E. coli* and actinobacteria	Integration of foreign DNA fragments at the bacterial chromosomal attachment sites
Transposition	Transposition	An intrinsic process involving insertion of transposable elements mediated by transposase	Almost all the three domains of life	Random genomic integration of foreign DNA fragments

number of industrial microbes, including *E. coli* (Bryant et al., 2014), *Bacillus subtilis* (Sauer et al., 2016), *Lactobacillus lactis* (Thompson and Gasson, 2001), *Streptomyces albus* (Bilyk et al., 2017) and *S. cerevisiae* (Chen and Zhang, 2016; Wu et al., 2017). For instance, GFP expression levels showed ~300-fold variations depending on its position on the chromosome, and GFP expression was sometimes completely silenced at several specific locations in *E. coli* (Bryant et al., 2014). Additionally, Luzhetskyy and colleagues demonstrated that the production titers of the antibiotic aranciamycin varied up to eight-fold when the associated BGC was introduced into different genomic loci (Bilyk et al., 2017). From a practical perspective, systematic examination of position effects will be helpful in guiding the design of integration sites for metabolic genes or pathways to maximize expression for enhanced biosynthesis.

2.4. Genetic stability of chromosomal integration

Compared with the plasmid-based systems, chromosomal integration and expression of foreign genes/pathways are usually highly stable due to the absence of copy number and segregation variations. However, chromosome-based expression can be highly unstable in two cases, namely, transposition-mediated integration and gene duplication-amplification (GDA). Although transposition can be an attractive strategy for chromosomal integration of target genes/pathways, transposons, as mobile genetic elements, possess intrinsic genetic instability (Hayes, 2003). For instance, due to unstable insertion of artificial *attB* sites via transposon mutagenesis, an engineered strain of *S. albus* carrying three copies of heterologous pamamycin BGCs exhibited large fluctuations in antibiotic production (Manderscheid et al., 2016). Moreover, chromosome-based expression is usually weaker than plasmid-based expression. Therefore, increasing the copy numbers of integrated genes/pathways, including via discrete, multi-locus insertion and GDA, has been an efficient strategy for optimization of the biosynthesis of target products. The GDA phenomenon in different microbial species has been widely described for strain adaptation and evolution (Andersson and Hughes, 2009; Reams and Roth, 2015). A variety of approaches for GDA have been developed for genetic engineering, such as RecA-mediated chemically inducible chromosomal evolution (CIChE) (Tyo et al., 2009), ZouA-RsA/B-based tandem amplification (Murakami et al., 2011a) and tyrosine recombinase-mediated chromosomal integration of genes with multiple copies (CIGMC) (Gu et al., 2015; Yin et al., 2015). However, the tandemly amplified genes/pathways are gradually lost in the absence of antibiotic selection due to the occurrence of high-frequency HR events (Zhou et al., 2014). Therefore, discrete, multi-copy integration of genes/pathways is a preferable choice for strain improvement in large-scale biotechnological application, such as multiplex site-specific genome engineering (MSGE) in actinomycetes (Li et al., 2017b) and the Di-CRISPR (delta integration CRISPR-Cas) platform in *S. cerevisiae* (Shi et al., 2016).

2.5. Detection of gene copy numbers

Some modern methods are capable of detecting copy numbers of homologous or heterologous genes/pathways integrated into the host chromosome, including real-time quantitative PCR (qPCR) (Isaksson et al., 2007), DNA microarrays (Gresham et al., 2008) and next-generation sequencing (Shendure et al., 2017). Generally, discrete integration of foreign genes/pathways enables the creation of a unique join point in the genome, which can be easily detected by routine PCR amplification and Sanger sequencing. In contrast, qPCR and Southern hybridization are better choices for rapid determination of gene dosage with high sensitivity in tandem amplification of target genes/pathways (Murakami et al., 2011a). Additionally, two newly developed methods, namely, microarrays and whole-genome sequencing, can be used for rapidly simultaneous identification of a large number of amplified genes in a high-throughput manner. Interested readers can refer to the detailed reviews of the methods used to assess and study DNA duplication and amplification (Andersson and Hughes, 2009; Elliott et al., 2013).

3. Major approaches and applications of genomic integration of genes and pathways in *E. coli*

E. coli is one of the best characterized microorganisms on earth and the most widely used industrial platform for the production of fuels, chemicals and pharmaceuticals (Pontrelli et al., 2018). In the past two decades, a large number of genetic manipulation tools, including λ-Red, tyrosine recombinases, multiplex automated genome engineering (MAGE), the CRISPR-Cas system and sRNA-mediated expression control, have been developed to efficiently modify the *E. coli* genome, such as deletion, replacement, point mutation and integration (Pontrelli et al., 2018). In particular, chromosomal expression of genes/pathways is favored in industry for efficient and stable phenotype construction (Ou et al., 2018). Here, we describe a panel of advanced genomic integration strategies and related applications for optimizing the expression of heterologous chemicals and drugs in this important platform strain.

3.1. Recombineering-mediated gene/pathway integration

3.1.1. λ-Red recombineering-mediated integration of large DNA constructs

Genetic engineering by phage recombination systems has been named “recombineering” and only requires very short homologous sequences (< 50 bp) (Court et al., 2002; Murphy, 1998). The λ-Red recombination system, consisting of three genetic components (Exo, Bet, and Gam), has been proven to promote HR between the host chromosome and foreign PCR fragments (Court et al., 2002). Gam inhibits the activity of the host RecBCD exonuclease, thus preventing degradation of foreign linear DNA fragments. Exo, a 5′-to-3′ dsDNA-dependent exonuclease, can degrade dsDNA to reveal a specific ssDNA. Finally, Bet binds to ssDNA to anneal complementary DNA strands (Court et al., 2002). A variety of integration strategies with the λ-Red recombination system have been developed, including I-*Sce*I endonuclease-mediated recombination (Yang et al., 2014b), knock-in/knock-out (KIKO) vector-assisted integration (Sabri et al., 2013) and pSB1K3(FRTK) vector-mediated insertion (Juhas et al., 2014). The *lacZ* gene and the flagellar genes *filT/filK* are usually chosen as the integration sites in *E. coli* (Juhas and Ajioka, 2016; Juhas et al., 2014; Kuhlman and Cox, 2010), and the antibiotic-resistant gene flanked by the FLP recombinase target (FRT) sites has to be flipped out for reutilization (Datsenko and Wanner, 2000). Notably, with increasing size of the inserted DNA fragments, the recombineering efficiency gradually decreases (Kuhlman and Cox, 2010). Generally, with the λ-Red recombination system, the maximum length of DNA fragments integrated into the *E. coli* genome is ~10 kb (Kuhlman and Cox, 2010; Sabri et al., 2013).

In 2016, Juhas and Ajioka have developed a flexible recombinase-based approach using overlapping DNA fragments and interlaced resistance markers, achieving high-efficiency, iterative integration of large-size DNA constructs (up to 50 kb) in *E. coli* (Juhas and Ajioka, 2016). In this novel method, two resistance-selection cassettes, namely, the kanamycin cassette (kan) and chloramphenicol-yellow fluorescent protein cassette, are used in the plasmids p14kan and pJScav, respectively. Large DNA constructs are first segmentally cloned into these two plasmids with the designed homologous sequences and then iteratively integrated into the genome of *E. coli* MG1655. As a case study, sucrose and lactose catabolism pathways (15 kb), divided into four DNA fragments (each ~3 kb), were integrated into the *fliK* locus with high efficiency. Furthermore, *B. subtilis* 168 DNA (50 kb), divided into seven DNA fragments (each ~6–7 kb), was successfully inserted into two target sites, namely, *filK* and *cheW* (Juhas and Ajioka, 2016).

3.1.2. CRISPR-based markerless, multi-locus gene integration

In the past five years, the CRISPR/Cas9 system has become the most

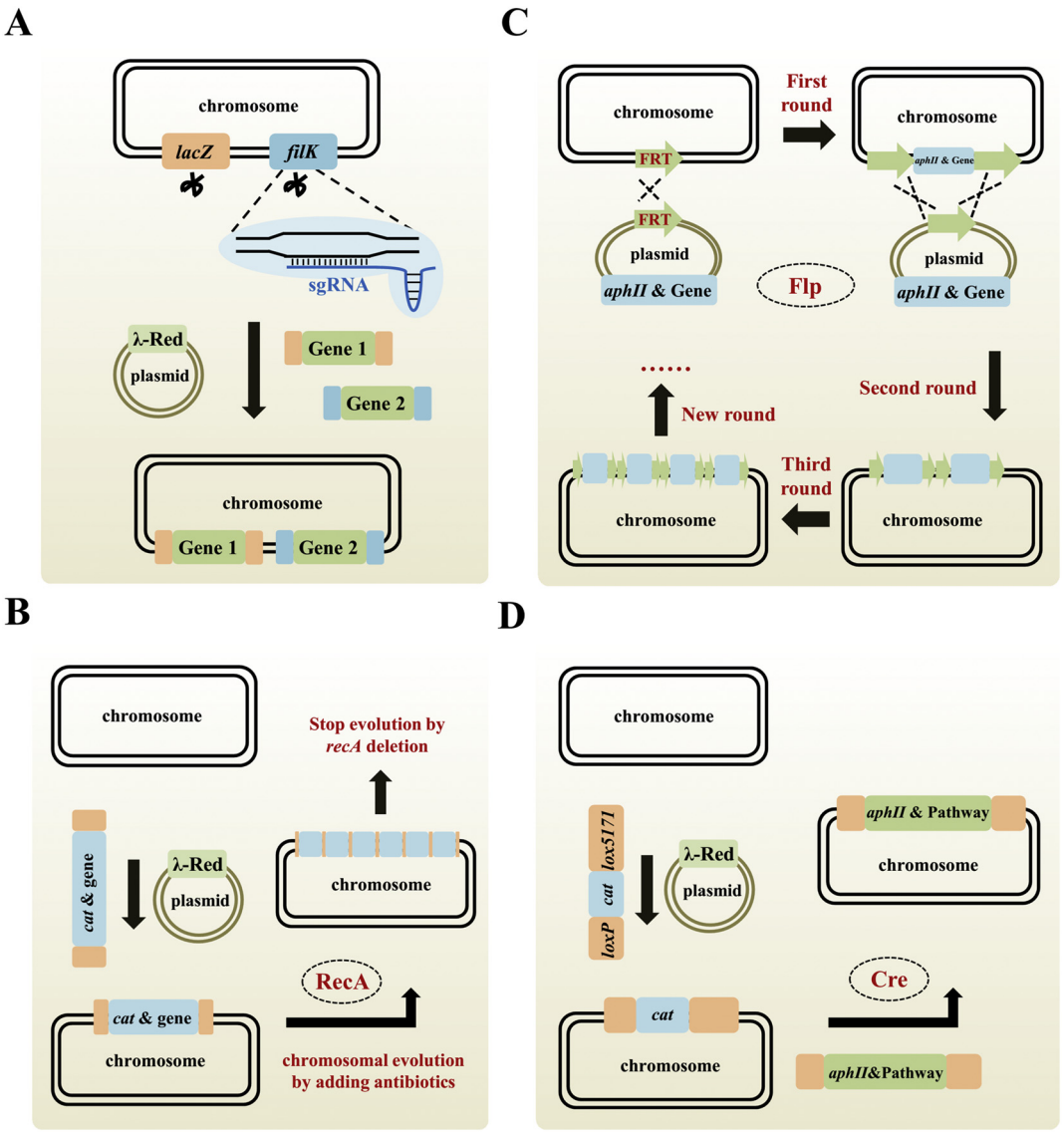

Fig. 2. Four major methods for genomic integration of genes and pathways in *E. coli*. (A) CRISPR-based markerless, multi-locus gene integration. The λ-Red system is used to increase homologous recombination efficiency. Here, the genes *lacZ* and *filK* are chosen as the integration sites. (B) Chemically inducible chromosomal evolution (CIChE) for tandem amplification of target genes or pathways by selection pressure (such as increasing antibiotic concentrations). The gene *recA* will be deleted to stop further evolution and stabilize the amplified genes. *cat*: chloramphenicol resistance gene. (C) CIGMC: Flp recombinase-mediated chromosomal integration of genes with multiple copies. FRT: Flp recombinase recognition target sites. *aphII*: kanamycin resistance gene. (D) RAGE: Cre recombinase-assisted genome engineering for targeted integration of genes/pathways. *loxP* and *lox5171*: Cre recombinase recognition sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

