Review

DOI: 10.5582/ddt.2018.01066

Discovery of novel bioactive natural products driven by genome mining

Zhongyue Li[1], Deyu Zhu[2], Yuemao Shen[1,*]

[1] *Key Laboratory of Chemical Biology, School of Pharmaceutical Sciences, Shandong University, Ji'nan, Shandong, China;*
[2] *School of Basic Medical Sciences, Shandong University, Ji'nan, Shandong, China.*

Summary **Microbial natural product is an important source for drug discovery. As more and more microbial genomes are sequenced, bioinformatics analysis shows that there are huge resources of novel natural products. Genome mining is a new strategy of natural product discovery based on gene cluster sequences and biosynthetic pathways. At the same time, it can directly associate the structures of natural products with synthetic pathways, and facilitate the study of biosynthesis and combinatorial biosynthesis. In this paper, the strategies of genome mining, including bioinformatics predictions, metabolomic comparisons and genetic manipulations, are reviewed, which shows a great advantage of this strategy in exploiting the potential of microbial natural products. With the development of genome mining methodology and techniques, it will be possible to realize rational exploitation of microbial natural product resources.**

Keywords: Genome mining, natural products, gene cluster, biosynthesis; microorganisms

1. Introduction

The discovery of streptomycin from *Streptomyces* has inspired the research of antibiotics from microorganisms. The discovery rate of new antibiotics has been increasing year by year, reaching its peak in the 1970s, and experienced a golden period of nearly 30 years. However, in the 1980s, due to the repeated isolation and screening of most actinomycetes, and the repeated isolation of bioactive natural products, the probability of finding new active compounds decreased rapidly and became more and more difficult. However, according to the most conservative estimates, the number of natural products found so far from any microbial strain is only 20% of the total encoded by its genome. Therefore, the decline in the rate of discovery of new active natural products is not due to the exhaustion of the biosynthetic gene resources of microbial natural products, but is limited by the methods of microbial isolation, cultivation and bioactivity screening (*1*). More importantly, since the birth of microbial natural product chemistry, the research methodology adapted has striking randomness, blindness and contingency, is insufficient in rationality. There is a lack of rational experimental design and theoretical guidance from isolation and culture of strains to extraction and separation of active components in addition to activity screening, so we can only expect to obtain the activity specified in the screening model but can't predict the structures of bioactive components.

In recent years, with decreasing sequencing costs and continuous improvements in sequencing and longer read technologies, more and more microorganism genomes have been sequenced. The emergence of these genomic information is gradually changing the research model of the whole life sciences and related disciplines. It is found that the natural product resources of conventional microorganisms are greatly underestimated (*2*). Taking bacterial genomes as an example, bioinformatics analysis shows that each strain of bacteria has the ability to synthesize 25-30 types of natural products, of which ca. 90% is still unknown. Even if only 20% of the 25-30 types of natural products produced by each strain are new structures, the number of new natural products we obtain will increase exponentially (*3*).

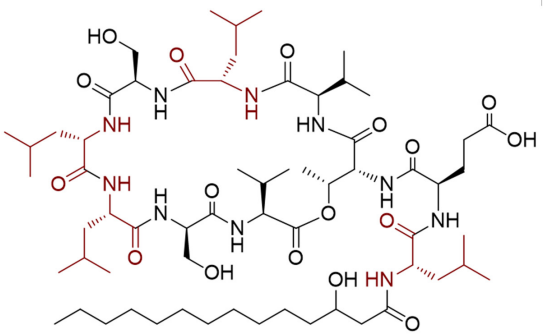
**Address correspondence to:*
Dr. Yuemao Shen, Key Laboratory of Chemical Biology, School of Pharmaceutical Sciences, Shandong University, No. 44 West Wenhua Road, Ji'nan, Shandong 250012, China.
E-mail: yshen@sdu.edu.cn

The Genomes Online Database (GOLD v.6) (*https://gold.jgi.doe.gov*) shows huge genomic data (*4*), which imposes a challenging and exciting research topic how to obtain novel natural products quickly and effectively by the method of genome mining. Compared with traditional methods, novel natural product discovery strategies based on gene cluster sequences and biosynthesis pathways have the following advantages: 1. Through gene "prediction" of natural product structure, "de-duplication" can be achieved. The possibility of finding new structures would be greatly increased by excluding known compounds. 2. Combined with transcriptional analysis and genetic manipulation of regulatory genes, the secondary metabolic potential of the strain could be released to the maximal extent. 3. The structures of natural products can be directly linked with the synthetic pathways, which can facilitate the study of biosynthesis and combinatorial biosynthesis, and further expand the structural diversity of natural products. Thus far, natural product discovery has been transformed to heavily rely on bioinformatics-based structure prediction, activation of cryptic gene clusters by genetic manipulation and heterologous expression, and potential engineering of natural product analogs (*5*).

