Synthetic and Systems Biotechnology 6 (2021) 20–22

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.sciencedirect.com/journal/synthetic-and-systems-biotechnology

Expression of fungal biosynthetic gene clusters in *S. cerevisiae* for natural product discovery

Zihe Liu [a], Zhenquan Lin [a], Jens Nielsen [a,b,c,d,*]

[a] College of Life Science and Technology, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, 100029, Beijing, China
[b] Department of Biology and Biological Engineering, Chalmers University of Technology, SE412 96, Gothenburg, Sweden
[c] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2800, Lyngby, Denmark
[d] BioInnovation Institute, Ole Maaløes Vej 3, DK2200, Copenhagen N, Denmark

ARTICLE INFO

Keywords:
Natural product discovery
Cross-species engineering
Spliceosome
High-throughput BGC cloning
Synthetic biology
Yeast

ABSTRACT

Fungi are well known for production of antibiotics and other bioactive secondary metabolites, that can be served as pharmaceuticals, therapeutic agents and industrially useful compounds. However, compared with the characterization of prokaryotic biosynthetic gene clusters (BGCs), less attention has been paid to evaluate fungal BGCs. This is partially because heterologous expression of eukaryotic gene constructs often requires replacement of original promoters and terminators, as well as removal of intron sequences, and this substantially slow down the workflow in natural product discovery. It is therefore of interest to investigate the possibility and effectiveness of heterologous expression and library screening of intact BGCs without refactoring in industrial friendly microbial cell factories, such as the yeast *Saccharomyces cerevisiae*. Here, we discuss the importance of developing new research directions on library screening of fungal BGCs in yeast without refactoring, followed by outlooking prominent opportunities and challenges for future advancement.

1. Introduction

Microorganisms provide us with a large number of biosynthetic gene clusters (BGCs) that produce bioactive secondary metabolites [1]. However, many of these are being produced by microorganisms that grow slowly, or are even unculturable at laboratory conditions [2]. Moreover, the majority of secondary metabolites are with complex structures, and are expressed at low levels, or even silent at most conditions [3]. Thus, extraction or chemical synthesis of natural products is often costly, and bioproduction through industrial friendly hosts is gaining increased attentions [4].

Recent microbial genome sequences studies showed that awaken the silent or cryptic BGCs would enlarge the reservoir of natural products and provide more opportunities for identifications of novel compounds [5,6]. However, the number of identified natural products is far less compared with the number of putative BGCs [7], and the low-throughput expression step, either by endogenous activation [8,9] or heterologous refactoring [10,11], is one of the rate limiting steps in the pipeline of natural product discovery. For example, efforts on activation of BGCs in native hosts have been successful in a variety of cases, however, these advancements will not be helpful for the 99% of microbial strains that are not readily cultivable at laboratory conditions [2]. Moreover, compared with prokaryotic BGCs, heterologous expression for eukaryotic BGCs requires refactoring each coding region of the whole gene clusters, which adds additional challenges.

Fungi produce a wide range of natural products. For example, many compounds from fungi possess antimicrobial activities such as the beta-lactam antibiotics penicillin and cephalosporins, as well as the antifungal griseofulvin [12]. However, compared with those from bacteria, less attentions have been paid to investigate fungal BGCs [13]. This is partially because heterologous refactoring of eukaryotic BGCs often requires replacement with native promoters and terminators, as well as removal of intron sequences [14]. Thus, it will substantially speed up the pipeline of natural product discovery if we could develop engineered yeast cell factories having improved intron splicing and recognition of promoters from target fungi.

