Title
Multi-chassis engineering for heterologous production of microbial natural products.

Authors
Ke, Jing
Yoshikuni, Yasuo

Publication Date
2020-04-01

DOI
10.1016/j.copbio.2019.09.005

Peer reviewed

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Multi-chassis engineering for heterologous production of microbial natural products

Jing Ke[1] and Yasuo Yoshikuni[1,2,3,4,5]

Microbial genomes encode numerous biosynthetic gene clusters (BGCs) that may produce natural products with diverse applications in medicine, agriculture, the environment, and materials science. With the advent of genome sequencing and bioinformatics, heterologous expression of BGCs is of increasing interest in bioactive natural product (NP) discovery. However, this approach has had limited success because expression of BGCs relies heavily on the physiology of just a few commonly available host chassis. Expanding and diversifying the chassis portfolio for heterologous BGC expression may greatly increase the chances for successful NP production. In this review, we first discuss genetic and genome engineering technologies used to clone, modify, and transform BGCs into multiple strains and to engineer chassis strains. We then highlight studies that employed the multi-chassis approach successfully to optimize NP production, discover previously uncharacterized NPs, and better understand BGC function.

Addresses
[1] US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Walnut Creek, CA 94598, USA
[2] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3] Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4] Center for Advanced Bioenergy and Bioproducts Innovation, Urbana, IL 61801, USA
[5] Global Institution for Collaborative Research and Education, Hokkaido University, Hokkaido, 060-8589, Japan

Corresponding author: Yoshikuni, Yasuo (yyoshikuni@lbl.gov)

Current Opinion in Biotechnology 2020, **62**:88–97

This review comes from a themed issue on **Environmental biotechnology**

Edited by **David R Johnson** and **Stephan Noack**

For a complete overview see the Issue and the Editorial

Available online 19th October 2019

https://doi.org/10.1016/j.copbio.2019.09.005

Introduction

Microorganisms provide a variety of bioactive natural products (NPs) used in medicine and agriculture (e.g. antibiotics and pesticides), as well as in protection of the environment and development of materials [1–3]. In bacteria, NP biosynthesis pathways are encoded by biosynthetic gene clusters (BGCs), which are often found in a single large, contiguous genomic region. With the advent of genome sequencing [4], bioinformatics programs [5], and genetic/genome engineering technologies, BGCs can now be transferred from native strains to various heterologous hosts more easily and successfully. However, about 90% of these BGCs are inactive or only partially activated under standard fermentation conditions, and regulatory triggers required for full activation of the cryptic genes are typically unknown [6]. Therefore, many of the microorganisms' chemical innovations are untapped.

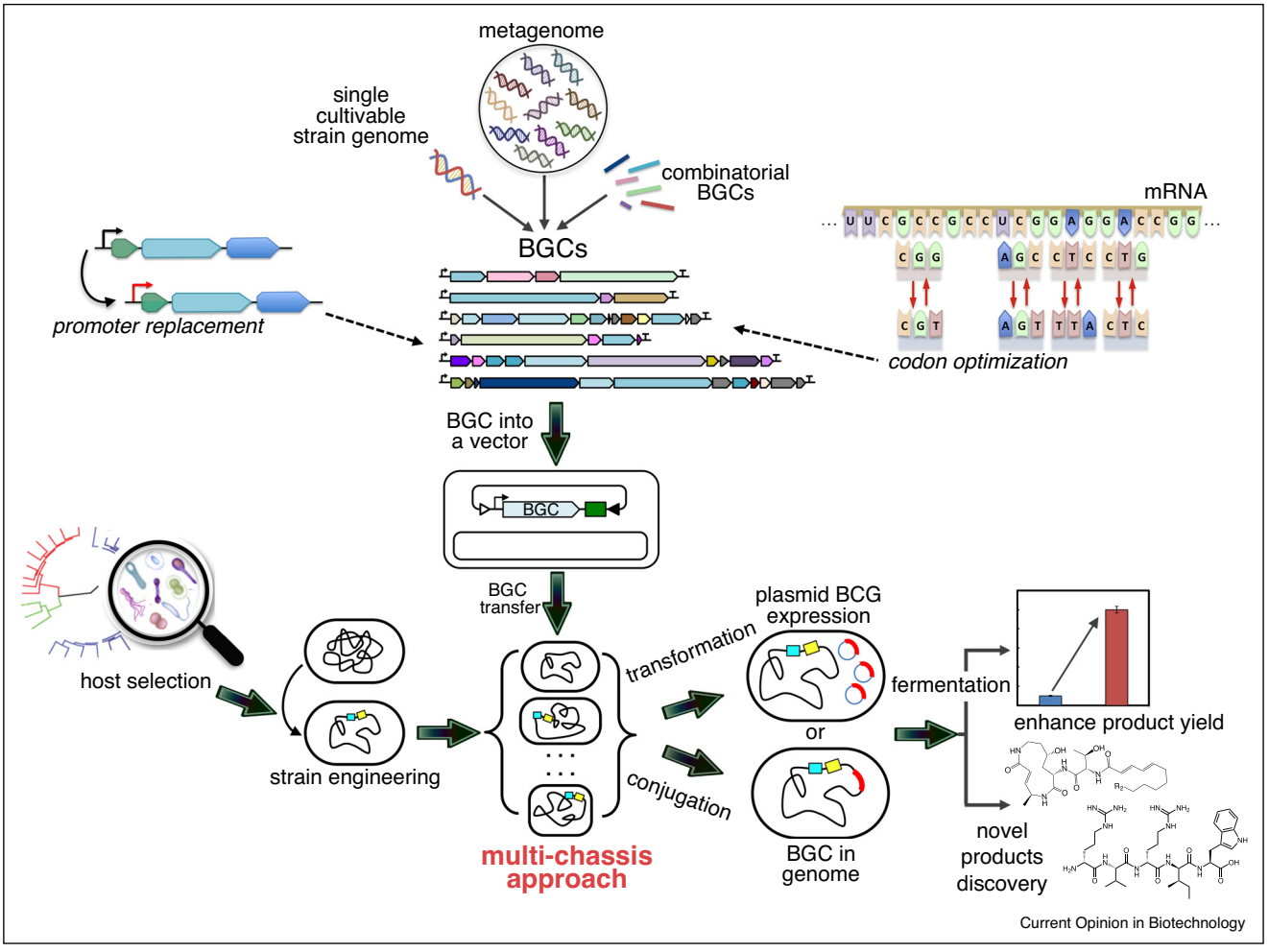
For functional expression of NPs, synthetic biology approaches such as codon optimization, refactoring, and/or DNA synthesis can theoretically be used to uncouple BGCs from their native regulatory constraints [7–10]. However, this approach requires carefully adapting the BGC to a species currently available for heterologous expression, such as *Escherichia coli*, *Bacillus subtilis*, or *Streptomyces lividans*. This approach also requires that all transferred genes be functionally expressed and all translated products fold properly and undergo appropriate post-translational modifications. The chassis strain itself must be able to supply all substrates and co-factors and tolerate all intermediates and products. Because of these complex requirements, BGC expression often fails to produce the corresponding product [11,12].

