

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Enhancement and Identification of Microbial Secondary Metabolites

*Ahmed M. Shuikan, Wael N. Hozzein,
Mohammed M. Alzharani, Maram N. Sandouka,
Sulaiman A. Al Yousef, Sulaiman A. Alharbi
and Eman Damra*

Abstract

Screening for microbial secondary metabolites (SMs) has attracted the attention of the scientific community since 1940s. In fact, since the discovery of penicillin, intensive researches have been conducted worldwide in order to detect and identify novel microbial secondary metabolites. As a result, the discovery of novel SMs has been decreased significantly by using traditional experiments. Therefore, searching for new techniques to discover novel SMs was one of the most priority objectives. However, the development and advances of omics-based techniques such as metabolomics and genomics have revealed the potential of discovering novel SMs which were coded in the microorganisms' DNA but not expressed in the lab media or might be produced in undetectable amount by detecting the biosynthesis gene clusters (BGCs) that are associated with the biosynthesis of secondary metabolites. Nowadays, the development and integration of gene editing tools such as CRISPR-Cas9 in metabolomics provide a successful platform for the identification and detection of known and novel SMs and also to increase the production of SMs.

Keywords: secondary metabolites, metabolomics, genomic, CRISPR-Cas9, secondary metabolites identification, production of secondary metabolites, microorganisms, gene editing

1. Introduction

The term secondary metabolites (SMs) was first mentioned in 1891 by A. Kossel. Microbial secondary metabolites have attracted the scientific world's attention, since the discovery of penicillin in 1940s. After that, the identification and characterization of SMs have reached the highest level between 1940s and 1960s, and this period is called as "the golden era of SMs discovery" [1, 2]. A lot of compounds have been characterized and reported during the golden era and are still utilized till now. Unfortunately, the discovery of approved novel chemical scaffolds of secondary metabolites has significantly decreased after the golden era [1]. The possible explanation of the decreasing in the SMs' identification might be due to the following: (1) using the biosynthetic modules that are used for SMs' production

by many bacteria, (2) focusing on some specific group of microorganisms such as actinobacteria, resulting in the isolation and identification of known compounds, and (3) almost 1% of the microbial community can be cultured in the laboratory [2] due to the difficulty in identifying their optimal medium compositions, resulting in the majority of SMs not being identified.

Microbial secondary metabolites play significant roles in our life. Because of SMs' unusual chemical structures, the microbial secondary metabolites show a variety of biological activities such as antimicrobial agents, antitumor agents, enzyme inhibitor, immunosuppressive agents, antiparasitic agents, herbicides, anthelmintic and food industry, etc. For instance, one of the huge successes in human medicine is the discovery of immunosuppression, such as cyclosporine A, which plays a significant role in establishing the organ transplant field.

All biochemical reactions carried out by organisms are called as metabolism and all products resulting from metabolism are called as metabolites. As a result of metabolism reactions, organisms produce primary and secondary metabolites. Primary metabolites are found in all living cells that are able to divide, while secondary metabolites are present only incidentally and do not affect the organism's life immediately. Nowadays, over 2,140,000 secondary metabolites (SMs) have been identified based on their vast diversity in function, structure, and biosynthesis (**Table 1**) [3]. The major sources of SMs are plants (about 80%) and microorganisms. Among microorganisms, bacteria, especially actinobacteria, and fungi have been reported to produce the majority of SMs that have been identified till now [4].

Microbial secondary metabolites (SMs) such as antibiotics, alkaloids, toxins, pigments, growth hormones, antitumor agents, and others are low molecular mass products that are produced by microorganisms, usually during the late growth phase. In fact, microbial secondary metabolites are not essential for the growth and development of microorganisms that produce them but are associated with some other functions such as competition, interactions, defense, and others [3, 5].

