Metabolic Engineering Communications 10 (2020) e00108

Contents lists available at ScienceDirect

Metabolic Engineering Communications

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

Emerging molecular biology tools and strategies for engineering natural product biosynthesis

Wei Xu [a,1], Evaldas Klumbys [a,1], Ee Lui Ang [a,**], Huimin Zhao [a,b,*]

[a] Institute of Chemical and Engineering Sciences, Agency for Science, Technology, and Research, Singapore
[b] Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA

ARTICLE INFO

Keywords:
Biosynthesis
Synthetic biology
Natural products
Enzyme engineering
Pathway engineering
Metabolic engineering

ABSTRACT

Natural products and their related derivatives play a significant role in drug discovery and have been the inspiration for the design of numerous synthetic bioactive compounds. With recent advances in molecular biology, numerous engineering tools and strategies were established to accelerate natural product synthesis in both academic and industrial settings. However, many obstacles in natural product biosynthesis still exist. For example, the native pathways are not appropriate for research or production; the key enzymes do not have enough activity; the native hosts are not suitable for high-level production. Emerging molecular biology tools and strategies have been developed to not only improve natural product titers but also generate novel bioactive compounds. In this review, we will discuss these emerging molecular biology tools and strategies at three main levels: enzyme level, pathway level, and genome level, and highlight their applications in natural product discovery and development.

1. Introduction

Natural products (NPs) have proven to be a rich source of bioactive compounds and drugs for thousands of years. Although numerous NPs, especially microbial NPs, have been isolated, identified, and engineered during the Golden Age of NPs in the 1950s, bioactivity guided NPs discovery strategies remain time consuming and labor intensive (Donadio et al., 2007; Newman and Cragg, 2012). With the rapid development of next-generation sequencing technologies, an increasing number of microbial genomes have been elucidated, enabling a new era of bioinformatics-guided NPs discovery. However, the number of biosynthetic gene clusters (BGCs) identified *in silico* far exceeds the number of natural products identified so far and most BGCs are silent or not expressed in native hosts under standard laboratory conditions (Banik and Brady, 2010; Katz et al., 2016; Ren et al., 2017b; M. M. Zhang et al., 2017a; M. M. Zhang et al., 2016). Moreover, the laboratory cultured microorganisms represent only a small part of the overall microbial populations in nature. Overall, the rate of discovering novel bioactive natural products has slowed down drastically after the Golden Age, necessitating the development of new molecular biology tools (Galm and Shen, 2006; Ren et al., 2017b; M. M. Zhang et al., 2017a).

The traditional paradigm for natural product discovery is typically bioactivity-guided whereas characterization of the corresponding BGC is usually carried out without whole genome sequence information. Consequently, known NPs are kept being re-discovered (Penesyan et al., 2010). In comparison, the modern paradigm for natural product discovery is based on genome sequencing, bioinformatics, and synthetic biology, and focuses on the direct identification and cloning of target BGCs (Ren et al., 2017b; M. M. Zhang et al., 2017a; M. M. Zhang et al., 2016), thereby avoiding the re-discovery of same NPs. Compared to random cloning using library-based approaches (Nah et al., 2017) or *in situ* manipulation (Tao et al., 2018), direct cloning is a much faster and more rational approach for microbial NP discovery. Moreover, a wide variety of enzyme engineering, pathway engineering, and genome engineering tools are used to either improve the production of target NPs or generate new analogs of target NPs (Fig. 1). In this review, we will discuss new molecular biology tools and strategies for natural product discovery and engineering, with a focus on recent advances (mainly between 2015 and 2019).

* Corresponding author. Institute of Chemical and Engineering Sciences, Agency for Science, Technology, and Research, Singapore.
** Corresponding author.
E-mail addresses: ang_ee_lui@ices.a-star.edu.sg (E.L. Ang), zhao5@illinois.edu (H. Zhao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.mec.2019.e00108
Received 18 September 2019; Received in revised form 5 November 2019; Accepted 5 November 2019

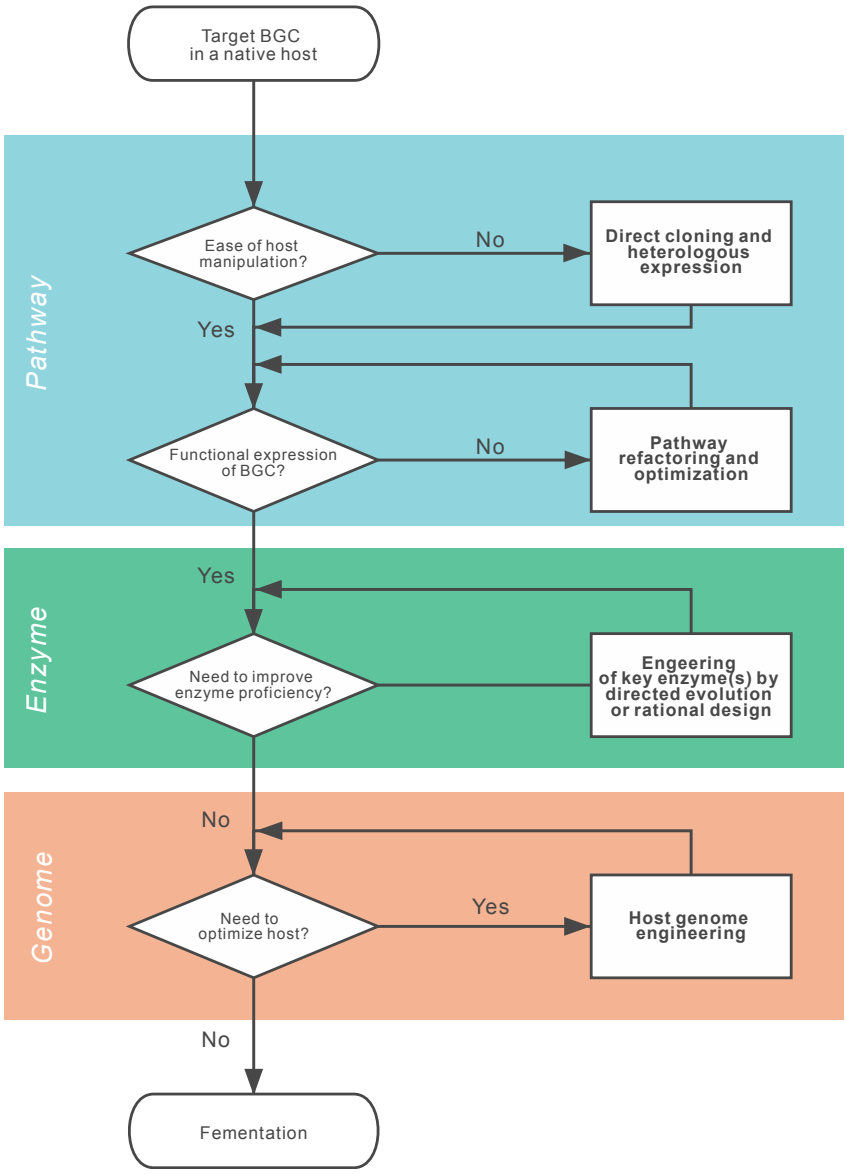

Fig. 1. Flowchart of engineering microbial natural product biosynthesis. Briefly, natural product discovery and engineering can be divided into three levels. (1) At the pathway level. A target BGC can be cloned and expressed in a heterologous host if the native host is not suitable for genetic engineering and the biosynthetic pathway can be refactored and optimized for its expression. (2) At the enzyme level. Key enzyme(s) is optimized via directed evolution or rational design if its proficiency needs to be improved. (3) At the genome level. The host genome can be edited if an optimized production host is needed.

2. Molecular biology tools for pathway engineering of BGCs

2.1. Direct cloning

Since the late 1980s, numerous natural product BGCs, especially relatively small BGCs, were cloned and expressed in heterologous hosts. The most widely used strategy was random cloning, in which genomic DNA is randomly sheared and cloned into vectors to create libraries for heterologous expression (Nah et al., 2017). However, this strategy is labor-intensive and time consuming. To overcome its limitations, a number of direct cloning strategies have been developed in recent years, in which a target natural product BGC is identified by genome sequencing and bioinformatics tools and then directly cloned for heterologous expression (Fig. 2).

For example, Moore and coworkers took advantage of the natural *in vivo* homologous recombination mechanism in *Saccharomyces cerevisiae* to directly capture a BGC of interest from genomic DNA (Yamanaka et al., 2014). The *S. cerevisiae/Escherichia coli* shuttle vector pCAP01 contained two homology arms with the target BGC and was co-transformed into *S. cerevisiae* with restriction enzyme (RE) digested genomic DNA harboring the target BGC to yield a large plasmid via homologous recombination. Using this Transformation-associated recombination (TAR) cloning method, a 67 kb nonribosomal peptide synthetase BGC which encodes the taromycin backbone from the marine actinomycete *Saccharomonospora* sp. CNQ-490 was successfully cloned and expressed in model actinomycete expression host *Streptomyces coelicolor*. In a follow-up study, a URA3 gene was inserted into pCAP01 under ADH1 promoter as a counter selectable marker so as to achieve high efficiency recombination with shorter capture arms and minimize non-homologous end joining (NHEJ). With the optimized vector pCAP03, two thiotetronic acids BGCs were directly cloned and heterologously expressed in a *Streptomyces* model host (Tang et al., 2015). Besides Gram-positive actinomycete natural product BGCs, the Moore group also expanded this TAR cloning method into Gram-negative species with the pCAP01 vector (Ross et al., 2015). The resulting plasmid pCAP05 combines yeast cloning elements with Gram-negative elements for heterologous expression. It was successfully used to capture and express the violacein BGC from *Pseudoalteromonas luteoviolacea* 2ta16 in two proteobacterial hosts, *Pseudomonas putida* KT2440 and *Agrobacterium tumefaciens* LBA4404 with robust production (J. J. Zhang et al., 2017b).

In addition to using yeast's native homologous recombination system, Leadlay and coworkers utilized Gibson assembly to recover digested