Clearly, the host organism is critical in heterologous production of NPs. However, research has primarily focused on a few model chassis strains as hosts. To demonstrate that exploring a wider range of chassis strains is worthwhile, this review describes engineering of non-conventional strains that has facilitated discovery of novel NPs by producing higher-than-native product yield. This review also highlights new tools and methodologies used to engineer diverse microbial strains to confirm the potential of the multi-chassis approach for heterologous production of NPs (Figure 1).

Genetic/genome engineering tools and techniques for heterologous NP production using multiple chassis

NP production is often regulated at multiple levels [13,14] that are difficult to determine. To facilitate robust expression of BGCs in diverse hosts, broadly applicable genetic parts, tools, and methodologies are important for functional BGC expression. In this section, we review those particularly important for multi-chassis approaches,

Figure 1



The general workflow of the multi-chassis engineering for heterologous expression of microbial natural products.

and all the disccused NPs and related approaches are summarized in Table 1.

Genome engineering techniques that enable the multi-chassis approach

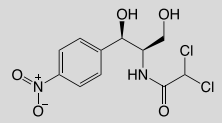
Several shuttle plasmids based on bacterial artificial chromosomes (BACs) have been developed [15–18], including some *E. coli-Streptomyces* shuttle BACs [19–21]. These shuttle BACs often contain an attP-int system from the *Streptomyces* phage [22], which catalyzes site-specific recombination of the phage attachment site (*attP*) with the bacterial attachment site (*attB*), forming two hybrid sites (*attL* and *attR*) [23,24]. Both ϕC31 and ϕBT1 attP-int loci have been well exploited to construct versatile vectors. These systems can integrate large BGCs of over 100 kb into the bacteria chromosome [23–27].

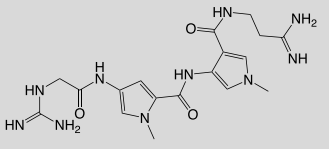
Use of integrative and conjugative elements (ICEs) is a similar approach. These modular mobile elements are integrated into chromosomes of firmicute species, mediated by the attP-int system [28,29]. ICEs undergo autonomous rolling-circle replication and are conjugally transferred between cell populations [30,31]. Brophy et al. recently developed a transposon mini-ICE from *B. subtilis* (ICEBs1) [32••]. This mini-ICEBs1 element achieved efficient transfer of large DNA constructs (~100 kb) into 35 firmicute strains isolated from humans and soil. While these attP-int systems are powerful tools to domesticate previously undomesticated strains, their host ranges are usually limited to a small group of bacteria. Integration efficiency also decreases exponentially as DNA construct size increases [32••].

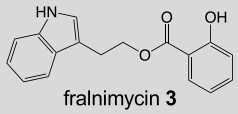
We recently developed 'chassis-independent recombinase-assisted genome engineering' (CRAGE), which can integrate large, complex BGCs into genomes of undomesticated strains with high accuracy and efficiency. A landing pad (LP) containing mutually exclusive *lox* sites is first

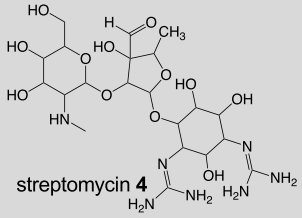
Table 1

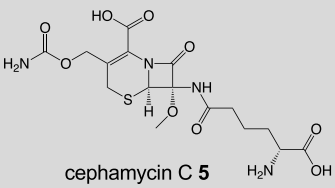
Summary of the heterologously produced natural products discussed in this paper