The development and advances of omics-based techniques such as genomics, metabolomics, proteomics, and trascryptomics have revealed that microorganisms have the potential to produce more secondary metabolites than were originally expected [6, 7]. These products are often coded by clustered genes present on the chromosomal DNA and rarely on plasmid DNA. In fact, most of these new SMs

Source	All known compounds	Bioactive
Plant kingdom	600,000–700,000	150,000–200,000
Microbes	Over 50,000	22,000–23,000
Higher plants	500,000–600,000	~100,000
Animal kingdom	300,000–400,000	50,000–100,000
Protozoa	Several hundreds	100–200
Vertebrates	200,000–250,000	50,000–70,000
Marine animals	20,000–25,000	7000–8000
Invertebrates	~100,000	NA
Algae, lichens	3000–5000	1500–2000
Insects, worms	8000–10,000	800–1000

NA—data not available.
Source: Bérdy [4].

Table 1.
Approximate number of identified natural metabolites.

have been predicted only by using bioinformatics analysis, which analyzes the putative SMs gene clusters in a sequenced genome. This is because, all of the new revealed SMs are not produced naturally under the lab conditions, or even though they are produced, this in very low amount that the traditional detection techniques are unable to detect them [8, 9]. Metabolomics approach aims to discover and characterize secondary metabolites in natural or engineered biosystems, and it can measure as many low molecular weight compounds as possible. Metabolomics-based technologies such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) have been identified as significant analytical methods to detect SMs produced under specific conditions [10]. The present chapter provides an overview of present-day metabolomic and genetic engineering approaches for secondary metabolites' enhancement and identification [11].

2. Genomic for screening and enhancement of SMs

2.1 Gene editing for metabolites discovery

Biosynthesis gene clusters (BGCs) are the genes associated with the biosynthesis of secondary metabolites. These BGCs include all genetic information necessary for SMs' biosynthesis, assembly, modification, and regulation of their export and transport [11]. Microorganisms' genome contains variety of cryptic or silent genes that are responsible for the production of secondary metabolites but are not expressed under laboratory conditions. It has been reported that most BGCs remain silent and cannot be fully expressed under standard laboratory conditions. These silent BGCs are potentially significant in the discovery of novel SMs [11–14].

Due to the development of genomic and bioinformatic field, we are able to access extensive sequencing data and genetic information and enable genome mining of relevant BGCs with the potential for valuable SM production [15]. Therefore, bio-synthetic biology and genetic engineering tools are now utilized for identification of novel BGCs. In fact, genetic engineering is now widely used and moving beyond traditional tools, which has opened a new era in the detection of novel secondary metabolites [16]. Genetic engineering for the production of SMs can be carried out in heterologous as well as homologous host. In fact, gene manipulation in heterologous host enables the activation of biosynthesis gene clusters (BGCs) obtained from unculturable organisms, whereas gene manipulation in homologous host allows the retention of all natural factors essential for the production of secondary metabolites [17]. While there is no single approach that will work for all genes of interest, a variety of techniques have been developed to induce the expression of these genes.

In fact, several genome techniques have emerged and are utilized in the metabolomic production field, including transcriptional activator-like effector nucleases (TALENs), zinc finger nucleases (ZFNs), and clustered regulatory interspaced short palindromic repeat (CRISPR-Cas9) [18, 19]. Each genome engineering technology has its own advantages and disadvantages (**Table 2**). For instance, ZFNs and TALENs have been successfully utilized in various microbes but still have limitation which includes the difficulty to engineer them [20]. Recently, CRISPR-Cas9 has been reported to be a significant and promising genome editing technology in the discovery and production of SMs [17, 16, 21].

2.1.1 Gene cloning

Direct cloning of the entire BGCs into the heterologous host is the most general and widely used approach for the activation of silent BGCs. Nowadays, many