popular genome editing tool in diverse organisms, including bacteria, fungi, plants and mammals (Donohoue et al., 2018; Hille et al., 2018; Hsu et al., 2014). Typically, unique recognition sites must be introduced beforehand into genomes via homonuclease I-SceI or tyrosine recombinase-mediated genetic engineering. However, Cas9 endonuclease can directly target any genomic locus adjacent to the PAM site (NGG) with the aid of synthetic guide RNA (sgRNA), thereby significantly reducing the time and labor required for genome modification (Hille et al., 2018). In *E. coli*, the CRISPR-Cas9 system is widely used in conjunction with the λ-Red recombination system for gene/pathway integration (Fig. 2A). For example, Li and colleagues demonstrated that the CRISPR-based technique could facilitate precise integration of DNA fragments as large as 7 kb with > 60% efficiency (Chung et al., 2017). In another example, researchers were able to scarlessly insert the 10 kb isobutanol biosynthetic pathway into the genome with a relatively high efficiency of 50% (Bassalo et al., 2016). As a representative metabolic engineering application, the 1-butanol synthetic pathway, consisting of five genes, namely, *atoB*, *hbd*, *crt*, *ter*, and *adhE2*, was successfully inserted into the genome of *E. coli*, thus creating a novel butanol cell factory. Notably, these genes were directly used to replace ethanol, lactate, acetate or succinate synthetic genes, thus simultaneously blocking the generation of multiple native by-products. Furthermore, by knockout of 28 other native genes and adaptive evolution, the vector-free *E. coli* strain could produce 20 g/L 1-butanol in batch fermentation without the addition of any antibiotics or expensive inducers, thus exhibiting great potential for industrial application (Dong et al., 2017). In total, the CRISPR-Cas9 genome editing tool will significantly simplify and accelerate efforts to chromosomally integrate genes/pathways into *E. coli* and holds great promise in strain improvement for large-scale biochemical production (Pontrelli et al., 2018).

3.2. CIChE: chemically inducible chromosomal evolution

In *E. coli*, most HR events depend on the universal bacterial enzyme RecA (Bell and Kowalczykowski, 2016). RecA can bind to ssDNA, forming protein-DNA filaments that search other DNA fragments for recombination. For example, tandem gene amplification can be efficiently achieved by RecA-mediated DNA crossover (Bell and Kowalczykowski, 2016). Based on this concept, a CIChE method was developed for high-copy chromosomal expression of recombinant genes/pathways (Fig. 2B) (Tyo et al., 2009). Briefly, the λ-Red recombination system is used to initially integrate the marker and gene of interest flanked by the same homologous regions into the host genome. Then, the host chromosome is subjected to evolution to generate a high gene copy number by selection pressure, which can be stopped by deletion of *recA*. The copy numbers of heterologous genes/pathways can be efficiently tuned with antibiotic strength and stabilized by removal of RecA. The final engineered strains do not require selection pressure and efficiently avoid genetic instability and metabolic burden. Compared with the plasmid-based systems, by using the CIChE method, the productivity of the biopolymer poly-3-hydroxybutyrate (PHB) and the yield of the nutraceutical lycopene increased significantly by 4-fold and 60%, respectively (Tyo et al., 2009).

3.3. SSR-mediated gene integration or tandem amplification

3.3.1. Clonetegration: one-step cloning and integration of multiple DNA segments

Bacteriophage-derived integrases, including tyrosine and serine recombinases, can efficiently promote site-specific recombination between an *attB* site in the bacterial chromosome and a cognate *attP* sequence on the plasmid (Fogg et al., 2014; Merrick et al., 2018). In one example, by using a panel of orthogonal integrases, the “clonetegration” method was developed for simple, multi-locus integration of target genes by combining DNA cloning and integration into a one-step procedure (St-Pierre et al., 2013). The one-step integration plasmid, pOSIP, can be readily modified to generate different variants, including the serine integrase gene from *Φ*C31 or the tyrosine integrase genes from phage 186, HK022, lambda, *Φ*80 and P21. The integration-resistance cassettes flanked by FRT sites are rapidly removed upon inducible expression of the Flp recombinase, thus achieving multi-round iterative integration of large DNA constructs. Using these orthogonal modular integrases, up to four expression cassettes (*cfp*, *gfp*, *rfp* and *lacZ*) were successively inserted into the corresponding genomic loci, and two cassettes (*gfp* and *rfp*) were chromosomally integrated in a single step. Clonetegration simplifies assembly and insertion of multiple DNA fragments, thereby facilitating rapid construction of metabolic pathways and synthetic networks.

3.3.2. CIGMC: Flp recombinase-mediated chromosomal integration of genes with multiple copies

In 2015, a reliable and flexible approach for chromosomal integration of genes with multiple copies (CIGMC) was developed based on Flp recombinase activity (Fig. 2C) (Gu et al., 2015). Briefly, multiple FRT sites are introduced beforehand into the genome of *E. coli*. Then, expression of the Flp protein is induced to mediate integration of the plasmid with the narrow-host-range replicon R6K into the FRT sites, thus achieving multi-locus insertion of target genes. Similar to the CIChE method, the gene *recA* is also deleted here to prevent subsequent recombination of amplified genes. Using this method, multiple copies of the shikimate kinase gene *aroK* and serine biosynthetic genes (*serA*, *serB* and *serC*) were inserted into the chromosome (up to 12 copies) for high-level production of L-tryptophan and serine, respectively. Although this iterative integration concept can be used in a variety of bacteria and fungi, the CIGMC method is laborious and time-consuming due to repeated insertion of artificial FRT sites.

3.3.3. RAGE: Cre recombinase-assisted genome engineering

In addition to Flp-mediated genomic integration, the Cre-*loxP* recombination system is also widely used to mediate duplication and amplification of genes/pathways. For example, to integrate large DNA fragments into the genome of *E. coli*, Yoshikuni and colleagues developed a recombinase-assisted genome engineering (RAGE) method using Cre recombinase and mutated *lox* site pairings (Fig. 2D) (Santos et al., 2013). First, the chloramphenicol resistance gene (*cat*) flanked by *loxP* and *lox5171* sites is inserted into the bacterial genome in advance using the standard λ-Red recombination system. Then, Cre recombinase expression is induced to mediate the transfer of a similarly *lox*-flanked large metabolic pathway and the kanamycin resistance marker (*kan*) into the genome. Finally, removal of *kan* by Flp recombinase leads to the generation of markerless integration strains and to iterative genetic modification. The utility of RAGE was demonstrated via one-step incorporation of a 34-kb alginate metabolic pathway and successive integration of two functional modules for alginate degradation and ethanol production. Notably, the final engineered strain could efficiently produce ethanol directly from macroalgae at higher titers and productivities compared to plasmid-based systems over up to 50 generations (Santos et al., 2013).

4. Multi-copy chromosomal integration of natural product biosynthetic gene clusters in actinomycetes

Actinomycetes are among the most prominent sources for the discovery of novel bioactive natural products (NPs) associated with a large number of drugs used in clinical practice, including antibiotics (e.g., pristinamycin), anticancer agents (e.g., bleomycin), immunosuppressants (e.g., rapamycin) and anthelmintics (e.g., avermectins) (Barka et al., 2016; Newman and Cragg, 2016). Recently, large-scale actinomycetal genome sequencing efforts have revealed that ~90% of NP BGCs are not expressed or are expressed at low levels under routine laboratory conditions and are thus regarded as rich sources for the discovery of novel drug leads (Baltz, 2016; Cimermancic et al., 2014; Doroghazi et al., 2014). Streptomycetes are the most extensively studied actinomycetes due to the propensity of these microbes to produce a large number of bioactive small molecules (Niu, 2017). Therefore, a variety of genetic modification tools have been developed for engineering *Streptomyces* genomes, which could be extended to other industrial actinomycetes, such as *Saccharopolyspora*, *Actinoplanes* and *Amycolatopsis* (Baltz, 2016; Deng et al., 2017).

Generally, two strategies have been used for chromosomal integration of target genes or BGCs into streptomycetes to generate stable recombinants, including site-specific integration and HR-based seamless insertion. A number of integrative plasmids, such as pSET152 with the *Φ*C31 integration system and pRT801 with the *Φ*BT1 integration system, are commonly used for chromosomal expression of target genes in different streptomycetes (Baltz, 2016). In particular, serine recombinase-mediated site-specific integration is highly suitable for chromosomal insertion of large DNA constructs in actinomycetes, such as secondary metabolite BGCs (10–100 kb) (Baltz, 2012). In addition, pSG5-based thermosensitive vectors with broad host ranges, such as pKC1139, can efficiently mediate foreign gene integration by a double crossover process (Muth, 2018). With the advent of the CRISRP-Cas9 or CRISPR-Cpf1 genome editing techniques, a number of plasmids derived from pKC1139, including pCRISPomyces (Cobb et al., 2015), pCRISPRdO (Huang et al., 2015) and pKCCpf1 (Li et al., 2018b), have been developed for rapid genetic modification in a single step. Interested readers could refer to the detailed reviews of the genetic engineering of streptomycetes (and other actinomycetes) (Deng et al., 2017; Tao et al., 2018; Tong et al., 2018). Here, we only focus on duplication or amplification of large BGCs for overproduction of bioactive compounds in actinomycetes.

Generally, industrial production of important drugs or genome mining of novel bioactive compounds requires enhanced production in

actinomycetes. In the last two decades, a variety of rational metabolic engineering strategies have been developed to facilitate target product biosynthesis in actinomycetes, including optimization of regulatory networks, increased precursor supply, modification of the transcription/translation apparatus and amplification of target BGCs (Olano et al., 2008; Weber et al., 2015). As a universal engineering strategy, amplification of natural product BGCs plays an important role in strain improvement. For example, some antibiotic-overproducing strains generated via random mutagenesis, such as *Streptomyces kanamyceticus* and *Streptomyces linolnensis*, exhibit tandem amplification of the corresponding BGCs (Peschke et al., 1995; Yanai et al., 2006). Furthermore, chromosomal tandem duplication of the tautomycetin (TMC) gene cluster resulted in a 14-fold increase in TMC productivity in *Streptomyces* sp. CK4412 (Nah et al., 2015). Here, we focus on these innovative strategies for high-order amplification of target BGCs in streptomycetes.

4.1. ZouA-RsA/B-mediated tandem amplification of target BGCs

Southern hybridization and DNA sequencing indicated that amplification of a large genomic segment (145 kb) harboring the kanamycin BGC occasionally exceeded 36 copies in the kanamycin-overproducing *S. kanamyceticus* (Yanai et al., 2006). Further functional analysis revealed that tandem amplification of the kanamycin BGC is mediated by a specific SSR system, which consists of two *oriT*-like recombination sites, namely, RsA and RsB, and a ZouA relaxase (Murakami et al., 2011a; Murakami et al., 2011b). To demonstrate the generality of the ZouA-RsA/B-mediated amplification system, an amplified DNA unit (35 kb) including the actinorhodin (ACT) BGC of 22 biosynthetic genes and a drug resistance marker was constructed in *S. coelicolor* M145. After selective enrichment, 4–12 tandem copies of the ACT BGC were obtained, resulting in a 20-fold increase in ACT production (Fig. 3A) (Murakami et al., 2011a). Similarly, 3–5 tandem copies of the entire validamycin A (VAL-A) gene cluster (40 kb) were also obtained using this amplification system, resulting in a 34% enhancement of VAL-A production in the industrial strain *Streptomyces hygroscopicus* 5008 (Zhou et al., 2014). Although the ZouA-RsA/B-mediated system can efficiently promote tandem amplification of target BGCs, this system has two obvious limitations that need to be addressed. First, the method requires the introduction of the ZouA-RsA/B system into both flanks of the target BGC in advance via two-round conjugal transfer before target BGC amplification, which is time-consuming and labor-intensive. Moreover, this prior introduction is very difficult and even impossible in genetically intractable actinomycetes. Second, the engineered strains may be genetically unstable in the absence of antibiotic selection due to tandem amplification of the target BGCs.