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author. College of Life Science and Technology, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, 100029, Beijing, China.
E-mail address: nielsenj@chalmers.se (J. Nielsen).

https://doi.org/10.1016/j.synbio.2021.01.003
Received 29 October 2020; Received in revised form 13 December 2020; Accepted 18 January 2021

2. Key research directions

In order to achieve high-throughput library expression of heterologous BGCs in a cut and paste fashion without refactoring (Fig. 1), there are at least three research directions that need to be optimized in *S. cerevisiae* (Fig. 2): 1) optimization of fungal BGC expression, 2) identification of target BGCs and 3) development of high-throughput BGC cloning. Powerful analytical platforms to identify novel natural products is also of great importance, but will not be covered further here as it is extensively covered in recent reviews [15–17].

2.1. Genetic elements and natural product discovery

Heterologous expression of eukaryotic BGCs usually requires removal or replacement of original transcription and translation regulation elements, such as introns, promoters, and terminators (Fig. 2). The spliceosome is host specific, and several widely used microbial cell factories such as *S. cerevisiae* generally cannot remove introns from distant fungal species [10,14,18]. This problem could be tackled by directly generating cDNAs of target BGCs [19], however, this could only be applied to expressed BGCs within the ~1% cultivable species [14]. For silent BGCs or the BGCs from the environment samples, another solution is *in vitro* assembly exons of a given gene or chemical synthesis. However, current *in silico* intron prediction tools have limitations, and mis-annotated exons, even with 1bp variation, can cause impaired protein translations and ruin the effort of the whole BGC characterization [20]. On the other hand, since a large number of fungal BGCs contain more than ten genes, it is challenging to replace promoters and terminators for each gene to generate expression cassettes, not to mention the high-throughput cloning. Thus, in order to multiplex expression of BGCs in yeast without refactoring and screen for novel natural products, it is important to engineer and evolve yeast to improve intron splicing and promoter/terminator recognition towards a given heterogeneous host [21,22]. Future research direction may therefore include expression of transcription and translation regulation elements from fungi in yeast or evolve yeast to recognize heterologous transcription units or BGCs.

2.2. Bioinformatic analysis and natural product discovery

Recent development in sequencing technologies and bioinformatic tools have greatly advanced BGC discovery. We now have tremendous genome sequences available on line for natural product discovery (Fig. 2). Databases and annotation tools are being developed and successfully applied for BGC annotation and natural product discovery. For example, we can use Bayesian statistics to perform phylogenetic analysis [23]; eukaryotic orthologous groups (KOGs) [24] and Kyoto Encyclopedia of Genes and Genome (KEGG, http://www.kegg.jp/) for annotation analysis; antiSMASH [25], BiG-SCAPE [26], SMURF [27] and

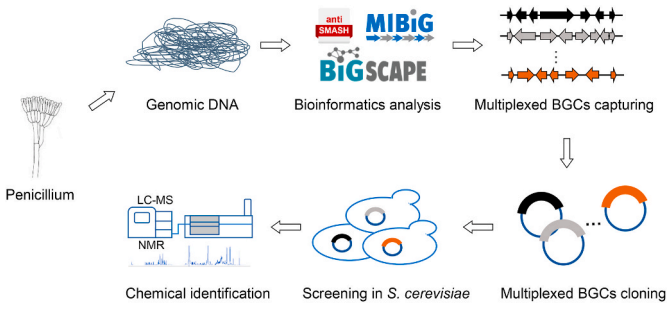

Fig. 1. Multiplexed screening of *Penicillum* BGCs in yeast. Putative BGCs can be analyzed by well-developed bioinformatics tools, and then captured through direct cloning strategies for heterologous expression and characterization in *S. cerevisiae*.

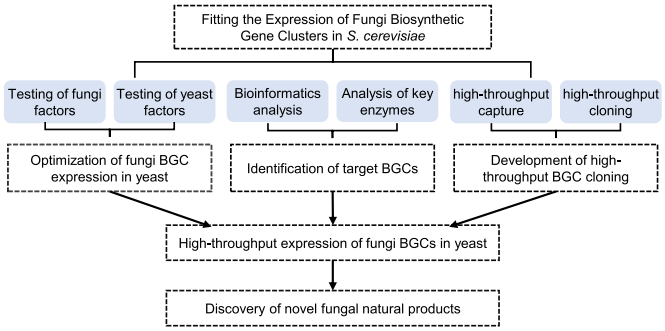

Fig. 2. Workflow of high-throughput fungal natural product discovery in yeast. Development of strategies of multiplexed expression of fungal BGCs in yeast could be divided optimization of BGC expression, identification of target BGCs and development of high-through BGC cloning.