		CRISPR/Cas9	Zinc finger nucleases (ZFNs)	Transcription factor-like effector nucleases (TALENs)
1	Protein engineering steps	It does not require protein engineering steps, very simple to test multiple gRNA	It requires complex steps to test gRNA	TALENs need protein engineering steps to test gRNA
2	Mode of action	It works by including double-strand breaks in target DNA or single-strand DNA nicks (Cas9 nickase)	It can induce double-strand breaks in target DNA	Induces DSBs in target DNA
3	Cloning	Not Required	Required	Required
4	Structural proteins	CRISPR consists of single monomeric protein and chimeric RNA	ZFNs work as dimeric and only protein component is required	TALENs also work as dimeric and require protein component
5	Mutation rate	Low mutation rate has been observed	High mutation rate has been observed in plants	Mutation rate is high as compared to CRISPR
6	Components	crRNA, Cas9 proteins	Zn-finger domains, nonspecific FOKI nuclease domain	Zn-finger domains, nonspecific folk nuclease domain
7	Length of target sequence (bp)	20–22	18–24	24–59
8	Target recognition efficiency	High	High	High
9	Level of experiment	Easy and very fast procedure	Complicated procedure and need for expertise in protein engineering	Relatively easy procedure
10	Methylated DNA cleavage	It can cleave methylated DNA in human cells. This aspect is of special concern for plants as this has not been much explored	Unable to do so	There are many question marks upon the capacity of TALENs to perform methylated DNA cleavage
11	Multiplexing	This is the main advantage of CRISPR, and several genes can be edited at same time. Only Cas9 is needed	Highly difficult to achieve this through ZFNs	Very difficult to obtain multiplexed genes by means of TALENs. Because it needs separate dimeric proteins specific for each target

Table 2.
Different genomic engineering techniques used in Metabolomic.

new cloning tools have been introduced, including Cas9-assisted targeting of chromosome segments (CATCH), transformation-assisted recombination (TAR), and TAR-CRISPR [22–24]. Basically, gene cloning steps include: determining the suitable heterologous host, cloning of the target BGC, transfer of the BGC into the chosen host, expression in chosen host system, and optimization of production.

Cas9-assisted targeting of chromosome segments (CATCH) is a cloning technique that utilizes the CRISPR-Cas9 gene editing system for direct BGCs cloning into the host. Comparing with traditional cloning tools such as PCR and restricted enzymes, CATCH is predicted to become a useful molecular tool for direct cloning of large gene clusters. Transformation-assisted recombination (TAR) has been used for cloning of large BGCs for about decades. However, the TAR approach is associated with a low cloning efficiency, which means it requires screening of hundreds of colonies to detect few positive clones [22, 24]. To address this challenge, TAR and CRISPR-Cas9 have been coupled resulting in a new approach called TAR-CRISPR [24]. By coupling TAR with CRISPR, a significant increase of the clone efficiency has been reported. Comparing with traditional TAR cloning, the advantages of TAR-CRISPR are that the positive clones could be achieved with secondary screening and lesser manpower and also it does not require a high experience of working with yeast [24]. In fact, the TAR-CRISTAR cloning will allow for the development of BGC cloning and SM production in the future.

2.1.2 Gene refactoring

Gene refactoring or replacement is useful not only in BGCs' activation but also for novel SMs' discovery. In fact, several silent BGCs have been refactored by replacing the BGC promoter to yield natural products such as secondary metabolites [25–28].

Another new tool in gene refactoring is multiplexed CRISPR-Cas9- and transformation-associated recombination (TAR)-mediated promoter engineering method (mCRISTAR) [21]. This new tool combines the advantages of the CRISPR-Cas9 system and TAR. It is different than the TAR-CRISPR that was discussed earlier. Comparing with TAR-CRISPR, which is a yeast-based method, basically mCRISTAR uses CRISPR-Cas9 to break the double-stranded in the promoter region of the BGC, and the fragments produced are reassembled by TAR with synthetic gene-cluster-specific promoter cassettes. Another gene refactoring tool that has aided in the faster cloning and refactoring of BGCs is the direct pathway cloning (DiPaC). Direct pathway cloning (DiPaC) depends on PCR amplification and in vitro DNA assembly for biosynthesis gene cluster capture and their expression. DiPaC was recently employed for the capture of biosynthesis gene cluster, which is small in size, followed by their activation and expression of novel natural products [29]. DiPaC was also able to successfully clone mid and large size of BGC [29].