4.2. Serine recombinase-mediated discrete, multi-copy integration of target BGCs

In the last three decades, phage-encoded large serine integrases have been widely used in genetic engineering of *Streptomyces* and other actinomycetes (Baltz, 2012). The integration process involves a single DNA crossover between the phage *attP* site and the bacterial *attB* site, leading to the generation of two hybrid sites, namely, *attL* and *attR* (Grindley et al., 2006). To date, at least nine SSR systems have been identified in actinomycetes, namely, the CBG73463, ΦBT1, ΦC31, ΦJoe, ΦK38-1, R4, SV1, TG1 and VWB integration systems, and some of these systems have been successfully used for strain development via target BGC amplification (Baltz, 2012; Fogg et al., 2017; Yang et al., 2014a). For example, Haginaka and coworkers sequentially used ΦC31 and TG1 integration systems to triplicate the goadsporin BGC, leading to a 2.3-fold increase in goadsporin production in *Streptomyces* sp. TP-A0584 (Haginaka et al., 2014). However, this method involves repeated rounds of plasmid construction and conjugal transfer and is limited by the number of selection markers. To overcome these bottlenecks, two innovative strategies were developed for efficient amplification of target BGCs in actinomycetes, including MSGE and aMSGE (Li et al., 2017b; Li et al., 2019).

4.2.1. MSGE: multiplex site-specific genome engineering

Typically, there is no or only one active *attB* site for each bacteriophage integrase in many actinomycetes (Baltz, 2012). In theory, prior introduction of artificial *attB* sites into the host chromosome will allow simultaneous, multi-locus insertion of target BGCs in a single step. In 2017, our group developed the MSGE method based on the "one integrase-multiple *attB* sites" concept for discrete integration of target BGCs in actinomycetes (Fig. 3B) (Li et al., 2017b). In this method, the genomic loci of non-target secondary metabolite BGCs are chosen for the insertion of artificial *attB* sites. Using the ΦC31 and ΦBT1 orthogonal integration systems, we successfully achieved five-copy, discrete genomic integration of the pristinamycin II BGC, resulting in elevated PII titers of 2.2 and 2 g/L under flask and batch fermentation conditions, respectively (Li et al., 2015; Li et al., 2017b). Then, the method was extended to construct a panel of heterologous expression superhosts with different numbers of ΦC31 *attB* sites based on *S. coelicolor* M1146 or M1152. Up to four copies of the BGCs for chloramphenicol or anti-tumor compound YM-216391 were high-efficiently inserted in a single step, resulting in significantly increased production titers. Similarly, Myronovskyi et al. developed a series of cluster-free *S. albus* heterologous hosts that contain up to four ΦC31 *attB* sites. The productivities of a number of antibiotics, including didesmethylmensacarcin, griseorhodin A, aloesaponarin II, didemethoxyaranciamycinone and cinnamycin, were higher in these superhosts with increasing numbers of the *attB* sites (Myronovskyi et al., 2018). We believe, these reliable and versatile *Streptomyces* hosts will significantly advance the discovery of novel drug leads via genome mining as well as combinatorial biosynthesis research (Baltz, 2016; Li et al., 2017a; Zhang et al., 2016b).

4.2.2. aMSGE: multi-locus integration by orthologous modular SSR systems

Although the MSGE method is highly suitable for the development of high-efficient expression hosts for the discovery of novel therapeutic compounds in actinomycetes, there remain some obvious limitations. On the one hand, repeated introduction of foreign *attB* sites is labor-intensive and time-consuming, even with the use of CRISPR-based genome editing tools. On the other hand, it is very difficult and even impossible to introduce artificial *attB* sites into the genomes of genetically intractable actinomycetes. To address these bottlenecks, an advanced MSGE method (aMSGE) was developed based on the innovative "multiple integrase-multiple *attB* sites" concept (Fig. 3C) (Li et al., 2019). In contrast to MSGE, this updated method takes advantage of native *attB* sites of different compatible SSR systems in the actinomycetal genome. Accordingly, a plug-and-play amplification toolkit consisting of 27 modular plasmids with single- or multi-integration modules was designed and developed. The procedure for BGC amplification takes only ~2 days for in vitro plasmid construction and ~18 days for the development of high-yield engineered strains (e.g., growth period is ~6 days). As a representative application, the 5-oxomilbemycin BGC was rapidly integrated into the parental strain with three extra copies using aMSGE, leading to a significant increase in production titers of 5-oxomilbemycin (from 2.23 to 6.37 g/L) (Li et al., 2019). Notably, compared with previously reported amplification methods, genetic modifications in host genomes are not required for aMSGE before multi-copy integration of target BGCs. This advantage will considerably simplify and accelerate efforts to optimize NP biosynthesis and will be applicable to genetically intractable actinomycetes. Additionally, given that SSR systems are widely distributed in a variety of prokaryotic microbes (Fogg et al., 2014; Stark, 2017), this novel methodology can be extended to other industrial strains to improve the activities of biosynthetic pathways for commercial applications.

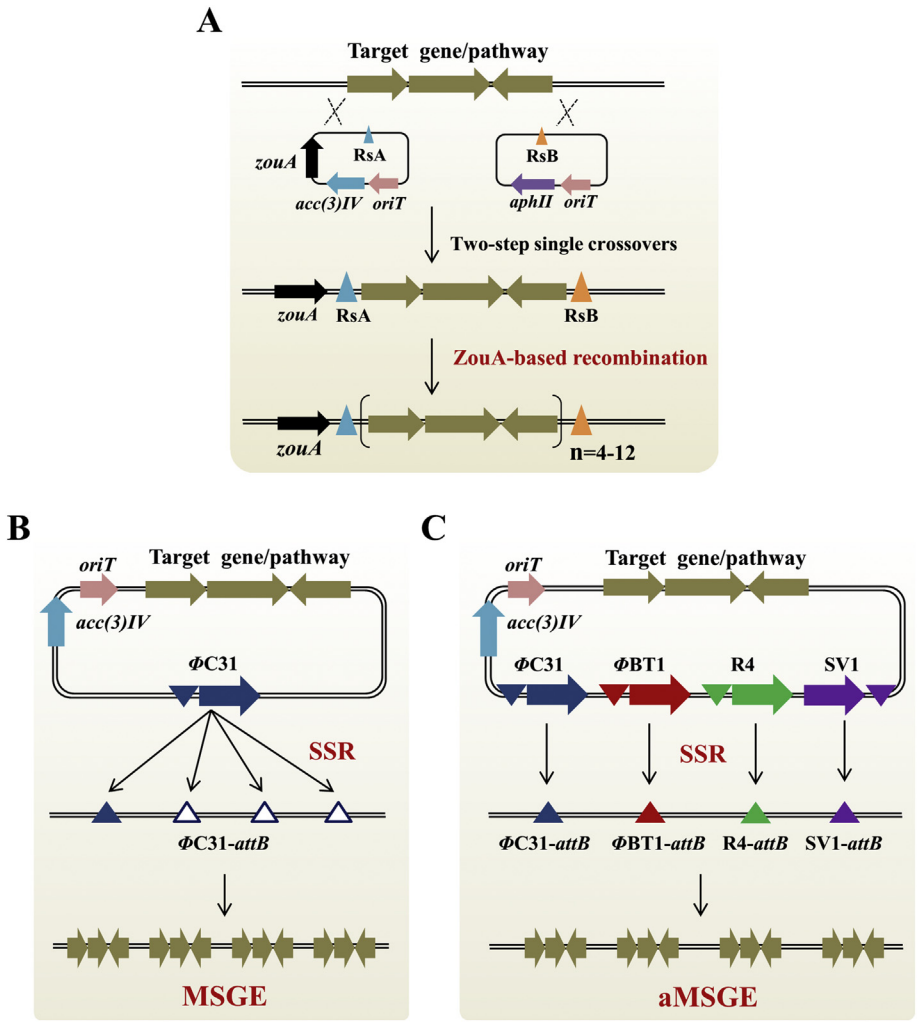

Fig. 3. Three major strategies for the amplification of natural product BGCs in actinomycetes. (A) ZouA-RsA/B-mediated site-specific recombination. This approach facilitates tandem BGC amplification. *acc (3)IV* and *aphII* represent apramycin and kanamycin resistance genes, respectively. (B) Multiplex site-specific genome engineering (MSGE) for multi-copy integration of target BGCs. This method is based on the "one integrase-multiple *attB* sites" concept. The blue and blank triangles represent the native and artificial *Φ*C31 *attB* sites, respectively. *oriT* represents the element of conjugal transfer. SSR: site-specific recombination. (C) Advanced multiplex site-specific genome engineering (aMSGE) for multi-locus chromosomal integration of target BGCs. This method is based on the "multiple integrases-multiple *attB* sites" concept. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Novel technologies and applications of genomic integration of genes and pathways in yeasts

Yeasts have been widely used in industrial biotechnology to produce biofuels, chemicals, pharmaceuticals and food additives (Liu et al., 2013). As unicellular eukaryotes, yeasts possess some key advantages compared to bacteria. For example, there are a number of organelles, such as mitochondria, peroxisomes, ER, the Golgi and vacuoles, for compartmentalization of metabolic pathways in different yeast species (Hammer and Avalos, 2017). In addition to the two well-known yeasts, baker's yeast (or *S. cerevisiae*) and fission yeast (or *Schizosaccharomyces pombe*), > 2000 yeasts have been discovered and characterized to date (Radecka et al., 2015). Some important non-conventional yeasts, such as *Y. lipolytica* and *Pichia pastoris*, are increasingly been used for industrial applications. For instance, the methylotrophic yeasts *P. pastoris* and *Hansenula polymorphas* have been widely used for heterologous protein production (Wagner and Alper, 2016). In current basic and biotechnological research, multiple rounds of genetic manipulation are required for efficient and stable production of chemicals and pharmaceuticals in yeasts. Here, we reviewed a panel of newly developed approaches for genomic integration of genes/pathways and discussed the relevant metabolic engineering applications in conventional and non-conventional yeasts.

5.1. S. cerevisiae: a powerful host for production of biofuels and pharmaceuticals

S. cerevisiae, as a platform microorganism for molecular biology and metabolic engineering, offers a large number of advantages, including high-efficiency HR ability, genetic tractability and fermentation robustness (Hong and Nielsen, 2012). Because of its generally recognized as safe (GRAS) status, *S. cerevisiae* has been widely used for the production of food, pharmaceuticals, biofuels and other value-added chemicals. These features have led to increasing and extensive research for the development of advanced genome engineering tools for both fundamental studies and biotechnological applications (Lian et al., 2018). In particular, high-efficiency integration of exogenous DNA with short (~50 bp) HAs can be easily achieved based on HDR (the preferred DSB repair pathway) in *S. cerevisiae*. In the past six years, the CRISPR-Cas system combined with HDR has become a revolutionary and versatile strategy for genome editing in *S. cerevisiae* (Raschmanova et al., 2018). Currently, a series of different approaches based on the CRISPR-Cas system have been developed for marker-free, multi-locus integration of large biochemical pathways, significantly advancing the development of high-efficient yeast cell factories (Jakociunas et al., 2016; Liu et al., 2017).