MIBiG [28] to identify putative BGCs; Softberry (http://www.softberry.com) to predict intron sequences; Mauve [29] to perform comparative genome analysis, etc. Moreover, we could further narrow down target BGCs through literature research, MIBiG database and FungiFun [30] to analyze key enzymes of selected BGCs. So far, millions putative BGCs have been predicted [31], however, the low throughput cloning and analytic techniques has substantially slowed down the discovery novel natural products and associated BGCs [7].

2.3. Cloning techniques and natural product discovery

A number of cloning techniques have been reported for heterologous expression of target BGCs (Fig. 2). These approaches could be divided into sequence-independent library cloning methods that screen natural products from random sheared genomes, and direct cloning methods that identify novel natural products based on precise bioinformatic annotations followed by cloning target BGCs in a low throughput fashion [32]. Current heterologous BGC cloning techniques have achieved vast progress [32,33], however, the process is usually laborious and time consuming. For example, regarding the sequence-independent library cloning method, in order to cover reasonable number of BGCs, 10-20 fold-coverage of the whole genome or metagenomics needs to be generated [34]. Moreover, current direct cloning methods need to replace or remove all heterologous transcription and translation regulation elements, e.g. promoters, terminators, introns, and have been limited to capture few clusters per round [10,11,18,19]. With CRISPR tools revolutionizing the field of genome editing, it would be interesting to combine the strength of both methods, and simultaneously clone and screen all putative BGCs of a give genome without refactoring. We anticipate that this could work through optimization of high-throughput capture and cloning techniques. The step of enrichment of target BGCs after CRISPR-assisted *in vitro* cleavage may also need to be optimized.

3. Summary

The ability of heterologous expression and library screening of intact fungal BGCs without refactoring will contribute substantially to the field of synthetic biology and natural production discovery. In order to fulfill the great demand of natural products, it is necessary to develop a range of novel technologies that can speed up the fungal natural products discovery pipeline. For example, as natural products are derived from a limited number of precursor metabolites and co-factors [1], such as short chain carboxylic acids, amino acids, NADH, we need to engineer the primary metabolism to ensure efficient provision of precursor supplies. We may also need to improve the yeast capability of transcription and post-transcriptional-modification of fungal BGCs, with focus primarily on promoter recognition and intron splicing. We may also need to develop CRISPR-based high-throughput and multiplexed BGC cloning

techniques and use it for screening for novel natural products. These results will not only lead to the fundamental understanding of eukaryotic cross-species intron splicing and promoter recognition, but shed lights in general on how species can be adapted to express heterologous gene clusters. These technologies may find wide applications also for optimization of intron splicing involved in pathways leading to completely different products, and substantially enlarge the scope as well as speed up the pipeline for natural product discovery.

This paper was written to honor late Professor Arnold Demain, who for was a pioneer in natural product discovery and their production. Besides being a scientific leader, he served as a role model, mentor and friend for the senior author of this paper. We will continuously remember his fantastic scientific contributions. Thank you Arni!

CRediT authorship contribution statement

Zihe Liu: Writing - original draft, drafted the outline and wrote the manuscript. **Zhenquan Lin:** Writing - original draft, drafted the outline and wrote the manuscript. **Jens Nielsen:** Writing - original draft, drafted the outline and wrote the manuscript, Supervision, supervised the research. All authors have read and approved the final manuscript.

Declaration of competing interest

The authors indicate that they have no conflict of interest.