2.1.3 Gene insertion or deletion

A large number of researches have documented the effect of gene knockout/in on BGC expression or levels of SM production. However, conventional methods of gene editing are time-intensive, while CRISPR-Cas9-based approach allows for much faster and efficient gene editing [30]. The emergence of CRISPR-Cas9 has opened up a new era in gene editing opportunities [31]. Recently, CRISPR gene editing approach has been used to insert promoter in order to activate microorganisms' SMs' production [32].

Nowadays, CRISPR-Cas9 is used to introduce promoter at multiple BGCs, and at the same time, resulting in the activation of BGCs followed by the production of SMs [32]. Multiplexed site-specific genome engineering (MSGE) was also used for multiple BGCs' editing [33]. MSGE has led to a significant increase in the secondary metabolites' production.

While, gene editing approaches provide a significant platform to manipulate the genetic machinery of microbes toward the production of novel, natural secondary

metabolites, the identification of secondary metabolites is also equally important. Metabolomics plays a significant role in the identification and characterization of secondary metabolites produced by native or genetically modified microorganisms.

2.2 Identification and characterization of secondary metabolites

Unlike all omics techniques, metabolomics often requires a broad array of instrumentation such as coulometric array detectors for detecting redox compounds, fluorescent spectrometers for detecting aromatic compounds, and ELSD for detecting lipids, whereas genomics, proteomics, or transcriptomics measurements are often conducted by a single instrument.

In general, microbial secondary metabolites' investigation is mainly conducted in two different approaches, the targeted and untargeted metabolites' identification [34]. Targeted metabolites' experiments aim to detect a specific group of compounds (about 20 compounds) that are already identified. Whereas, untargeted secondary metabolites' investigation aims to detect and identify a large scale of metabolites that are produced by microorganisms, including known and novel metabolites [35].

Over the past decade, two general technologies have emerged as the primary tools in metabolomics, the nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) [36]. Some of the common MS-based analyses are (GC-MS), (CE-MS), and (LC-MS) [37, 38]. These high-throughput tools provide a broad coverage of many classes of secondary metabolites, including amino acids, lipids, sugars, organic acids, and others.

2.2.1 Detection of secondary metabolites

Mass spectrometry (MS) is a technique that measures the mass-to-charge ratio of molecule. The principle of chromatography is to detect the retention time of the constituents that travel at different speeds under a specific condition. Therefore, various constituents take different time to pass from the inlet to the detector of the chromatography system [36].

Nuclear magnetic resonance (NMR) spectroscopy principle is based on using the magnetic properties of atomic nuclei to determine the chemical and physical properties of atoms or molecules in which they are contained. NMR's mechanism of action is that the magnetic nuclei in magnetic field absorb, resulting in reemitting of electromagnetic radiation at a specific resonance frequency depending on the magnetic properties of the isotope of atom as well as the strength of the magnetic field [36].

Both MS and NMR can be utilized to identify targeted and untargeted metabolomics. In fact, MS and NMR are often complementary techniques to each other. While NMR can be used to differentiate between structural isomers, MS provides information on the formula of the molecule [39]. Comparing to NMR, mass spectrometry is more sensitive and is able to detect a large scale of metabolites. On the other hand, nuclear magnetic resonance (NMR) spectroscopy is highly quantitative and reproducible. Unlike MS, NMR requires a larger sample amount for analysis [40, 41].

2.2.2 Data analysis

In fact, the complexity and huge amount of information that are obtained from either NMR spectroscopy or MS are considered to be one of the major challenges in metabolomics experiments [7]. The extraction of the important information that is generated by MS or NMR spectroscopy depends on using computer software

in order to organize the vast amount of data [40]. First, the raw data acquired from the NMR spectroscopy or MS must be first converted into computer formats compatible with software packages. In fact, the goal of metabolomics data analysis is to compare and identify the differences between hundreds or thousands of SMs. It is impractical to visualize changes between groups of metabolites by analyzing metabolites individually; therefore, univariate and multivariate statistical techniques can then be used to interpret the data. One of the most widely used statistical methods is the principal component analysis (PCA) [39, 42, 43]. By using PCA, the data can be simplified without losing their main features. Generally, the PCA principle is based on reducing the dimensionality of the data set, while keeping characteristics participating most to the variance. In fact, PCA provides information on multivariate differences among metabolites. It is usually conducted at the early stages of data analysis.