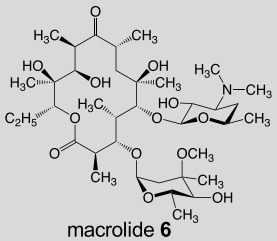

chloramphenicol **1**

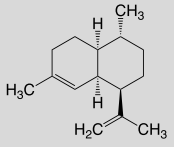

congocidine **2**

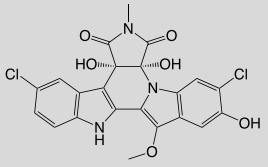

fralnimycin **3**

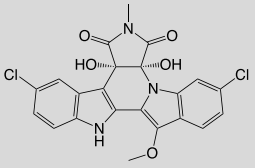

streptomycin **4**

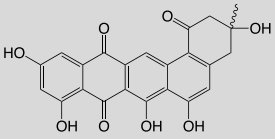

cephamycin C **5**

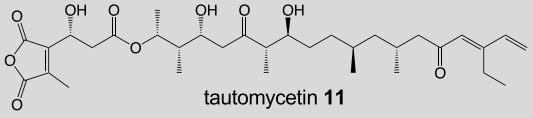

macrolide **6**

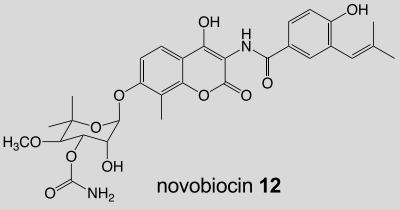

amorpha-4,11-diene **7**

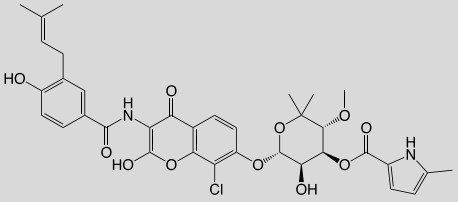

lazarimide A **8**

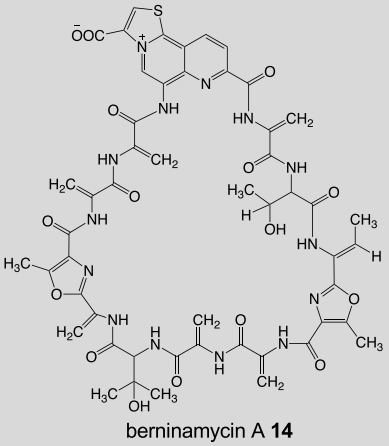

lazarimide B **9**

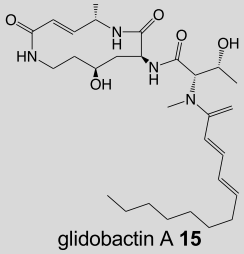

pentangular type II polyketide **10**

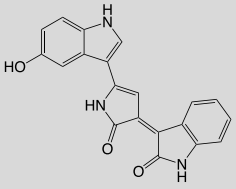

tautomycetin **11**

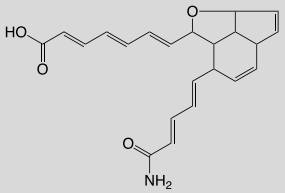

novobiocin **12**


clorobiocin **13**


berninamycin A **14**


glidobactin A **15**


violacein **16**


metatricycloene **17**

No	NP	BGC source	Heterologous host	Strategy	Comments	Ref.
1	Chloramphenicol	*S. venezuelae* ATCC 10712	*S. coelicolor* M145, M1146, M1152, M1154	Deletion of four BGCs competing for a common precursor, *rpoB* & *rpsL* mutations.	Point mutations to transcription (*rpoB*) and translation (*rpsl*) machinery in the host with quadruple BGC deletion increased productivity by 40-fold	[35]

Table 1 (*Continued*)

No	NP	BGC source	Heterologous host	Strategy	Comments	Ref.
2	Congocidine	*S. ambofaciens* ATCC23877	*S. coelicolor* M145, M1146, M1152, M1154	Deletion of four BGCs competing for a common precursor, *rpoB* & *rpsL* mutations	Point mutations to transcription (*rpoB*) and translation (*rpsl*) machinery in the host with quadruple BGC deletion increased productivity by 40-fold	[35]
3	Fralnimycin	*Frankia alni* ACN14a	*S. albus* J1074, Del14, M1152, M1154	Deletion of 15 BGCs	This approach improved compound detection limit	[37]
4	Streptomycin	*S. griseus* IFO 13350	*S. avermitilis*, *S. avermitilis* SUKA17	Reduced genome with deletion of 1.4 Mb non-essential genes.	Productivity higher than that of the native strain	[38]
5	Cephamycin C	*S. clavuligerus* ATCC 27064	*S. avermitilis*, *S. avermitilis* SUKA17	Reduced genome with deletion of 1.4 Mb non-essential genes. Expression of an extra copy of *ccaR*	Productivity higher than that of the native strain	[38]
6	Macrolide	*S. platensis* Mer-11107, 75 kb	*S. avermitilis, S. avermitilis SUKA17*	Reduced genome with deletion of 1.4 Mb non-essential genes. Expression of an extra copy of the regulatory gene *pldR*	PKS production at the industrial production level	[38]
7	Amorpha-4, 11-diene	*Artemisia annua*	*S. avermitilis, S. avermitilis* SUKA17	Reduced genome with deletion of 1.4 Mb non-essential genes. A synthetic gene optimized for *Streptomyces* codon usage	S. avermitilis could efficiently produce a plant terpenoid	[38]
8, **9**	lazarimide A and B	Environmental DNA	*S. albus*	Replaced native BGC promoter with bi-directional promoter-RBS cassettes	This approach can streamline the discovery of NPs from silent BGCs	[49]
10	Pentangular type II polyketide	*S. roseosporus* NRRL15998	*S. albus*, *S. lividans*, *S. roseosporus*, *S. venezuelae*, *S. viridochromogenes*	One-step CRISPR-Cas9 knock-in strategy to activate silent BGCs in five *Streptomyces* species	This polyketide was produced only in *S. viridochromogenes*; potentially a scalable approach facilitating novel NP discovery	[67••]
11	Tautomycetin	*S. spiroverticillatus*, 80 kb	*S. coelicolor* M145, *S. lividans* TK21, *Streptomyces* sp. CK4412, *S. coelicolor* TMC003	pSBAC was used along with cluster tandem repeat integration	*S. coelicolor* M145 produced >4-fold higher yield than that of the native strain	[21]
12	Novobiocin	novobiocin BGC	*S. coelicolor* M512,			[71]
13	Clorobiocin	clorobiocin BGC	*S. lividans* TK24, *S. spheroides* NCIMB 11891, *S. roseochromogenes var. oscitans* DS 12.976	λ-Red-mediated homologous recombination, phase attP-int system	While *S. coelicolor* M512 produced NPs comparable levels to their native hosts, *S. lividans* TK24 was ~5 times less productive	[71]
14	Berninamycin A	*S. bernensis* UC 5144	*S. lividans* TK24, *S. venezuelae* ATCC 10712	pSET152+bern plasmid, conjugative transfer	*S. lividans* produced this NP 2.4-fold greater than in the native species	[74]
15	Glidobactin A	*Photorhabdus luminescens* subsp. laumondii TTO1	25 of γ-proteobacteria chassis strains	CRAGE	This strategy enables integration of BGCs into chromosomes of diverse bacteria species	[33••]
16	Violacein	*Pseudoalteromonas luteoviolacea* 2ta16	*Pseudomonas putida* KT2440, *Agrobacterium tumefaciens* LBA4404, *E.coli*	TAR-cloning and heterologous BGC expression platform	A nonclustered LuxR-type quorum-sensing receptor from *P. luteoviolacea* 2ta16, PviR, was identified to increase pathway transcription and violacein production	[60]
17	Metatricycloene	Uncultured environmental bacteria	Diverse *Streptomyces* species, including *S. albus*	Optimized methods for constructing high quality metagenomic libraries in the best *Streptomyces* host	Multi-chassis approach improved the efficiency of novel NP discovery	[85]

integrated into chromosomes of recipient strains, mediated through a transposase. The BGCs, also flanked by mutually exclusive *lox* sites, are subsequently transferred into recipient strains via conjugation. With the activity of a Cre recombinase, we demonstrated integration of BGCs ranging from 10 to 48 kb with high efficiency [33••]. The utility of CRAGE, however, is currently limited to species within a few bacterial phyla including proteobacteria and actinobacteria. We are currently modifying the system to further extend the utility of CRAGE.