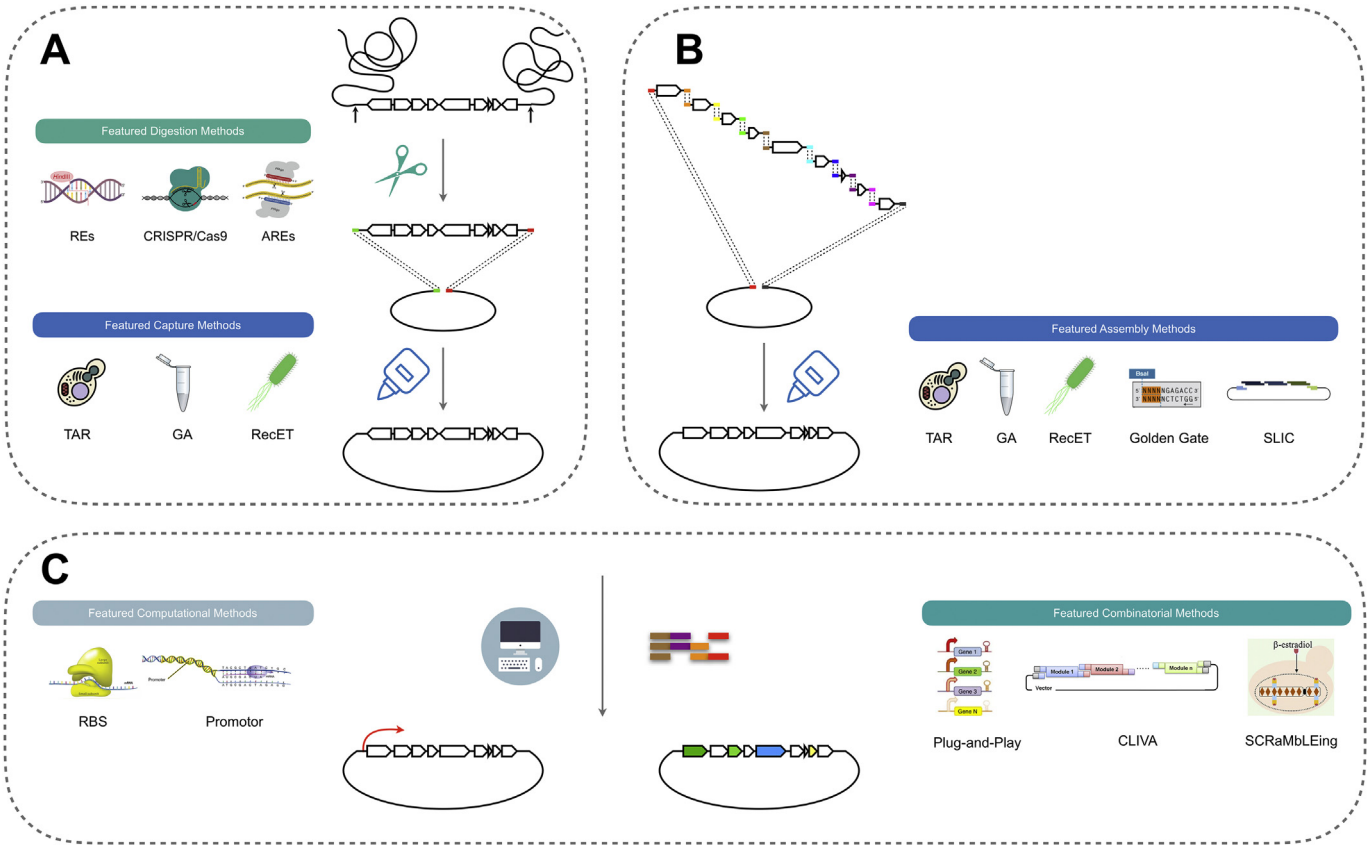

Fig. 2. Scheme for various molecular biology tools for pathway engineering. (A) Direct cloning: a target BGC is identified by bioinformatics and the boundary of the BGC (indicated by small arrows) is confirmed. Various nucleases such as restriction enzymes (RE), CRISPR/Cas9, and the PfAgo-based artificial restriction enzymes (AREs) can be used to excise the target BGC from genomic DNA. A capture vector recovers the target BGC by various methods such as transformation associated recombination (TAR), Gibson assembly (GA), and RecET recombination. (B) Pathway refactoring: every biological element can be amplified and tagged with two homologous ends. All the elements and plasmid backbone can be ligated together with different methods such as TAR, GA, RecET recombination, Golden Gate assembly, and Sequence and Ligation Independent Cloning (SLIC). (C) Pathway optimization: the captured or assembled BGC can be optimized with computational or combinatorial methods. The computational methods can optimize RBSs and promoters of target genes while the combinatorial methods can change some key enzymes to achieve high NP production yield or novel chemical structure.

genomic DNA fragments. In this approach, the genomic DNA was digested by two REs and the desired 41 kb BGC fragment was cloned with Gibson assembly (Zhou et al., 2015). Similar to the TAR cloning method, this method also required the use of REs to digest the high-quality genomic DNA and the removal of the small DNA fragments (less than 20 kb) to improve the cloning efficiency.

A major limitation of the RE-based cloning methods is that the RE recognition sequences have to be avoided within the target BGCs. Thus, selection of appropriate REs can be difficult or even impossible, especially for large size BGCs such as Type I polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS) BGCs. Programmable nucleases, exemplified by the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein (Cas) endonuclease system (Deltcheva et al., 2011; Hsu et al., 2014; Jinek et al., 2012), may offer an alternative to circumvent this constraint. In 2015, Lockey and coworkers linearized a plasmid using CRISPR/Cas9 coupled with a programmed sgRNA and seamlessly inserted a DNA fragment into the plasmid using Gibson assembly (Wang et al., 2015). In the same year, this same strategy was extended to genomic DNA (Jiang et al., 2015). The CRISPR/Cas9 endonuclease system was first used *in vitro* to digest the high-quality genomic DNA in an agarose matrix. Due to the special digestion method which minimizes mechanical shearing and the specific sgRNA programmed restriction sites, up to 100 kb long bacterial genomic DNA could be cut from the whole genome and completely captured by Gibson assembly with a PCR amplified cloning vector. Similar to traditional RE systems, the CRISPR/Cas9 endonuclease system has also been incorporated with TAR cloning to capture the desired DNA fragment or BGC. Kouprina and coworkers reported that double strands breaks near the target recombination region instead of random breaks dramatically increased the capture efficiency from 0.5% to up to 32% (Lee et al., 2015). In the same year, Keasling and coworkers utilized CRISPR/Cas9-mediated double-strand breaks to assemble tens of DNA parts into *S. cerevisiae* genome with high efficiency (Jakočiunas et al., 2015). This method, named CasEMBLR, was used to assemble fifteen DNA fragments into a carotenoid pathway, which was integrated into three targeted loci. Another class of programmable nucleases is the artificial restriction enzymes based on a hyperthermophilic *Pyrococcus furiosus* Argonaute (PfAgo) (Enghiad and Zhao, 2017). The PfAgo system is similar to the CRISPR/Cas system but uses a short DNA guide instead of an sgRNA guide to locate the target cleavage site. Compared to the CRISPR/Cas system, the PfAgo system is more flexible (it can target virtually any sequence) and more active (PfAgo is a multiple turnover enzyme while CRISPR/Cas is a single turnover enzyme). However, this system has not been demonstrated to clone large BGCs yet.

Another strategy to overcome the main limitation of the RE-based cloning methods is to use the linear plus linear homologous recombination (LLHR) mechanism mediated by the prophage recombinase RecET (Wang et al., 2016a). The full-length RecE and RecT could mediate highly efficient LLHR for cloning of the target BGCs from genomic DNA. The capture backbone harbors standardized cassettes for horizontal gene transfer and different replicator for selection in different hosts. Subsequently, an improved RecET method (Wang et al., 2018), termed

Exonuclease Combined with RecET recombination (ExoCET), was used to directly clone a BGC of >50 kb from bacterial genomic DNA with high precision. After digesting the genomic DNA with REs or CRISPR/Cas endonucleases, the ExoCET system could capture the 106 kb salinomycin BGC.

2.2. Pathway refactoring

As regulatory pathways may differ from organism to organism, the productivity of a BGC can significantly decrease when introduced into a heterologous host. In such cases, refactoring the BGCs through decoupling and rewiring of the native regulatory system may be required to achieve initial production (Tan and Liu, 2017). In addition, during refactoring design, BGCs can be dissected into modular units for easy and efficient swap of any required genetic parts (Smanski et al., 2016). The complex nature of NP biosynthesis usually involves large gene clusters, necessitating the development of high-throughput refactoring systems for simultaneous adjustment of multiple genetic elements. To aid in BGC refactoring, computational tools have been developed to help predict genetic element behaviour (Brewster et al., 2012; Salis et al., 2009). Novel molecular biology tools and their applications in natural product biosynthesis will be discussed below.

2.2.1. Homologous recombination based refactoring tools

The homologous recombination mechanism in *S. cerevisiae* was explored to refactor the BGCs of rebeccamycin, tetarimycin and lazarimides A and B (Montiel et al., 2015), and bottromycin-related metabolites (Eyles et al., 2018). In the first study, yeast homologous recombination was used together with an auxotrophic complementation-based yeast selection system to replace native promoters with orthogonal actinomycetes constitutive promoters and ribosomal binding sites (RBSs). Each bidirectional promoter exchange cassette was amplified with primers containing 40 bp homologous sequences for recombination. The target BGC was co-transformed with PCR-amplified cassette and plated in cassette-specific media for selection. However, the decrease in homologous recombination efficiency with increasing number of desired genetic replacement, and the need to couple selectable markers to the promoter cassettes limit the types of modification achievable by this method. In the second study, TAR-based cloning was employed together with yeast-mediated assembly of DNA fragments to refactor and remodel the bottromycin BGC from *Streptomyces scabies* (Eyles et al., 2018). The bottromycin BGC was first captured in yeast/*E. coli* shuttle vector pCAP01 using TAR. Subsequently, yeast-mediated recombination was used to assemble combinations of DNA fragments created by REs targeting naturally occurring restriction sites in the BGC, PCR products, and single stranded oligonucleotides to create remodelled BGCs with refactored promoters, gene deletions and targeted mutations in a single step. This method is not only highly efficient, not hindered by high GC content, marker free, and inexpensive, but also enables flexible modifications to the BGCs.

Due to its programmability and specificity, the CRISPR/Cas9 system has opened new avenues for BGC refactoring. Yeast-based promoter engineering platform named multiplexed-CRISPR-TAR (mCRISTAR) was developed to enable single-marker multiplexed promoter engineering (Kang et al., 2016). This approach involves fragmentation of a BGC of interest using a CRISPR/Cas9 system targeting the native promoter sequences, followed by TAR mediated reassembly to incorporate synthetic promoters. Similarly, multiplex *in vitro* Cas9-TAR (*mi*CASTAR) was engineered for activation and refactoring of the atolypenes BGC (Kim et al., 2019). Compared to mCRISTAR, *mi*CASTAR does not require construction and transformation of unique CRISPR/Cas9 plasmids into yeast for each refactoring experiment.

A strategy similar to the TAR method was developed in *E. coli*. This method is based on the Red/ET homologous recombination system, involving a λ phage-derived protein pair, Redα/Redβ or RecE/RecT, and 50 bp homology arms (Zhang et al., 2000). It has been widely applied for genetic manipulation purposes and is gaining attention as a refactoring platform (Horbal et al., 2018). For example, Red/ET was used to refactor the bottromycin BGC to increase the production titer by up to 50 fold by simply decoupling the native regulatory system with artificial promoters (Horbal et al., 2018).

2.2.2. Non-homologous recombination based refactoring

Discovery and application of rapid multi-fragment cloning methods is essential for efficient refactoring process. Golden Gate cloning was initially described in 2008 (Engler et al., 2008), and it has since been successfully applied to the refactoring and activation of various BGCs, including the discovery of the BGC for the production of phosphonoacetic acid derivatives (Freestone et al., 2017). In this case, three successive Golden Gate assemblies were required for complete refactoring of the whole BGC from *Streptomyces* sp. strain NRRL F-525 to the *Streptomyces lividans* expression platform. An improved two-tiered Golden Gate assembly employing a plug-and-play approach was used to refactor 96 combinatorial carotenoid pathways, simplifying the aforementioned method (Ren et al., 2017a). In the first tier of this approach, genes from a BGC were first individually cloned into helper plasmids preassembled with promoters and terminators to form separate expression cassettes. In the second tier, the expression cassettes were then assembled using Golden Gate assembly to yield refactored BGCs. In addition, the use of spacer plasmids to help bridge sequence gaps in the constructs increases the flexibility of this approach, and facilitates gene deletion and replacement studies.

Golden Gate assembly-based refactoring approach requires the removal of unwanted Type II RE recognition sites, which can be tedious. Therefore, RE independent cloning and refactoring offers promising alternatives. ExoCET was used for simultaneous assembly of 12 PCR-amplified products with refactoring of natural spinosad BGC (Song et al., 2019). The efficiency of this cloning method was much higher than Gibson with RecET assembly, whereas Gibson cloning on its own did not yield any successful clones. Another strategy is to combine *in vitro* and *in vivo* tools for successful assembly of BGCs. For example, sequence- and ligation-independent cloning (SLIC) harnessing a bacteriophage T4 DNA polymerase for *in vitro* DNA hybridization and *E. coli* DNA repair system can be used for assembling refactored BGCs (Jeong et al., 2012; Li and Elledge, 2012). It was applied for direct pathway cloning (DiPaC) (Greunke et al., 2018) using long-amplification PCR for refactoring of the cyanobacterial hapalosin BGC (D'Agostino and Gulder, 2018). This method is superior to the HiFi DNA assembly for cloning, assembly and simultaneous promoter and terminator refactoring of small- and mid-sized BGCs.

