Proceedings of the 2nd International Conference on Biological Engineering and Medical Science (ICBioMed 2022), Part I
DOI: 10.54254/2753-8818/3/20220291

Discovery of new microbial natural products by genome mining

Rui Zhang[1,*]

[1]Lanzhou University, Lanzhou, China, 730000

[*]ruizhang2019@lzu.edu.cn

Abstract. Microbial natural products are an important source of new drug creation. As more and more microbial whole genomes are sequenced, bioinformatics analysis shows that the natural product resources contained therein are greatly underestimated. Genome mining is a new strategy for natural product discovery guided by gene cluster sequences. At the same time, it can directly associate the structure of natural products with synthesis pathways, which is convenient for biosynthesis and combinatorial biosynthesis research. This paper reviews the representative examples of successful discovery of novel natural products using genome mining strategy in recent years, which fully shows the great advantage of this strategy in exploring the potential of microbial natural metabolites. It is believed that with the continuous development and improvement of genome mining methods and technologies, it will be possible to realize the in-depth development of microbial natural product resources.This paper describes the application of combinatorial biosynthesis in new drugs, points out the existing problems and puts forward some feasible suggestions.

Keywords: genome mining, Natural products, Gene clusters, Biosynthesis, Microorganisms

1. Introduction

Natural compounds embody complex and rich diversity in chemical structure. Their unique chemical structures give many natural products the ability to specifically combine with specific targets. Although the proportion of natural products in the whole known compounds is very small, the proportion of new drugs based on them is very large, especially in the aspects of anti-bacterial and anti-tumor drugs, which reach 79% and 74% respectively. Therefore, the research and development of new drugs based on natural compounds has always been a key area of concern in the chemical and pharmaceutical circles. However, the discovery and development of new drugs from natural products is also a long and complex process [1].

The rapid development of modern biotechnology has injected new vitality into the discovery and development of new drugs. With the decoding of human genome information and the deepening of corresponding proteomic research, the functions of a large number of protein products related to various diseases have been determined. Small molecular compounds specifically bound to them can affect the biological effects of these proteins, thus providing a good starting point for the treatment of diseases. At the same time, the determination of the whole genome sequence of some pathogenic microorganisms has also been completed successively [2], finding genes closely related to their survival and clarifying the function of coding proteins, creating conditions for screening leading

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/3/20220291

compounds that specifically inhibit the growth and reproduction of pathogenic microorganisms and developing high-efficiency and low toxicity chemotherapy drugs.

The separation of active substances from biological materials from driven plants or microorganisms is the main research method of natural product pharmacy. However, as more and more natural products are identified, it becomes more and more difficult to find compounds with new structures and activities from microbial metabolites. Another effective way to increase the structural diversity of compounds is chemical synthesis [3]. As an important part of organic chemistry research, the total synthesis of complex natural products has invested the hard work of many chemists, and has made great achievements. Almost all natural products that clarify the structure can be obtained by chemical synthesis. However, because most natural products have very complex chemical structures and many chiral centers and active groups, it often requires complicated technological approaches and harsh reaction conditions to synthesize a large number of natural products by chemical methods, and their practical application value in the pharmaceutical industry is relatively limited; At the same time, due to the need for multiple steps of protection and deprotection [4], the strategy of chemical modification based on natural products is not always possible. The significance of this paper is to summarize the domestic and foreign research on the discovery of new microbial natural products through combinatorial biosynthesis, explain the existing problems and put forward feasible suggestions

2. Development status and trends of genome scanning at home and abroad

2.1. Development status

The development of human society and the improvement of living standards, the continuous emergence of new diseases, and the increasingly serious resistance of pathogenic microorganisms to clinical drugs are all calling for new structures, new active antibiotics and the improvement of the production of existing antibiotics [5]. The strategies to obtain new structures and new active compounds mainly include direct screening from nature, chemical semi synthetic modification of known antibiotics, chemical total synthesis, and the emerging combinatorialbiosynthesis based on molecular genetics.