Genome-reduced strains used to improve heterologous NP production

Engineering chassis strains by deleting endogenous pathways competing for the same substrates can improve NP production and simplify NP identification [34]. For example, eliminating a few BGCs from *S. coelicolor* significantly enhanced production of **1** and **2** [35,36]. Myronovskyi *et al.* [37] expanded this approach and eliminated 15 BGCs from *S. albus* Del14, which led to enhanced production and discovery of compound **3**, which was previously uncharacterized.

Several studies have shown that strains with minimized genomes are useful NP production hosts. By deleting 1.4 Mb of nonessential *S. avermitilis* genes, Ikeda *et al.* [38] facilitated efficient production of **4**, **5**, **6**, and **7**. Choi *et al.* [39] used a similar approach with *Pseudomonas putida* to produce four proof-of-concept bioproducts (a protein, a polyketide, an isoprenoid, and an amino acid derivative). Baumgart *et al.* [40] created a series of genome-reduced *Corynebacterium glutamicum* strains. The final strain, with 412 genes (13.4% of the genome) deleted, showed wild-type-like growth behavior and tolerance of multiple stresses, such as oxygen limitation.

Notably, a variant of *E. coli* with a 4 Mb synthetic genome was created recently through high-fidelity convergent total synthesis [41]. Design included genome-wide replacement of synonymous codons (two sense codons and a stop codon). This synthetic *E. coli* with fewer synonymous sense codons may be easier to manipulate as a heterologous chassis for *in vivo* biosynthesis of non-canonical biopolymers [42].

Pathway refactoring and codon optimization based on established heterologous expression systems

Pathway refactoring refers to rewriting the genetic parts with defined function in a certain BGC for finer control [43,44]. BGCs can undergo modifications from simple promoter replacement to extensive swapping of genetic parts to optimize expression levels and increase production titers [44–47].

Replacing native BGC promoters with well-characterized promoters can streamline NP discovery from transcriptionally silent/cryptic BGCs. Diverse promoters with different expression characteristics are now available for this purpose [48•]. Montiel *et al.* [49] integrated bi-directional promoter-RBS cassettes into several BGCs, which activated rebeccamycin, tetarimycin, and lazarimide BGCs in *S. albus*. Promoter replacement also led to discovery of potent antiproliferative agents, **8** and **9** [49].

The efficiency and accuracy of codon usage is different for different processes and characteristics such as mRNA folding [50], gene expression [51,52], translatability [53], and co-translational folding and protein levels [50,54,55], as well as for fitness through RNA toxicity [56]. Rare codons potentially slow down the translation rate [57,58]. Translation is sometimes more efficient when genes are expressed in genetically distant strains [59,60]. Alternatively, translation efficiency can significantly improve when codons are optimized for the host organism [61,62]. Decreasing cost for de novo DNA synthesis makes codon-optimization an increasingly amenable strategy for increased heterologous NP production [63,64].

Pathway refactoring and codon optimization are therefore effective approaches, especially when BGC expression is largely controlled at the transcriptional and/or translational levels.

CRISPR-Cas9-mediated genome editing

CRISPR-Cas9 has rapidly become a promising genome-editing tool. Through sgRNA guidance, CRISPR-Cas9 cuts target double-strand DNA precisely 3 bp upstream of a specific protospacer adjacent motif (PAM). Subsequent repair of the break allows modification of the target locus. CRISPR-assisted targeted deletions can efficiently eliminate endogenous BGCs in hosts [65,66]. Additionally, CRISPR technology has enabled expression of multiple BGCs of different classes in five *Streptomyces* species and triggered production of unique metabolites, including a novel pentangular type II polyketide **10** in *Streptomyces viridochromogenes* [67••]. Furthermore, Jiang and Zhu used CRISPR-Cas9 in vitro and developed a new cloning method, Cas9-Assisted targeting of chromosome segments (CATCH) [68••]. From intact bacterial chromosomes, CATCH allows single-step targeted cloning of near-arbitrary genomic sequences of up to 100 kb [69].

Applying the multi-chassis approach — adapting a BGC to available chassis strains

Ideally, an expression host suite would include several species, perhaps including industrial strains that could produce NPs at very high levels [34]. Although uncertainty remains because of insufficient systematic analysis, in general NP researchers conclude that chassis strains closely related to a native BGC strain can efficiently activate the functions of the transferred BGCs. Regulatory elements that silence the BGC might not be conserved in these chassis strains, but other factors important

for BGC expression, such as promoters [12] and ribosome binding sites (RBSs) [70], may be similar to those of native strains. Although it is hard to predict specific host compatibility from phylogeny, parallel expression of NP constructs in multiple chassis strains provides a way to screen and optimize NP production levels and thereby facilitate discovery of novel NPs.

The multi-chassis approach helps optimize and enhance the titer of NPs produced heterologously

A versatile *E. coli-Streptomyces* shuttle BAC plasmid such as pSBAC was used along with cluster tandem integration to express an 80-kb tautomycetin (TMC) BGC from *S. spiroverticillatus* in several *Streptomyces* strains, including *Streptomyces coelicolor* M145 and *S. lividans* TK21 [21]. A tandem repeat of the TMC BGC was integrated into *S. coelicolor*. This strain produced 13.31 mg/L of **11**, a >4-fold higher yield than that of the native strain. Eustáquio *et al.* also demonstrated varied NP expression levels in different heterologous hosts. When the entire novobiocin and clorobiocin BGC was transformed into *S. coelicolor* M512 and *S. lividans* TK24, *S. coelicolor* M512 derivatives produced antibiotics in yields (31 mg/L **12** and 26 mg/L **13**, respectively) comparable to those of natural producer strains (35 mg/L **12** by *S. spheroides* and 25 mg/L **13** by *S. roseochromogenes*), whereas *S. lividans* TK24 derivatives were at least five times less productive [71].

