Engineering Microbiology 3 (2023) 100060

Contents lists available at ScienceDirect

Engineering Microbiology

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

Review

Next-generation synthetic biology approaches for the accelerated discovery of microbial natural products

Lei Li

State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, China

ARTICLE INFO

Keywords:
Natural products
Synthetic biology
Silent BGCs
Large-scale discovery
Peptide synthesis

ABSTRACT

Microbial natural products (NPs) and their derivates have been widely used in health care and agriculture during the past few decades. Although large-scale bacterial or fungal (meta)genomic mining has revealed the tremendous biosynthetic potentials to produce novel small molecules, there remains a lack of universal approaches to link NP biosynthetic gene clusters (BGCs) to their associated products at a large scale and speed. In the last ten years, a series of emerging technologies have been established alongside the developments in synthetic biology to engineer cryptic metabolite BGCs and edit host genomes. Diverse computational tools, such as antiSMASH and PRISM, have also been simultaneously developed to rapidly identify BGCs and predict the chemical structures of their products. This review discusses the recent developments and trends pertaining to the accelerated discovery of microbial NPs driven by a wide variety of next-generation synthetic biology approaches, with an emphasis on the *in situ* activation of silent BGCs at scale, the direct cloning or refactoring of BGCs of interest for heterologous expression, and the synthetic-bioinformatic natural products (syn-BNP) approach for the guided rapid access of bioactive non-ribosomal peptides.

1. Introduction

Microbial natural products (NPs) have intrinsic chemical complexity and intriguing biological profiles and have been widely used in human medicine, agriculture and animal health during the last few decades [1–3]; however, the discovery of NPs with novel structural scaffolds presents a unique challenge to the continued discovery and development of useful microbial NPs due to low production and high rediscovery rates. Antimicrobial resistance has become a critical healthcare problem worldwide that contributes to ~700,000 deaths annually [4,5]. New chemical entities with distinct modes of action are urgently needed to combat the growing challenge of antimicrobial resistance and other public problems.

The continuous development of low-cost, high-throughput DNA sequencing technology has allowed for large-scale microbial genome sequencing that has revealed highly diverse sets of biosynthetic gene clusters (BGCs) in bacteria and fungi that contain a large portion of previously uncharacterized metabolites. Complex microbes, such as *Streptomycetes* and *Eurotiomycetes* species, can have upwards of 30 BGCs within a single genome, most of which encode for previously unknown metabolites. In the last ten years, a variety of powerful bioinformatics tools, such as PRISM or antiSMASH, have been developed to rapidly predict BGCs and their encoded chemical entities [6,7]. Repositories for BGCs with known compounds, such as IMG-ABC v.5.0 and MIBiG 2.0, have also been established to further facilitate the *in silico* analysis of novel BGCs [8,9]. Meanwhile, a panel of computational frameworks, including BiG-SCAPE and BiG-SLiCE, have created new possibilities to explore large-scale biosynthetic diversity [10,11]. These improvements in genome sequencing technologies and bioinformatic analysis algorithms have led to a renaissance in the discovery of microbially bioactive NPs; however, a large number of BGCs are cryptic or silent under standard laboratory conditions, which remains a challenge. Thus, finding a method to quickly link silent BGCs to small molecules on a large scale is the key to revive NP discovery in the post-genomic era.

Advances in synthetic biology have enabled a wide variety of novel technologies that can engineer BGCs *in vitro* or *in vivo* for next-generation drug discovery from cultivatable microorganisms and metagenomes [12–14]. On one hand, CRISPR-based promoter drop-in or chemical elicitor libraries have been broadly used for large-scale BGC activation in native producers. In addition, BGC cloning or refactoring combined with heterologous expression in versatile surrogate hosts has provided a powerful pipeline to unlock the chemical potential of silent BGCs

E-mail address: lei.li@sjtu.edu.cn

https://doi.org/10.1016/j.engmic.2022.100060
Received 23 September 2022; Received in revised form 15 November 2022; Accepted 17 November 2022
Available online 19 November 2022