5.1.1. Multi-locus integration without assembly

5.1.1.1. Di-CRISPR: native delta repeat sequences as integrated sites. Although generation of DSBs using the CRISPR-Cas system can increase HR efficiency, most previous studies have only focused on

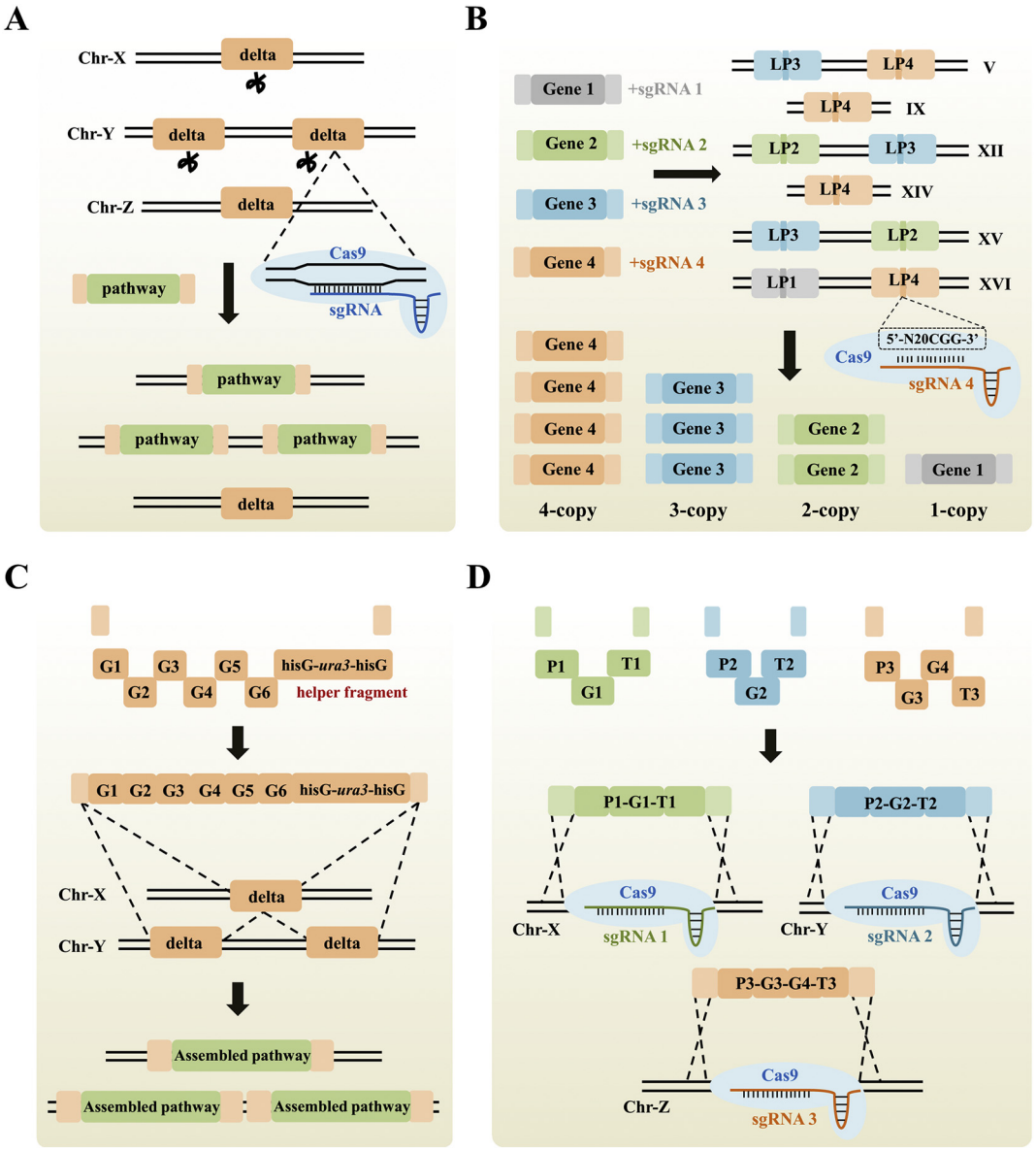

Fig. 4. Novel technologies for chromosomal integration of genes and pathways in *S. cerevisiae*. (A) Di-CRISPR (delta integration CRISPR-Cas) for multi-locus integration of large biochemical pathways by using the native repeated delta sequences as integrated sites. (B) Synthetic DNA landing pads (LPs) as docks for strict control of copy numbers of integrated genes. (C) DNA assembler for one-step DNA assembly and integration of multiple DNA fragments with the aid of selection markers. (D) CasEMBLR: CRISPR-mediated marker-free, multi-locus DNA assembly and integration. P1-P3: promoters 1–3, G1-G4: genes 1–4, T1-T3: terminators 1–3.

genomic integration of single genes/pathways (Horwitz et al., 2015; Mans et al., 2015; Ronda et al., 2015). In 2015, a CRISPR-based method, CRISPRm, was developed to achieve multiplex integration of metabolic pathways in *S. cerevisiae*. In that study, a muconic acid biosynthesis pathway (24 kb, 11 genes) was integrated into three genomic loci in a single step. Although less time and labor are required to realize individual strain designs, the triple integration efficiency was very low (only 4.2%) (Horwitz et al., 2015). Subsequently, a high-efficiency, single-step and markerless platform, Di-CRISPR (delta integration CRISPR-Cas), was developed for multi-locus chromosomal integration of large metabolic pathways by using the native repetitive Ty retrotransposon delta sequences as integrated sites (Shi et al., 2016) (Fig. 4A). As an example, a large pathway responsible for both xylose utilization and (*R*,*R*)-2,3-butanediol (BDO) production (up to 24 kb) was high-efficiently introduced into the chromosome with up to 18 copy numbers in a single step. Owing to its ease of use and high efficiency, this Di-CRISPR platform will be invaluable for synthetic biology and genome engineering studies in *S. cerevisiae*.

5.1.1.2. Synthetic DNA landing pads as docks for gene integration. Although the Di-CRISPR method can mediate multi-copy integration of large biochemical pathways, the Ty retrotransposon delta sites are unusually unstable genomic regions for the construction of engineered strains. In 2018, an alternative tool, “Wicket”, was developed for high-efficiency integration and rapid optimization of foreign pathways in *S. cerevisiae* (Hou et al., 2018). Briefly, a short DNA cassette, consisting of two 50-bp HAs and a 23-bp CRISPR/Cas9 target sequence, was simultaneously introduced into multiple designated loci, such as *ADE2*, *CAN1* and *URA3*, to accept exogenous DNA fragments upon Cas9-mediated excision. To demonstrate the application of “Wicket”, the entire, preassembled β-carotene biosynthesis pathway, including the three genes *crtE*, *crtI*, and *crtYB*, were multicopy integrated into the genome without the need for selective markers in a single step. In the same study, the three genes from the β-carotene biosynthesis pathway were also independently integrated with variable

copy numbers, resulting in a large variety of engineered strains for enhanced production. Recently, a similar CRISPR-Cas9 integration system was also developed by introducing a series of synthetic DNA landing pads (LPs) (Fig. 4B) (Bourgeois et al., 2018). Unlike the "Wicket" method, orthogonal LP modules with one to four DNA copies are designated to achieve precise control of gene copy number, which efficiently expedites pathway engineering campaigns for strain improvement. As an application case, the production titers of (*S*)-norcoclaurine were significantly increased by integrating and selecting ten norcoclaurine synthase (NCS) orthologs with four gene copies each.

5.1.2. Multi-locus integration and assembly

5.1.2.1. DNA assembler. HR has been widely harnessed for plasmid construction, library creation and foreign DNA integration in *S. cerevisiae*. However, the use of HR for the assembly of large biochemical pathways in a single step on the yeast chromosome had not been explored. In 2009, a one-step method, called 'DNA assembler', was developed for in vivo assembly and targeted integration of multigene biochemical pathways into the genomic delta loci with the help of a selection marker (Fig. 4C) (Shao et al., 2009). First, the interlaced overlaps are introduced into each gene expression cassette in a pathway by PCR amplification. For efficient DNA assembly and integration, the lengths of all the overlaps are > 40 bp. Second, a helper integration fragment *hisG-ura3-hisG* is designed, and the repeated *hisG* motifs can be introduced for removal of the inserted gene *ura3*, thus recycling the selection marker for multi-round integration. Finally, multiple expression cassettes and helper fragment are co-transformed into *S. cerevisiae* via electroporation, leading to the assembly of the entire pathway into the chromosome.

Using the DNA assembler method, a three-gene D-xylose utilization pathway (~9 kb) and a five-gene zeaxanthin biosynthesis pathway (~11 kb) could be integrated into the yeast chromosome with a high efficiency of 80–100%. By extending the overlap length (from 50 to 125–430 bp), the efficiency of one-step DNA assembly and integration of an eight-gene hybrid pathway (~19 kb) for both D-xylose utilization and zeaxanthin biosynthesis was also significantly increased to 40–70%. The engineered yeast with the hybrid pathway could grow in the medium with D-xylose as the sole carbon source and produce zeaxanthin. In principle, this method may be applied to iteratively integrate a DNA molecule as large as an entire microbial genome by recycling the *ura3* gene using 5-fluororotic acid for counter-selection. However, marker recycling is also time-consuming and is associated with genome instability.

5.1.2.2. CasEMBLR. As mentioned above, Cas9-induced DSBs can significantly facilitate HDR in *S. cerevisiae*. Notably, Cas9-induced DSBs exert negative screening pressure for efficient genome engineering, which enables markerless integration of metabolic pathways without the aid of selection markers. In 2015, Jay Keasling's group developed a new method, called CasEMBLR, for multi-locus chromosomal integration of in vivo-assembled DNA parts based on the CRISPR-Cas9 system in *S. cerevisiae* (Jakociunas et al., 2015) (Fig. 4D). Three genes, namely, *URA3*, *HIS3* and *ADE2*, were chosen as the integration sites. To validate its applicability for the development of yeast cell factories, this method was used to assemble and integrate a carotenoid biosynthesis pathway from 15 DNA parts into three targeted genomic loci and construct a tyrosine production strain with ten parts inserted into two loci in a single step. Notably, average engineering efficiencies of 30% and 58% were obtained when the target parts were assembled into three and two loci, respectively, without any phenotypic selection. Then, the same group characterized 23 Cas9 targeting integration loci in different chromosomes, thereby expanding the application scope of CasEMBLR for one-step or iterative integration of a large number of genetic parts (Apel et al., 2017). In the future, the CasEMBLR method will significantly advance multiplex genome engineering for enhanced biosynthesis of value-added products.

5.2. Non-conventional yeasts: promising alternative biomanufacturing platforms

Although the well-characterized model organism *S. cerevisiae* is widely used in industrial biotechnology and molecular genetics, other non-conventional yeasts, such as *Y. lipolytica*, *P. pastoris*, *K. lactis*, *S. stipitis* and *H. polymorpha*, have been also developed as expression hosts due to their desirable features, including thermotolerance, high levels of protein secretion and wide substrate scopes (Lobs et al., 2017; Raschmanova et al., 2018; Wagner and Alper, 2016). However, efficient metabolic engineering and synthetic biology tools require further development in non-conventional yeasts (Markham and Alper, 2018; Wagner and Alper, 2016). In addition, in contrast to *S. cerevisiae*, which exhibits efficient HR ability, non-conventional yeasts usually do not possess high-efficiency HR machinery (Wagner and Alper, 2016). Therefore, other DSB repair systems, including NHEJ and MMEJ/HMEJ systems, should be considered for genomic integration of genes/pathways in non-conventional yeasts.

5.2.1. Agrobacterium-mediated transformation (ATMT)-based functional integration in Rhodosporidium toruloides

The basidiomycetous yeast *R. toruloides* is an excellent microbial host for production of carotenoids and lipids (Park et al., 2018; Wu et al., 2011). For example, this species can accumulate lipids up to 60% cell dry weight under certain fermentation conditions (Li et al., 2007; Wu et al., 2011). However, genetic engineering of *R. toruloides* remains challenging due to the absence of an efficient transformation method. Previously, HR-based transformation was tested to integrate several genes into the host chromosome with large HAs (up to 1300 bp), but no positive recombinants were obtained (Lin et al., 2014). To address this limitation, Zhao and colleagues developed an *Agrobacterium*-mediated transformation (ATMT) protocol for effective chromosomal integration of multiple genes in both haploid and diploid *R. toruloides* strains (Lin et al., 2014). With the aid of T-DNA, this ATMT method efficiently mediated random functional integration and expression of resistance genes, which could be verified by colony PCR, western blotting and genome walking. Using the ATMT method, a carotenoid biosynthetic gene, *CRTI,* was efficiently inserted into the genome to produce red carotenoid in the non-model yeast species (Sun et al., 2017).