Nah et al. listed about 90 actinomycetes NP BGCs that have been successfully expressed in various *Streptomyces* hosts [72••]. *S. coelicolor* and *S. lividans* were two competent strains for heterologous expression, thanks to their well-characterized genetic and biochemical properties. About 12% of the BGCs were expressed in another heterologous host, *S. albus*, which has fast growth and an efficient genetic system [73]. Compared with the original NP-producing strains, expression levels were higher for approximately 14% of the NPs and lower for 12% when they were expressed in the *Streptomyces* hosts. Production yield of the BGC of **14** in *S. lividans* was 2.4-fold greater than that in the native species of *S. bernensis*; there was no production in *S. venezuelae* [74]. In brief, species within the *Streptomyces* genus exhibit varied NP production yields from heterologous BGC expression, which further demonstrates the significance of the heterologous host screening and selection process for successful NP BGC expression.

A previous study by authors of this review demonstrated the efficacy of the multi-chassis approach [33••]. The same BGC was introduced in parallel into a phylogenetically diverse portfolio of 25 γ-proteobacteria strains using CRAGE [33••]. This approach substantially increased successful BGC expression. For example, the BGC for **15** (*plu1877–plu1881*) was inactive in its native strain even after homologous expression, suggesting strict regulation. In contrast, most of the strains harboring this BGC produced at least one of the three NPs previously identified. Interestingly, the production ratios of the three metabolites shifted dramatically (~10 000-fold) among the chassis strains. This result indicates that product specificity of BGCs can also change significantly, depending on the background physiology of each strain. Notably, the heterologous production titer of **15** reached as high as 177 mg/L, which is 700-fold higher than the production titer previously reported for the same BGC expressed heterologously in *E. coli* [75]. This suggests greater enhancement of NP production when BGCs are expressed in species more closely related to the native strain than in distantly related species.

However, some exceptions exist. Zhang et al. expressed the BGC of **16** from *Pseudoalteromonas luteoviolacea* 2ta16 and revealed robust production in two proteobacterial chassis strains, *P. putida* KT2440 and *Agrobacterium tumefaciens* LBA4404, but achieved very little production in various laboratory strains of *E. coli*, despite their closer phylogenetic relationship to the native strain than *A. tumefaciens*' relationship to the native strain [60]. This result suggests optimal heterologous NP producers exist among distantly related strains that can be identified only empirically.

The multi-chassis approach facilitates discovery of novel metabolites

Discovery of NPs through activation of putative BGCs often relies on suitable host strains providing yields sufficient for NP identification, purification, and NMR characterization [76]. In our previous study [33••], among nine heterologously expressed BGCs, eight previously uncharacterized NPs were identified from four BGCs. Notably, none of these newly detected NPs were detected in *E. coli*. An additional advantage of the multi-chassis strategy was also demonstrated. Untargeted searches for novel metabolites by LC–MS tend to yield high rates of false positives [77,78]. However, when the same pathway is expressed simultaneously in multiple hosts, we can use independent occurrence of the same feature under different physiological conditions to narrow down the search space considerably. This concept is analogous to FAC-MS scoring [79].

Cultivable microbes have historically been used to discover bioactive NPs. However, they make up as little as 1% of environmental microbiomes. Functional metagenomics screening has revealed the presence of large numbers of unexplored BGCs, suggesting the biosynthetic potential of the uncultured majority. A prerequisite for identifying NPs is successful expression of BGCs from metagenome libraries. Some attempts at metagenomic screening in *Streptomyces* using a single host have been proven insufficient [80–83,84•]. On the other hand, Brady *et al.* successfully used diverse *Streptomyces* species as chassis to improve metagenomic screening [85]. *S. albus*, in particular, exhibited the highest

propensity for heterologous expression of BGCs, leading to the discovery of **17**.

Guo *et al.* found a widely distributed family of NRPS gene clusters in human gut microbiomes [8]. By heterologously expressing a subset of these clusters in *E. coli* or *B. subtilis*, they showed the BGCs encoded biosynthesis of pyrazinones and dihydropyrazinones. Guo *et al.* subsequently demonstrated that the active form of these molecules is the peptide aldehyde, which bears potent protease inhibitory activity and selectively targets a subset of cathepsins in human cell proteomes [8]. Their findings show that an approach combining bioinformatics, synthetic biology, and heterologous BGC expression can rapidly expand knowledge of the metabolic potential of the microbiota while avoiding the challenges of cultivating fastidious commensals.

Rational design of novel NPs via combinatorial assembly of BGC modules [86,87] is another important area to which the multi-chassis approach can contribute, because the regulatory complexity of combinatorial BGCs requires appropriate expression hosts [88]. For example, through expression of combinatorial BGCs in two different hosts, Pérez *et al.* generated seven new mithramycin derivatives with antitumor activity against different tumor cell lines [89]. The expressing hosts were the wild type *S. argillaceus* and its mutant M7U1, with an inactivated *mtmU* gene coding a 4-ketoreductase involved in the biosynthesis of D-oliose [90]. *S. argillaceus* Δ*mtmU* produced twice as many compounds as did the wild type. Similar results may be readily obtained if the multi-chassis approach is employed for this BGC.

Concluding remarks and future perspectives

NPs and their derivatives continue to provide high value in medicine, agriculture, environmental protection, and materials development. The multi-chassis approach for heterologous BGC expression can facilitate improvement of NP production, discovery of NPs from diverse natural resources, and even rational design and engineering of bioactive 'unnatural' NPs. Continued technological and conceptual advances in strain engineering as well as genetic and genome editing techniques will open up opportunities to fully realize the multi-chassis approach and to explore and harness the immensely diverse chemical repertoire of nature.'

Conflict of interest statement

We filed a patent application for the method described in this manuscript.

Acknowledgements

Work conducted by the U.S. Department of Energy Joint Genome Institute , a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. We thank Anita Wahler for professional editing support.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Dayan FE, Cantrell CL, Duke SO: **Natural products in crop protection**. *Bioorg Med Chem* 2009, **17**:4022-4034.

2. Gu J, Gui Y, Chen L, Yuan G, Lu HZ, Xu X: **Use of natural products as chemical library for drug discovery and network pharmacology**. *PLoS One* 2013, **8**:e62839.

3. Flüchter S, Follonier S, Schiel-Bengelsdorf B, Bengelsdorf FR, Zinn M, Dürre P: **Anaerobic production of poly(3-hydroxybutyrate) and its precursor 3-hydroxybutyrate from synthesis gas by *Autotrophic Clostridia***. *Biomacromolecules* 2019 http://dx.doi.org/10.1021/acs.biomac.9b00342.

4. Mukherjee S, Seshadri R, Varghese NJ, Eloe-Fadrosh EA, Meier-Kolthoff JP, Göker M, Coates RC, Hadjithomas M, Pavlopoulos GA, Paez-Espino D *et al.*: **1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life**. *Nat Biotechnol* 2017, **35** 676-003683.

5. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R: **antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences**. *Nucleic Acids Res* 2011, **39**:W339-346 (Web Server issue).

6. Rutledge PJ, Challis GL: **Discovery of microbial natural products by activation of silent biosynthetic gene clusters**. *Nat Rev Microbiol* 2015, **13**:509-523.

7. Cohen LJ, Esterhazy D, Kim SH, Lemetre C, Aguilar RR, Gordon EA, Pickard AJ, Cross JR, Emiliano AB, Han SM *et al.*: **Commensal bacteria make GPCR ligands that mimic human signaling molecules**. *Nature* 2017, **549**:48-53.

8. Guo CJ, Chang FY, Wyche TP: **Discovery of reactive microbiota-derived metabolites that inhibit host proteases**. *Cell* 2017, **168**:517-526.

9. Smanski MJ, Zhou H, Claesen J, Shen B, Fischbach MA, Voigt CA: **Synthetic biology to access and expand nature's chemical diversity**. *Nat Rev Microbiol* 2016, **14**:135-149.

10. Shao Z, Rao G, Li C, Abil Z, Luo Y, Zhao H: **Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold**. *ACS Synth Biol* 2013, **2**:662-669.

11. Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, Zhang Y: **Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting**. *Nat Biotechnol* 2012, **30**:440-446.

12. Wenzel SC, Müller R: **Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways**. *Curr Opin Biotechnol* 2005, **16**:594-606.

13. Bibb MJ: **Regulation of secondary metabolism in *Streptomycetes***. *Curr Opin Microbiol* 2005, **8**:208-215.

14. Bibb MJ, Hesketh A: **Analyzing the regulation of antibiotic production in *Streptomycetes***. *Methods Enzymol* 2009, **458**:93-116.

15. Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S: **Artificial chromosomes for antibiotic-producing *actinomycetes***. *Nat Biotechnol* 2000, **18**:343-345.

16. Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, Osburne MS: **Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts**. *Appl Environ Microbiol* 2004, **70**:2452-2463.

17. Fuji K, Koyama T, Kai W, Kubota S, Yoshida K, Ozaki A, Aoki JY, Kawabata Y, Araki K, Tsuzaki T *et al.*: **Construction of a high-coverage bacterial artificial chromosome library and comprehensive genetic linkage map of yellowtail *Seriola quinqueradiata***. *BMC Res Notes* 2014, **7**:200.

18. Varshney RK, Mir RR, Bhatia S, Thudi M, Hu Y, Azam S, Zhang Y, Jaganathan D, You FM, Gao J *et al.*: **Integrated physical, genetic and genome map of chickpea (*Cicer arietinum* L.)**. *Funct Integr Genomics* 2014, **14**:59-73.

19. Miao V, Coëffet-Legal MF, Brian P, Brost R, Penn J, Whiting A, Martin S, Ford R, Parr I, Bouchard M *et al.*: **Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry**. *Microbiology* 2005, **151**:1507-1523.

20. Liu H, Jiang H, Haltli B, Kulowski K, Muszynska E, Feng X, Summers M, Young M, Graziani E, Koehn F *et al.*: **Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile *Escherichia coli-Streptomyces* artificial chromosome vector, pSBAC**. *J Nat Prod* 2009, **72**:389-395.

21. Nah HJ, Woo MW, Choi SS, Kim ES: **Precise cloning and tandem integration of large polyketide biosynthetic gene cluster using *Streptomyces* artificial chromosome system**. *Microb Cell Fact* 2015, **14**:140 http://dx.doi.org/10.1186/s12934-015-0325-2.

22. Thorpe HM, Smith MC: **In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family**. *Proc Natl Acad Sci U S A* 1998, **95**:5505-5510.

23. Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, Schoner BE: **Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to Streptomyces spp**. *Gene* 1992, **116**:43-49.

24. Gregory MA, Till R, Smith MCM: **Integration site for *Streptomyces* phage ϕBT1 and development of site-specific integrating vectors**. *J Bacteriol* 2003, **185**:5320-5323.

25. Fogg PCM, Colloms S, Rosser S, Stark M, Smith MCM: **New applications for phage integrases**. *J Mol Biol* 2014, **426**:2703-2716.

26. Du D, Wang L, Tian Y, Liu H, Tan H, Niu G: **Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in *Streptomyces***. *Sci Rep* 2015, **5**:8740.

27. Engelhardt K, Degnes KF, Zotchev SB: **Isolation and characterization of the gene cluster for biosynthesis of the thiopeptide antibiotic TP-1161**. *Appl Environ Microbiol* 2010, **76**:7093-7101.

28. Hickman AB, Chandler M, Dyda F: **Integrating prokaryotes and eukaryotes: DNA transposases in light of structure**. *Crit Rev Biochem Mol Biol* 2010, **45**:50-69.

29. Johnson CM, Grossman AD: **Integrative and conjugative elements (ICEs): what they do and how they work**. *Annu Rev Genet* 2015, **49**:577-601.

30. Lee CA, Babic A, Grossman AD: **Autonomous plasmid-like replication of a conjugative transposon**. *Mol Microbiol* 2010, **75**:268-279.

31. Wozniak RAF, Waldor MK: **Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow**. *Nat Rev Microbiol* 2010, **8**:552-563.

32. •• Brophy JAN, Triassi AJ, Adams BL, Renberg RL, Stratis-Cullum DN, Grossman AD, Voigt CA: **Engineered integrative and conjugative elements for efficient and inducible DNA transfer to undomesticated bacteria**. *Nat Microbiol* 2018, **3**:1043-1053.
The authors developed an approach based on ICEBs1 to overcome the problem of transferring DNA into the cell of undomesticated bacteria. This breakthrough technology makes it easier to harness undomesticated microbes and expand the range of heterologous chassis strains for NP production.

33. •• Wang G, Zhao Z, Ke J, Engel Y, Shi YM, Zhang Z, Bingol K, Robinson D, Wang B, Evans R *et al.*: **Chassis-independent recombinase-assisted genome engineering enables rapid activation of biosynthetic gene clusters**. *Nat Microbiol* 2019. in preparation.
One of the best examples of the success of multi-chassis engineering for discovery and heterologous production of microbial natural products.