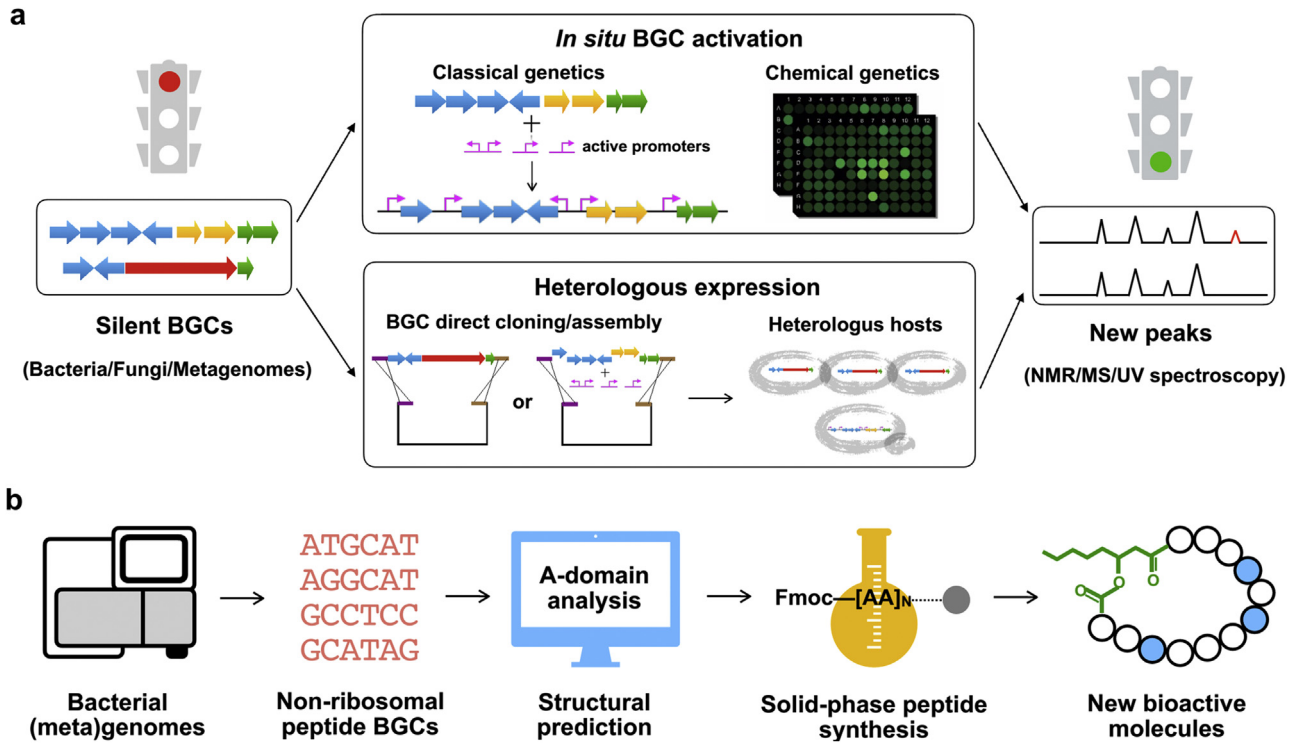

Fig. 1. Outline for the accelerated discovery of microbial natural products. **a.** Activation of silent biosynthetic gene clusters (BGCs) in native or heterologous hosts; **b**. synthetic-bioinformatic natural product (syn-BNP) approach, which includes structural prediction followed by chemical synthesis.

(Fig. 1**a**). On the other hand, the increasing accuracy of BGC structural predictions provides a potential route to rapidly access small molecules encoded by BGCs via the chemical synthesis of their predicted products in the synthetic-bioinformatic natural products approach (syn-BNP; Fig. 1**b**). The syn-BNP pipeline completely skips bacterial cultures and BGC expression and has been widely used to rapidly mine bioactive non-ribosomal peptides (NRPs). This review aims to provide an overview of the high-throughput discovery of novel drug leads driven by advanced synthetic biology approaches and concepts. The innovative discovery platforms discussed in this review could also facilitate the effective, large-scale unearthing of NP structural diversity.

2. Rapid activation of silent BGCs in natural producers

As microbial genome sequencing data accumulates, BGC prioritization will be critical to the continued discovery of diversity within both known and unknown NPs. Although the simplest strategy for BGC prioritization is to directly predict NP structures from genetic sequences, only a few classes of NPs with linear structures, including NRPs and ribosomally synthesized and post-translationally modified peptides, can be partially or completely predicted with high confidence. Thus, other strategies have been developed to access biosynthetic novelty and discover new chemical entities, such as bioactive feature-guided and resistance gene-targeting (meta)genomic mining. Interested readers are referred to some recent reviews and articles on BGC prioritization using such approaches [15–18]. Identifying potential BGCs of interest is only one of the hurdles to overcome–more than 90 % of BGCs are functionally inaccessible under laboratory conditions due to tight control in their native hosts. Thus, the development of approaches that activate the expression of silent BGCs is key to the discovery of novel microbial NPs. The progress toward this important goal is summarized below and recent trends in the strategies used for targeted BGC activation in natural producers at scale are also highlighted.

2.1. CRISPR-Cas9 knock-in strategy

During the last decade, the universal genome editing tool known as the class 2 CRISPR-Cas system, which includes CRISPR-Cas9 and CRISPR-Cas12, was developed and used for a broad spectrum of microorganisms due to its simplicity and versatility [19–21]. The Zhao group first used this CRISPR technology in 2017 to activate multiple silent BGCs in *Streptomycetes*, which encodes for the largest biosynthetic diversity by far (Fig. 2**a**) [22]. CRISPR-Cas9-mediated promoter knock-ins enable upstream regulatory regions of pathway-specific activators or main biosynthetic operons from multiple BGCs of different classes to be replaced with constitutive promoters in a single step. For example, a novel pigmented compound was identified using this method in the engineered, not wild-type *Streptomyces viridochromogenes* [22]. It is worth noting that this strategy is potentially scalable to access the biosynthetic potential of *Streptomycetes*; however, the challenge remains that the CRISPR-Cas system cannot be applied to some genetically intractable strains, such as many non-*Streptomycetes* actinobacteria and filamentous fungi.

2.2. Transcription factor decoy strategy

Most BGCs are expressed either at low levels or not at all in standard culture conditions. One potential reason is that naturally occurring global or cluster-situated regulators may bind to BGC promoter regions and repress their transcription. Based on this hypothesis, the Zhao group developed a transcription factor decoy (TFDs) strategy to unlock silent BGCs (Fig. 2**b**) [23]. Since TFDs are designed DNA molecules, they can interfere with gene expression by mimicking regulatory DNA and preventing regulators from binding to their cognate DNA targets, which results in the de-repression of cryptic metabolite BGCs. The TFD strategy successfully activated eight silent NRP synthase (NRPS) and polyketide synthase (PKS) BGCs in multiple *Actinomycetes* and characterized a novel oxazole family compound [23]. The TFD strategy is distinct from

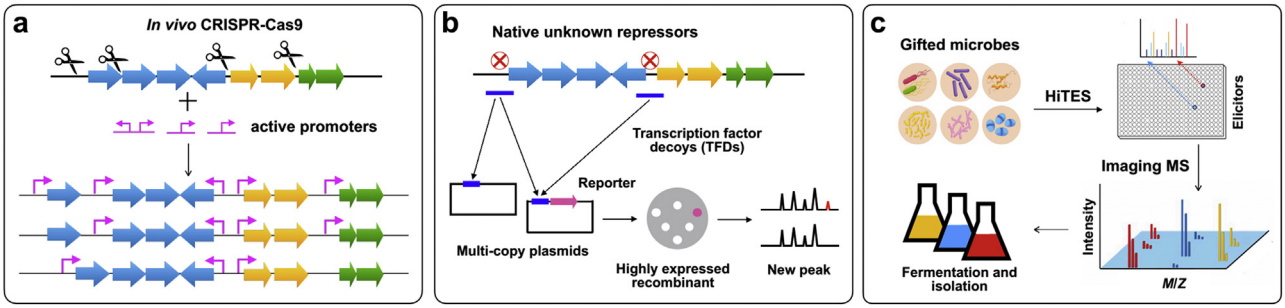

Fig. 2. *In situ* activation of silent or cryptic biosynthetic gene clusters (BGCs) in native hosts. **a**. CRISRP-Cas9-mediated knock-in of active promoters; **b**. rapid activation of silent BGCs by transcription factor decoy (TFDs) sequestration of repressors; **c.** HiTES-imaging mass spectrometry workflow for the genetics-free discovery of genetically encoded small molecules.

the CRISPR-Cas9 knock-in strategy and does not require the introduction of recombinant DNA into native hosts, thus it has potential for use in the large-scale activation of silent BGCs.