5.2.2. HR-mediated markerless integration in Y. lipolytica, K. lactis and other non-model yeast species

Distinct from *S. cerevisiae*, the oleochemical-producing yeast *Y. lipolytica* typically requires 0.5-1-kb HAs to allow low integration efficiency with the use of selective markers (Markham and Alper, 2018). Recently, with the advent of CRISPR-Cas9 technology, a CRISPR-based HR strategy has been demonstrated for targeted multi-locus integration without the need for marker recovery in *Y. lipolytica* (Schwartz et al., 2017b). To identify accessible integration sites without impairing cell growth, a set of genomic loci was screened. The insertion efficiency of a GFP expression cassette reached 40–60% but varied across different inserted loci. Meanwhile, except for the *MFE1* integration site, the GFP expression levels were similar when integrated into the other four sites. Furthermore, a semisynthetic lycopene biosynthesis pathway containing four genes was successfully integrated into different genomic loci via a single transformation. To further increase integration efficiency, the same group developed a CRISPR interference (CRISPRi) system to repress the expression of *ku70* and *ku80*, and a high rate of HR was achieved (up to 90%) (Schwartz et al., 2017a). In a second industrially relevant yeast, *K. lactis*, a large muconic acid biosynthetic pathway comprising six genes (totally 24 kb) was separately integrated into three different loci in a single step based on the CRISPRm method (Horwitz et al., 2015). Unfortunately, triple integration efficiency was very low (only 2.1%), and the engineered strain could not produce

muconic acid. Until now, on the basis of CRISPR-Cas9 system, HDR-mediated gene markerless integration has also established in other important non-conventional yeasts, such as *Kluyveromyces marxianus* (Cernak et al., 2018), *Ogataea polymorpha* (Wang et al., 2018a), *P. pastoris* (Weninger et al., 2018), and *S. stipitis* (Cao et al., 2018).

5.2.3. Homology-independent random insertion in Y. lipolytica

In addition to HR-mediated genomic integration, the inherent NHEJ pathway could be used to integrate DNA into *Y. lipolytica*. For instance, Hou and colleagues demonstrated that multiple (up to three) and large fragments (up to 12.5 kb) could be tandemly integrated in a single step without HAs in an NHEJ-dependent manner (Cui et al., 2019). In particular, the expression levels of integrated genes varied due to different locations and copy numbers, facilitating rapid construction of modular libraries for pathway optimization. In one example, a β-carotene biosynthetic pathway consisting of three functional modules was rapidly introduced into *Y. lipolytica*, leading to a nearly 27-fold increase in β-carotene titers (Cui et al., 2019).

5.2.4. HMEJ-mediated in-frame integration of exogenous genes in Y. lipolytica

Although CRISPR-based HDR offers precise integration of target DNA, the efficiency of this process is relatively low in *Y. lipolytica*. Recently, a new DSB repair pathway, HMEJ, was exploited for targeted integration of transgenes in animal cells, particularly non-dividing cells (Yao et al., 2017). Briefly, the CRISPR-Cas9 system was used to induce site-specific cleavage in the genome as well as a template donor vector with sgRNA target sequences flanking 800-bp HAs. HMEJ harbors longer and perhaps more stable homologous regions with higher knock-in efficiencies in cultured cells and animal embryos than MMEJ. Furthermore, the HMEJ-based method also exhibited a 6-15-fold higher insert efficiency than the HR-based strategy. Subsequently, this innovative strategy was extended for complete gene knockout and integration in *Y. lipolytica* (Gao et al., 2018). A 1-kb HA with two identical sgRNA target sequences flanking the HA was designed in the HMEJ donor plasmid. Similarly, a higher integration efficiency (~37.5%) was obtained using the HMEJ-based method than by using the HR-based method. Although the precise mechanism of HMEJ-mediated DSB repair requires further investigation, HMEJ-mediated targeted integration could be widely used in many HR-deficient industrial bacteria and fungi.

6. Concluding remarks

In modern society, engineering microbial cell factories for overproduction of value-added compounds is a promising approach to overcome a wide variety of challenges in the food, pharmaceutical, and biochemical industries (Lee and Kim, 2015; Li et al., 2018a; Nielsen and Keasling, 2016). Over the last three decades, the inherent genetic instability of plasmid-based expression has led researchers to develop a series of powerful genomic integration approaches for pathway optimization and efficient microbial biosynthesis of target products (Liu et al., 2017; Ou et al., 2018; Yamamoto and Gerbi, 2018). This review aimed to provide an inspiring insight into recent advances and applications in the strategies and approaches for the genomic integration of genes/pathways in representative industrial microbes (Table 2).

With the advent of CRISPR-Cas genome editing technology, markerless, multi-locus integration of large metabolic pathways has become increasingly feasible on the basis of traditional HR in *E. coli* and *S. cerevisiae*, further promoting the use of these two model strains as preferred microbial cell classiss for the production of complex pharmaceuticals and chemicals (Donohoue et al., 2018; Pontrelli et al., 2018; Raschmanova et al., 2018). Notably, with the aid of different numbers of orthogonal synthetic LPs, precise control of gene copy number in the chromosome will possibly achieve the most optimized metabolic flux balance in specific pathways (Bourgeois et al., 2018; Hou et al., 2018). In the meantime, the ability to generate DSBs at almost any site of interest by using the programmable Cas nucleases provides an opportunity for NHEJ- or MMEJ/HMEJ-mediated gene integration in genetically intractable or HR-deficient industrial microbes, such as non-conventional yeasts and filamentous fungi (Raschmanova et al., 2018; Shi et al., 2017). In particular, this novel HMEJ pathway combined with the CRIPSR-Cas system may offer broad applications in targeted pathway integration in some important non-model microbes that suffer from low HR efficiency (Gao et al., 2018; Yao et al., 2017). In addition, the SSR concept has been continuously applied to the development of advanced multiplex integration approaches in *E. coli* and actinomycetes. Notably, by the newly developed MSGE or aMSGE methods, multi-copy integration of large natural product BGCs (~20–100 kb) could be easily achieved in a single step, which will significantly advance genome mining of novel bioactive small molecules and the construction of high-yield industrial actinomycetes (Li et al., 2017b; Li et al., 2019; Zhang et al., 2016b).

However, most studies on the development of approaches for multicopy integration of large DNA constructs have only focused on model microbes, including *E. coli*, *Streptomyces* and *S. cerevisiae*. Although the programmable CRISPR-Cas system has been adapted for targeted gene integration in other important industrial microbes, such as *Clostridia* species (Xue et al., 2017), *Lactobacilli* (Hidalgo-Cantabrana et al., 2017), *Corynebacteria* (Wang et al., 2018b), non-conventional yeasts (Raschmanova et al., 2018) and filamentous fungi (Shi et al., 2017), the insertion sizes and copy numbers of foreign DNA constructs remain limited. Given that SSR is very suitable for genomic integration of large DNA constructs (Grindley et al., 2006; Merrick et al., 2018), the MSGE methodology, in which multiple orthogonal artificial *attB* sites are introduced beforehand into the host chromosome, could be possibly be used to engineer large pathways for improved biochemical synthesis in many non-model industrial strains (Li et al., 2017b; Myronovskyi et al., 2018). For instance, our group recently adapted the innovative methodology into the gas-fermenting *Clostridium ljungdahlii* for chromosome-based stable expression of the whole heterologous butyrate biosynthetic pathway (8.5 kb), which efficiently overcomes the deficiency of currently available synthetic biology approaches in the genomic integration of large DNA constructs in the industrially important anaerobe (Huang et al., 2019).

Finally, although chromosomal integration of foreign DNA constructs will efficiently overcome the obstacle of plasmid-based systems for the biosynthesis of target products, increased copy numbers of chromosomally integrated genes/pathways may lead to severe metabolic burden and growth defects (Li et al., 2017b; Li et al., 2019). Therefore, an integrated approach, including promoter engineering, precursor engineering, amplification of biosynthetic genes and dynamic control of metabolic pathways, should be combinatorially applicable to microbial strain improvement (Cho et al., 2015; Lee and Kim, 2015; Nielsen and Keasling, 2016). Nevertheless, the strategies for targeted genomic integration highlighted herein and continued advances in this area will play increasingly important roles in industrial biotechnology for large-scale production of value-added products.

Acknowledgements

We gratefully acknowledge the financial support from the National Natural Science Foundation of China (31630003), the National Postdoctoral Program for Innovative Talents (BX201700265), the China Postdoctoral Science Foundation (2017M621545), the Science and Technology Commission of Shanghai Municipality (18ZR1446700), the National Mega-project for Innovative Drugs (2018ZX09711001-006-012) and the Derivative Bank of Chinese Biological Resources, CAS (ZSYS-016).

Table 2
Summary and comparison of novel enabling technologies for chromosomal integration of genes and pathways in three important industrial microbes.

Species	Principles	Methods	Single- or multi-copy	Selection or markerless	Largest successful integration size	Efficiency	Advantages	Disadvantages	References
E. coli	HR	Integration of large DNA constructs by recombineering	Single-copy	Selection	50 kb (*B. subtilis* 168 DNA)	NA	Stability; Suitable to iterative integration of large-size DNA	Time-consuming	(Juhas and Ajioka, 2016)
		CRISPR-based gene multi-locus integration	Single-copy	Markerless	10 kb (Isobutanol pathway)	50%	Stability; Ease of use	Unsuitable to integration of large-size DNA	(Bassalo et al., 2016)
		CIChE	Up to 40 tandem copies	Selection	5 kb (PHB pathway)	NA	Ease of use; High copy numbers	Instability	(Tyo et al., 2009)
	SSR	Clonetegration by site-specific recombinases	Single-copy	Selection	~6 kb (*vioE-ABDE* cassette)	NA	Suitable to one-step DNA cloning and integration	Time-consuming	(St-Pierre et al., 2013)
		RAGE by Cre recombinase	Single-copy	Selection	59 kb (Alginate pathway)	65–85%	Suitable to iterative integration of large-size DNA	Time-consuming	(Santos et al., 2013)
		CIGMC by Flp recombinase	Up to 10 tandem copies	Selection	1.2 kb (*serA* gene)	NA	Ease of use; High copy numbers	Instability; Time-consuming	(Gu et al., 2015)
Actinomycetes	HR	HR-mediated BGC duplification	Single-copy	Selection	80 kb (TMC BGC)	75%	Ease of use	Instability; Low copy numbers	(Nah et al., 2015)
	SSR	ZouA-RsA/B system-mediated tandem amplification	4–12 tandem copies (five-round)	Selection	35 kb (ACT BGC)	NA	High copy numbers	Instability; Difficult to use; Time-consuming	(Murakami et al., 2011a)
		MSGE via "one integrase-multiple *attB* sites"	5 discrete copies (two-round)	Selection	67 kb (PII BGC)	80–100%	Stability; Suitable to a wide variety of bacteria and fungi	Time-consuming	(Li et al., 2017b)
		aMSGE via "multiple integrases-multiple *attB* sites"	4 discrete copies (two-round)	Selection	72 kb (5-oxomilbemycin BGC)	35–100%	Stability; Suitable to genetically tractable and intractable actinomycetes	Host genome must harbor multiple native *attB* sites	(Li et al., 2019)
S. cerevisiae	HR	DI-CRISPR	Up to 18 discrete copies	Markerless	24 kb (xylose utilization and BDO pathway)	NA	High copy number; Ease of use; Suitable to multi-locus integration	Instability; Random copy numbers	(Shi et al., 2016)
		Synthetic DNA landing pads system	1–4 discrete copies	Markerless	1–2 kb (norcoclaurine synthase genes)	39–81%	Stability; Ease of use; Strict control of gene copy numbers	Low copy numbers	(Bourgeois et al., 2018)
		DNA assembler	> 10 discrete copies with assembly	Selection	19 kb (xylose utilization and zeaxanthin pathway)	70–100%	High copy numbers; Ease of use; Suitable to integration of large-size DNA	Time-consuming	(Shao et al., 2009)
		CasEMBLR	Up to 3 discrete copies with assembly	Markerless	~4 kb (beta-carotene pathway)	30%	Ease of use; Simultaneous assembly and integration	Low copy numbers	(Jakociunas et al., 2015)
Non-conventional yeasts	HR	CRISPRm method for multi-locus integration in *K.lactis*	Single-copy	Markerless	24 kb (muconic acid pathway)	2%	Ease of use; Suitable to integration of large-size DNA	Low-efficiency; Low copy numbers	(Horwitz et al., 2015)
		CRISPR-based multi-locus integration in *Y. lipolytica*	Up to 4 discrete copies	Markerless	~4 kb (4 genes from lycopene pathway)	NA	Stability; Ease of use	Low copy numbers; Unsuitable to integration of large-size DNA	(Schwartz et al., 2017b)
	NHEJ	Homology-independent integration in *Y. lipolytica*	Up to 3 discrete copies	Selection	7.1 + 12.5 kb (*hrGFP* and *PYC/FRD*)	NA	Ease of use; Without the need for homology templates	Randem insertion; Time-consuming	(Yao et al., 2017)
	HMEJ	HMEJ-based integration in *Y. lipolytica*	Single-copy	Markerless	1.4 kb (*hph* gene)	37%	Ease of use; High efficiency compared with HR	Unsuitable to integration of large-size DNA	(Gao et al., 2018)

Notes: HR, homologous recombination; SSR, site-specific recombination; NHEJ: non-homologous end joining; HMEJ: homology-mediated end joining; BGC, biosynthetic gene cluster; PHB, poly-β-hydroxybutyrate; TMC, tautomycetin; ACT, actinorhodin; PII, pristinamycin II; BDO, (*R*,*R*)-2,3-butanediol. NA: not available.