34. Baltz RH: ***Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters**. *J Ind Microbiol Biotechnol* 2010, **37**:759-772.

35. Gomez-Escribano JP, Bibb MJ: **Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters**. *Microb Biotechnol* 2011, **4**:207-215.

36. Zhang MM, Wang Y, Ang EL, Zhao H: **Engineering microbial hosts for production of bacterial natural products**. *Nat Prod Rep* 2016, **33**:963-987.

37. Myronovskyi M, Rosenkränzer B, Nadmid S, Pujic P, Normand P, Luzhetskyy A: **Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters**. *Metab Eng* 2018, **49**:316-324.

38. Komatsu M, Uchiyama T, Ōmura S, Cane DE, Ikeda H: **Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism**. *Proc Nat Acad Sci U S A* 2010, **107**:2646-2651.

39. Choi KR, Cho JS, Cho IJ, Park D, Lee AY: **Markerless gene knockout and integration to express heterologous biosynthetic gene clusters in *Pseudomonas putida***. *Metab Eng* 2018, **47**:463-474.

40. Baumgart M, Unthan S, Kloß R, Radek A: ***Corynebacterium glutamicum* chassis C1*: building and testing a novel platform host for synthetic biology and industrial biotechnology**. *ACS Synth Biol* 2018, **7**:132-144.

41. Fredens J, Wang K, de la Torre D, Funke LFH, Robertson WE, Christova Y, Chia T, Schmied WT, Dunkelmann DL, Beránek V *et al.*: **Total synthesis of *Escherichia coli* with a recoded genome**. *Nature* 2019, **569**:514-518.

42. Chin JW: **Expanding and reprogramming the genetic code**. *Nature* 2017, **550**:53-60.

43. Smanski MJ, Zhou H, Claesen J, Shen B, Fischbach MA, Voigt CA: **Synthetic biology to access and expand nature's chemical diversity**. *Nat Rev Microbiol* 2016, **14**:135-149.

44. Temme K, Zhao D, Voigt CA: **Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca***. *Proc Natl Acad Sci U S A* 2012, **109**:7085-7090.

45. Shao Z, Rao G, Li C, Abil Z, Luo Y, Zhao H: **Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold**. *ACS Synth Biol* 2013, **2**:662-669.

46. Osswald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, Wenzel SC: **Modular construction of a functional artificial epothilone polyketide pathway**. *ACS Synth Biol* 2014, **3**:759-772.

47. Luo Y, Huang H, Liang J, Wang M: **Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster**. *Nat Commun* 2013, **4**:2894.

48. • Lale R, Brautaset T, Valla S: **Broad-host-range plasmid vectors for gene expression in bacteria**. In *Strain Engineering. Methods in Molecular Biology (Methods and Protocols)*. Edited by Williams J. Humana Press; 2011:327-343.
This chapter provides methods for the use of broad-host-range plasmid vectors for expression of genes in a variety of bacteria, which provides technique support for the employment of multi-chassis hosts for heterologous NP production.

49. Montiel D, Kang HS, Chang FY, Charlop-Powers Z, Brady SF: **Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters**. *Proc Natl Acad Sci U S A* 2015, **112**:8953-8958.

50. Kudla G, Murray AW, Tollervey D, Plotkin JB: **Coding-sequence determinants of gene expression in *Escherichia coli***. *Science* 2009, **324**:255-258.

51. Cho BK, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, Gao Y, Palsson BØ: **The transcription unit architecture of the *Escherichia coli* genome**. *Nat Biotechnol* 2009, **27**:1043-1049.

52. Li GW, Oh E, Weissman JS: **The anti–Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria**. *Nature* 2012, **484**:538-541.

53. Cambray G, Guimaraes JC, Arkin AP: **Evaluation of 244,000 synthetic sequences reveals design principles to optimize**

translation in *Escherichia coli*. *Nat Biotechnol* 2018, **36**:1005-0001015.

54. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM: **A "silent" polymorphism in the MDR1 gene changes substrate specificity**. *Science* 2007, **315**:525-528.

55. Zhang G, Hubalewska M, Ignatova Z: **Transient ribosomal attenuation coordinates protein synthesis and co-translational folding**. *Nat Struct Mol Biol* 2009, **16**:274-280.

56. Mittal P, Brindle J, Stephen J, Plotkin JB, Kudla G: **Codon usage influences fitness through RNA toxicity**. *Proc Natl Acad Sci U S A* 2018, **115**:8639-8644.

57. Lavner Y, Kotlar D: **Codon bias as a factor in regulating expression via translation rate in the human genome**. *Gene* 2005, **345**:127-138.

58. Nilsson G, Belasco JG, Cohen SN, von Gabain A: **Effect of premature termination of translation on mRNA stability depends on the site of ribosome release**. *Proc Natl Acad Sci U S A* 1987, **84**:4890-4894.

59. Ongley SE, Bian X, Neilan BA, Müller R: **Recent advances in the heterologous expression of microbial natural product biosynthetic pathways**. *Nat Prod Rep* 2013, **30**:1121-1138.

60. Zhang JJ, Tang X, Zhang M, Nguyen D, Moore BS: **Broad-host-range expression reveals native and host regulatory elements that influence heterologous antibiotic production in Gram-negative bacteria**. *mBio* 2017, **8**:e01291-17.

61. Berg OG, Kurland CG: **Growth rate-optimised tRNA abundance and codon usage**. *J Mol Biol* 1997, **270**:544-550.

62. Gustafsson C, Govindarajan S, Minshull J: **Codon bias and heterologous protein expression**. *Trends Biotechnol* 2004, **22**:346-353.

63. Lanza AM, Curran KA, Rey LG, Alper HS: **A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae***. *BMC Syst Biol* 2014, **8**:33-42.

64. Gould N, Hendy O, Papamichail D: **Computational tools and algorithms for designing customized synthetic genes**. *Front Bioeng Biotechnol* 2014, **2**:41.

65. Cobb RE, Wang Y, Zhao H: **High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system**. *ACS Synth Biol* 2015, **4**:723-728.

66. Tong Y, Charusanti P, Zhang L, Weber T, Lee SY: **CRISPR-Cas9 based engineering of *Actinomycetal* genomes**. *ACS Synth Biol* 2015, **4**:1020-1029.