2.3. High-throughput elicitor screening (HiTES)-imaging MS (IMS) platform

Although the CRISPR-Cas9 knock-in and TFD strategies could be widely used in genetically tractable bacteria, there are a large variety of strains with genetic alterations that are too difficult for the use of either strategy. To overcome this obstacle, the Seyedsayamdost group recently developed a genetics-free method for the large-scale discovery of cryptic secondary metabolites in diverse microorganisms (Fig. 2**c**) [24]. The HiTES–IMS method combines the HiTES platform with imaging mass spectrometry (IMS) to avoid the genetic manipulation of natural producers (i.e., the introduction of a reporter system) and can access small molecules encoded by silent BGCs in a rapid and untargeted fashion. For example, the method has been used to visualize the global secondary metabolome of *Pseudomonas protegens* in response to ~500 conditions [24]. In another example, nine cryptic secondary metabolites with potentially therapeutic bioactivities were found in *Streptomycetes* and rare *Actinomycetes*, further demonstrating the power of this method [24]. This method could be widely used in both sequenced and unsequenced strains given its genetics-free characteristic. Recently, this general method was extended to activate diverse cryptic metabolites in fungi, including *Rhizoctonia solani* and *Sclerotinia sclerotiorum* [25]. This method undoubtedly holds great promise in unearthing the vast biosynthetic potential of microorganisms.

3. Large-scale cloning and refactoring of BGCs for heterologous expression

Studying microbial BGCs in their native context is often difficult. Heterologous expression provides an alternative route to access nature's abundance of small molecules. Advances in synthetic biology have enabled a variety of versatile surrogate heterologous bacterial or fungal hosts to be developed for the high-efficient expression of cloned or refactored BGCs, which reduces the need for new genetic tools to engineer native BGCs from each new species or genera [26,27]. For example, the genome-reduced *Schlegelella brevitalea* chassis exhibited alleviated cell autolysis and was better at producing cryptic metabolites from proteobacteria than the two commonly used Gram-negative chassis *Escherichia coli* and *Pseudomonas putida* [28]. Roux and Chooi developed a panel of *Aspergillus nidulans* heterologous hosts harboring a chromosomal landing pad for Cre/loxP-mediated, site-specific recombination and demonstrated the high-efficient integration of a 21 kb DNA fragment [29]. Critically, heterologous expression provides a culture-independent strategy for discovery of bioactive NPs from obligate symbionts and environmental DNA (e.g., soil or ocean metagenomes). The development of this method to identify BGCs of interest at a large scale and speed is key for the high-throughput characterization of novel NPs. Two distinct strategies have generally been applied, including direct cloning and bottom-up assembly, with the latter directly applied to refactor BGCs via the introduction of diverse regulatory elements, such as constitute promoters, ribosomal binding sites, and terminators.

3.1. Direct cloning and heterologous expression of BGCs

The direct cloning of BGCs remains an overwhelming challenge due to their repetitive sequences, large size, and/or high GC contents. With advents in recombinant DNA technologies, a variety of approaches have been developed to clone microbial BGCs, including yeast transformation-assisted recombination cloning and RecE/T-mediated linear-linear homologous recombination. Interested readers are referred to a recent review on BGC cloning [30]. These strategies use homologous recombination to capture target genomic regions, but the repetitive DNA sequences of many BGCs can interfere with recombination events, thus resulting in the unsuccessful acquisition of the correct DNA fragments. Recently, a variety of Cas12a-assisted approaches have been developed for targeted and precise BGC cloning, including CAT-FISHING and CAPTURE [31,32]. In 2021, the Zhao group developed a Cas12a-assisted precise targeted cloning method based on the *in vivo* Cre-*loxP* recombination system (CAPTURE) (Fig. 3**a**) [32]. Cre-*loxP* site-specific recombination was used instead of homologous recombination to create circular DNA molecules using the CAPTURE method, which allows for the cloning of targeted BGCs with repetitive regions. By combining Cas12a digestion and T4 polymerase exo + fill-in DNA assembly, large (up to 113 kb) and high-GC BGCs can be successfully captured into vectors in only 3–4 days. Using this robust approach, 47 BGCs from *Actinobacteria* and *Bacilli* were cloned with ~100 % efficiency and 15 new small molecules were characterized in heterologous hosts [32]. Intriguingly, Hao et al. developed a natural competence-based large DNA fragment cloning approach in the facultative anaerobe *Streptococcus mutans* UA159 and successfully identified an anti-infiltration compound from anaerobic bacteria in human oral microbiomes [33].

3.2. Bottom-up BGC refactoring and cross-kingdom expression

Although direct cloning and heterologous expression of novel BGCs provides important access routes to underexplored microbial sources, the majority of BGCs are either expressed at a low level or not at all, even in engineered heterologous hosts. The rapid advances in synthetic biology have enabled BGC refactoring to become the generalized strategy to access these silent pathways under standard laboratory conditions. In 2022, Patel et al., developed a series of synthetic genetic elements for the cross-kingdom expression of biosynthetic pathways in Gram-negative and -positive bacteria and eukaryotes by decoupling biosynthetic potential from host-range constraints to unlock cryptic BGCs. They used computer-aided design of the targeted BGCs to identify a new class of human microbiome-derived nucleotide metabolites [34]. As the cost of