2.3. Pathway optimization

Cloning and refactoring of target BGCs is essential for the initial production of NPs. However, further optimization of the refactored pathways of interest is almost always necessary in order to obtain high enough yields of the products. For example, expression of BGC enzymes in a heterologous host can negatively affect the host by impairing its growth, disrupting its central metabolism thus reducing yields of desired compounds, or even killing the host (Keasling, 2010). Furthermore, most enzymes in BGCs have not been extensively studied, and usually there is no solubility, stability and activity data available. Therefore, enzymes and pathways may have to be engineered in concert with the host in order to obtain a highly productive functional pathway (Nielsen and Keasling, 2016). The development of standardized parts in synthetic biology has allowed researchers to dissect natural product pathways into genetic elements, such as promoters, RBSs, terminators, and genes of interest, and assemble the elements into redesigned pathways for screening in the design-build-test-learn cycle, while metabolic engineering of the host strain is required to ensure flux balance and host viability. In this section, new combinatorial and computational tools for pathway and host optimization will be discussed.

2.3.1. Computational tools for pathway optimization

As the number of genetic elements increases in a redesigned pathway for BGC refactoring, the number of possible permutations increases exponentially, making it intractable to identify the optimal combination of genetic elements through conventional pathway construction and screening approaches. This drives the use of computational approaches to help reduce the screening load in such endeavours.

A machine learning algorithm was applied for RBS optimization in the limonene biosynthetic pathway consisting of 8 genes (Jervis et al., 2019). Initially, the two rate-limiting enzymes, trAg-GPPS and trMs-LimS, were paired with 12 RBS variants each with translation initiation rates spanning over two orders of magnitude in a combinatorial fashion to create a library of 144 possible permutations in total. The resulting RBS sequence-product titer data was used to train a machine learning algorithm, which led to the development of a model capable of predicting limonene titer from RBS sequences. This model was then further applied to a novel reconstituted mevalonate pathway, allowing a reduced library to be screened for improved limonene producers. Similarly, a Reduced Libraries (RebLibs) algorithm was used to facilitate rational minimization of experimental library screening for violacein biosynthesis (Jeschek et al., 2016). It produced a reduced set of uniformly distributed sub-libraries of RBSs for different screening purposes spanning across a range of accessible expression levels and contained degenerate sequences that allow one-pot restriction cloning.

In another example, a combinatorial library consisting of 2592 potential pathway configurations was reduced to 16 using a design of experiment (DoE) approach, which were automatically assembled using robot-assisted ligase cycling reaction method (Carbonell et al., 2018). Promoters, order of the four genes, and plasmid copy numbers were varied in the first and second design-build-test-learn cycles. The selected constructs showed up to 500-fold improvement of pinocembrin production. The DoE approach was also used to guide combinatorial pathway engineering containing five genes for violacein production (P. Xu et al., 2017). A T7 promoter library was applied in combination with Plackett–Bruman and Box–Behnken designs to probe the gene expression levels and discriminate between potential high and low producers. This approach led to identification of optimal strength promoters for VioAB, VioD, and VioEC protein expression levels and overall 3-fold increase in violacein titer. The sub-libraries of genetic elements and pathway assembly were constructed using ePathBrick vector system comprising of four compatible restriction sites allowing multiple enzyme manipulation simultaneously (Xu et al., 2012).

2.3.2. Combinatorial tools for pathway optimization

Combinatorial pathway optimization involves modifications of target BGCs by introducing genetic elements for easy replacement or deletion of regulators, RBSs, and enzymes. A 'Plug and Play' platform was developed for synthesis of C10 monoterpenoids on the basis of the common substrate geranyl pyrophosphate (GPP) and product-determining monoterpene cyclases and synthases (Leferink et al., 2016). The modular nature of the monoterpenoid pathway allows the flexible switch between a library of monoterpene cyclases and synthases using traditional restriction cloning. This approach produced more than 30 different terpenes of various classes from glucose, including several compounds never produced before in engineered microbes, demonstrating the efficacy of standardized part-based system in *E. coli*.

In another related study, a method named Multidimensional Heuristic Process (MHP) in combination with Cross-Lapping *In Vitro* Assembly (CLIVA) (Zou et al., 2013) was used to optimize the biosynthetic pathways of lycopene related compounds (Zhang et al., 2018). This modular pathway optimization approach allows screening of a library of diverse promoters, RBSs and enzyme variants in combination or individually to identify most efficient *E. coli* producers. Moreover, the Synthetic Chromosome Rearrangement and Modification by LoxP-mediated Evolution system (SCRaMbLE) was applied in a two-step manner for optimization of β-carotene and violacein BGCs (Liu et al., 2018). In this method, termed SCRaMBLE-in, pathway diversity was first created using purified recombinases to integrate genetic regulatory elements into the corresponding pathways. This is followed by random integration of the whole BGC into synthetic yeast chromosomes while creating combinatorial genome rearrangements in the host through an *in vivo* genome reshuffling system. This method enabled pathway optimization and host evolution to be carried out simultaneously to generate strains with 2-fold and 17-fold improved β-carotene and violacein titers, respectively.

3. Molecular biology tools for engineering enzymes in a pathway

3.1. New strategies and tools for enzyme engineering

The titers of natural products in heterologous hosts usually are low due to many reasons such as specific substrate and product inhibition, enzyme instability, lack of high flux of intermediates or even promiscuity of the enzymes. This arises from the intrinsic nature of natural product synthesis as the native organisms do not require natural products at high concentrations and can generate many derivatives starting from the same intermediates. In the past few years, a number of enzyme engineering strategies have been used to enhance natural product production. The paradigm for optimizing natural product synthesis shifted towards dissecting each cascade into single enzyme units to find the rate limiting steps. In this section, enzyme engineering tools and approaches in natural product pathways will be discussed, with emphasis on enzyme engineering in the context of a pathway.

3.1.1. Improving enzyme activity

Transfer of substrate between enzymes in a natural product biosynthetic pathway in a heterologous host can be a rate limiting step due to potential side reactions, diffusion issues and instability of intermediates. To avoid these issues, a fusion protein strategy may be used. For example, four novel linkers were developed and studied including natural CrtB/CrtY linker, short segment connecting the FMN and FAD/NADPH domains of cytochrome P450 reductase (CPR), and two synthetic flexible linkers for a tridomain CrtB/CrtY/CrtI system producing β-carotene and neurosporene (Rabeharindranto et al., 2019). However, physical linkers can affect enzymatic activity and modular scaffolds can become a heavy burden for protein expressing cells (Lee et al., 2012). An alternative strategy involves naturally occurring self-assembly systems such as PDZ domains and the corresponding PDZ ligands with their interaction facilitated by formation of additional β strand of ligand in the groove between α and β helices of PDZ domain (Lee and Zheng, 2010). The integration of PDZ and the corresponding ligand in enzyme assembly was successfully demonstrated using squalene epoxidase (ERG1) and dammarenediol synthase (PgDDS) showing 2-fold increase in production of ginsenoside intermediates (Zhao et al., 2016).

Rational enzyme segregation and engineering followed by reintegration to optimize necessary cascade reactions is an emerging strategy for optimizing and diverging natural product biosynthesis. Current enzyme engineering tools can be categorized into sequence based and structure-mechanism driven (Davids et al., 2013), and their application in this strategy will be briefly mentioned. Primary approaches involve homology modelling based on sequence or structure similarities followed by genetic manipulation using traditional molecular biology tools. This strategy in combination with site directed mutagenesis (SDM) was applied for fine tuning of Erg20p enzyme producing either GPP, farnesyl pyrophosphate (FPP) (Ignea et al., 2014), or geranylgeranyl pyrophosphate (GGPP) (Ignea et al., 2015) substrates for terpene synthases, which improved terpene yields by more than 10 fold. In addition, SDM integration with alanine scanning allowed engineering of S-limonene synthase active site plasticity (Srividya et al., 2015; J. Xu et al., 2017) and combination of the latter technique with saturation mutagenesis was shown to be effective in regulating stereoselectivity of taxadiene synthase (Edgar et al., 2017).

3.1.2. Functional expression of pathway genes

Reconstitution of plant BGCs in a microbial heterologous host is complicated due to different cellular compartmentalization of the pathways in the native hosts. In particular, efforts in enabling functional expression of P450 enzymes with their CPR partners in heterologous hosts will be highlighted, which has been the focus of many studies in the past few years. These P450 enzymes are crucial for formation of various enantiomeric centers giving rise to unique properties of natural products, but expression of some of the membrane-bound plant P450 enzymes in yeast or *E. coli* is one of the limiting steps in reconstitution of functional BGCs in heterologous hosts (Schlegel et al., 2009). Current strategies rely on expression of the rate limiting or insoluble enzymes from higher copy number; however this approach could only be applied in a few cases showing limited efficiency in natural product production (Brown et al., 2015). The emerging strategies for expression of functional P450 enzyme systems within BGCs include traditional tools for single enzyme systems such as a mix-and-match approach using CPR libraries to find the right partner enzymes (Trenchard and Smolke, 2015) as shown in the reconstitution of the sanguinarine branch of the benzylisoquinoline alkaloids. This strategy is not applicable in cases where P450 enzymes do not show soluble expression at any level, thereby making targeted N-terminal signal peptide engineering the primary approach. Target signal peptide replacement strategies involve swapping with either homologous P450 domain as shown in the case of salutaridine synthase enzyme (Galanie et al., 2015) or homologous non-P450 membrane bound enzyme signal peptide (Fossati et al., 2014). Advances in plant enzyme expression platforms are required because current strategies are limited to each specific BGC.

3.2. Screening methods and tools for enzyme engineering

Optimization of the enzymes responsible for natural product biosynthesis relies on the ability to detect intermediates or final products. Traditional analytical chemistry detection methods such as gas chromatography (GC), liquid chromatography (LC), their combination with mass spectrometry (MS) and nuclear magnetic resonance (NMR) require time-consuming sample preparation and have limited screening capacity (Xu et al., 2007). Within this context, *in vivo* biosensors for detecting small molecules as an alternative to analytical chemistry is gaining attention in the natural product biosynthesis field. Typical biosensors involve enzyme coupled responses, transcriptional factors, riboswitches and RNA (Liu et al., 2015; Yang Liu et al., 2017). In this section, the development of novel biosensor systems for the natural products and their intermediate detection will be discussed. Moreover, the development and versatility of molecular sensing circuits will be emphasized and only recently developed biosensors will be mentioned in this section.

3.2.1. Enzyme biosensor systems

Enzyme biosensors are based on conversion of one of the intermediates or final products to a detectable molecule with either fluorescent or colorimetric properties. An enzyme coupled biosensor was successfully applied in identifying a candidate tyrosine hydroxylase and improving its activity using error-prone PCR (epPCR) for the synthesis of L-3,4-dihydroxyphenylalanine (L-DOPA) from L-tyrosine in yeast (DeLoache et al., 2015). The assay is based on expression of plant DOPA dioxygenase that converts L-DOPA to betalamic acid, which undergoes spontaneous reaction with free amines forming fluorescent betaxanthins. The L-DOPA biosensor was further combined with fluorescent-activated cell sorting (FACS) and used for screening transposon-mediated gene disruption library in yeast (Savitskaya et al., 2019). Random mutations allowed identification of key deletions that improved final yields of L-DOPA.