Acknowledgment

This work was supported by National Key Research and Development Program (2018YFA0901800), National Natural Science Foundation of China (22078012), Chinese Postdoctoral Science Foundation (2020M670115), the Novo Nordisk Foundation (NNF10CC1016517), the Knut and Alice Wallenberg Foundation and Beijing Advanced Innovation Center for Soft Matter Science and Engineering.

References

[1] Nielsen J. Cell factory engineering for improved production of natural products. Nat Prod Rep 2019;36:1233–6. https://doi.org/10.1039/C9NP00005D.
[2] Daniel R. The metagenomics of soil. Nat Rev Microbiol 2005;3:470–8. https://doi.org/10.1038/nrmicro1160.
[3] Cook TB, Pfleger BF. Leveraging synthetic biology for producing bioactive polyketides and non-ribosomal peptides in bacterial heterologous hosts. Medchemcomm 2019;10:668–81. https://doi.org/10.1039/C9MD00055K.
[4] Bian G, Deng Z, Liu T. Strategies for terpenoid overproduction and new terpenoid discovery. Curr Opin Biotechnol 2017;48:234–41. https://doi.org/10.1016/j.copbio.2017.07.002.
[5] Zhang JJ, Tang X, Moore BS. Genetic platforms for heterologous expression of microbial natural products. Nat Prod Rep 2019;36:1313–32. https://doi.org/10.1039/C9NP00025A.
[6] Hussain A, Hassan QP, Shouche YS. New approaches for antituberculosis leads from Actinobacteria. Drug Discov Today 2020;25:2335–42. https://doi.org/10.1016/j.drudis.2020.10.005.
[7] Dejong CA, Chen GM, Li H, et al. Polyketide and nonribosomal peptide retro-biosynthesis and global gene cluster matching. Nat Chem Biol 2016;12:1007–14. https://doi.org/10.1038/nchembio.2188.
[8] Zhang MM, Wong FT, Wang Y, et al. CRISPR-Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. Nat Chem Biol 2017;13:607–9. https://doi.org/10.1038/nchembio.2341.
[9] Pfannenstiel BT, Keller NP. On top of biosynthetic gene clusters: how epigenetic machinery influences secondary metabolism in fungi. Biotechnol Adv 2019;37:107345. https://doi.org/10.1016/j.biotechadv.2019.02.001.
[10] Harvey CJB, Tang M, Schlecht U, et al. HEx: a heterologous expression platform for the discovery of fungal natural products. Sci Adv 2018;4:eaar5459. https://doi.org/10.1126/sciadv.aar5459.
[11] Kim S-H, Lu W, Ahmadi MK, et al. Atolypenes, tricyclic bacterial sesterterpenes discovered using a multiplexed in vitro cas9-TAR gene cluster refactoring approach. ACS Synth Biol 2019;8:109–18. https://doi.org/10.1021/acssynbio.8b00361.
[12] Grijseels S, Nielsen JC, Nielsen JC, et al. Penicillium arizonense, a new, genome sequenced fungal species, reveals a high chemical diversity in secreted metabolites. Sci Rep 2016;6:1–13. https://doi.org/10.1038/srep35112.
[13] Park D, Swayambhu G, Pfeifer BA. Heterologous biosynthesis as a platform for producing new generation natural products. Curr Opin Biotechnol 2020;66:123–30. https://doi.org/10.1016/j.copbio.2020.06.014.
[14] Billingsley JM, DeNicola AB, Tang Y. Technology development for natural product biosynthesis in Saccharomyces cerevisiae. Curr Opin Biotechnol 2016;42:74–83. https://doi.org/10.1016/j.copbio.2016.02.033.
[15] Covington BC, McLean JA, Bachmann BO. Comparative mass spectrometry-based metabolomics strategies for the investigation of microbial secondary metabolites. Nat Prod Rep 2017;34:6–24. https://doi.org/10.1039/C6NP00048G.