However, different univariate statistical tests can be used to analyze isolated metabolites such as ANOVA, nonparametric Wilcoxon signed-rank test, Kruskal-Wallis test, and the parametric Student's t-test [44]. Furthermore, other univariate analysis can be used to validate the analysis such as false discovery rate calculations or Bonferroni correction [44].

2.2.3 Metabolites' identification

Due to the development of various bioinformatics software, most of metabolites can be identified. Two types of metabolites' identification are applied, including (a) putative identification and (b) definitive identification [7]. In putative identification, one or two molecular properties are utilized for identification. However, in definitive identification, two properties such as the retention time and accurate mass and/or fragmentation mass spectrum and/or NMR spectrum are used and compared with authentic chemical standard. Comparing to putative identification, definitive identification is a more accurate form of identification, while definitive identification uses the authentic chemical standard. Usually, the definitive identification is performed after the putative identification.

Nowadays, a variety of different metabolomics' databases are available online [45, 46].

Some are spectral-based databases as well as chemical structure-based databases for metabolites' identification. Generally, spectra generated during analysis are compared with reference compounds in databases, and then similarity is assigned to each other. Even though, metabolome databases are updated daily, still significant numbers of secondary metabolites in biological system are unidentified.

Some of the common databases used in nuclear magnetic resonance (NMR) spectroscopy are METLIN (<http://metlin.scripps.edu>), the Human Metabolome Database (HMDB, <http://www.hmdb.ca>), and Biological Magnetic Resonance Datbank (<http://www.bmrb.wisc.edu/metabolomics/>), whereas commonly used databases for mass spectrometry are NIST (<http://www.nist.gov/srd/nist1a.htm>), MassBank (<http://www.massbank.jp>), the Golm Metabolite Database (GMD, <http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html>), METLI, and MMCD (<http://mmcd.nmr.fam.wisc.edu>) [45].

3. Conclusion

Microorganisms are a rich source of secondary metabolites which have significant pharmaceutical, biomedical, and food applications. Nowadays, the development and integration of gene editing tools, especially CRISPR-Cas9

(gene cloning, gene refactoring, and gene insertion or deletion) in metabolomics, provide a successful platform for the identification and detection of known and novel SMs and also to increase the production of SMs. However, there are still some challenges associated with the application of metabolomics and gene editing, including that complete identification of novel SMs requires a combination of different methods which also result in increase in the screening cost. Thus, a comprehensive and sensitive technique is the need of hour, which has the ability to provide comprehensive information of any SMs under any conditions. Also, the off-target effect of CRISPR-Cas9 is a significant problem. However, the integration of metabolomics and CRISPR-Cas9-based gene editing tools may improve the efficiency of microbial secondary metabolites' discovery.

Author details

Ahmed M. Shuikan^{1*}, Wael N. Hozzein^{2,3}, Mohammed M. Alzharani⁴, Maram N. Sandouka⁵, Sulaiman A. Al Yousef⁶, Sulaiman A. Alharbi¹ and Eman Damra¹

1 Botany and Microbiology Department, College of Science, King Saud University, Riyadh, Saudi Arabia

2 Bioproducts Research Chair, Zoology Department, College of Science, King Saud University, Riyadh, Saudi Arabia

3 Botany and Microbiology Department, Faculty of Science, Beni-Suef University, Beni Suef, Egypt

4 Biology Department, College of Science, Al Imam Mohammad Ibn Saud University, Riyadh, Saudi Arabia

5 Department of Chemistry, Université du Québec à Montréal, Montréal, Quebec, Canada

6 Department of Clinical Laboratory Sciences, College Applied Medical Science, University of Hafr Albatin, Hafar Al-batin, Saudi Arabia