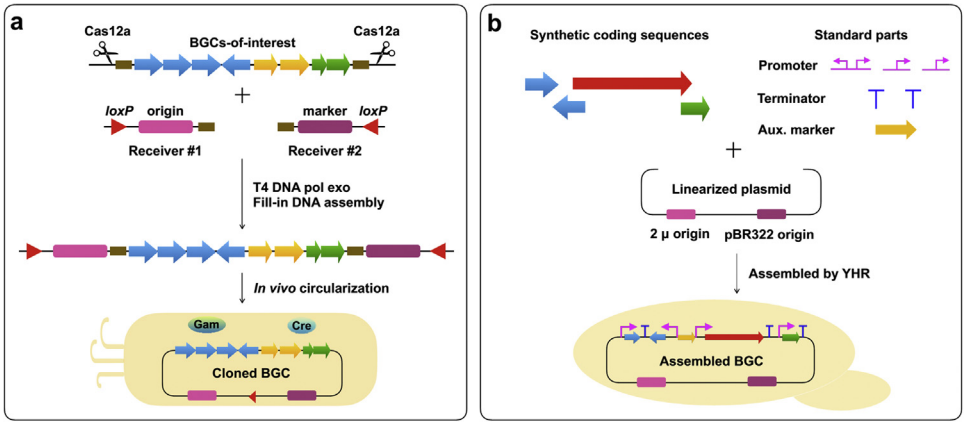

Fig. 3. Direct cloning or de novo assembly of biosynthetic gene clusters (BGCs) for heterologous expression. **a**. Cre-*loxP* recombination-based Cas12a-assisted targeted cloning; **b**. heterologous expression (Hex) platform for the discovery of fungal natural products in yeast.

gene synthesis continues to decrease and DNA assembly capacities become more powerful, it is expected that scientists will shift toward the bottom-up BGC refactoring strategy to rapidly activate BGCs of interest.

3.3. BGC refactoring for the large-scale discovery of fungal natural products

Bottom-up DNA assembly provides is a powerful method to acquire target BGCs, generate complex libraries, and even reprogram entire microbial genomes. In 2018, the Hillenmeyer and Tang groups developed a heterologous expression (HEx) discovery platform for the high-efficient expression of large-scale fungal BGCs in engineered *Saccharomyces cerevisiae* (Fig. 3**b**) [35]. This platform consisted of three parts: (1) a set of bioinformatic tools for BGC prioritization; (2) novel synthetic biology tools for BGC refactoring and assembly; and (3) a series of optimized *S. cerevisiae* chassis strains. Using the integrated platform, 41 cryptic BGCs from diverse fungal species were heterologously expressed. Of these, 22 produced detectable levels of compounds. Recently, the bottom-up assembly strategy was further improved using an automated biofoundry workflow for the high-efficient identification of bioactive fungal terpenoids using *Aspergillus oryzae* as a chassis [36,37]. Using this workflow, >180 distinct terpenoids, including a type of novel non-squalene triterpene, were identified by refactoring 38 BGCs into 208 engineered strains [36,37]. These automated high-throughput platforms provide an innovative access route to the vast number of bioactive NPs with novel structural scaffolds found in bacteria and fungi.

4. Synthetic-Bioinformatic Natural Products (syn-BNP) facilitated high-throughput discovery of bioactive peptides

Large-scale microbial genomic sequencing has revealed a large amount of previously unidentified BGCs from both cultured bugs and diverse metagenomes. The traditional methods that involves the fermentation of producer microorganisms is too laborious and time-consuming to identify BGC-encoded compounds. Additionally, even the best activation approaches can only unlock a small fraction of silent BGCs. Furthermore, when interesting bioactivity is observed, the target compounds must be isolated at a large scale for further study, which is very difficult due to the low yield found in the native producers. In fact, low fermentation titers have caused many promising antibiotics to fall by the wayside without undergoing further development. With the increased understanding of biosynthetic enzymes, more accurate structural predictions based on BGC sequences becomes possible, thus providing an alternative route to access BGC-encoded products via prediction-guided total chemical synthesis. This new methodology completely avoids the issues of isolation and scalability of target compounds. Here, we discuss examples that use this approach to discover new antimicrobial or anti-tumor chemical entities from sequenced bacterial genomes or metagenomes, with an emphasis on the de novo discovery of novel NRPs and targeted (meta)genomic mining to access biosynthetic diversity among a family of clinically relevant antibiotics (Fig. 4**a**).

4.1. De novo syn-BNP

4.1.1. Linear peptides

Large-scale genomic sequencing indicates that NRPs are one of most diverse families of complex bacterial secondary metabolites [38]. NRPs are synthesized by modular enzyme complexes (e.g., NRP synthases; NRPSs), exhibit high structural diversity, and include many clinically-used drugs, such as colistin, daptomycin and bleomycin [39]. In the last ten years, increasingly accurate bioinformatics algorithms have resulted in the successful in silico prediction of linear structures encoded by NRP BGCs based solely on the primary NRPS sequence. The predicted products can be quickly accessed by solid-phase peptide synthesis. In 2016, the Brady lab first explored NRPS BGCs from human microbiomes as test cases for the syn-BNP approach [40]. Different filters were used to generate predicted peptides with high confidence, including the removal of incomplete BGCs, at least five building blocks, and less than two PKS modules. This resulted in 25 NRP BGCs that were chosen for structural prediction, with 25 of 30 predicted linear syn-BNPs successfully synthesized. After bioactivity screening, two syn-BNPs (humimycin A and B), showed promising activities against *Staphylococcus aureus* by inhibiting lipid II flippase. Interestingly, the two peptides-related BGCs were not expressed in native hosts, which highlights the culture-independent feature of this syn-BNP approach [40]. In a second study, the approach was expanded to access 280 NRP BGCs from cultured bacteria and diverse metagenomes. Two novel peptides were identified to be active against diverse bacterial or fungal pathogens, thus further demonstrating that artificial libraries made by syn-BNPs can provide new resources to discover antimicrobial agents [41].