[16] Pye CR, Bertin MJ, Lokey RS, et al. Retrospective analysis of natural products provides insights for future discovery trends. Proc Natl Acad Sci Unit States Am 2017;114:5601–6. https://doi.org/10.1073/pnas.1614680114.
[17] Salem MA, Perez de Souza L, Serag A, et al. Metabolomics in the context of plant natural products research: from sample preparation to metabolite analysis. Metabolites 2020;10:37. https://doi.org/10.3390/metabo10010037.
[18] Bond C, Tang Y, Li L. Saccharomyces cerevisiae as a tool for mining, studying and engineering fungal polyketide synthases. Fungal Genet Biol 2016;89:52–61. https://doi.org/10.1016/j.fgb.2016.01.005.
[19] Ishiuchi K, Nakazawa T, Ookuma T, et al. Establishing a new methodology for genome mining and biosynthesis of polyketides and peptides through yeast molecular genetics. Chembiochem 2012;13:846–54. https://doi.org/10.1002/cbic.201100798.
[20] Tsunematsu Y, Ishiuchi K, Hotta K, Watanabe K. Yeast-based genome mining, production and mechanistic studies of the biosynthesis of fungal polyketide and peptide natural products. Nat Prod Rep 2013;30:1139. https://doi.org/10.1039/c3np70037b.
[21] DeNicola A. Engineering expanded spliceosome function in Saccharomyces cerevisiae. 2018.
[22] Vassaux A, Meunier L, Vandenbol M, et al. Nonribosomal peptides in fungal cell factories: from genome mining to optimized heterologous production. Biotechnol Adv 2019;37:107449. https://doi.org/10.1016/j.biotechadv.2019.107449.
[23] Nascimento FF, Reis M Dos, Yang Z. A biologist's guide to Bayesian phylogenetic analysis. Nat Ecol Evol 2017;1:1446–54. https://doi.org/10.1038/s41559-017-0280-x.
[24] Nielsen JC, Grijseels S, Prigent S, et al. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in Penicillium species. Nat Microbiol 2017;2:17044. https://doi.org/10.1038/nmicrobiol.2017.44.
[25] Blin K, Shaw S, Steinke K, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 2019;47:W81–7. https://doi.org/10.1093/nar/gkz310.
[26] Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, et al. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol 2020;16:60–8. https://doi.org/10.1038/s41589-019-0400-9.
[27] Khaldi N, Seifuddin FT, Turner G, et al. SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol 2010;47:736–41. https://doi.org/10.1016/j.fgb.2010.06.003.
[28] Kautsar SA, Blin K, Shaw S, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 2019:1–5. https://doi.org/10.1093/nar/gkz882.
[29] Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004. https://doi.org/10.1101/gr.2289704.
[30] Priebe S, Linde J, Albrecht D, et al. FungiFun: a web-based application for functional categorization of fungal genes and proteins. Fungal Genet Biol 2011;48:353–8. https://doi.org/10.1016/j.fgb.2010.11.001.
[31] Keller NP. Fungal secondary metabolism: regulation, function and drug discovery. Nat Rev Microbiol 2019;17:167–80. https://doi.org/10.1038/s41579-018-0121-1.
[32] Lin Z, Nielsen J, Liu Z. Bioprospecting through cloning of whole natural product biosynthetic gene clusters. Front Bioeng Biotechnol 2020;8. https://doi.org/10.3389/fbioe.2020.00526.
[33] Xu W, Klumbys E, Ang EL, et al. Emerging molecular biology tools and strategies for engineering natural product biosynthesis. Metab Eng Commun 2020;10:e00108. https://doi.org/10.1016/j.mec.2019.e00108.
[34] Bok JW, Ye R, Clevenger KD, et al. Fungal artificial chromosomes for mining of the fungal secondary metabolome. BMC Genom 2015;16:343. https://doi.org/10.1186/s12864-015-1561-x.