References

Andersson, D.I., Hughes, D., 2009. Gene amplification and adaptive evolution in bacteria. Annu. Rev. Genet. 43, 167–195.

Apel, A.R., d'Espaux, L., Wehrs, M., Sachs, D., Li, R.A., Tong, G.J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N.J., Keasling, J.D., Mukhopadhyay, A., 2017. A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. Nucleic Acids Res. 45 (1), 496–508.

Ayora, S., Carrasco, B., Cardenas, P.P., Cesar, C.E., Canas, C., Yadav, T., Marchisone, C., Alonso, J.C., 2011. Double-strand break repair in bacteria: a view from *Bacillus subtilis*. FEMS Microbiol. Rev. 35 (6), 1055–1081.

Aziz, R.K., Breitbart, M., Edwards, R.A., 2010. Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res. 38 (13), 4207–4217.

Baltz, R.H., 2012. *Streptomyces* temperate bacteriophage integration systems for stable genetic engineering of actinomycetes (and other organisms). J. Ind. Microbiol. Biotechnol. 39 (5), 661–672.

Baltz, R.H., 2016. Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. J. Ind. Microbiol. Biotechnol. 43 (2–3), 343–370.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.P., Clement, C., Ouhdouch, Y., van Wezel, G.P., 2016. Taxonomy, physiology, and natural products of actinobacteria. Microbiol. Mol. Biol. Rev. 80 (1), 1–43.

Barrangou, R., Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. Nat. Biotechnol. 34 (9), 933–941.

Bassalo, M.C., Garst, A.D., Halweg-Edwards, A.L., Grau, W.C., Domaille, D.W., Mutalik, V.K., Arkin, A.P., Gill, R.T., 2016. Rapid and efficient one-step metabolic pathway integration in *E. coli*. ACS Synth. Biol. 5 (7), 561–568.

Bell, J.C., Kowalczykowski, S.C., 2016. RecA: regulation and mechanism of a molecular search engine. Trends Biochem. Sci. 41 (6), 491–507.

Bian, X.Y., Tang, B., Yu, Y.C., Tu, Q., Gross, F., Wang, H.L., Li, A.Y., Shen, Y.M., Li, Y.Z., Stewart, A.F., Zhao, G.P., Ding, X.M., Muller, R., Zhang, Y.M., 2017. Heterologous production and yield improvement of epothilones in *Burkholderiales* strain DSM 7029. ACS Chem. Biol. 12 (7), 1805–1812.

Bilyk, B., Horbal, L., Luzhetskyy, A., 2017. Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074. Microb. Cell Factories 16 (1), 5.

Blazeck, J., Hill, A., Liu, L.Q., Knight, R., Miller, J., Pan, A., Otoupal, P., Alper, H.S., 2014. Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production. Nat. Commun. 5, 3131.

Bourgeois, L., Pyne, M.E., Martin, V.J.J., 2018. A highly characterized synthetic landing pad system for precise multicopy gene integration in yeast. ACS Synth. Biol. 7 (11), 2675–2685.

Bowater, R., Doherty, A.J., 2006. Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. PLoS Genet. 2 (2), 93–99.

Bryant, J.A., Sellars, L.E., Busby, S.J.W., Lee, D.J., 2014. Chromosome position effects on gene expression in *Escherichia coli* K-12. Nucleic Acids Res. 42 (18), 11383–11392.

Cameron, D.E., Bashor, C.J., Collins, J.J., 2014. A brief history of synthetic biology. Nat. Rev. Microbiol. 12 (5), 381–390.

Cao, M.F., Gao, M.R., Ploessl, D., Song, C.J., Shao, Z.Y., 2018. CRISPR-mediated genome editing and gene repression in *Scheffersomyces stipitis*. Biotechnol. J. 13 (9), e1700598.

Carrier, T., Jones, K.L., Keasling, J.D., 1998. mRNA stability and plasmid copy number effects on gene expression from an inducible promoter system. Biotechnol. Bioeng. 59 (6), 666–672.

Cernak, P., Estrela, R., Poddar, S., Skerker, J.M., Cheng, Y.F., Carlson, A.K., Chen, B., Glynn, V.M., Furlan, M., Ryan, O.W., Donnelly, M.K., Arkin, A.P., Taylor, J.W., Cate, J.H.D., 2018. Engineering *Kluyveromyces marxianus* as a robust synthetic biology platform host. mBio 9 (5), e01410-18.

Chaikind, B., Bessen, J.L., Thompson, D.B., Hu, J.H., Liu, D.R., 2016. A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells. Nucleic Acids Res. 44 (20), 9758–9770.

Chen, X., Zhang, J., 2016. The genomic landscape of position effects on protein expression level and noise in yeast. Cell Syst. 2 (5), 347–354.

Chevalier, B.S., Stoddard, B.L., 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res. 29 (18), 3757–3774.

Cho, C., Choi, S.Y., Luo, Z.W., Lee, S.Y., 2015. Recent advances in microbial production of fuels and chemicals using tools and strategies of systems metabolic engineering. Biotechnol. Adv. 33 (7), 1455–1466.

Cho, J.S., Choi, K.R., Prabowo, C.P.S., Shin, J.H., Yang, D., Jang, J., Lee, S.Y., 2017. CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. Metab. Eng. 42, 157–167.

Choi, K.R., Lee, S.Y., 2016. CRISPR technologies for bacterial systems: current achievements and future directions. Biotechnol. Adv. 34 (7), 1180–1209.

Choi, K.R., Cho, J.S., Cho, I.J., Park, D., Lee, S.Y., 2018. Markerless gene knockout and integration to express heterologous biosynthetic gene clusters in *Pseudomonas putida*. Metab. Eng. 47, 463–474.

Chung, M.E., Yeh, I.H., Sung, L.Y., Wu, M.Y., Chao, Y.P., Ng, I.S., Hu, Y.C., 2017. Enhanced integration of large DNA into *E. coli* chromosome by CRISPR/Cas9. Biotechnol. Bioeng. 114 (1), 172–183.

Cimermancic, P., Medema, M.H., Claesen, J., Kurita, K., Brown, L.C.W., Mavrommatis, K., Pati, A., Godfrey, P.A., Koehrsen, M., Clardy, J., Birren, B.W., Takano, E., Sali, A., Linington, R.G., Fischbach, M.A., 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 158 (2), 412–421.

Cobb, R.E., Wang, Y.J., Zhao, H.M., 2015. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. ACS Synth. Biol. 4 (6), 723–728.

Court, D.L., Sawitzke, J.A., Thomason, L.C., 2002. Genetic engineering using homologous recombination. Annu. Rev. Genet. 36, 361–388.

Cromie, G.A., Connelly, J.C., Leach, D.R.F., 2001. Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. Mol. Cell 8 (6), 1163–1174.

Cui, Z., Jiang, X., Zheng, H., Qi, Q., Hou, J., 2019. Homology-independent genome integration enables rapid library construction for enzyme expression and pathway optimization in *Yarrowia lipolytica*. Biotechnol. Bioeng. 116 (2), 354–363.

Da Silva, N.A., Srikrishnan, S., 2012. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 12 (2), 197–214.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97 (12), 6640–6645.

Deng, Y., Zhang, X., Zhang, X.J., 2017. Recent advances in genetic modification systems for actinobacteria. Appl. Microbiol. Biotechnol. 101 (6), 2217–2226.

Domingues, S., Harms, K., Fricke, W.F., Johnsen, P.J., da Silva, G.J., Nielsen, K.M., 2012. Natural transformation facilitates transfer of transposons, integrons and gene cassettes between bacterial species. PLoS Pathog. 8 (8), e1002837.

Dong, H.J., Zhao, C.H., Zhang, T.R., Zhu, H.W., Lin, Z., Tao, W.W., Zhang, Y.P., Li, Y., 2017. A systematically chromosomally engineered *Escherichia coli* efficiently produces butanol. Metab. Eng. 44, 284–292.

Donohoue, P.D., Barrangou, R., May, A.P., 2018. Advances in industrial biotechnology using CRISPR-Cas systems. Trends Biotechnol. 36 (2), 134–146.

Doroghazi, J.R., Albright, J.C., Goering, A.W., Ju, K.S., Haines, R.R., Tchalukov, K.A., Labeda, D.P., Kelleher, N.L., Metcalf, W.W., 2014. A roadmap for natural product discovery based on large-scale genomics and metabolomics. Nat. Chem. Biol. 10 (11), 963–968.

Elliott, K.T., Cuff, L.E., Neidle, E.L., 2013. Copy number change: evolving views on gene amplification. Future Microbiol. 8 (7), 887–899.

Fogg, P.C.M., Colloms, S., Rosser, S., Stark, M., Smith, M.C.M., 2014. New applications for phage integrases. J. Mol. Biol. 426 (15), 2703–2716.

Fogg, P.C.M., Haley, J.A., Stark, W.M., Smith, M.C.M., 2017. Genome integration and excision by a new *Streptomyces* bacteriophage, φJoe. Appl. Environ. Microbiol. 83 (5) (e02767–16).

Fu, J., Wenzel, S.C., Perlova, O., Wang, J.P., Gross, F., Tang, Z.R., Yin, Y.L., Stewart, A.F., Muller, R., Zhang, Y.M., 2008. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. Nucleic Acids Res. 36 (17), e113.

Gaj, T., Gersbach, C.A., Barbas, C.F., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31 (7), 397–405.

Gao, D., Smith, S., Spagnuolo, M., Rodriguez, G., Blenner, M., 2018. Dual CRISPR-Cas9 cleavage mediated gene excision and targeted integration in *Yarrowia lipolytica*. Biotechnol. J. 13 (9), e1700590.

Gietz, R.D., Woods, R.A., 2001. Genetic transformation of yeast. BioTech. 30 (4), 816–820.

Gresham, D., Dunham, M.J., Botstein, D., 2008. Comparing whole genomes using DNA microarrays. Nat. Rev. Genet. 9 (4), 291–302.

Grindley, N.D.F., Whiteson, K.L., Rice, P.A., 2006. Mechanisms of site-specific recombination. Annu. Rev. Biochem. 75, 567–605.

Gu, P.F., Yang, F., Su, T.Y., Wang, Q., Liang, Q.F., Qi, Q.S., 2015. A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. Sci. Rep. 5, 9684.

Haginaka, K., Asamizu, S., Ozaki, T., Igarashi, Y., Furumai, T., Onaka, H., 2014. Genetic approaches to generate hyper-producing strains of goadsporin: the relationships between productivity and gene duplication in secondary metabolite biosynthesis. Biosci. Biotechnol. Biochem. 78 (3), 394–399.

Hammer, S.K., Avalos, J.L., 2017. Harnessing yeast organelles for metabolic engineering. Nat. Chem. Biol. 13 (8), 823–832.

Hayes, F., 2003. Transposon-based strategies for microbial functional genomics and proteomics. Annu. Rev. Genet. 37, 3–29.

Hidalgo-Cantabrana, C., O'Flaherty, S., Barrangou, R., 2017. CRISPR-based engineering of next-generation lactic acid bacteria. Curr. Opin. Microbiol. 37, 79–87.

Hille, F., Richter, H., Wong, S.P., Bratovic, M., Ressel, S., Charpentier, E., 2018. The biology of CRISPR-Cas: backward and forward. Cell 172 (6), 1239–1259.

Hong, K.K., Nielsen, J., 2012. Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. Cell. Mol. Life Sci. 69 (16), 2671–2690.

Horwitz, A.A., Walter, J.M., Schubert, M.G., Kung, S.H., Hawkins, K., Platt, D.M., Hernday, A.D., Mahatdejkul-Meadows, T., Szeto, W., Chandran, S.S., Newman, J.D., 2015. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. Cell Syst. 1 (1), 88–96.