4.1.2. Cyclic peptides

Previous studies based on the de novo syn-BNP platform only focused on the synthesis of linear NRPs [40, 41]; however, large numbers of NRPS-derived natural products are cyclized through head-to-tail (cHT), nucleophilic amino acid side chain (cSC), or the hydroxyl group of fatty acids (cFA) (Fig. 4**b**). To better mimic natural products, the Brady lab expanded upon their initial work to NRPS-inspired cyclic structures. In the first study, 72 syn-BNP cyclic peptides inspired by 25 NRPS BGCs from the human microbiome were synthesized [42]. Five cyclic peptides elicited an obvious response in the HeLa cell-based MTT assay. In the second study, a larger scale analysis of cyclic syn-BNPs was performed using 96 NRPS BGCs from ~3000 bacterial genomes [43]. Instead of random cyclization for each cyclic syn-BNP, a panel of rules was used by examining authentic NRPs. In total, 157 of 171 designed peptides were successfully synthesized and tested against a series of important

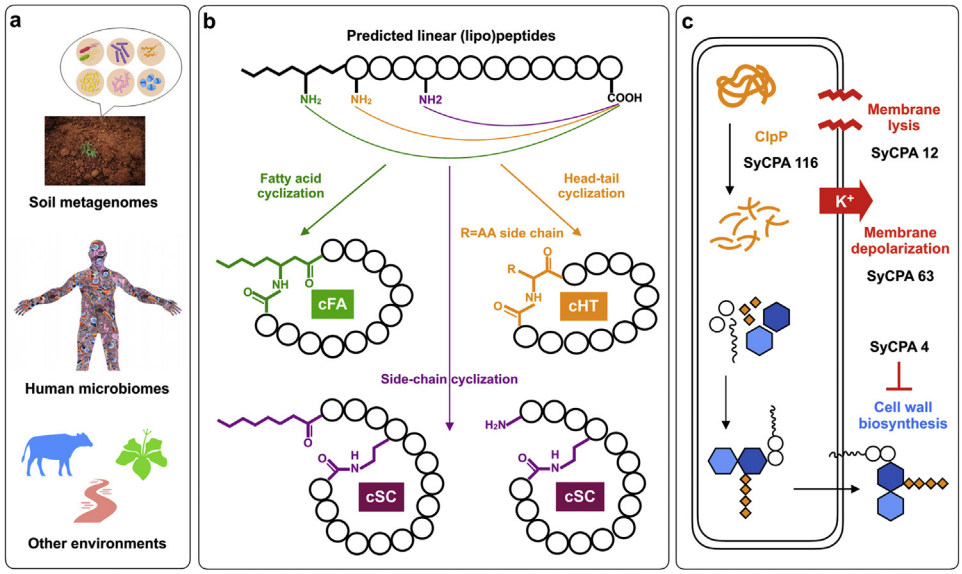

Fig. 4. De novo synthetic-bioinformatic natural products (syn-BNP). **a**. Different sources of non-ribosomal peptide biosynthetic gene clusters (BGCs); **b.** linear or cyclized syn-BNPs; **c.** syn-BNP cyclic peptide antibiotics with different modes of action.

pathogens, with nine of the cyclic syn-BNPs showed promising antibacterial activities. These bioactive syn-BNPs inhibited bacterial growth via diverse modes of action, including the inhibition of cell wall biosynthesis, membrane lysis or depolarization, and the dysregulation of the ClpP protease (Fig. 4**c**). It is expected that further improvements in bioinformatic algorithms for the structural prediction of different BGC types will expand the syn-BNP approach to access diverse classes of NPs.

4.2. Targeted syn-BNP

Congener structures for known NRPs can generally be predicted with high confidence due to our understanding of the biosynthesis of these known NRPs, including tailoring modification and cyclization. Recently, the Brady lab used this approach to unlock the biosynthetic diversity of a panel of known antibiotics and a new antibiotic family from previously sequenced bacteria and metagenomes, including ρ-aminobenzoic acid (PABA)-based antibiotics [44,45], menaquinone-binding antibiotics (MBAs) [46], and cationic lipopeptides (CATs) [47–49].

4.2.1. Menaquinone-binding antibiotics (MBAs)

Menaquinone (MK) is essential in electron transport and ATP generator in most Gram-positive bacteria. Humans are not capable of de novo MK biosynthesis, making it an appealing target for the development of therapeutically effective agents. Three known MBAs, including WAP-8294A2, WBP-29479A1, and lysocin E, have been shown to rapidly kill bacteria by MK-dependent membrane lysis [50]. In 2022, the Brady lab used two orthogonal bioinformatic approaches, including sequence-based metagenomic mining and the motif search of the predicted NRP database to discover any potential BGCs that encode for structurally diverse classes of MBAs (Fig. 5**a**) [46]. The predicted NRP database contains >36,000 linear peptides and was built with improved bioinformatic algorithms from >10,000 bacterial genomes, which provides a new source for the discovery of bioactive peptides. Six new MBA BGCs were identified and their encoded products were rapidly synthesized via two specific cyclization modes that were previously observed in known MBAs. Among these six novel MBAs, two were active against methicillin-resistant *S. aureus* (MRSA) in a mouse infection model. Notably, MBAs were first found to display potent activities against a panel of multi-drug resistant *Mycobacterium tuberculosis* strains *in vitro* and in macrophages with the same mode of action. Identifying NPs with a desired mode of action by finding a structure or substructure among a database of predicted structures provides a generalizable and high-throughput approach for the discovery of novel bioactive NPs.

4.2.2. Cationic lipopeptides

Cationic lipopeptides (CATs) are one of the most widespread and structurally diverse NPs. Since they carry the unique chemical features of at least two positively charged amino acids, CATs are generally capable of penetrating the outer membranes of Gram-negative bacteria and interacting with different anionic intracellular targets. Thus, most CATs (e.g., polymyxin/colistin, octapeptin, and tridecaptin) have been shown to be effective against a variety of MDR Gram-negative pathogens. Intriguingly, some CATs (e.g., paenipeptins) show broad-spectrum activities against both Gram-negative and Gram-positive bacteria. A few CATs, such as lysocin E and WAP-8294A2, have been shown to possess narrow-spectrum activity against Gram-positive pathogens by unique modes of action [50]. Their diverse modes of action further make CATs an appealing class of antibacterial agents for in-depth genome mining. In 2017, the Qian group first performed the large-scale genomic mining of novel bioactive CATs. They identified two new molecules (brevicidine and laterocidine) from 5585 complete bacterial genomes via traditional fermentation and showed that brevicidine was active against colistin-resistant *E. coli* in a mouse infection model [51].