Biosynthetic pathways for isoprene and lycopene diverge at isopentenyl diphosphate and the latter pigment can be produced with commercially available pAC-LYC system harbouring CrtE, CrtB, and CrtI downstream enzymes (Cunningham Jnr et al., 1994). High-throughput screening based on visual detection of lycopene with directed co-evolution of 1-deoxyxylulose-5-phosphate synthase, 1-deoxyxylulose-5-phosphate reductoisomerase and isopentenyldiphosphate (IPP)-dimethylallyldiphosphate (DMAPP) isomerase enzymes using epPCR was investigated allowing identification of enzyme variants with enhanced activity (Lv et al., 2016). The same approach employing cascade enzymatic steps for production of colorimetric molecules using either CrtE/CrtB/CrtI for diterpene or CrtM/CrtN for mono and sesquiterpene detection was used in designing terpene synthase (TPS) biosensor (Furubayashi et al., 2014). In this case, lack of pigmentation in colonies would indicate higher TPS activity due to lower amounts of residual shared isoprenyl diphosphates (GPP, FPP and GGPP) substrates. The biosensor was optimized using taxadiene synthase (diterpene) and further applied for the improvement of 5-epi-aristolochene (sesquiterpene) and geraniol (monoterpene) synthases. The developed CrtM/CrtN sensory system producing diaponeurosporene (yellow pigment) was also applied for screening of pinene synthase mutants in *E. coli* (Tashiro et al., 2016). Colorimetric high-throughput screening was used in combination with epPCR targeting active site residues. The mutated enzyme was superior to wild-type in competition experiments with limonene synthase.

3.2.2. Genetic biosensors

Genetic biosensors employ the natural transcription and translation machinery for producing reporter signals as opposed to enzyme biosensors producing response molecules in cascade reactions. Phenylalanine is a crucial intermediate in the synthesis of various flavonoids including pinocembrin, galangin, resveratrol and others (Pandey et al., 2016). TyrR transcription factor was shown to induce *tyrP* promoter in the presence of tyrosine and phenylalanine, and was used as a chassis for the development and optimization of the phenylalanine biosensor (Liu et al., 2017a). As a reporter, yellow fluorescent protein (YFP) and resistance gene *strA* were chosen for FACS and as a selection marker, respectively.

A natural flavonoid naringenin responsive regulatory system was engineered in *E. coli* based on the PfdeAR-FdeR promoter-transcription factor pair that is naturally responsible for the activation of the naringenin degradation pathway (De Paepe et al., 2018). Essential parameters for naturally occurring sensing circuits were optimized including decoupling and engineering of the detector and effector modules, and RBS optimization for fluorescent signal such as the far-red fluorescent protein mkate2. Another flavonoid molecular sensory tool was developed using a combination of two sets of natural regulatory systems: the PfdeAR-FdeR and a PnodAD1-NodD1 promoter-transcription factor pairs with mkate2 fluorescent reporter (De Paepe et al., 2019). Further, the double biosensor consisting of the previously mentioned FdeR transcription factor with a pG promoter and a cyan fluorescent protein (CFP) reporter gene, and transcription repressor QdoR with p441 promoter controlling GFP expression was engineered for detection of naringenin chalcone, naringenin, kaempferol and quercetin (Siedler et al., 2014). This designed molecular tool in combination with SDM and FACS was successfully used to detect efficient kaempferol producing mutants even at single cell level.

The versatility of transcription factor based biosensors was exploited for detection of different classes of natural products including stilbenoids and macrolides. A molecular sensor based on the *Bacillus subtilis* PadR transcriptional repressor was engineered to detect *p*-coumaric acid, which is an intermediate of the resveratrol biosynthetic pathway (Siedler et al., 2017). This biosensor in *E. coli* was combined with *p*-coumaric acid overproducing *S. cerevisiae* in a co-culture system to identify the high producers, which were isolated using FACS. The subsequent ligation of *p*-coumaric acid with coenzyme A to form its CoA thioester performed by *p*-coumarate:CoA ligase (4CL) is also a crucial step in synthesis of resveratrol (Halls and Yu, 2008) and phenylpropanoids in plants (Ehlting et al., 1999). A resveratrol biosensor was engineered using random mutagenesis of TtgR transcriptional repressor in combination with a LacZ reporter system to optimize 4CL activity (Xiong et al., 2017).

A natural macrolide sensory circuit which is used as a host defence mechanism against antibiotics was exploited for application in biosensor design (Rogers et al., 2015). The sensory system was based on the coupling of the MphR transcriptional repressor capable of binding a broad range of macrolides with a GFP reporter gene, yielding a biosensor that produces a macrolide concentration-dependent fluorescence. The sensitivity of MphR towards erythromycin A was improved by more than 10-fold using multisite saturation mutagenesis and epPCR in combination with positive-negative FACS (Kasey et al., 2018). To demonstrate its utility, the MphR-based biosensor was used to rapidly identify erythromycin A producing *Aeromicrobium erythreum* strains from knockout strains that do not produce erythromycin A in a 96-well plate.

Finally, the systematic evolution of ligands by exponential enrichment (SELEX) approach was used to develop naringenin responsive RNA elements fused with tetracycline/H^+ antiporter (TetA) and superfolder green fluorescent protein (sGFP) reporters (Xiu et al., 2017). The sensory device was used in conjunction with FACS to establish a platform for flavonoid biosensors. The efficacy of the naringenin biosensors in this study was demonstrated in co-cultures of the naringenin producing strains and the biosensor strains, which showed a positive correlation between naringenin titre and fluorescence signal. Identification of novel natural product responsive RNA elements using strategies such as Parallel Analysis of RNA Conformations Exposed to Ligand binding (PARCEL) (Tapsin et al., 2018) will be essential for development of this type of biosensors. A summary of described biosensors is listed in Table 1.

4. Molecular biology tools for engineering the host genome

As the most important source for natural products discovery, *Actinobacteria*, particularly those of genus *Streptomyces*, has been the focus of attention in the genome engineering field for decades. Compared with other model organisms such as *E. coli* and *S. cerevisiae*, *Actinobacteria* shows poverty in genetic manipulation tools and a lot of valuable NP producer strains are not genetically tractable (Alberti and Corre, 2019). Recently, the CRISPR/Cas9 system has been successfully established in different species for genome editing. Zhao and coworkers (Cobb et al., 2015) designed the pCRISPomyces-2 system, which harbors a codon-optimized spCas9 endonuclease, a sgRNA expressing cassette and two traditional 1 kb homology repairing donor arms. After conjugated into *Streptomyces*, the Cas9 endonuclease could generate a double-strand-break (DSB) at the targeted site with the guidance of sgRNA. Subsequently, the chromosome DSB was repaired by homology-dependent-repairing system with DNA donor fragments possessing two homology repair donor arms, resulting in chromosome deletions ranging from 20 bp to 31 kb at 70–100% efficiency. Furthermore, this system could be equipped with multiplexed sgRNA cassettes and corresponding homology repair donors to achieve simultaneous multiplex genome editing. In the same year, three independent groups also successfully developed a CRISPR/Cas9-based genome engineering tool in *Streptomyces* (Huang et al., 2015; Tong et al., 2015; Zeng et al., 2015). In recent years, the CRISPR/Cas9 technology is applied not only in model *Streptomyces* strains, but also in more and more other NP producers, such as non-model *Streptomyces* (Jia et al., 2017; Qin et al., 2017), *Micromonospora* (Braesel et al., 2019; Cohen and Townsend, 2018; Wolf et al., 2016), *Myxococcus* (Yang et al., 2017), *Clostridium* (Q. Li et al., 2016; Wang et al., 2017; Y. Wang et al., 2016b), and even *Yarrowia* (Gao et al., 2016) and *Cyanobacteria* (Li et al., 2016a; Wendt et al., 2016).

Although the CRISPR/Cas9 system has been widely applied to different hosts, it also has some limitations. The toxicity of Cas9 expression in some industrial *Streptomyces* strains and the requirement of complex expression constructs for targeting multiple genomic loci have limited its scope of application. Lu and coworkers employed another CRISPR system from *Francisella novicida* to *Streptomyces* and successfully developed a high-efficiency CRISPR/Cpf1 system for multiplex genome editing and transcriptional repression (Li et al., 2018). The authors engineered BGCs of three pigments (cryptic polyketone, prodiginine and actinorhodin) with CRISPR/Cpf1 technology and proved that it is a powerful alternative to CRISPR/Cas9 for industrial strains that cannot utilize the CRISPR/Cas9 system.

A combination of CRISPR/Cas9 technology and λ Red recombineering based Multiplex Automated Genomic Engineering technology, termed CRMAGE, was shown to achieve extremely high recombineering efficiency between 96.5% and 99.7% for the gene recoding of three genomic targets, while traditional recombineering could only achieve between 0.68% and 5.4% efficiency (Ronda et al., 2016). CRMAGE utilizes two USER-cloned plasmids to quickly and cost-efficiently introduce more than two mutations in a single recombineering round. Besides, degenerate codon usage could be used to expand the CRISPR target range beyond loci with PAM sequence, making CRMAGE capable of targeting virtually any site to generate single point mutations as well as larger mutation libraries. This technology has the potential of automating genome wide engineering.

During the engineering of NP producers with CRISPR systems, the toxic effects of Cas9 or Cpf1 expression and unwanted off-target effects are non-negligible problems (Tong et al., 2015). Furthermore, DNA double-strand breaks (DSB) are essential for high-efficiency HR or NHEJ, but DSB in the actinomycetes chromosomal arm region could trigger instability and mutagenesis (Hoff et al., 2018). CRISPR-Base Editing SysTem (CRISPR-BEST) (Tong et al., 2019) and cytidine base editors (CBEs) and adenine base editors (ABEs) (Zhong et al., 2019) enable precise C to T or A to G nucleotide substitutions without double-strand breaks and repairing donor. The cytidine or adenine deaminase was guided by the CRISPR system with dCas9 (catalytically defective Cas9) and performed base deamination reaction to change C to U or A to I. After DNA repair or replication, the U or I could change to T or G and complete C to T or A to G editing. As proof of concept, this base editing technology was applied to engineer different Streptomyces species and could realize simultaneous disruptions of nine different PKS gene clusters in one step. This technology provides a rapid, effective and multiplex genetic engineering approach for *Streptomyces*, and promises to be a powerful tool for other NP producers in future.

Silent BGCs are a treasure grove of novel NPs, which may be activated

Table 1
Recently developed biosensors for optimization of natural product pathways.

Transducer		Pathway	Detection		Reference
Class	Name		Signal	Response	
Enzyme	DOPA dioxygenase	L-DOPA	Betaxanthins	Fluorescence	DeLoache et al. (2015)
	Crt enzymes	Terpenes	Carotenoids	Pigmentation	Lv et al. (2016); Tashiro et al. (2016)
Transcription factor	TyrR	Flavonoid	YFP/*strA*	Fluorescence/antibiotic	Liu et al. (2017a)
	PfdeAR-FdeR	Flavonoid	mkate2	Fluorescence	De Paepe et al. (2018)
	PfdeAR-FdeR/PnodAD1-NodD1	Flavonoid	mkate2	Fluorescence	De Paepe et al. (2019)
	FdeR-pG/QdoR-p441	Flavonoid	GFP	Fluorescence	Siedler et al. (2014)
	PadR	Stilbenoid	YFP	Fluorescence	Siedler et al. (2017)
	TtgR	Stilbenoid	LacZ	Pigmentation	Xiong et al. (2017).
	MphR	Macrolide	GFP	Fluorescence	Kasey et al. (2018)
RNA	Riboswitch	Flavonoid	GFP/*tetA*	Fluorescence/antibiotic	Xiu et al. (2017)

by genome engineering tools. Zhao and coworkers developed a CRISPR-Cas9 mediated promoter knock-in strategy to activate silent BGCs in their native hosts (M. M. Zhang et al., 2017c). In this strategy, an inactive promoter region was targeted with the CRISPR-Cas9 nuclease and the resulting DSB was repaired through homologous recombination with a donor cassette harboring a constitutive promoter. In this way, the silent gene or genes could be activated, and the entire BGC could recover the ability to produce a NP. Similar to this strategy, Corre and coworkers triggered the expression of a silent gene cluster by inactivating a key transcriptional repressor with the CRISPR/Cas9 system (Alberti et al., 2019). The authors captured a silent BGC via TAR cloning and expressed this BGC in a validated *Streptomyces* host. However, the BGC was repressed by a transcriptional repressor. Inactivation of this repressor gene with CRISPR/Cas9, triggered the activation of the BGC to produce a novel NP, scleric acid. Another repressor related activation strategy was also developed by Zhao and coworkers (Wang et al., 2019a). This so-called transcription factor decoy strategy uses the repressor binding DNA fragments that are cloned from the target BGC into a plasmid to sequester the cognate repressor(s), resulting in the de-repression of the targeted BGC. This strategy has been successfully used to activate eight large BGCs ranging from 50 to 134 kilobases in multiple streptomycetes.