Despite the shortcomings of high repeated screening rate, high cost and heavy workload, the separation and cultivation of microorganisms from different habitats and the separation of small molecular compounds have always been the mainstream method for the screening of new structure and new activity antibiotics. As the difficulty of screening increases, researchers all over the world have gradually turned their attention to special habitats to separate rare microorganisms and improve the opportunities to obtain new structural compounds. These special habitats include oceans, marine animals and plants, terrestrial animals and plants, mountains, polar regions, saline alkali lands, etc. Due to the vast territory and diverse habitats in China, microbiologists and chemists in China have made remarkable achievements in the mining of microorganisms and their product resources. Tens of thousands of microbial strains dominated by actinomycetes have been isolated and purified, and some of their metabolic products have been analyzed in depth for their structure and activity.

2.2. Development status in China

Abundant microbial antibiotic producing bacteria resources indicate more abundant antibiotic biosynthesis gene resources. Through whole genome sequencing, 20 and 32 possible antibiotic biosynthesis gene clusters were found in Streptomyces coeruleus and Streptomyces avermitis, respectively. Therefore, in the face of such rich microbial antibiotic producing bacteria and their biosynthetic gene resources, how to use the latest research results and means of microbial antibiotic biosynthesis to carry out rapid and efficient functional gene mining has become a top priority.

The cloning of a variety of antibiotic biosynthesis genes reveals that the structural genes, regulatory genes, resistance genes and transport genes related to the biosynthesis of specific antibiotics are clustered and concentrated in specific positions on the chromosome [6]. Therefore, in most cases, the cloning of any related gene may lead to the cloning of the whole biosynthesis gene cluster. There are

DOI: 10.54254/2753-8818/3/20220291

many cloning strategies for antibiotic biosynthetic gene clusters, mainly including mutant complementation, hybridization using heterologous DNA with high homology as a probe, cloning of resistance genes, reverse inheritance from protein separation to reverse coding DNA, heterologous expression of the whole biosynthetic gene cluster, and annexation primer amplification method designed according to the conserved region of homologous proteins. Since the cloning of actinomycin biosynthetic gene clusters, about 200 microbial antibiotic biosynthetic gene clusters have been cloned and / or sequenced [7], mainly polyketones and non ribosomal synthetic peptides, of which 80 biological gene clusters have applied for patents. The biosynthetic gene clusters cloned in China include Jinggangmycin, Nanchang mycin [8], clindamycin, erzomycin, Meilingmycin, nanoligomycin from Deng Zixin's research group, chloromycetin and safromycin a from Liu Wen and Tang Gongli's research group, nicolamycin a from Tan Huarong's research group, geldensu from Wang Yiguang's research group, and apolanamycin from Xia Huanzhang's research group, and have carried out gene function analysis and regulatory mechanism research to varying degrees.

2.3. Development status abroad

The cloning of many biosynthetic gene clusters reveals that genes responsible for the same structural synthesis in different antibiotics have high homology and specific conserved domains, such as 3-amino-5-hydroxy-p-aminobenzoic acid starting unit (corresponding to AHBA synthase gene) common to ansa antibiotics, 4,6-dehydration reaction (corresponding to 4,6-dehydratase gene) common to 6-deoxysugar synthesis, i211 Enediyne antibiotics share a 9-membered ring or 10 membered ring enediyne structure (the corresponding gene is enediyne synthase gene) F2L. Combined with the increasingly mature PCR amplification technology, degenerate primers are designed according to the conservative domain, and the total DNA of microorganisms is used as a template for amplification, it is possible to clone the biosynthetic gene cluster of antibiotics carrying the same structural group, or the biosynthetic gene cluster with coding potential. Taking the genes encoding the synthesis of specific structural groups as the starting point, the whole biosynthetic gene cluster can be obtained, and the potential new functional genes can be found through sequencing. For the possible new functional genes in the biosynthetic gene cluster of known antibiotic products, their functions can be identified in the original producing bacteria by genetic, biochemical and chemical means such as in vivo gene knockout and replenishment, in vitro enzyme catalytic analysis, and then directional transformation of antibiotics with similar structures can be carried out by combining biosynthetic means. However, the functional identification of new genes in the biosynthetic gene cluster of unknown products needs to be carried out in microorganisms that produce antibiotics with similar structures, that is, to identify and mine the function of genes directly through combinatorial biosynthesis [9].