Instead of the time-consuming, culture-dependent fermentation method, the Brady lab recently used the targeted syn-BNP approach to discover two novel CATs against MDR pathogens [47–49] (Fig. 5**b**). To overcome polymyxin/colistin resistance, they first set out to identify BGCs that encode for the polymyxin/colistin congeners from the predicted NRP database used in the MBA study [47], because nature may figure out how to diversify the antibiotic structures to overcome antimicrobial resistance. For instance, rifamycin congeners kanglemycins have been shown to be active against rifamycin-resistant pathogens via a distinct mechanism [52]. In total, 35 polymyxin/colistin-like BGCs were identified, with the predicted product encoded by one of these BGCs rapidly synthesized via Fmoc-based chemistry, which the authors named macolacin. Macolacin possessed potent activity against a variety of colistin-resistant clinical isolates. To facilitate its interaction with the bacterial membrane, the authors further optimized macolacin's lipid moiety and found that biphenyl-macolacin, one of the improved derivatives, was active against a panel of highly colistin-resistant pathogens, such as intrinsically colistin-resistant *Neisseria gonorrhoeae*. More importantly, biphenyl-macolacin has been shown to be active against the extensively drug-resistant *Acinetobacter baumannii*, which contained mcr-1 in a mouse thigh infection model, thus providing an important drug lead for further clinical development. In the second study, the same research group identified a distinct lipopeptide BGC via phylogenetic analysis of the condensation start domain that is involved in the incorporation

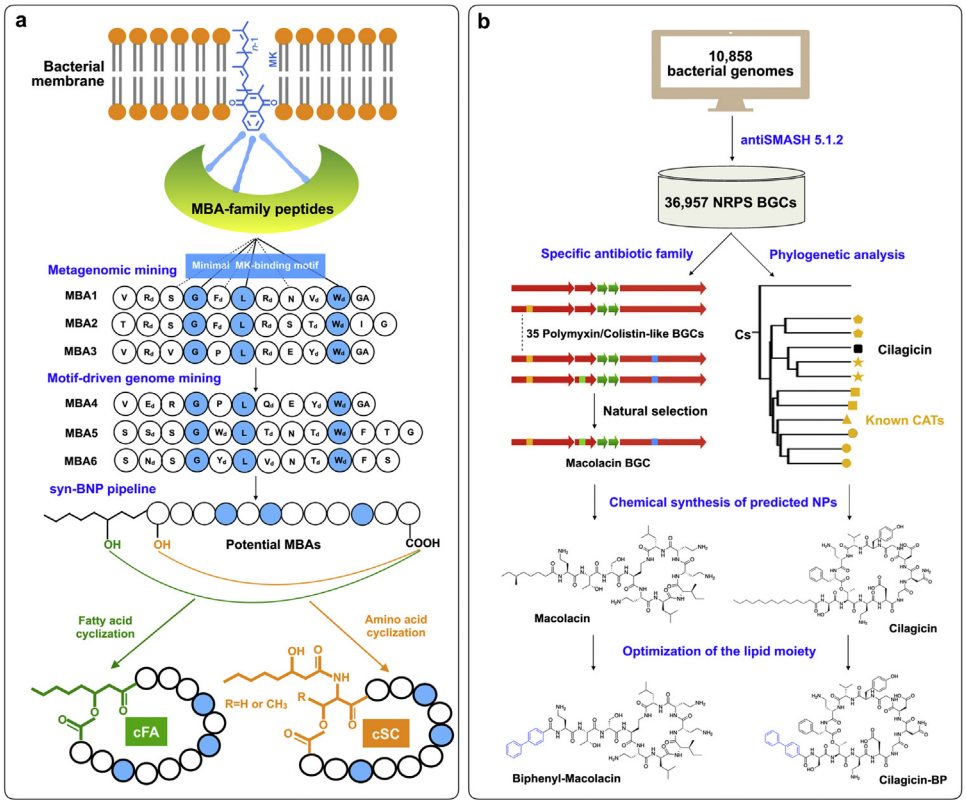

Fig. 5. Targeted synthetic-bioinformatic natural products (syn-BNP). **a.** High-throughput discovery of menaquinone-binding antibiotics (MBAs) based on two orthogonal (meta)genomic mining approaches; **b.** targeted discovery of cationic lipopeptides (CATs; e.g., macolacin and dilagicin) and synthesis of their analogs. Known CATs include polymyxin, paenibacterin, octapeptin, pelgipeptin, tridecaptin, brevicidine, synBNP1, and mucilasyn.

of the *N*-terminal lipid in lipopeptide biosynthesis [49]. The predicted product, which they named cilagicin, was also synthesized and found to be active against a series of Gram-positive pathogens. The authors also found that cilagicin causes cell death by directly binding to both undecaprenyl phosphate (C55-P) and undecaprenyl pyrophosphate (C55-PP). Notably, they did not detect cilagicin resistance in a 25-day serial passage experiment due to its dual mode of action. To reduce high levels of serum binding to cilagicin, the biphenyl moiety was also used to generate cilagicin-BP, which was proved to be effective against MDR *S. aureus* and *Streptococcus pyogenes* in a mouse study. The targeted syn-BNP approach provides an inspirational interdisciplinary and high-throughput pipeline for future antibiotic discovery to overcome antimicrobial resistance.

5. Conclusions and outlook

The study of microbial NPs can not only help to better understand their functions in the native environment, but also to discover additional drug leads with novel structural scaffolds or distinct modes of action. During the last two decades, large-scale microbial genome sequencing combined with powerful bioinformatics tools have birthed a renaissance in microbial NP discovery. In addition to the re-investigation of well-known producers, more underexplored sources, such as human microbiomes or obligate symbionts, have exhibited promising chemical potential [53,54]. Major bottlenecks in accessing the hidden chemistry of microbial NPs will be overcome by the increasing development of synthetic biology approaches, innovative culturing strategies, and improved chemical analytics [55–57].