*Address all correspondence to: ahmedshuikan@gmail.com

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Li JWH, Vederas JC. Drug discovery and natural products: End of an era or an endless frontier? *Biomeditsinskaya Khimiya*. 2011;**57**(2):148-160
- [2] Pelaez F. The historical delivery of antibiotics from microbial natural products - Can history repeat? *Biochemical Pharmacology*. 2006;**71**(7):981-990
- [3] McMurry JE. Organic chemistry with biological applications. In: *Secondary Metabolites: An Introduction to Natural Products Chemistry*. Stamford, USA: Cengage Learning Ltd; 2015. pp. 1016-1046
- [4] Bérdy J. Bioactive microbial metabolites. *The Journal of Antibiotics*. 2005;**58**(1):1-26
- [5] Berg M, Vanaerschot M, Jankevics A, Cuypers B, Breitling R, Dujardin J-C. LC-MS metabolomics from study design to data-analysis – Using a versatile pathogen as a test case. *Computational and Structural Biotechnology Journal*. 2013;**4**:e201301002
- [6] Putri SP, Nakayama Y, Matsuda F, Uchikata T, Kobayashi S, Matsubara A, et al. Current metabolomics: Practical applications. *Journal of Bioscience and Bioengineering*. 2013;**115**:579-589
- [7] Go EP. Database resources in metabolomics: An overview. *Journal of Neuroimmune Pharmacology*. 2010;**5**(1):18-30
- [8] Lim FY, Sanchez JF, Wang CCC, Keller NP. Toward awakening cryptic secondary metabolite gene clusters in filamentous fungi. *Methods in Enzymology*. 2012;**517**:303-324
- [9] Gottelt M, Kol S, Gomez-Escribano JP, Bibb M, Takano E. Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology*. 2010;**156**:2343-2353
- [10] Lenders J, Frédérick M, De Tullio P. Nuclear magnetic resonance: A key metabolomics platform in the drug discovery process. *Drug Discovery Today: Technologies*. 2015;**13**:39-46
- [11] Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, et al. Potential of metabolomics as a functional genomics tool. *Trends in Plant Science*. 2004;**9**(9):418-425
- [12] Tran PN, Yen MR, Chiang CY, Lin HC, Chen PY. Detecting and prioritizing biosynthetic gene clusters for bioactive compounds in bacteria and fungi. *Applied Microbiology and Biotechnology*. 2019;**103**:3277-3287
- [13] Valayil JM. Activation of microbial silent gene clusters: Genomics driven drug discovery approaches. *Journal of Biochemistry & Analytical Biochemistry*. 2016;**5**:276
- [14] Rutledge PJ, Challis GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews. Microbiology*. 2015;**13**:509-523
- [15] Blin K, Kim HU, Medema MH, Weber T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Briefings in Bioinformatics*. 2017;**11**:103-1113. DOI: 10.1093/bib/bbx146
- [16] Tong Y, Weber T, Lee SY. CRISPR/Cas-based genome engineering in natural product discovery. *Natural Product Reports*. 2019;**36**:1262-1280
- [17] Zhang MM, Wang Y, Ang EL, Zhao H. Engineering microbial hosts for production of bacterial natural

products. Natural Product Reports. 2016;**33**:963-987

[18] Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nature Biotechnology. 2007;**25**:778-785

[19] Jankele R, Svoboda P. TAL effectors: Tools for DNA targeting. Briefings in Functional Genomics. 2014;**13**:409-419

[20] Voytas DF. Plant genome engineering with sequence-specific nucleases. Annual Review of Plant Biology. 2013;**64**:327-350

[21] Kang HS, Charlop-Powers Z, Brady SF. Multiplexed CRISPR/Cas9-and TAR-mediated promoter engineering of natural product biosynthetic gene clusters in yeast. ACS Synthetic Biology. 2016;**5**:1002-1010

[22] Yamanaka K, Reynolds KA, Kersten RD, Ryan KS, Gonzalez DJ, Nizet V, et al. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. Proceedings of the National Academy of Sciences of the United States of America. 2014;**111**:1957-1962