2.4. Development trend

Since the essence of combinatorial biosynthesis is the artificial combination between different genes based on gene recombination, which is the successful application of the concept of gene recombination in the structural transformation of microbial antibiotics, its technology itself has no independent system, which can be reflected in the patents related to combinatorial biosynthesis. Therefore, in view of the abundant microbial antibiotic producing bacteria and their biosynthetic gene resources in China, the in-depth mining of new functional genes will not be limited by relevant international patents. As long as we carefully learn from the successful experience at home and abroad and rely on our own high-efficiency genetic operation system, it is entirely possible to organically combine our own resources with the structural and activity oriented transformation of antibiotics, Produce drug lead compounds and improve the efficacy of existing antibiotic drugs.

3. Research on biosynthesis and inatorial biosynthesis of natural products

Based on the above problems in the discovery and development of natural products drugs by using traditional screening and chemical synthesis methods, it also benefits from the rapid development of

DOI: 10.54254/2753-8818/3/20220291

biotechnology based on genomics and proteomics research [10], especially the DNA recombination technology in microorganisms. Some scientists have put forward the concept of "combinatorial biosynthesis" or "bio combinatorial approach", that is to use microorganisms as "cell factories" to synthesize new complex compounds through genetic control of the metabolic pathway of natural products, And the purpose of mass production is achieved by microbial fermentation on the one hand, the biosynthetic pathway of natural products is specifically genetically modified to obtain recombinant strains, natural products and their structural analogues required for biosynthesis; On the other hand, the biosynthesis genes of natural products from different sources are recombined to establish a combined new metabolic pathway in microorganisms. Therefore, the analog library composed of new natural products produced by the recombinant microbial library is conducive to the discovery and development of more valuable drugs II (Figure 1). Kosan biosciences Inc., it is a professional biological company that creates "unnatural" natural compounds based on the principle of combinatorial biosynthesis, and then provides molecular diversity for the development of new drugs. Compared with chemical synthesis, the target metabolites can be produced by the obtained recombinant strains, which can reduce production costs and reduce environmental pollution.

Genes related to the biosynthesis of natural products (structural genes, regulatory genes and resistance genes) are the genetic basis and materials for the operation of combinatorial biosynthesis technology. By 2003, incomplete statistics showed that more than 150 biosynthetic gene clusters of natural products had been cloned and sequenced 1121. Although it is still a big challenge to clone, locate and confirm the target gene cluster from the genome up to 5-10mbp (the average size of natural product producing bacteria), with the improvement of cloning strategy, the progress of DNA sequencing technology and the rapid development of genetic Informatics, it will no longer be the main difficulty of combinatorial biosynthesis research. The key to the successful application of combinatorial biosynthesis is to understand and clarify the biosynthesis and regulation mechanism of complex natural products, which is the biochemical basis of recombinant metabolic pathways in micro

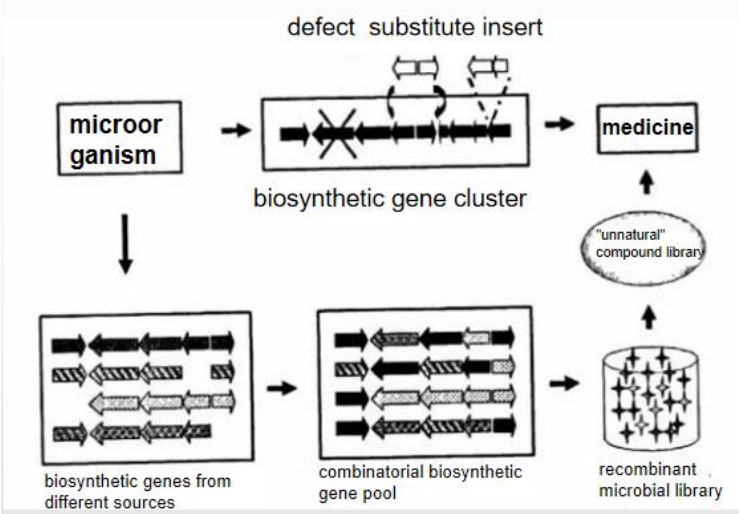

Figure 1. Research content and basic principle of combinatorial biosynthesis.

organisms [11].