This review focused on three synthetic biology strategies used by researchers for the high-throughput discovery of microbial NPs. Two of these strategies, *in situ* BGC activation and heterologous expression, require previously isolated chemistry to screen for potentially interesting bioactivity. Although a variety of strategies have been developed to activate silent BGCs, compound isolation and structure elucidation are still major bottlenecks in the discovery of novel NPs [58]. As an important complement, the syn-BNP strategy only relies on sequence data to directly predict a peptide structure ultimately made via chemical synthesis, which provides a culture-independent manner to rapidly access biosynthetic diversity. The high-throughput approaches discussed in this review combined with continued future developments will facilitate the discovery of unlimited chemical entities and enable us to expand the repertoire of therapeutic effective agents in the years to come.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Research on natural product discovery, bioengineering and synthetic biology in the Li group is currently supported by Shanghai Jiao Tong University Start-up Funds and Shanghai Pujiang Program (22PJ1406000).

References

[1] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019, J. Nat. Prod. 83 (2020) 770–803.
[2] M.H. Medema, T. de Rond, B.S. Moore, Mining genomes to illuminate the specialized chemistry of life, Nat. Rev. Genet. 22 (2021) 553–571.
[3] A.G. Atanasov, S.B. Zotchev, V.M. Dirsch, et al., Natural products in drug discovery: advances and opportunities, Nat. Rev. Drug Discov. 20 (2021) 200–216.
[4] Antimicrobial Resistance Collaborators, Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis, Lancet 399 (2022) 629–655.
[5] M.C. Fisher, A. Alastruey-Izquierdo, J. Berman, et al., Tackling the emerging threat of antifungal resistance to human health, Nat. Rev. Microbiol. 20 (2022) 557–571.
[6] K. Blin, S. Shaw, A.M. Kloosterman, et al., antiSMASH 6.0: improving cluster detection and comparison capabilities, Nucleic Acids Res. 49 (2021) W29–W35.
[7] M.A. Skinnider, N.J. Merwin, C.W. Johnston, et al., PRISM 3: expanded prediction of natural product chemical structures from microbial genomes, Nucleic Acids Res. 45 (2017) W49–W54.