Hou, S., Qin, Q., Dai, J., 2018. Wicket: a versatile tool for the integration and optimization of exogenous pathways in *Saccharomyces cerevisiae*. ACS Synth. Biol. 7 (3), 782–788.

Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157 (6), 1262–1278.

Huang, H., Zheng, G.S., Jiang, W.H., Hu, H.F., Lu, Y.H., 2015. One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. Acta Biochim. Biophys. Sin. 47 (4), 231–243.

Huang, H., Chai, C.S., Yang, S., Jiang, W.H., Gu, Y., 2019. Phage serine integrase-mediated genome engineering for efficient expression of chemical biosynthetic pathway in gas-fermenting *Clostridium ljungdahlii*. Metab. Eng. 52, 293–302.

Isaksson, M., Stenberg, J., Dahl, F., Thuresson, A.C., Bondeson, M.L., Nilsson, M., 2007. MLGA-a rapid and cost-efficient assay for gene copy-number analysis. Nucleic Acids

Res. 35 (17), e115.
Jakociunas, T., Rajkumar, A.S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C.B., Skjodt, M.L., Nielsen, A.T., Borodina, I., Jensen, M.K., Keasling, J.D., 2015. CasEMBLR: Cas9-facilitated multiloci genomic integration of *in vivo* assembled DNA parts in *Saccharomyces cerevisiae*. ACS Synth. Biol. 4 (11), 1226–1234.
Jakociunas, T., Jensen, M.K., Keasling, J.D., 2016. CRISPR/Cas9 advances engineering of microbial cell factories. Metab. Eng. 34, 44–59.
Jiang, D.W., Zhu, W., Wang, Y.C., Sun, C., Zhang, K.Q., Yang, J.K., 2013. Molecular tools for functional genomics in filamentous fungi: recent advances and new strategies. Biotechnol. Adv. 31 (8), 1562–1574.
Johnston, C., Martin, B., Fichant, G., Polard, P., Claverys, J.P., 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat. Rev. Microbiol. 12 (3), 181–196.
Juhas, M., Ajioka, J.W., 2016. Lambda Red recombinase-mediated integration of the high molecular weight DNA into the *Escherichia coli* chromosome. Microb. Cell Factories 15 (1), 172.
Juhas, M., Ajioka, J.W., 2017. High molecular weight DNA assembly *in vivo* for synthetic biology applications. Crit. Rev. Biotechnol. 37 (3), 277–286.
Juhas, M., Evans, L.D.B., Frost, J., Davenport, P.W., Yarkoni, O., Fraser, G.M., Ajioka, J.W., 2014. *Escherichia coli* flagellar genes as target sites for integration and expression of genetic circuits. PLoS ONE 9 (10), e111451.
Kawai, S., Hashimoto, W., Murata, K., 2010. Transformation of *Saccharomyces cerevisiae* and other fungi: methods and possible underlying mechanism. Bioeng. Bugs 1 (6), 395–403.
Kieser, T., Bibb, M.J., Butter, M.J., Chater, K.F., Hopwood, D.A., 2000. Practical *Streptomyces* Genetics. The John Innes Foundation, Norwich, United Kingdom.
Kim, H., Kim, J.S., 2014. A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. 15 (5), 321–334.
Kosuri, S., Church, G.M., 2014. Large-scale de novo DNA synthesis: technologies and applications. Nat. Methods 11 (5), 499–507.
Kotopka, B.J., Li, Y., Smolke, C.D., 2018. Synthetic biology strategies toward heterologous phytochemical production. Nat. Prod. Rep. 35 (9), 902–920.
Kuhlman, T.E., Cox, E.C., 2010. Site-specific chromosomal integration of large synthetic constructs. Nucleic Acids Res. 38 (6), e92.
Lampe, D.J., Churchill, M.E., Robertson, H.M., 1996. A purified mariner transposase is sufficient to mediate transposition *in vitro*. EMBO J. 15 (19), 5470–5479.
Lee, F.W., Da Silva, N.A., 1997. Improved efficiency and stability of multiple cloned gene insertions at the delta sequences of *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 48 (3), 339–345.
Lee, S.Y., Kim, H.U., 2015. Systems strategies for developing industrial microbial strains. Nat. Biotechnol. 33 (10), 1061–1072.
Li, Y.H., Zhao, Z.B., Bai, F.W., 2007. High-density cultivation of oleaginous yeast *Rhodosporidium toruloides* Y4 in fed-batch culture. Enzym. Microb. Technol. 41 (3), 312–317.
Li, L., Zhao, Y.W., Ruan, L.J., Yang, S., Ge, M., Jiang, W.H., Lu, Y.H., 2015. A stepwise increase in pristinamycin II biosynthesis by *Streptomyces pristinaespiralis* through combinatorial metabolic engineering. Metab. Eng. 29, 12–25.
Li, L., Jiang, W.H., Lu, Y.H., 2017a. New strategies and approaches for engineering biosynthetic gene clusters of microbial natural products. Biotechnol. Adv. 35 (8), 936–949.
Li, L., Zheng, G.S., Chen, J., Ge, M., Jiang, W.H., Lu, Y.H., 2017b. Multiplexed site-specific genome engineering for overproducing bioactive secondary metabolites in actinomycetes. Metab. Eng. 40, 80–92.
Li, S., Li, Y., Smolke, C.D., 2018a. Strategies for microbial synthesis of high-value phytochemicals. Nat. Chem. 10 (4), 395–404.
Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W., Lu, Y., 2018b. CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. Appl. Environ. Microbiol. 84 (18) (e00827–12).
Li, L., Wei, K., Liu, X., Wu, Y., Zheng, G., Chen, S., Jiang, W., Lu, Y., 2019. aMSGE: advanced multiplex site-specific genome engineering with orthogonal modular recombinases in actinomycetes. Metab. Eng. 52, 153–167.
Lian, J., Mishra, S., Zhao, H., 2018. Recent advances in metabolic engineering of *Saccharomyces cerevisiae*: new tools and their applications. Metab. Eng. 50, 85–108.
Lin, X.P., Wang, Y.N., Zhang, S.F., Zhu, Z.W., Zhou, Y.J.J., Yang, F., Sun, W.Y., Wang, X.Y., Zhao, Z.B.K., 2014. Functional integration of multiple genes into the genome of the oleaginous yeast *Rhodosporidium toruloides*. FEMS Yeast Res. 14 (4), 547–555.
Liu, L., Redden, H., Alper, H.S., 2013. Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. Curr. Opin. Biotechnol. 24 (6), 1023–1030.
Liu, Z.H., Liang, Y.Y., Ang, E.L., Zhao, H.M., 2017. A new era of genome integration-simply cut and paste. ACS Synth. Biol. 6 (4), 601–609.
Lobs, A.K., Schwartz, C., Wheeldon, I., 2017. Genome and metabolic engineering in non-conventional yeasts: current advances and applications. Synth. Syst. Biotechnol. 2 (3), 198–207.
Ma, S.Y., Tang, N., Tian, J.D., 2012. DNA synthesis, assembly and applications in synthetic biology. Curr. Opin. Chem. Biol. 16 (3–4), 260–267.
Manderscheid, N., Bilyk, B., Busche, T., Kalinowski, J., Paululat, T., Bechthold, A., Petzke, L., Luzhetskyy, A., 2016. An influence of the copy number of biosynthetic gene clusters on the production level of antibiotics in a heterologous host. J. Biotechnol. 232, 110–117.
Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J., Daran, J.M., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 15 (2), fov004.
Markham, K.A., Alper, H.S., 2018. Synthetic biology expands the industrial potential of *Yarrowia lipolytica*. Trends Biotechnol. 36 (10), 1085–1095.
Mercer, A.C., Gaj, T., Fuller, R.P., Barbas, C.F., 2012. Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res. 40 (21), 11163–11172.
Mermelstein, L.D., Papoutsakis, E.T., 1993. *In vivo* methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 59 (4), 1077–1081.
Merrick, C., Zhao, J., Rosser, S., 2018. Serine integrases: advancing synthetic biology. ACS Synth. Biol. 7 (2), 299–310.
Murakami, T., Burian, J., Yanai, K., Bibb, M.J., Thompson, C.J., 2011a. A system for the targeted amplification of bacterial gene clusters multiplies antibiotic yield in *Streptomyces coelicolor*. Proc. Natl. Acad. Sci. U. S. A. 108 (38), 16020–16025.
Murakami, T., Sumida, N., Bibb, M., Yanai, K., 2011b. ZouA, a putative relaxase, is essential for DNA amplification in *Streptomyces kanamyceticus*. J. Bacteriol. 193 (8), 1815–1822.
Murphy, K.C., 1998. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J. Bacteriol. 180 (8), 2063–2071.
Muth, G., 2018. The pSG5-based thermosensitive vector family for genome editing and gene expression in actinomycetes. Appl. Microbiol. Biotechnol. 102 (21), 9067–9080.
Myronovskyi, M., Luzhetskyy, A., 2013. Genome engineering in actinomycetes using site-specific recombinases. Appl. Microbiol. Biotechnol. 97 (11), 4701–4712.
Myronovskyi, M., Rosenkranzer, B., Nadmid, S., Pujic, P., Normand, P., Luzhetskyy, A., 2018. Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. Metab. Eng. 49, 316–324.
Nah, H.J., Woo, M.W., Choi, S.S., Kim, E.S., 2015. Precise cloning and tandem integration of large polyketide biosynthetic gene cluster using *Streptomyces* artificial chromosome system. Microb. Cell Factories 14, 140.
Nakade, S., Tsubota, T., Sakane, Y., Kume, S., Sakamoto, N., Obara, M., Daimon, T., Sezutsu, H., Yamamoto, T., Sakuma, T., Suzuki, K.T., 2014. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat. Commun. 5, 5560.
Narula, J., Kuchina, A., Lee, D.Y.D., Fujita, M., Suel, G.M., Igoshin, O.A., 2015. Chromosomal arrangement of phosphorelay genes couples sporulation and DNA replication. Cell 162 (2), 328–337.
Newman, D.J., Cragg, G.M., 2016. Natural products as sources of new drugs from 1981 to 2014. J. Nat. Prod. 79 (3), 629–661.
Nielsen, J., Keasling, J.D., 2016. Engineering cellular metabolism. Cell 164 (6), 1185–1197.
Niu, G., 2017. Genomics-driven natural product discovery in actinomycetes. Trends Biotechnol. 36 (3), 238–241.
Olano, C., Lombo, F., Mendez, C., Salas, J.A., 2008. Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering. Metab. Eng. 10 (5), 281–292.
Oliveira, P.H., Touchon, M., Rocha, E.P., 2014. The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. Nucleic Acids Res. 42 (16), 10618–10631.
Orr-Weaver, T.L., Szostak, J.W., Rothstein, R.J., 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. U. S. A. 78 (10), 6354–6358.
Ou, B.M., Garcia, C., Wang, Y.J., Zhang, W.P., Zhu, G.Q., 2018. Techniques for chromosomal integration and expression optimization in *Escherichia coli*. Biotechnol. Bioeng. 115 (10), 2467–2478.
Park, Y.K., Nicaud, J.M., Ledesma-Amaro, R., 2018. The engineering potential of *Rhodosporidium toruloides* as a workhorse for biotechnological applications. Trends Biotechnol. 36 (3), 304–317.
Peschke, U., Schmidt, H., Zhang, H.Z., Piepersberg, W., 1995. Molecular characterization of the lincomycin-production gene-cluster of *Streptomyces lincolnensis*-78-11. Mol. Microbiol. 16 (6), 1137–1156.
Pitcher, R.S., Brissett, N.C., Doherty, A., 2007. Nonhomologous end joining in bacteria: a microbial perspective. Annu. Rev. Microbiol. 61, 259–282.
Pontrelli, S., Chiu, T.Y., Lan, E.I., Chen, F.Y., Chang, P., Liao, J.C., 2018. *Escherichia coli* as a host for metabolic engineering. Metab. Eng. 50, 16–46.
Qin, X.L., Li, R.J., Luo, X., Lin, Y.M., Feng, J.X., 2017. Deletion of *ligD* significantly improves gene targeting frequency in the lignocellulolytic filamentous fungus *Penicillium oxalicum*. Fungal Biol. 121 (6–7), 615–623.
Radecka, D., Mukherjee, V., Mateo, R.Q., Stojiljkovic, M., Foulquie-Moreno, M.R., Thevelein, J.M., 2015. Looking beyond *Saccharomyces*: the potential of non-conventional yeast species for desirable traits in bioethanol fermentation. FEMS Yeast Res. 15 (6), fov053.
Raschmanova, H., Weninger, A., Glieder, A., Kovar, K., Vogl, T., 2018. Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: Current state and future prospects. Biotechnol. Adv. 36 (3), 641–665.
Reams, A.B., Roth, J.R., 2015. Mechanisms of gene duplication and amplification. Cold Spring Harb. Perspect. Biol. 7 (2), a016592.
Ronda, C., Maury, J., Jakociunas, T., Jacobsen, S.A., Germann, S.M., Harrison, S.J., Borodina, I., Keasling, J.D., Jensen, M.K., Nielsen, A.T., 2015. CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. Microb. Cell Factories 14, 97.
Rutledge, P.J., Challis, G.L., 2015. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat. Rev. Microbiol. 13 (8), 509–523.
Sabri, S., Steen, J.A., Bongers, M., Nielsen, L.K., Vickers, C.E., 2013. Knock-in/Knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the *Escherichia coli* chromosome at well-characterised loci. Microb. Cell Factories 12, 60.
Santos, C.N.S., Regitsky, D.D., Yoshikuni, Y., 2013. Implementation of stable and complex biological systems through recombinase-assisted genome engineering. Nat. Commun. 4, 2503.
Sauer, C., Syvertsson, S., Bohorquez, L.C., Cruz, R., Harwood, C.R., van Rij, T., Hamoen,