More recently, inspired by the mechanisms of evolution and diversification of NP BGCs, Yoshikuni and coworkers developed a chassis-independent recombinase-assisted genome engineering (CRAGE) technology to activate a BGC via single-step integration of the entire BGC directly into different bacterial chromosomes (Wang et al., 2019b). As proof-of-concept, the authors expressed three known and six previously identified but experimentally elusive NRPS and NRPS-PKS hybrid BGCs. All nine BGCs from *Photorhabdus luminescens* were expressed in 25 diverse γ-Proteobacteria species with the CRAGE technology. The successful activation of six BGCs showed that the diversity and yield of products from BGCs expressed in closely related strains were greater than these from BGCs expressed in either native or more distantly related strains.

Finally, to improve the yield, titer, and productivity of NPs in the heterologous hosts, multiple modes of optimization at the genome level such as gene overexpression, knock-down and knock-out are usually required. For example, Zhao and coworkers reported an orthogonal trifunctional CRISPR system, which combines transcriptional activation, transcriptional interference, and gene deletion (CRISPR-AID) in *S. cerevisiae* for modular, parallel and high-throughput combinatorial engineering (Lian et al., 2017). As proof of concept, the BGC of β-carotene was engineered with CRISPR-AID and the production was increased by 3-fold in a single step.

In another example, Lu and coworkers developed an advanced multiplex site-specific genome engineering method (aMSGE) for BGC integration in actinomycetes (Li et al., 2019). Based on the multiple integrases-multiple *attB* sites notion, the authors established a plug-and-play toolkit to realize high-efficiency, single-step, multi-locus integration of BGCs. As proof of concept, acetyl-CoA carboxylase genes were introduced in *S. coelicolor* to produce actinorhodin and up to four copies of the entire 72 kb 5-oxomilbemycin BGC were integrated into *Streptomyces hygroscopicus* genome to achieve 185.6% increase in titer (from 2.23 to 6.37 g L^{-1}).

5. Selected examples

Construction and optimization of BGCs for high-level production of complex NPs can be extremely challenging, requiring the use of cloning, refactoring, optimization, enzyme and host engineering tools simultaneously. In this section, we will discuss several selected examples that successfully integrated multiple molecular biology tools for production of industrially relevant compounds.

Opioids such as morphine and codeine are important psychoactive compounds used clinically for pain management and palliative care. Currently, opium poppy plants are the main source of opioids, but the production of opioids from plants raises a lot of regulatory issues due to their potential for abuse in society. In addition, production in plants suffers from low yield and susceptibility to external environmental conditions. Therefore, targeted biosynthesis of active opioid compounds in enclosed bioreactor environment was envisioned to be big leap forward for application of opioids for medicinal purposes. Biosynthesis of thebaine, a crucial intermediate of opioids and hydrocodone were reconstituted in *E. coli* (Nakagawa et al., 2016) and *S. cerevisiae* (Galanie et al., 2015) production platforms. Both systems were designed using previous knowledge of intermediate synthesis, while enzyme and strain engineering demonstrated the importance of combinatorial pathway optimization. Bacterial and yeast central metabolism was modified and exploited initially to create tyrosine over-producing strains, followed by production of L-DOPA using tyrosine hydroxylase (TYR), dopamine using DOPA decarboxylase (DoDC) before taking diverging pathways for synthesis of (*R*)-reticuline in both systems.

The *S. cerevisiae* production platform was built using a modular approach containing 17 biosynthetic enzymes for the synthesis of (*S*)-reticuline, a key intermediate in production of benzylisoquinoline alkaloids (Galanie et al., 2015). The pathway was subdivided into precursor overproduction, tetrahydrobiopterin recycling, as well as (*S*)-norcoclaurine and (*S*)-reticuline synthesis modules. It required the expression of mammalian, plant, bacterial and yeast enzymes in soluble and active forms to produce detectable levels of (*S*)-reticuline (20.7 μg L^{-1}). An additional module encoding various rate limiting enzymes was integrated to obtain 82 μg L^{-1} of (*S*)-enantiomer for further conversion.

Subsequent bioinformatics search identified DRS-DRR epimerase for (*R*)-reticuline production, which was further combined with yeast codon-optimized *P. somniferum* salutaridine synthase (yPsSalSyn), *P. bracteatum* lutaridine reductase (PbSalR), and salutaridinol acetyltransferase (PbSalAT) for thebaine synthesis. SalSyn was incorrectly processed and therefore its N-terminus had to be replaced with cheilanthifoline synthase (CFS) to generate another module containing PbDRS-DRR, $yEcCFS^{1–83}$-$yPbSalSyn^{92–504}$, PbSalR, and PsSalAT for conversion of (*S*)-reticuline to thebaine. The final module encoding the expression of thebaine 6-*O*-demethylase (T6ODM) from *P. somniferum* and morphine reductase (MorB) from *P. putida* M10 was added to yield a strain producing up to ~0.3 μg L^{-1} of hydrocodone.

Alternatively, stepwise fermentation was employed for *E. coli* production of thebaine and hydrocodone (Nakagawa et al., 2016). The total biosynthesis system using a four-step culture was investigated using strains AN1126 (dopamine), AN1055 ((*R, S*)-THP), AN1600 ((*R, S*)-reticuline) and AN1829 (thebaine and hydrocodone). Combinatorial (*R, S*)-THP and (*S*)-reticuline producing *E. coli* strains have been engineered in single batch and stepwise systems demonstrating its applicability (Nakagawa et al., 2014, 2012). Interestingly, absence of the norcoclaurine 6-*O*-methyltransferase (6OMT) allowed detection of racemic reticuline mixture rather than (*S*)-enantiomer in the reconstituted pathway (Fig. 3) (Nakagawa et al., 2016). The downstream biosynthesis of salutaridine was catalysed by another P450, SalSyn (SalS – salutaridine synthase), which could only be functionally expressed in an N-terminal truncation form. In comparison with the yeast system, the same enzyme faced issues due to incorrect sorting. Combination with recombinantly expressed plant salutaridine reductase (SalR) and salutaridinol 7-*O*-acetyltransferase (SalAT) yielded production of thebaine at 2.1 mg mL^{-1} (Fig. 3). The final two genes encoding thebaine 6-*O*-demethylase (T6ODM) and morphinone reductase (MorB) (Thodey et al., 2014) were introduced into the thebaine producing strain to generate 360 μg L^{-1} hydrocodone. The *E. coli* platform showed more than 1000-fold higher hydrocodone production compared to the yeast system, however neither of the strains could produce industrially relevant amounts. With the development of novel molecular biology tools for functional expression of plant enzymes, especially P450s, and greater understanding of metabolic flux and rapid refactoring of gene clusters, more medically important plant and bacterial metabolites will be produced in heterologous hosts. Comparison between *E. coli* and *S. cerevisiae*

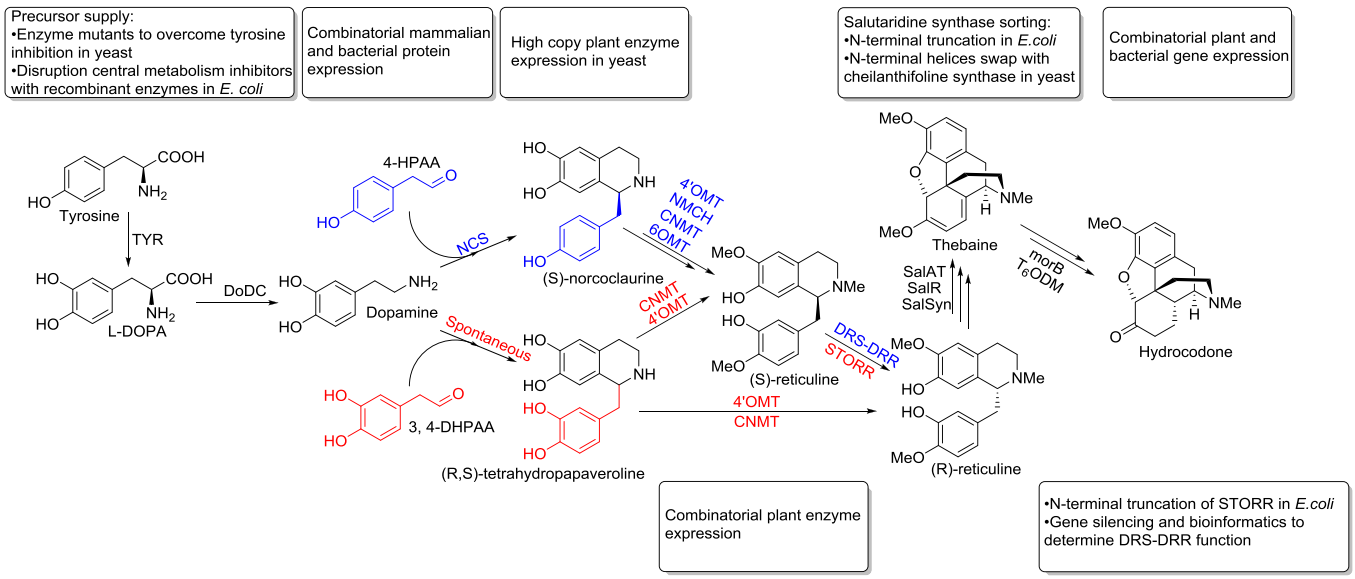

Fig. 3. Comparison of engineered biosynthetic pathways for *de novo* production of thebaine and hydrocodone in *E. coli* (red) and *S. cerevisiae* (blue). Enzymes and substrates are labelled in different colours for variation of pathway reconstitution in different microorganisms. TYR, tyrosine hydroxylase; DoDC, dopa decarboxylase; NCS, (*S*)-norcoclaurine synthase; 6OMT, norcoclaurine 6-*O*-methyltransferase; CNMT, coclaurine *N*-methyltransferase; NMCH, *N*-methylcoclaurine hydroxylase; 4′OMT, 3′-hydroxy-*N*-methylcoclaurine 4′-*O*-methyltransferase; DRS-DRR/STORR, 1,2-dehydroreticuline synthase-1,2-dehydroreticuline reductase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; T6ODM, thebaine 6-*O*-demethylase; morB, morphinone reductase. The strategies and tools to optimize different parts of pathways are highlighted in boxes above and below figure. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

platforms for biosynthesis of thebaine and hydrocodone demonstrates an immense interest in combinatorial engineering approaches. Such approaches have also been applied in the biosynthesis of noscapine (Li and Smolke, 2016) and cannabinoid derivatives (Luo et al., 2019). Both studies harness yeast central metabolism for precursor production and further integrate enzyme discovery and characterization, and unknown intermediate identification tools.

6. Future prospects and conclusions

A variety of emerging molecular biology tools and strategies have been successfully developed and applied for natural product biosynthesis at pathway, enzyme, and genome engineering levels. However, despite promising successes, many challenges still remain in engineering NP biosynthesis.

For example, with the rapid development of genome sequencing technology and bioinformatics, BGCs encoding NPs are much easier to be identified. But it is still a challenge to identify and even clone target BGCs from uncultured microorganisms, which are the majority in the environment. Metagenomics may provide a good platform for identification of novel or cryptic BGCs from soil bacteria while new direct cloning technologies and even DNA synthesis technologies may be needed for the investigation and engineering of these BGCs. Moreover, as the other main source of NPs, BGCs in plants are less investigated, and the biosynthetic genes are usually not organized in clusters. New computational and experimental tools are needed for identification, cloning, expression and optimization of the BGCs from plants in microbial hosts.