67. •• Zhang MM, Wong FT, Wang Y, Luo S, Lim YH, Heng E, Yeo WL, Cobb RE, Enghiad B, Ang EL *et al.*: **CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters**. *Nat Chem Biol* 2017, **13**:607-609.
The authors introduced a one-step CRISPR-Cas9 knock-in strategy to activate multiple silent BGCs in *Streptomycetes* and trigger production of unique metabolites, including a novel pentangular type II polyketide. This breakthrough strategy complements existing activation approaches and facilitates discovery of more NPs.

68. •• Jiang W, Zhu TF: **Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments**. *Nat Protoc* 2016, **11**:960-975.
The authors report an optimized protocol that potentially simplifies and accelerates efforts to isolate and clone large gene clusters from micro-organisms by Cas9-assisted targeting of chromosome segments.

69. Jiang W, Zhao X, Gabrieli T, Lou C, Ebenstein Y, Zhu TF: **Cas9-assisted targeting of chromosome segments CATCH enables one-step targeted cloning of large gene clusters**. *Nat Commun* 2015, **6**:8101 http://dx.doi.org/10.1038/ncomms9101.

70. Eppelmann K, Doekel S, Marahiel MA: **Engineered biosynthesis of the peptide antibiotic bacitracin in the surrogate host *Bacillus subtilis***. *J Biol Chem* 2001, **276**:34824-34831.

71. Eustáquio AS, Gust B, Galm U, Li SM, Chater KF, Heide L: **Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters**. *Appl Environ Microbiol* 2005, **71**:2452-2459.

72. •• Nah HJ, Pyeon HR, Kang SH, Choi SS, Kim ES: **Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species**. *Front Microbiol* 2017, **8**:394-408.
The authors summarize a technique for producing a great amount of actinomycetes NP through BGC heterologous expression systems as well as recent strategies specialized for large NP BGCs in *Streptomyces* heterologous hosts.

73. Zaburannyi N, Rabyk M, Ostash B, Fedorenko V, Luzhetskyy A: **Insights into naturally minimised *Streptomyces albus* J1074 genome**. *BMC Genomics* 2014, **15**:97.

74. Malcolmson SJ, Young TS, Ruby JG, Skewes-Cox P, Walsh CT: **The posttranslational modification cascade to the thiopeptide berninamycin generates linear forms and altered macrocyclic scaffolds**. *Proc Natl Acad Sci U S A* 2013, **110**:8483-8488.

75. Bian X, Plaza A, Zhang Y, Muller R: **Luminmycins A-C, cryptic natural products from *Photorhabdus luminescens* identified by heterologous expression in *Escherichia coli***. *J Nat Prod* 2012, **75**:1652-1655.

76. Binz TM, Wenzel SC, Schnell HJ, Bechthold A, Müller R: **Heterologous expression and genetic engineering of the phenalinolactone biosynthetic gene cluster by using red/ET recombineering**. *ChemBioChem* 2019, **9**:447-454.

77. Mehta SS, Samra S, Čajka T, Wancewicz B, Fahrmann JF, Fiehn O: **Mass spectral feature list optimizer (MS-FLO): a tool to minimize false positive peak reports in untargeted liquid chromatography-mass spectroscopy (LC-MS) data processing**. *Anal Chem* 2017, **89**:3250-3255.

78. Myers OD, Sumner SJ, Li S, Barnes S, Du X: **One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks**. *Anal Chem* 2017, **89**:8696-8703.

79. Clevenger KD, Bok JW, Ye R, Miley GP, Verdan MH, Velk T, Chen C, Yang K, Robey MT, Gao P *et al.*: **A scalable platform to identify fungal secondary metabolites and their gene clusters**. *Nat Chem Biol* 2017, **13**:895-901.

80. Wang GY, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J: **Novel natural products from soil DNA libraries in a *streptomycete* host**. *J Org Lett* 2000, **2**:2401-2404.

81. Courtois S, Cappellano CM, Ball M, Francou FX, Normand P, Helynck G, Martinez A, Kolvek SJ, Hopke J, Osburne MS *et al.*: **Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products**. *Appl Environ Microbiol* 2003, **69**:49-55.

82. Martinez A, Kolvek SJ, Hopke J, Tiong Yip CL, Osburne MS: **Environmental DNA fragment conferring early and increased sporulation and antibiotic production in *Streptomyces* species**. *Appl Environ Microbiol* 2005, **71**:1638-1641.

83. Lussier FX, Chambenoit O, Côté A, Hupé JF, Denis F, Juteau P, Beaudet R, Shareck FJ: **Construction and functional screening of a metagenomic library using a T7 RNA polymerase-based expression cosmid vector**. *J Ind Microbiol Biotechnol* 2011, **38**:1321-1328.

84. • McMahon MD, Guan C, Handelsman J, Thomas MG: **Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression**. *Appl Environ Microbiol* 2012, **78**:3622-3629.
The authors developed tools and methods for constructing and screening metagenomic libraries in *S. lividans* that enable discovery and characterization of bioactive clones that could not be found using *E. coli* as a host.

85. Iqbal HA, Low-Beinart L, Obiajulu JU, Brady SF: **Natural product discovery through improved functional metagenomics in *Streptomyces***. *J Am Chem Soc* 2016, **138**:9341-9344.

86. Fischer C, Lipata F, Rohr J: **The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the**

biosynthesis of the gilvocarcins. *J Am Chem Soc* 2003, **125**:7818-7819.

87. Castro JF, Razmilic V, Gomez-Escribano JP, Andrews B, Asenjo JA, Bibb MJ: **Identification and heterologous expression of the chaxamycin biosynthesis gene cluster from *Streptomyces leeuwenhoekii***. *Appl Environ Microbiol* 2015, **81**:5820-5831.

88. Pickens LB, Tang Y, Chooi YH: **Metabolic engineering for the production of natural products**. *Annu Rev Chem Biomol Eng* 2011, **2**:211-236.

89. Pérez M, Baig I, Braña AF, Salas JA: **Generation of new derivatives of the antitumor antibiotic mithramycin by altering the glycosylation pattern through combinatorial biosynthesis**. *ChemBioChem* 2008, **9**:2295-2304.

90. González A, Remsing LL, Lombó F, Fernández MJ, Prado L, Braña AF, Künzel E, Rohr J, Méndez C, Salas JA: **The mtmVUC genes of the mithramycin gene cluster in *Streptomyces argillaceus* are involved in the biosynthesis of the sugar moieties**. *Mol Gen Genet* 2001, **264**:827-835.