4. Main contents and difficulties to be solved

4.1. Main contents

Taking nearly 100 strains of actinomycetes collected in the laboratory and provided by academician Deng Zixin of Jiaotong University, including strains isolated from Shennongjia, plant symbiosis, sponge symbiosis, and important industrial production strains of antibiotics, as research materials, by analyzing the homology and specific conservative domains of genes synthesized with the same

DOI: 10.54254/2753-8818/3/20220291

structure in different antibiotics, degenerate primers were designed to amplify the synthetic genes of key groups, Taking this as the starting point [12], we cloned the biosynthetic gene cluster of antibiotics carrying the same structural group [13], or the biosynthetic gene cluster with coding potential.

(1) Taking advantage of the differences between atypical polyketide synthase and typical polyketide synthase in domain composition and amino acid sequence, degenerate primers were designed, atypical polyketide antibiotic biosynthesis gene clusters were cloned and gene functions were clarified [14], and the genetic operating system of producer bacteria and combinatorial biosynthesis were used for structural modification.

(2) Using the conserved sequence of sugar N, n-dimethyl transferase, degenerate primers were designed to clone the gene cluster encoding its enzyme antibiotic biosynthesis and clarify the gene function.

(3) Degenerate primers were designed by using the conserved sequence of thiopeptide cyclic dehydratase to clone the biosynthetic gene cluster of thiopeptide antibiotics and clarify the gene function.

4.2. Difficulties to be solved

(1) Degenerate primer design: designing primers for the conservative region of the synthetic gene of specific structural groups is an important guarantee for screening specificity and maximum coverage, which determines the research key of this topic. When designing primers, we should try to collect all relevant homologous proteins. On the one hand, we should use different software to compare and design, on the other hand, we should design multiple pairs of degenerate primers according to the distribution of conservative regions [15], and use known antibiotic producing bacteria to evaluate the effect of primers and optimize the amplification conditions of 100pcr.

(2) Establishment of genetic operation system of new antibiotic producing strains: when the chemical isolation and identification of new structure and new activity antibiotics are carried out synchronously with the cloning of biosynthetic gene clusters, and the cloning of biosynthetic gene clusters of important antibiotic industrial production strains, an effective genetic operation system needs to be established. On the one hand, the corresponding relationship between antibiotics and cloned gene clusters can be established; on the other hand, the obtained genetic information can be directly used for the structural transformation and metabolic engineering transformation of important antibiotics. To establish a new genetic operation system, we can try conjugation transfer, protoplast transformation, electro transformation, transduction and other methods [16]. This study has established conjugation transfer and protoplast transformation systems for a variety of antibiotic producing bacteria, and have accumulated a series of related vectors and mature experience.

(3) Functional identification of new genes: after obtaining the sequence of new genes, we must first determine the catalytic function of gene coding products. For the biosynthetic gene of antibiotics with known structure [17], the correlation between the gene and antibiotic synthesis was first confirmed by gene knockout in vivo and the complementation of mutant strains, and the accumulated intermediate products were isolated, and then its catalytic function was confirmed in vitro by protein expression in vitro and using the isolated intermediate products as substrates. For the gene functions in the biosynthetic gene cluster of unknown products, on the one hand, the possible domains or conserved regions are determined according to bioinformatics analysis, and then the corresponding basic catalytic properties of the proteins expressed in vitro are analyzed [18]. On the other hand, the catalytic properties of antibiotics with similar structures are studied through in vitro reaction.

5. Conclusion

This project is a large-scale systematic attempt to clone unique types of antibiotics. The research results will not only stop at obtaining new genes for antibiotic biosynthesis, but also include deep functional mining of directional structural transformation of known antibiotics through combinatorial biosynthesis on the basis of comprehensive functional analysis.

DOI: 10.54254/2753-8818/3/20220291

Combinatorial biosynthesis technology is in the ascendant abroad, but only a few laboratories in China have tried it, which requires us to catch up. The cloning of many new genes and the mature genetic operation system of known antibiotic producing bacteria provide a guarantee for the large-scale use of combinatorial biosynthesis technology. Moreover, the new structure and new activity antibiotics obtained from this will probably become new drug lead compounds and an important auxiliary means to screen new antibiotics directly from nature.