At the pathway engineering level, the existing methods are still time-consuming and labor intensive. Novel technologies, such as machine learning and automation, are increasingly applied to NP biosynthesis. With the help of computational methods, pathway engineering and optimization could be sped up and more rationally used to reduce screening efforts often associated with combinatorial strategies.

At the enzyme engineering level, the engineering strategy strongly depends on the availability of biochemical and structural data. In the future, computational enzyme engineering, especially equipped with deep machine learning technology, could accelerate BGC optimization for NPs production. Moreover, *de novo* artificial enzyme design is another promising direction. The artificially designed enzymes may catalyze unnatural chemical reactions or achieve remarkable catalytic efficiencies.

At the genome engineering level, the CRISPR/Cas-based technologies are the main emerging approach. However, CRISPR/Cas systems have some issues, such as toxicity and off-target effects. Thus, new molecular biology tools for genome engineering are still needed. For microbial host engineering, *E. coli*, yeast and *Streptomyces*, which are well studied in the laboratory for many years, are the most commonly used hosts. However, these hosts also have their respective limitations at industrial settings, and development of new non-model organisms as industrial production hosts is one of the future directions.

In conclusion, numerous molecular biology tools and strategies have been developed for natural product discovery and engineering in recent years. However, given the vast number of uncharacterized BGCs and the ever-growing importance of NPs in the pharmaceutical industry, more powerful molecular biology tools and strategies are still needed for large-scale and more efficient natural product discovery and engineering. With these tools and strategies, we expect a renaissance of natural product research in the coming years.

Declaration of competing interest

The authors declare no competing financial interests.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM077596 and AI144967 to H.Z.), the A*STAR Visiting Investigator Program (H.Z.), and the National Research Foundation, Singapore (NRF2017NRF-CRP001-044) (E.A.).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2019.e00108.

References

Alberti, F., Corre, C., 2019. Editing streptomycete genomes in the CRISPR/Cas9 age. Nat. Prod. Rep. https://doi.org/10.1039/C8NP00081F.

Alberti, F., Leng, D.J., Wilkening, I., Song, L., Tosin, M., Corre, C., 2019. Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. Chem. Sci. 10, 453–463. https://doi.org/10.1039/c8sc03814g.

Banik, J.J., Brady, S.F., 2010. Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. Curr. Opin. Microbiol. 13, 603–609. https://doi.org/10.1016/j.mib.2010.08.012.

Braesel, J., Lee, J.H., Arnould, B., Murphy, B.T., Eustáquio, A.S., 2019. Diazaquinomycin biosynthetic gene clusters from marine and freshwater actinomycetes. J. Nat. Prod. 82, 937–946. https://doi.org/10.1021/acs.jnatprod.8b01028.

Brewster, R.C., Jones, D.L., Phillips, R., 2012. Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. PLoS Comput. Biol. 8, e1002811. https://doi.org/10.1371/journal.pcbi.1002811.

Brown, S., Clastre, M., Courdavault, V., O'Connor, S.E., 2015. De novo production of the plant-derived alkaloid strictosidine in yeast. Proc. Natl. Acad. Sci. 112, 3205–3210. https://doi.org/10.1073/pnas.1423555112.

Carbonell, P., Jervis, A.J., Robinson, C.J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K.A., Currin, A., Rattray, N.J.W., Taylor, S., Spiess, R., Sung, R., Williams, A.R., Fellows, D., Stanford, N.J., Mulherin, P., Le Feuvre, R., Barran, P., Goodacre, R., Turner, N.J., Goble, C., Chen, G.G., Kell, D.B., Micklefield, J., Breitling, R., Takano, E., Faulon, J.-L., Scrutton, N.S., 2018. An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. Commun. Biol. 1, 66. https://doi.org/10.1038/s42003-018-0076-9.

Cobb, R.E., Wang, Y., Zhao, H., 2015. High-Efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. ACS Synth. Biol. 4, 723–728. https://doi.org/10.1021/sb500351f.

Cohen, D.R., Townsend, C.A., 2018. A dual role for a polyketide synthase in dynemicin enediyne and anthraquinone biosynthesis. Nat. Chem. 10, 231–236. https://doi.org/10.1038/nchem.2876.

Cunningham Jnr, F.X., Sun, Z., Chamovitz, D., Hirschberg, J., Gantt, E., 1994. Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium *Synechococcus* sp strain PCC7942. Plant Cell 6, 1107–1121. https://doi.org/10.1105/tpc.6.8.1107.

D'Agostino, P.M., Gulder, T.A.M., 2018. Direct pathway cloning combined with sequence- and ligation-independent cloning for fast biosynthetic gene cluster refactoring and heterologous expression. ACS Synth. Biol. 7, 1702–1708. https://doi.org/10.1021/acssynbio.8b00151.

Davids, T., Schmidt, M., Böttcher, D., Bornscheuer, U.T., 2013. Strategies for the discovery and engineering of enzymes for biocatalysis. Curr. Opin. Chem. Biol. 17, 215–220. https://doi.org/10.1016/J.CBPA.2013.02.022.

De Paepe, B., Maertens, J., Vanholme, B., De Mey, M., 2019. Chimeric LysR-type transcriptional biosensors for customizing ligand specificity profiles toward flavonoids. ACS Synth. Biol. 8, 318–331. https://doi.org/10.1021/acssynbio.8b00326.

De Paepe, B., Maertens, J., Vanholme, B., De Mey, M., 2018. Modularization and response curve engineering of a naringenin-responsive transcriptional biosensor. ACS Synth. Biol. 7, 1303–1314. https://doi.org/10.1021/acssynbio.7b00419.

DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J.J., Dueber, J.E., 2015. An enzyme-coupled biosensor enables (*S*)-reticuline production in yeast from glucose. Nat. Chem. Biol. 11, 465–471. https://doi.org/10.1038/nchembio.1816.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602–607. https://doi.org/10.1038/nature09886.

Donadio, S., Monciardini, P., Sosio, M., 2007. Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. Nat. Prod. Rep. 24, 1073–1079. https://doi.org/10.1039/b514050c.

Edgar, S., Li, F.S., Qiao, K., Weng, J.K., Stephanopoulos, G., 2017. Engineering of taxadiene synthase for improved selectivity and yield of a key taxol biosynthetic intermediate. ACS Synth. Biol. 6, 201–205. https://doi.org/10.1021/acssynbio.6b00206.

Ehlting, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., Kombrink, E., 1999. Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. Plant J. 19, 9–20. https://doi.org/10.1046/j.1365-313x.1999.00491.x.

Enghiad, B., Zhao, H., 2017. Programmable DNA-guided artificial restriction enzymes. ACS Synth. Biol. 6, 752–757. https://doi.org/10.1021/acssynbio.6b00324.

Engler, C., Kandzia, R., Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3, e3647. https://doi.org/10.1371/journal.pone.0003647.

Eyles, T.H., Vior, N.M., Truman, A.W., 2018. Rapid and robust yeast-mediated pathway refactoring generates multiple new bottromycin-related metabolites. ACS Synth. Biol. 7, 1211–1218. https://doi.org/10.1021/acssynbio.8b00038.

Fossati, E., Ekins, A., Narcross, L., Zhu, Y., Falgueyret, J.P., Beaudoin, G.A.W., Facchini, P.J., Martin, V.J.J., 2014. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in *Saccharomyces cerevisiae*. Nat. Commun. 5, 3283. https://doi.org/10.1038/ncomms4283.

Freestone, T.S., Ju, K.S., Wang, B., Zhao, H., 2017. Discovery of a phosphonoacetic acid derived natural product by pathway refactoring. ACS Synth. Biol. 6, 217–223. https://doi.org/10.1021/acssynbio.6b00299.

Furubayashi, M., Ikezumi, M., Kajiwara, J., Iwasaki, M., Fujii, A., Li, N., Saito, K., Umeno, D., 2014. A high-throughput colorimetric screening assay for terpene synthase activity based on substrate consumption. PLoS One 9, e93317. https://doi.org/10.1371/journal.pone.0093317.

Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F., Smolke, C.D., 2015. Complete biosynthesis of opioids in yeast. Science 349, 1095–1100. https://doi.org/10.1126/science.aac9373.

Galm, U., Shen, B., 2006. Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis. Expert Opin. Drug Discov. 1, 409–437. https://doi.org/10.1517/17460441.1.5.409.

Gao, S., Tong, Y., Wen, Z., Zhu, L., Ge, M., Chen, D., Jiang, Y., Yang, S., 2016. Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. J. Ind. Microbiol. Biotechnol. 43, 1085–1093. https://doi.org/10.1007/s10295-016-1789-8.

Greunke, C., Duell, E.R., D'Agostino, P.M., Glöckle, A., Lamm, K., Gulder, T.A.M., 2018. Direct Pathway Cloning (DiPaC) to unlock natural product biosynthetic potential. Metab. Eng. 47, 334–345. https://doi.org/10.1016/j.ymben.2018.03.010.

Halls, C., Yu, O., 2008. Potential for metabolic engineering of resveratrol biosynthesis. Trends Biotechnol. 26, 77–81. https://doi.org/10.1016/j.tibtech.2007.11.002.

Hoff, G., Bertrand, C., Piotrowski, E., Thibessard, A., Leblond, P., 2018. Genome plasticity is governed by double strand break DNA repair in *Streptomyces*. Sci. Rep. 8, 5272. https://doi.org/10.1038/s41598-018-23622-w.

Horbal, L., Marques, F., Nadmid, S., Mendes, M.V., Luzhetskyy, A., 2018. Secondary metabolites overproduction through transcriptional gene cluster refactoring. Metab. Eng. 49, 299–315. https://doi.org/10.1016/j.ymben.2018.09.010.

Hsu, P.D., Lander, E.S., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010.

Huang, H., Zheng, G., Jiang, W., Hu, H., Lu, Y., 2015. One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. Acta Biochim. Biophys. Sin. (Shanghai). 47, 231–243. https://doi.org/10.1093/abbs/gmv007.

Ignea, C., Pontini, M., Maffei, M.E., Makris, A.M., Kampranis, S.C., 2014. Engineering monoterpene production in yeast using a synthetic dominant negative geranyl diphosphate synthase. ACS Synth. Biol. 3, 298–306. https://doi.org/10.1021/sb400115e.

Ignea, C., Trikka, F.A., Nikolaidis, A.K., Georgantea, P., Ioannou, E., Loupassaki, S., Kefalas, P., Kanellis, A.K., Roussis, V., Makris, A.M., Kampranis, S.C., 2015. Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. Metab. Eng. 27, 65–75. https://doi.org/10.1016/j.ymben.2014.10.008.

Jakočiunas, T., Rajkumar, A.S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C.B., Skjødt, 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, 1126–1134. https://doi.org/10.1021/acssynbio.5b00007.

Jeong, J.Y., Yim, H.S., Ryu, J.Y., Lee, H.S., Lee, J.H., Seen, D.S., Kang, S.G., 2012. One-step sequence-and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies. Appl. Environ. Microbiol. 78, 5440–5443. https://doi.org/10.1128/AEM.00844-12.

Jervis, A.J., Carbonell, P., Vinaixa, M., Dunstan, M.S., Hollywood, K.A., Robinson, C.J., Rattray, N.J.W., Yan, C., Swainston, N., Currin, A., Sung, R., Toogood, H., Taylor, S., Faulon, J.L., Breitling, R., Takano, E., Scrutton, N.S., 2019. Machine learning of designed translational control allows predictive pathway optimization in *Escherichia coli*. ACS Synth. Biol. 8, 127–136. https://doi.org/10.1021/acssynbio.8b00398.