[23] 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. Nature Communications. 2015;**6**:810

[24] Bonet B, Teufel R, Crüsemann M, Ziemert N, Moore BS. Direct capture and heterologous expression of *Salinispora* natural product genes for the biosynthesis of enterocin. Journal of Natural Products. 2015;**78**:539-542

[25] Lee NC, Larionov V, Kouprina N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal

loci from complex genomes in yeast. Nucleic Acids Research. 2015;**43**:e55-e55

[26] Horbal L, Marques F, Nadmid S, Mendes MV, Luzhetskyy A. Secondary metabolites overproduction through transcriptional gene cluster refactoring. Metabolic Engineering. 2018;**49**:299-315

[27] 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 Synthetic Biology. 2013;**2**:662-669

[28] Bauman KD, Li J, Murata K, Mantovani SM, Dahesh S, Nizet V, et al. Refactoring the cryptic streptophenazine biosynthetic gene cluster unites phenazine, polyketide, and nonribosomal peptide biochemistry. Cell Chemical Biology. 2019;**26**:724-736

[29] Greunke C, Duell ER, D'Agostino PM, Glöckle A, Lamm K, Gulder TAM. Direct pathway cloning (DiPaC) to unlock natural product biosynthetic potential. Metabolic Engineering. 2018;**47**:334-345

[30] Pohl C, Kiel JAKW, Driessen AJM, Bovenberg RAL, Nygard Y. CRISPR/Cas9 based genome editing of *Penicillium chrysogenum*. ACS Synthetic Biology. 2016;**5**:754-764

[31] Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology. 2014;**32**:347-355

[32] Zhang MM, Wong FT, Wang Y, Luo S, Lim YH, Heng E, et al. CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. Nature Chemical Biology. 2017;**13**:607

[33] Li L, Zheng G, Chen J, Ge M, Jiang W, Lu Y. Multiplexed sitespecific genome engineering for overproducing bioactive secondary metabolites in

actinomycetes. *Metabolic Engineering*. 2017;**40**:80-92

[34] Breitling R, Cenicerros A, Jankevics A, Takano E. Metabolomics for secondary metabolite research. *Metabolites*. 2013;**3**:1076-1083

[35] Wu C, Kim HK, van Wezel GP, Choi YH. Metabolomics in the natural products field—A gateway to novel antibiotics. *Drug Discovery Today: Technologies*. 2015;**13**:11-17

[36] Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: The roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews*. 2011;**40**(1):387-426

[37] Lu W, Bennett BD, Rabinowitz JD. Analytical strategies for LC-MS-based targeted metabolomics. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*. 2008;**871**(2):236-242

[38] Dettmer K, Aronov PA, Hammock BD. Mass spectrometry based metabolomics. *Mass Spectrometry Reviews*. 2007;**26**(1):51-78

[39] Alia A, Ganapathy S, de Groot HJ. Magic angle spinning (MAS) NMR: A new tool to study the spatial and electronic structure of photosynthetic complexes. *Photosynthesis Research*. 2009;**102**(2-3):415-425

[40] Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*. 2011;**6**(7):1060-1083

[41] Johnson CH, Gonzalez FJ. Challenges and opportunities of metabolomics. *Journal of Cellular Physiology*. 2012;**227**(8):2975-2981

[42] Hotelling H. Analysis of a complex of statistical variables into principal components. *Journal of Education & Psychology*. 1933;**24**:417-441

[43] Young SP, Wallace GR. Metabolomic analysis of human disease and its application to the eye. *Journal of Ocular Biology, Diseases, and Informatics*. 2009;**2**(4):235-242

[44] Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*. 2006;**2**(4):171-196

[45] Brown M, Dunn WB, Dobson P, Patel Y, Winder CL, Francis-McIntyre S, et al. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst*. 2009;**134**(7):1322-1332

[46] Midelfart A, Dybdahl A, Gribbestad IS. Metabolic analysis of the rabbit cornea by proton nuclear magnetic resonance spectroscopy. *Ophthalmic Research*. 1996;**28**(5):319-329