L.W., 2016. Effect of genome position on heterologous gene expression in *Bacillus subtilis*: an unbiased analysis. ACS Synth. Biol. 5 (9), 942–947.
Scherer, S., Davis, R.W., 1979. Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. Proc. Natl. Acad. Sci. U. S. A. 76 (10), 4951–4955.
Schwartz, C., Frogue, K., Ramesh, A., Misa, J., Wheeldon, I., 2017a. CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*. Biotechnol. Bioeng. 114 (12), 2896–2906.
Schwartz, C., Shabbir-Hussain, M., Frogue, K., Blenner, M., Wheeldon, I., 2017b. Standardized markerless gene integration for pathway engineering in *Yarrowia lipolytica*. ACS Synth. Biol. 6 (3), 402–409.
Sfeir, A., Symington, L.S., 2015. Microhomology-mediated end loining: a back-up survival mechanism or dedicated pathway? Trends Biochem. Sci. 40 (11), 701–714.
Shao, Z.Y., Zhao, H., Zhao, H.M., 2009. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. Nucleic Acids Res. 37 (2), e16.
Shendure, J., Balasubramanian, S., Church, G.M., Gilbert, W., Rogers, J., Schloss, J.A., Waterston, R.H., 2017. DNA sequencing at 40: past, present and future. Nature 550 (7676), 345–353.
Shi, S.B., Valle-Rodriguez, J.O., Siewers, V., Nielsen, J., 2014. Engineering of chromosomal wax ester synthase integrated *Saccharomyces cerevisiae* mutants for improved biosynthesis of fatty acid ethyl esters. Biotechnol. Bioeng. 111 (9), 1740–1747.
Shi, S.B., Liang, Y.Y., Zhang, M.Z.M., Ang, E.L., Zhao, H.M., 2016. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. Metab. Eng. 33, 19–27.
Shi, T.Q., Liu, G.N., Ji, R.Y., Shi, K., Song, P., Ren, L.J., Huang, H., Ji, X.J., 2017. CRISPR/Cas9-based genome editing of the filamentous fungi: the state of the art. Appl. Microbiol. Biotechnol. 101 (20), 7435–7443.
Shuman, S., Glickman, M.S., 2007. Bacterial DNA repair by non-homologous end joining. Nat. Rev. Microbiol. 5 (11), 852–861.
Slager, J., Veening, J.W., 2016. Hard-wired control of bacterial processes by chromosomal gene location. Trends Microbiol. 24 (10), 788–800.
Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., Kucherlapati, R.S., 1985. Insertion of DNA-sequences into the human chromosomal beta-globin locus by homologous recombination. Nature 317 (6034), 230–234.
Stark, W.M., 2017. Making serine integrases work for us. Curr. Opin. Microbiol. 38, 130–136.
St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B., Shearwin, K.E., 2013. One-step cloning and chromosomal integration of DNA. ACS Synth. Biol. 2 (9), 537–541.
Sun, W.Y., Yang, X.B., Wang, X.Y., Lin, X.P., Wang, Y.A., Zhang, S.F., Luan, Y.S., Zhao, Z.B.K., 2017. Homologous gene targeting of a carotenoids biosynthetic gene in *Rhodosporidium toruloides* by *Agrobacterium*-mediated transformation. Biotechnol. Lett. 39 (7), 1001–1007.
Sung, P., Klein, H., 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat. Rev. Mol. Cell Biol. 7 (10), 739–750.
Tao, W., Yang, A., Deng, Z., Sun, Y., 2018. CRISPR/Cas9-based editing of *Streptomyces* for discovery, characterization, and production of natural products. Front. Microbiol. 9, 1660.
Thomas, K.R., Folger, K.R., Capecchi, M.R., 1986. High-frequency targeting of genes to specific sites in the mammalian genome. Cell 44 (3), 419–428.
Thompson, A., Gasson, M.J., 2001. Location effects of a reporter gene on expression levels and on native protein synthesis in *Lactococcus lactis* and *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 67 (8), 3434–3439.
Tong, Y., Weber, T., Lee, S.Y., 2018. CRISPR/Cas-based genome engineering in natural product discovery. Nat. Prod. Rep. https://doi.org/10.1039/c8np00089a.
Tyo, K.E.J., Ajikumar, P.K., Stephanopoulos, G., 2009. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. Nat. Biotechnol. 27 (8) (760-U115).
Vasu, K., Nagaraja, V., 2013. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiol. Mol. Biol. Rev. 77 (1), 53–72.
Verbeke, J., Beopoulos, A., Nicaud, J.M., 2013. Efficient homologous recombination with short length flanking fragments in *Ku70* deficient *Yarrowia lipolytica* strains. Biotechnol. Lett. 35 (4), 571–576.
Wagner, J.M., Alper, H.S., 2016. Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances. Fungal Genet. Biol. 89, 126–136.
Wang, L.Y., Deng, A.H., Zhang, Y., Liu, S.W., Liang, Y., Bai, H., Cui, D., Qiu, Q.D., Shang, X.L., Yang, Z., He, X.P., Wen, T.Y., 2018a. Efficient CRISPR-Cas9 mediated multiplex genome editing in yeasts. Biotechnol. Biofuels 11, 277.
Wang, B., Hu, Q.T., Zhang, Y., Shi, R.L., Chai, X., Liu, Z., Shang, X.L., Zhang, Y., Wen, T.Y., 2018b. A RecET-assisted CRISPR-Cas9 genome editing in *Corynebacterium glutamicum*. Microb. Cell Factories 17 (1), 63.
Weber, T., Charusanti, P., Musiol-Kroll, E.M., Jiang, X.L., Tong, Y.J., Kim, H.U., Lee, S.Y., 2015. Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. Trends Biotechnol. 33 (1), 15–26.
Weninger, A., Fischer, J.E., Raschmanova, H., Kniely, C., Vogl, T., Glieder, A., 2018. Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. J. Cell. Biochem. 119 (4), 3183–3198.
Wu, S.G., Zhao, X., Shen, H.W., Wang, Q.A., Zhao, Z.B.K., 2011. Microbial lipid production by *Rhodosporidium toruloides* under sulfate-limited conditions. Bioresour. Technol. 102 (2), 1803–1807.
Wu, X.L., Li, B.Z., Zhang, W.Z., Song, K., Qi, H., Dai, J.B., Yuan, Y.J., 2017. Genome-wide landscape of position effects on heterogeneous gene expression in *Saccharomyces cerevisiae*. Biotechnol. Biofuels 10, 189.
Xiao, Y., Bowen, C.H., Liu, D., Zhang, F.Z., 2016. Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. Nat. Chem. Biol. 12 (5), 339–344.
Xue, C., Zhao, J.B., Chen, L.J., Yang, S.T., Bai, F.W., 2017. Recent advances and state-of-the-art strategies in strain and process engineering for biobutanol production by *Clostridium acetobutylicum*. Biotechnol. Adv. 35 (2), 310–322.
Yamada, R., Taniguchi, N., Tanaka, T., Ogino, C., Fukuda, H., Kondo, A., 2010. Cocktail delta-integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains. Microb. Cell Factories 9, 32.
Yamamoto, Y., Gerbi, S.A., 2018. Making ends meet: targeted integration of DNA fragments by genome editing. Chromosoma 127 (4), 405–420.
Yanai, K., Murakami, T., Bibb, M., 2006. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. Proc. Natl. Acad. Sci. U. S. A. 103 (25), 9661–9666.
Yang, L., Nielsen, A.A.K., Fernandez-Rodriguez, J., McClune, C.J., Laub, M.T., Lu, T.K., Voigt, C.A., 2014a. Permanent genetic memory with > 1-byte capacity. Nat. Methods 11 (12), 1261–1266.
Yang, J.J., Sun, B.B., Huang, H., Jiang, Y., Diao, L.Y., Chen, B., Xu, C.M., Wang, X., Liu, J.L., Jiang, W.H., Yang, S., 2014b. High-efficiency scarless genetic modification in *Escherichia coli* by using lambda red recombination and I-SceI cleavage. Appl. Environ. Microbiol. 80 (13), 3826–3834.
Yang, X., Xu, M., Yang, S.T., 2016. Restriction modification system analysis and development of *in vivo* methylation for the transformation of *Clostridium cellulovorans*. Appl. Microbiol. Biotechnol. 100 (5), 2289–2299.
Yao, X., Wang, X., Hu, X.D., Liu, Z., Liu, J.L., Zhou, H.B., Shen, X.W., Wei, Y., Huang, Z.J., Ying, W.Q., Wang, Y., Nie, Y.H., Zhang, C.C., Li, S.L., Cheng, L.P., Wang, Q.F., Wu, Y., Huang, P.Y., Sun, Q., Shi, L.Y., Yang, H., 2017. Homology-mediated end joining-based targeted integration using CRISPR/Cas9. Cell Res. 27 (6), 801–814.
Yao, R.L., Liu, D., Jia, X., Zheng, Y., Liu, W., Xiao, Y., 2018. CRISPR-Cas9/Cas12a biotechnology and application in bacteria. Synth. Syst. Biotechnol. 3 (3), 135–149.
Yin, J., Wang, H., Fu, X.Z., Gao, X., Wu, Q., Chen, G.Q., 2015. Effects of chromosomal gene copy number and locations on polyhydroxyalkanoate synthesis by *Escherichia coli* and *Halomonas* sp. Appl. Microbiol. Biotechnol. 99 (13), 5523–5534.
Yoshida, N., Sato, M., 2009. Plasmid uptake by bacteria: a comparison of methods and efficiencies. Appl. Microbiol. Biotechnol. 83 (5), 791–798.
Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., Court, D.L., 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 97 (11), 5978–5983.
Zhang, Z., MooYoung, M., Chisti, Y., 1996. Plasmid stability in recombinant *Saccharomyces cerevisiae*. Biotechnol. Adv. 14 (4), 401–435.
Zhang, G., Wang, W., Deng, A., Sun, Z., Zhang, Y., Liang, Y., Che, Y., Wen, T., 2012. A mimicking-of-DNA-methylation-patterns pipeline for overcoming the restriction barrier of bacteria. PLoS Genet. 8 (9), e1002987.
Zhang, C., Meng, X.H., Wei, X.L., Lu, L., 2016a. Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in *Aspergillus fumigatus*. Fungal Genet. Biol. 86, 47–57.
Zhang, M.M., Wang, Y.J., Ang, E.L., Zhao, H.M., 2016b. Engineering microbial hosts for production of bacterial natural products. Nat. Prod. Rep. 33 (8), 963–987.
Zhou, T.C., Kim, B.G., Zhong, J.J., 2014. Enhanced production of validamycin A in *Streptomyces hygroscopicus* 5008 by engineering validamycin biosynthetic gene cluster. Appl. Microbiol. Biotechnol. 98 (18), 7911–7922.