Jeschek, M., Gerngross, D., Panke, S., 2016. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. Nat. Commun. 7, 11163. https://doi.org/10.1038/ncomms11163.

Jia, H., Zhang, L., Wang, T., Han, J., Tang, H., Zhang, L., 2017. Development of a CRISPR/Cas9-mediated gene-editing tool in *Streptomyces rimosus*. Microbiology 163, 1148–1155. https://doi.org/10.1099/mic.0.000501.

Jiang, W., Zhao, X., Gabrieli, T., Lou, C., Ebenstein, Y., Zhu, T.F., 2015. Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. Nat. Commun. 6, 8101. https://doi.org/10.1038/ncomms9101.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821. https://doi.org/10.1126/science.1225829.

Kang, H.-S.S., Charlop-Powers, Z., Brady, S.F., 2016. Multiplexed CRISPR/Cas9- and TAR-mediated promoter engineering of natural product biosynthetic gene clusters in yeast. ACS Synth. Biol. 5, 1002–1010. https://doi.org/10.1021/acssynbio.6b00080.

Kasey, C.M., Zerrad, M., Li, Y., Cropp, T.A., Williams, G.J., 2018. Development of transcription factor-based designer macrolide biosensors for metabolic engineering and synthetic biology. ACS Synth. Biol. 7, 227–239. https://doi.org/10.1021/acssynbio.7b00287.

Katz, M., Hover, B.M., Brady, S.F., 2016. Culture-independent discovery of natural products from soil metagenomes. J. Ind. Microbiol. Biotechnol. 43, 129–141. https://doi.org/10.1007/s10295-015-1706-6.

Keasling, J.D., 2010. Manufacturing molecules through metabolic engineering. Science 330, 1355–1358. https://doi.org/10.1126/science.1193990.

Kim, S.H., Lu, W., Ahmadi, M.K., Montiel, D., Ternei, M.A., Brady, S.F., 2019. Atolypenes, tricyclic bacterial sesterterpenes discovered using a multiplexed in vitro Cas9-TAR

gene cluster refactoring approach. ACS Synth. Biol. 8, 109–118. https://doi.org/10.1021/acssynbio.8b00361.
Lee, H.-J., Zheng, J.J., 2010. PDZ domains and their binding partners: structure, specificity, and modification. Cell Commun. Signal. 8, 8. https://doi.org/10.1186/1478-811X-8-8.
Lee, H., DeLoache, W.C., Dueber, J.E., 2012. Spatial organization of enzymes for metabolic engineering. Metab. Eng. 14, 242–251. https://doi.org/10.1016/J.YMBEN.2011.09.003.
Lee, N.C.O., Larionov, V., Kouprina, N., 2015. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. Nucleic Acids Res. 43, e55. https://doi.org/10.1093/nar/gkv112.
Leferink, N.G.H., Jervis, A.J., Zebec, Z., Toogood, H.S., Hay, S., Takano, E., Scrutton, N.S., 2016. A 'Plug and Play' platform for the production of diverse monoterpene hydrocarbon scaffolds in *Escherichia coli*. ChemistrySelect 1, 1893–1896. https://doi.org/10.1002/slct.201600563.
Li, H., Shen, C.R., Huang, C.H., Sung, L.Y., Wu, M.Y., Hu, Y.C., 2016a. CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. Metab. Eng. 38, 293–302. https://doi.org/10.1016/j.ymben.2016.09.006.
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. https://doi.org/10.1016/j.ymben.2018.12.001.
Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W., Lu, Y., 2018. CRISPR-Cpf1-Assisted multiplex genome editing and transcriptional repression in *Streptomyces*. Appl. Environ. Microbiol. 84, e00827–18. https://doi.org/10.1128/AEM.00827-18.
Li, Q., Chen, J., Minton, N.P., Zhang, Y., Wen, Z., Liu, J., Yang, H., Zeng, Z., Ren, X., Yang, J., Gu, Y., Jiang, W., Jiang, Y., Yang, S., 2016b. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. Biotechnol. J. 11, 961–972. https://doi.org/10.1002/biot.201600053.
Li, M.Z., Elledge, S.J., 2012. SLIC: a method for sequence- and ligation-independent cloning. Methods Mol. Biol. 852, 51–59. https://doi.org/10.1007/978-1-61779-564-0_5.
Li, Y., Smolke, C.D., 2016. Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nat. Commun. 7, 12137. https://doi.org/10.1038/ncomms12137.
Lian, J., HamediRad, M., Hu, S., Zhao, H., 2017. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. Nat. Commun. 8, 1688. https://doi.org/10.1038/s41467-017-01695-x.
Liu, D., Evans, T., Zhang, F., 2015. Applications and advances of metabolite biosensors for metabolic engineering. Metab. Eng. 31, 35–43. https://doi.org/10.1016/J.YMBEN.2015.06.008.
Liu, W., Luo, Z., Wang, Y., Pham, N.T., Tuck, L., Pérez-Pi, I., Liu, L., Shen, Y., French, C., Auer, M., Marles-Wright, J., Dai, J., Cai, Y., 2018. Rapid pathway prototyping and engineering using in vitro and in vivo synthetic genome SCRaMbLE-in methods. Nat. Commun. 9, 1936. https://doi.org/10.1038/s41467-018-04254-0.
Liu, Y., Zhuang, Y., Ding, D., Xu, Y., Sun, J., Zhang, D., 2017a. Biosensor-based evolution and elucidation of a biosynthetic pathway in *Escherichia coli*. ACS Synth. Biol. 6, 837–848. https://doi.org/10.1021/acssynbio.6b00328.
Liu, Yang, Liu, Ye, Wang, M., 2017b. Design, optimization and application of small molecule biosensor in metabolic engineering. Front. Microbiol. 8, 2012. https://doi.org/10.3389/fmicb.2017.02012.
Luo, X., Reiter, M.A., d'Espaux, L., Wong, J., Denby, C.M., Lechner, A., Zhang, Y., Grzybowski, A.T., Harth, S., Lin, W., Lee, H., Yu, C., Shin, J., Deng, K., Benites, V.T., Wang, G., Baidoo, E.E.K., Chen, Y., Dev, I., Petzold, C.J., Keasling, J.D., 2019. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. Nature 567, 123–126. https://doi.org/10.1038/s41586-019-0978-9.
Lv, X., Gu, J., Wang, F., Xie, W., Liu, M., Ye, L., Yu, H., 2016. Combinatorial pathway optimization in *Escherichia coli* by directed co-evolution of rate-limiting enzymes and modular pathway engineering. Biotechnol. Bioeng. 113, 2661–2669. https://doi.org/10.1002/bit.26034.
Montiel, D., Kang, H.-S., Chang, F.-Y., Charlop-Powers, Z., Brady, S.F., 2015. Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters. Proc. Natl. Acad. Sci. 112, 8953–8958. https://doi.org/10.1073/pnas.1507606112.
Nah, H.J., Pyeon, H.R., Kang, S.H., Choi, S.S., Kim, E.S., 2017. Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. Front. Microbiol. 8, 394. https://doi.org/10.3389/fmicb.2017.00394.
Nakagawa, A., Matsumura, E., Koyanagi, T., Katayama, T., Kawano, N., Yoshimatsu, K., Yamamoto, K., Kumagai, H., Sato, F., Minami, H., 2016. Total biosynthesis of opiates by stepwise fermentation using engineered *Escherichia coli*. Nat. Commun. 7, 10390. https://doi.org/10.1038/ncomms10390.
Nakagawa, A., Matsuzaki, C., Matsumura, E., Koyanagi, T., Katayama, T., Yamamoto, K., Sato, F., Kumagai, H., Minami, H., 2014. (*R*, *S*)-Tetrahydropapaveroline production by stepwise fermentation using engineered *Escherichia coli*. Sci. Rep. 4, 6695. https://doi.org/10.1038/srep06695.
Nakagawa, A., Minami, H., Kim, J.-S., Koyanagi, T., Katayama, T., Sato, F., Kumagai, H., 2012. Bench-top fermentative production of plant benzylisoquinoline alkaloids using a bacterial platform. Bioengineered 3, 49–53. https://doi.org/10.4161/bbug.3.1.18446.
Newman, D.J., Cragg, G.M., 2012. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J. Nat. Prod. 75, 311–335. https://doi.org/10.1021/np200906s.
Nielsen, J., Keasling, J.D., 2016. Engineering cellular metabolism. Cell 164, 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004.
Pandey, R.P., Parajuli, P., Koffas, M.A.G., Sohng, J.K., 2016. Microbial production of natural and non-natural flavonoids: pathway engineering, directed evolution and systems/synthetic biology. Biotechnol. Adv. 34, 634–662. https://doi.org/10.1016/j.biotechadv.2016.02.012.
Penesyan, A., Kjelleberg, S., Egan, S., 2010. Development of novel drugs from marine surface associated microorganisms. Mar. Drugs 8, 438–459. https://doi.org/10.3390/md8030438.
Qin, Z., Munnoch, J.T., Devine, R., Holmes, N.A., Seipke, R.F., Wilkinson, K.A., Wilkinson, B., Hutchings, M.I., 2017. Formicamycins, antibacterial polyketides produced by *Streptomyces formicae* isolated from African *Tetraponera* plant-ants. Chem. Sci. 8, 3218–3227. https://doi.org/10.1039/c6sc04265a.
Rabeharindranto, H., Castaño-Cerezo, S., Lautier, T., Garcia-Alles, L.F., Treitz, C., Tholey, A., Truan, G., 2019. Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in *S. cerevisiae*. Metab. Eng. Commun. 8, e00086. https://doi.org/10.1016/j.mec.2019.e00086.
Ren, H., Hu, P., Zhao, H., 2017a. A plug-and-play pathway refactoring workflow for natural product research in *Escherichia coli* and *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 114, 1847–1854. https://doi.org/10.1002/bit.26309.
Ren, H., Wang, B., Zhao, H., 2017b. Breaking the silence: new strategies for discovering novel natural products. Curr. Opin. Biotechnol. 48, 21–27. https://doi.org/10.1016/j.copbio.2017.02.008.
Rogers, J.K., Guzman, C.D., Taylor, N.D., Raman, S., Anderson, K., Church, G.M., 2015. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. Nucleic Acids Res. 43, 7648–7660. https://doi.org/10.1093/nar/gkv616.
Ronda, C., Pedersen, L.E., Sommer, M.O.A., Nielsen, A.T., 2016. CRMAGE: CRISPR optimized MAGE recombineering. Sci. Rep. 6, 19452. https://doi.org/10.1038/srep19452.
Ross, A.C., Gulland, L.E.S., Dorrestein, P.C., Moore, B.S., 2015. Targeted capture and heterologous expression of the pseudoalteromonas alterochromide gene cluster in escherichia coli represents a promising natural product exploratory platform. ACS Synth. Biol. 4, 414–420. https://doi.org/10.1021/sb500280q.
Salis, H.M., Mirsky, E.A., Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. 27, 946–950. https://doi.org/10.1038/nbt.1568.
Savitskaya, J., Protzko, R.J., Li, F.-Z., Arkin, A.P., Dueber, J.E., 2019. Iterative screening methodology enables isolation of strains with improved properties for a FACS-based screen and increased L-DOPA production. Sci. Rep. 9, 5815. https://doi.org/10.1038/s41598-019-41759-0.
Schlegel, S., Klepsch, M., Gialama, D., Wickström, D., Slotboom, D.J., De Gier, J.-W., 2009. Revolutionizing membrane protein overexpression in bacteria. Microb. Biotechnol. 3, 403–411. https://doi.org/10.1111/j.1751-7915.2009.00148.x.
Siedler, S., Khatri, N.K., Zsohár, A., Kjærbølling, I., Vogt, M., Hammar, P., Nielsen, C.F., Marienhagen, J., Sommer, M.O.A., Joensson, H.N., 2017. Development of a bacterial biosensor for rapid Screening of yeast p-Coumaric acid production. ACS Synth. Biol. 6, 1860–1869. https://doi.org/10.1021/acssynbio.7b00009.
Siedler, S., Stahlhut, S.G., Malla, S., Maury, J., Neves, A.R., 2014. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. Metab. Eng. 21, 2–8. https://doi.org/10.1016/j.ymben.2013.10.011.
Smanski, M.J., Zhou, H., Claesen, J., Shen, B., Fischbach, M.A., Voigt, C.A., 2016. Synthetic biology to access and expand nature's chemical diversity. Nat. Rev. Microbiol. 14, 135–149. https://doi.org/10.1038/nrmicro.2015.24.
Song, C., Luan, J., Cui, Q., Duan, Q., Li, Z., Gao, Y., Li, R., Li, A., Shen, Y., Li, Y., Stewart, A.F., Zhang, Y., Fu, J., Wang, H., 2019. Enhanced heterologous spinosad production from a 79-kb synthetic multioperon assembly. ACS Synth. Biol. 8, 137–147. https://doi.org/10.1021/acssynbio.8b00402.
Srividya, N., Davis, E.M., Croteau, R.B., Lange, B.M., 2015. Functional analysis of (4*S*)-limonene synthase mutants reveals determinants of catalytic outcome in a model monoterpene synthase. Proc. Natl. Acad. Sci. 112, 3332–3337. https://doi.org/10.1073/pnas.1501203112.
Tan, G.-Y., Liu, T., 2017. Rational synthetic pathway refactoring of natural products biosynthesis in actinobacteria. Metab. Eng. 39, 228–236. https://doi.org/10.1016/J.YMBEN.2016.12.006.
Tang, X., Li, J., Millán-Aguiñaga, N., Zhang, J.J., O'Neill, E.C., Ugalde, J.A., Jensen, P.R., Mantovani, S.M., Moore, B.S., 2015. Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. ACS Chem. Biol. 10, 2841–2849. https://doi.org/10.1021/acschembio.5b00658.
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. https://doi.org/10.3389/fmicb.2018.01660.
Tapsin, S., Sun, M., Shen, Y., Zhang, H., Lim, X.N., Susanto, T.T., Yang, S.L., Zeng, G.S., Lee, J., Lezhava, A., Ang, E.L., Zhang, L.H., Wang, Y., Zhao, H., Nagarajan, N., Wan, Y., 2018. Genome-wide identification of natural RNA aptamers in prokaryotes and eukaryotes. Nat. Commun. 9, 1289. https://doi.org/10.1038/s41467-018-03675-1.
Tashiro, M., Kiyota, H., Kawai-Noma, S., Saito, K., Ikeuchi, M., Iijima, Y., Umeno, D., 2016. Bacterial production of pinene by a laboratory-evolved pinene-synthase. ACS Synth. Biol. 5, 1011–1020. https://doi.org/10.1021/acssynbio.6b00140.
Thodey, K., Galanie, S., Smolke, C.D., 2014. A microbial biomanufacturing platform for natural and semisynthetic opioids. Nat. Chem. Biol. 10, 837–844. https://doi.org/10.1038/nchembio.1613.
Tong, Y., Charusanti, P., Zhang, L., Weber, T., Lee, S.Y., 2015. CRISPR-Cas9 based engineering of actinomycetal genomes. ACS Synth. Biol. 4, 1020–1029. https://doi.org/10.1021/acssynbio.5b00038.
Tong, Y., Robertsen, H.L., Blin, K., Klitgaard, A.K., Weber, T., Lee, S.Y., 2019. CRISPR-BEST: a Highly Efficient DSB-free Base Editor for Filamentous Actinomycetes. https://doi.org/10.1101/582403 bioRxiv 582403.