[8] S.A. Kautsar, K. Blin, S. Shaw, et al., MIBiG 2.0: a repository for biosynthetic gene clusters of known function, Nucleic Acids Res. 48 (2020) D454–D458.
[9] K. Palaniappan, I-M.A. Chen, K. Chu, et al., IMG-ABC v.5.0: an update to the IMG/Atlas of biosynthetic gene clusters knowledgebase, Nucleic Acids Res. 48 (2020) D422–D430.
[10] J.C. Navarro-Muñoz, N. Selem-Mojica, M.W. Mullowney, et al., A computational framework to explore large-scale biosynthetic diversity, Nat. Chem. Biol. 16 (2020) 60–68.
[11] S.A. Kautsar, J.J.J. van der Hooft, D. de Ridder, et al., BiG-SLiCE: A highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters, Gigascience 10 (2021) giaa154.
[12] K. Scherlach, C. Hertweck, Mining and unearthing hidden biosynthetic potential, Nat. Commun. 12 (2021) 3864.
[13] F. Hemmerling, J. Piel, Strategies to access biosynthetic novelty in bacterial genomes for drug discovery, Nat. Rev. Drug. Discov. 21 (2022) 359–378.
[14] E. Kalkreuter, G.H. Pan, A.J. Cepeda, et al., Targeting bacterial genomes for natural product discovery, Trends Pharmacol. Sci. 41 (2020) 13–26.
[15] K.D. Bauman, K.S. Butler, B.S. Moore, et al., Genome mining methods to discover bioactive natural products, Nat. Prod. Rep. 38 (2021) 2100–2129.
[16] F. Panter, C.D. Bader, R. Müller, Synergizing the potential of bacterial genomics and metabolomics to find novel antibiotics, Chem. Sci. 12 (2021) 5994–6010.
[17] Y. Yan, Q.K. Liu, X. Zang, et al., Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action, Nature 559 (2018) 415–418.
[18] E.J. Culp, D. Sychantha, C. Hobson, et al., ClpP inhibitors are produced by a widespread family of bacterial gene clusters, Nat. Microbiol. 7 (2022) 451–462.
[19] J.Y. Wang, P. Pausch, J.A. Doudna, Structural biology of CRISPR-Cas immunity and genome editing enzymes, Nat. Rev. Microbiol. (2022), doi:10.1038/s41579-022-00739-4.
[20] Y.J. Tong, T. Weber, S.Y. Lee, CRISPR/Cas-based genome engineering in natural product discovery, Nat. Prod. Rep. 36 (2019) 1262–1280.
[21] C.M. Jiang, G.B. Lv, Y.Y. Tu, et al., Applications of CRISPR/Cas9 in the synthesis of secondary metabolites in filamentous fungi, Front. Microbiol. 12 (2021) 638096.
[22] M.Z.M. Zhang, F.T. Wong, Y.J. Wang, et al., CRISPR-Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters, Nat. Chem. Biol. 13 (2017) 607–609.
[23] B. Wang, F. Guo, S.H. Don, et al., Activation of silent biosynthetic gene clusters using transcription factor decoys, Nat. Chem. Biol. 15 (2019) 111–114.
[24] F. Xu, Y.H. Wu, C. Zhang, et al., A genetics-free method for high-throughput discovery of cryptic microbial metabolites, Nat. Chem. Biol. 15 (2019) 161–168.
[25] S.R. Lee, M.R. Seyedsayamdost, Induction of diverse cryptic fungal metabolites by steroids and channel blockers, Angew. Chem. Int. Ed. Engl. 61 (2022) e202204519.
[26] J.Q. Liu, X. Wang, G.Z. Dai, et al., Microbial chassis engineering drives heterologous production of complex secondary metabolites, Biotechnol. Adv. 59 (2022) 107966.
[27] X.F. Meng, Y. Fang, M.Y. Ding, et al., Developing fungal heterologous expression platforms to explore and improve the production of natural products from fungal biodiversity, Biotechnol. Adv. 54 (2022) 107866.
[28] J.Q. Liu, H.B. Zhou, Z.Y. Yang, et al., Rational construction of genome-reduced *Burkholderiales* chassis facilitates efficient heterologous production of natural products from proteobacteria, Nat. Commun. 12 (2021) 4347.
[29] I. Roux, Y.H. Chooi, Cre/lox-mediated chromosomal integration of biosynthetic gene clusters for heterologous expression in *Aspergillus nidulans*, ACS Synth. Biol. 11 (2022) 1186–1195.
[30] L. Li, W.H. Jiang, Y.H. Lu, New strategies and approaches for engineering biosynthetic gene clusters of microbial natural products, Biotechnol. Adv. 35 (2017) 936–949.
[31] M.D. Liang, L.S. Liu, F. Xu, et al., Activating cryptic biosynthetic gene cluster through a CRISPR-Cas12a-mediated direct cloning approach, Nucleic Acids Res. 50 (2022) 3581–3592.
[32] B. Enghiad, C.S. Huang, F. Guo, et al., Cas12a-assisted precise targeted cloning using *in vivo* Cre-lox recombination, Nat. Commun. 12 (2021) 1171.
[33] T.T. Hao, Z.J. Xie, M. Wang, et al., An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters, Nat. Commun. 10 (2019) 3665.
[34] J.R. Patel, J. Oh, S.Q. Wang, et al., Cross-kingdom expression of synthetic genetic elements promotes discovery of metabolites in the human microbiome, Cell 185 (2022) 1487–1505 e14.
[35] C.J.B. Harvey, M.C. Tang, U. Schlecht, et al., HEx: A heterologous expression platform for the discovery of fungal natural products, Sci. Adv. 4 (2018) eaar5459.
[36] Y.J. Yuan, S. Cheng, G.K. Bian, et al., Efficient exploration of terpenoid biosynthetic gene clusters in filamentous fungi, Nat. Catal. 5 (2022) 277–287.
[37] H. Tao, L. Lauterbach, G.K. Bian, et al., Discovery of non-squalene triterpenes, Nature 606 (2022) 414–419.
[38] Y. Liu, S.Y. Ding, J.Z. Shen, et al., Nonribosomal antibacterial peptides that target multidrug-resistant bacteria, Nat. Prod. Rep. 36 (2019) 573–592.
[39] R.D. Süssmuth, A. Mainz, Nonribosomal peptide synthesis-principles and prospects, Angew. Chem. Int. Ed. Engl. 56 (2017) 3770–3821.
[40] J. Chu, X. Vila-Farres, D. Inoyama, et al., Discovery of MRSA active antibiotics using primary sequence from the human microbiome, Nat. Chem. Biol. 12 (2016) 1004–1006.
[41] X. Vila-Farres, J. Chu, D. Inoyama, et al., Antimicrobials inspired by nonribosomal peptide synthetase gene clusters, J. Am. Chem. Soc. 139 (2017) 1404–1407.
[42] J. Chu, X. Vila-Farres, S.F. Brady, Bioactive synthetic-bioinformatic natural product cyclic peptides inspired by nonribosomal peptide synthetase gene clusters from the human microbiome, J. Am. Chem. Soc. 141 (2019) 15737–15741.
[43] J. Chu, B. Koirala, N. Forelli, et al., Synthetic-bioinformatic natural product antibiotics with diverse modes of action, J. Am. Chem. Soc. 142 (2020) 14158–14168.
[44] Z.Q. Wang, A. Kasper, R. Mehmood, et al., Metagenome-guided analogue synthesis yields improved Gram-negative-active albicidin- and cystobactamid-type antibiotics, Angew. Chem. Int. Ed. Engl. 60 (2021) 22172–22177.
[45] Z.Q. Wang, N. Forelli, Y. Hernandez, et al., Lapcin, a potent dual topoisomerase I/II inhibitor discovered by soil metagenome guided total chemical synthesis, Nat. Commun. 13 (2022) 842.
[46] L. Li, B. Koirala, Y. Hernandez, et al., Identification of structurally diverse menaquinone-binding antibiotics with *in vivo* activity against multidrug-resistant pathogens, Nat. Microbiol. 7 (2022) 120–131.
[47] Z.Q. Wang, B. Koirala, Y. Hernandez, et al., Discovery of Paenibacillaceae family Gram-negative-active cationic lipopeptide antibiotics using evolution-guided chemical synthesis, Org. Lett. 24 (2022) 4943–4948.
[48] Z.Q. Wang, B. Koirala, Y. Hernandez, et al., A naturally inspired antibiotic to target multidrug-resistant pathogens, Nature 601 (2022) 606–611.
[49] Z.Q. Wang, B. Koirala, Y. Hernandez, et al., Bioinformatic prospecting and synthesis of a bifunctional lipopeptide antibiotic that evades resistance, Science 376 (2022) 991–996.
[50] H. Yue, A.L. Miller, V. Khetrapal, et al., Biosynthesis, regulation, and engineering of natural products from *Lysobacter*, Nat. Prod. Rep. 39 (2022) 842–874.
[51] Y.X. Li, Z. Zhong, W.P. Zhang, et al., Discovery of cationic nonribosomal peptides as Gram-negative antibiotics through global genome mining, Nat. Commun. 9 (2018) 3273.
[52] J. Peek, M. Lilic, D. Montiel, et al., Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism, Nat. Commun. 9 (2018) 4147.
[53] A. Crits-Christoph, S. Diamond, C.N. Butterfield, et al., Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis, Nature 558 (2018) 440–444.
[54] S. Nayfach, S. Roux, R. Seshadri, et al., A genomic catalog of Earth's microbiomes, Nat. Biotechnol. 39 (2021) 499–509.
[55] T.A. Scott, J. Piel, The hidden enzymology of bacterial natural product biosynthesis, Nat. Rev. Chem. 3 (2019) 404–425.
[56] M. Miethke, M. Pieroni, T. Weber, et al., Towards the sustainable discovery and development of new antibiotics, Nat. Rev. Chem. 5 (2021) 726–749.
[57] F.I. Saldívar-González, V.D. Aldas-Bulos, J.L. Medina-Franco, et al., Natural product drug discovery in the artificial intelligence era, Chem. Sci. 13 (2021) 1526–1546.
[58] B.C. Covington, F. Xu, M.R. Seyedsayamdost, et al., A natural product chemist's guide to unlocking silent biosynthetic gene clusters, Annu. Rev. Biochem. 90 (2021) 763–788.