References

[1] Michael R. Boyd. The position of intellectual property rights in drug discovery and development from natural products. Journal of Ethnopharmacology, 1996, pp.17-27.
[2] Stefania Bambini, Rino Rappuoli. The use of genomics in microbial vaccine development. Drug Discovery Today, 2009, pp.252-260.
[3] S. L. Schreiber. Target-Oriented and Diversity-Oriented Organic Synthesis in Drug Discovery. Science, 2000, pp.1964-1969.
[4] P. Hamm, L. Ellermann, P. Jauffret, G. Kaufmann. Perception and generalization of protective groups in the COSYMA system. Recueil des Travaux Chimiques des Pays-Bas, 1993, pp.503-506.
[5] Olga Genilloud. The re-emerging role of microbial natural products in antibiotic discovery. Antonie van Leeuwenhoek, 2014, pp.173-188.
[6] J. A. Robinson. Enzymes of secondary metabolism in microorganisms. Chemical Society Reviews, 1988, 383.
[7] Helge B. Bode, Rolf Müller. The Impact of Bacterial Genomics on Natural Product Research. Angewandte Chemie International Edition, 2005, pp.6828-6846.
[8] Andriy Luzhetskyy, Carmen Mendez, Jose Salas, Andreas Bechthold. Glycosyltransferases, Important Tools for Drug Design. Current Topics in Medicinal Chemistry, 2008, pp.680-709.
[9] Philipp Arendt, Jacob Pollier, Nico Callewaert, Alain Goossens. Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. The Plant Journal, 2016, pp.16-37.
[10] Mark V. Boggess, John D. Lippolis, William J. Hurkman, Clifton K. Fagerquist, Steve P. Briggs, Aldrin V. Gomes, Pier Giorgio Righetti,Kumar Bala.The need for agriculture phenotyping: “Moving from genotype to phenotype”. Journal of Proteomics, 2013, pp.20-39.
[11] Stephen W. Brown, K. Kamogawa. The Production of Biotin by Genetically Modified Micro-organisms. Biotechnology and Genetic Engineering Reviews, 1991, pp.295-326.
[12] Nobutaka Fujieda, Atsuko Satoh, Noriaki Tsuse, Kenji Kano, Tokuji Ikeda.6-S-Cysteinyl Flavin Mononucleotide-Containing Histamine Dehydrogenase from Nocardioides simplex: Molecular Cloning, Sequencing, Overexpression, and Characterization of Redox Centers of Enzyme. Biochemistry, 2004, pp.10800-10808.
[13] S. Alessandra, Eustáquio, Bertolt Gust, Shu-Ming Li, Stefan Pelzer, Wolfgang Wohlleben, Keith F. Chater, Lutz Heide. Production of 8'-Halogenated and 8'-Unsubstituted Novobiocin Derivatives in Genetically Engineered Streptomyces coelicolor Strains.Chemistry & Biology, 2004, pp.1561-1572.
[14] Xuegong Li, Xiaomin Tang, Jing Xiao, Guanghui Ma, Li Xu, Shujie Xie, Minjuan Xu, Xiang Xiao, Jun Xu. Harnessing the Potential of Halogenated Natural Product Biosynthesis by Mangrove-Derived Actinomycetes. Marine Drugs, 2013, pp.3875-3890.
[15] Yixiong Zheng, Chunjuan Li, Yu Liu, Caixia Yan, Tingting Zhang, Weijian Zhuang, Shihua Shan. Cloning and Characterization of a NBS-LRR Resistance Gene from Peanut (Arachis hypogaea L. Journal of Agricultural Science, 2012.
[16] Hedan Li, Lirong Zhang, Wei Guo, Daqing Xu. Development of a genetically engineered Escherichia coli strain for plasmid transformation in Corynebacterium glutamicum. Journal of Microbiological Methods, 2016, pp.156-160.

DOI: 10.54254/2753-8818/3/20220291

[17] L. Heide, B. Gust, C. Anderle, S. M. Li. Combinatorial Biosynthesis, Metabolic Engineering and Mutasynthesis for the Generation of New Aminocoumarin Antibiotics. Current Topics in Medicinal Chemistry, 2008, pp.667-679.
[18] Shuyi Chen, Hongchen Chen. Direct Interaction of Focal Adhesion Kinase (FAK) with Met Is Required for FAK To Promote Hepatocyte Growth Factor-Induced Cell Invasion. Molecular and Cellular Biology, 2006, pp.5155-5167.