Trenchard, I.J., Smolke, C.D., 2015. Engineering strategies for the fermentative production of plant alkaloids in yeast. Metab. Eng. 30, 96–104. https://doi.org/10.1016/j.ymben.2015.05.001.

Wang, B., Guo, F., Dong, S.H., Zhao, H., 2019a. Activation of silent biosynthetic gene clusters using transcription factor decoys. Nat. Chem. Biol. 15, 111–114. https://doi.org/10.1038/s41589-018-0187-0.

Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B., Yoshikuni, Y., 2019b. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. Nat. Microbiol. https://doi.org/10.1038/s41564-019-0573-8.

Wang, H., Li, Z., Jia, R., Hou, Y., Yin, J., Bian, X., Li, A., Müller, R., Stewart, A.F., Fu, J., Zhang, Y., 2016a. RecET direct cloning and Redαβ recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression. Nat. Protoc. 11, 1175–1190. https://doi.org/10.1038/nprot.2016.054.

Wang, H., Li, Z., Jia, R., Yin, J., Li, A., Xia, L., Yin, Y., Müller, R., Fu, J., Stewart, A.F., Zhang, Y., 2018. ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. Nucleic Acids Res. 46, 2697. https://doi.org/10.1093/nar/gkx1296, 2697.

Wang, J.W., Wang, A., Li, K., Wang, B., Jin, S., Reiser, M., Lockey, R.F., 2015. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. Biotechniques 58, 161–170. https://doi.org/10.2144/000114261.

Wang, S., Dong, S., Wang, P., Tao, Y., Wang, Y., 2017. Genome editing in *Clostridium saccharoperbutylacetonicum* N1-4 with the CRISPR-Cas9 system. Appl. Environ. Microbiol. 83, e00233. https://doi.org/10.1128/AEM.00233-17, 17.

Wang, Y., Zhang, Z.T., Seo, S.O., Lynn, P., Lu, T., Jin, Y.S., Blaschek, H.P., 2016b. Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. ACS Synth. Biol. 5, 721–732. https://doi.org/10.1021/acssynbio.6b00060.

Wendt, K.E., Ungerer, J., Cobb, R.E., Zhao, H., Pakrasi, H.B., 2016. CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. Microb. Cell Factories 15, 115. https://doi.org/10.1186/s12934-016-0514-7.

Wolf, T., Gren, T., Thieme, E., Wibberg, D., Zemke, T., Pühler, A., Kalinowski, J., 2016. Targeted genome editing in the rare actinomycete *Actinoplanes* sp. SE50/110 by using the CRISPR/Cas9 System. J. Biotechnol. 231, 122–128. https://doi.org/10.1016/J.JBIOTEC.2016.05.039.

Xiong, D., Lu, S., Wu, J., Liang, C., Wang, Wei, Wang, Wenzhao, Jin, J.-M., Tang, S.-Y., 2017. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. Metab. Eng. 40, 115–123. https://doi.org/10.1016/J.YMBEN.2017.01.006.

Xiu, Y., Jang, S., Jones, J.A., Zill, N.A., Linhardt, R.J., Yuan, Q., Jung, G.Y., Koffas, M.A.G., 2017a. Naringenin-responsive riboswitch-based fluorescent biosensor module for *Escherichia coli* co-cultures. Biotechnol. Bioeng. 114, 2235–2244. https://doi.org/10.1002/bit.26340.

Xu, J., Ai, Y., Wang, J., Xu, Jingwei, Zhang, Y., Yang, D., 2017b. Converting S-limonene synthase to pinene or phellandrene synthases reveals the plasticity of the active site. Phytochemistry 137, 34–41. https://doi.org/10.1016/j.phytochem.2017.02.017.

Xu, P., Rizzoni, E.A., Sul, S.Y., Stephanopoulos, G., 2017c. Improving metabolic pathway efficiency by statistical model-based multivariate regulatory metabolic engineering. ACS Synth. Biol. 6, 148–158. https://doi.org/10.1021/acssynbio.6b00187.

Xu, P., Vansiri, A., Bhan, N., Koffas, M.A.G., 2012. ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. ACS Synth. Biol. 1, 256–266. https://doi.org/10.1021/sb300016b.

Xu, R.N., Fan, L., Rieser, M.J., El-Shourbagy, T.A., 2007. Recent advances in high-throughput quantitative bioanalysis by LC–MS/MS. J. Pharm. Biomed. Anal. 44, 342–355. https://doi.org/10.1016/J.JPBA.2007.02.006.

Yamanaka, K., Reynolds, K.A., Kersten, R.D., Ryan, K.S., Gonzalez, D.J., Nizet, V., Dorrestein, P.C., Moore, B.S., 2014. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. Proc. Natl. Acad. Sci. 111, 1957–1962. https://doi.org/10.1073/pnas.1319584111.

Yang, Y., Wang, Y., Li, Z., Gong, Y., Zhang, P., Hu, W., Sheng, D., Li, Y., 2017. Increasing on-target cleavage efficiency for CRISPR/Cas9-induced large fragment deletion in *Myxococcus xanthus*. Microb. Cell Factories 16, 142. https://doi.org/10.1186/s12934-017-0758-x.

Zeng, H., Wen, S., Xu, W., He, Z., Zhai, G., Liu, Y., Deng, Z., Sun, Y., 2015. Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. Appl. Microbiol. Biotechnol. 99, 10575–10585. https://doi.org/10.1007/s00253-015-6931-4.

Zhang, C., Seow, V.Y., Chen, X., Too, H.-P., 2018. Multidimensional heuristic process for high-yield production of astaxanthin and fragrance molecules in *Escherichia coli*. Nat. Commun. 9, 1858. https://doi.org/10.1038/s41467-018-04211-x.

Zhang, J.J., Tang, X., Zhang, M., Nguyen, D., Moore, B.S., 2017b. Broad-host-range expression reveals native and host regulatory elements that influence heterologous antibiotic production in Gram-negative bacteria. mBio 8, e01291. https://doi.org/10.1128/mBio.01291-17, 17.

Zhang, M.M., Qiao, Y., Ang, E.L., Zhao, H., 2017a. Using natural products for drug discovery: the impact of the genomics era. Expert Opin. Drug Discov. 12, 475–487. https://doi.org/10.1080/17460441.2017.1303478.

Zhang, M.M., Wang, Y., Ang, E.L., Zhao, H., 2016. Engineering microbial hosts for production of bacterial natural products. Nat. Prod. Rep. 33, 963–987. https://doi.org/10.1039/C6NP00017G.

Zhang, Y., Muyrers, J.P.P., Testa, G., Stewart, A.F., 2000. DNA cloning by homologous recombination in Escherichia coli. Nat. Biotechnol. 18, 1314–1317. https://doi.org/10.1038/82449.

Zhang, M.M., Wong, F.T., Wang, Y., Luo, S., Lim, Y.H., Heng, E., Yeo, W.L., Cobb, R.E., Enghiad, B., Ang, E.L., Zhao, H., 2017c. CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. Nat. Chem. Biol. 13, 607–609. https://doi.org/10.1038/nchembio.2341.

Zhao, C., Gao, X., Liu, X., Wang, Y., Yang, S., Wang, F., Ren, Y., 2016. Enhancing biosynthesis of a ginsenoside precursor by self-assembly of two key enzymes in *Pichia pastoris*. J. Agric. Food Chem. 64, 3380–3385. https://doi.org/10.1021/acs.jafc.6b00650.

Zhong, Z., Guo, J., Deng, L., Chen, L., Wang, J., Li, S., Xu, W., Deng, Z., Sun, Y., 2019. Base Editing in Streptomyces with Cas9-Deaminase Fusions. https://doi.org/10.1101/630137 bioRxiv 630137.

Zhou, Y., Murphy, A.C., Samborskyy, M., Prediger, P., Dias, L.C., Leadlay, P.F., 2015. Iterative mechanism of macrodiolide formation in the anticancer compound conglobatin. Chem. Biol. 22, 745–754. https://doi.org/10.1016/j.chembiol.2015.05.010.

Zou, R., Zhou, K., Stephanopoulos, G., Too, H.P., 2013. Combinatorial engineering of 1-deoxy-D-xylulose 5-phosphate pathway using cross-lapping *in vitro* assembly (CLIVA) method. PLoS One 8, e79557. https://doi.org/10.1371/journal.pone.0079557.