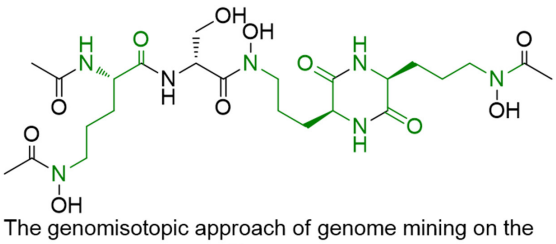
2. Discovery of novel natural products with the aid of bioinformatics prediction and comparative metabolomic analysis

Until the late 1980s, biosynthetic pathways of microbial natural products could only be inferred by feeding selectively labelled precursors, and by analyzing the MS and NMR data of the products to obtain the labelling patterns of the products. In 1984, Malpartida and Hopwood reported the cloning and heterologous expression of the biosynthetic gene cluster of actinorhodin (*6*). Katz and his colleagues from Abbott and Leadlay from Cambridge University reported independently the cloning of polyketide synthase (PKS) gene of 6-deoxyerythronolide B (6-DEB) in 1990 and 1991, respectively, which greatly changed the research mode of biosynthetic pathway of natural products. Subsequently, it was found that the nonribosomal peptides such as penicillin/cephalosporin, vancomycin, cyclosporin and daptomycin were also synthesized by the assembly lines, nonribosomal peptide synthetases (NRPSs), similar to that of 6-DEB PKS. These studies have changed our understanding of biosynthetic mechanisms of peptide natural products. With the further study of biosynthetic genes of rapamycin and FK506, it has been found that their syntheses have the compatibility of biosynthetic system with assembly line characteristic by hybrid of PKS and NRPS. This seemingly simple "assembly line enzymology" can explain the syntheses of thousands of complex and diverse natural products.

It is precisely because of the discovery of those biosynthetic pathways, the number of modules contained and the composition of the domain of each module can be determined by analyzing the sequences of PKS and NRPS genes of multiple modules, and the compound structures encoded by them can be predicted. By analyzing the predicted structures of the compounds, we can identify the assembly modules and the related physicochemical properties needed for the synthesis, so as to direct the isolation procedure and finally obtain the natural products encoded by the corresponding biosynthetic gene clusters (*7*). Specifically, the genomisotopic strategy includes discovering novel natural products on the basis of predicting structural moieties and biosynthetic precursors (*8*), which is carried out by bioinformatics analysis, subsequently, feeding isotope labelled precursors and isotope-guided isolation of the targeted natural products. E.g., orfamides A-C were isolated from *Pseudomonas fluorescens* Pf-5 by this genomisotopic approach as bioinformatics analysis indicated that these cyclopeptides contain four L-leucine residues (Figure 1) (*8*). Erythrochelin was obtained from *Saccharopolyspora erythraea* by feeding ^{14}C-L-ornithine followed by radio-LC-MS-guided isolation (Figure 1) (*9,10*). Additionally, bioinformatics analysis revealed that the *cch* gene cluster could be responsible



The isolation of orfamide A from *Pseudomonas fluorescens* Pf-5 represents a proof of concept example of genomisotopic approach of genome mining to obtain the natural products encoded by cryptic gene clusters through isotope-guided fractionation by NMR after the feeding of ^{13}C-L-leucine (red) (ref. 8).



The genomisotopic approach of genome mining on the basis of the feeding of ^{14}C-L-ornithine (green) and radio-LC-MS-guided fractionation resulted in the isolation of erythrochelin (ref. 9).

Figure 1. Two example compounds, orfamides A and erythrochelin, isolated through the genomisotopic approach.

for the biosynthesis of the tetrapeptides with the presence of a hydroxamate moiety. Hydroxamate readily conjugates with ferric ion, which produces characteristic UV absorbance. Through UV comparison of the products of *Streptomyces coelicolor* M145 mutant and wild type strains, coelichelin was obtained (*11*). UV spectra-guided isolations were also successful in the discovery of many chromophore-containing natural products, *e.g*., aspoquinolones A-D from *Aspergillus nidulans* (*12*), salinilactam A from *Salinispora tropica* (*13*), thailandamide A from *Burkholderia thailandensis* (*14*). The feasibility of structural moiety-based genome mining was demonstrated in screening a collection of over 10,000 actinomycetes for the genetic potential of producing phosphonic acids. Before this study no genome mining experiments were performed on a large number of bacterial strains (*15*).

For the natural products encoded by unknown gene clusters in genomes, it is difficult to predict their structures or physiochemical properties (*16*). Usually, it is useful to obtain these natural products through the approach of gene inactivation and metabolomic comparison of the mutant and the wild type strains by HPLC or LC-MS analysis (*17-20*). HPLC profile comparison is the most straightforward method used in genome mining. However, its limitation is obvious in that low resolution and rough identities of products. Once coupled with bioassays, this method could be feasible. WAP-8294As are cyclic depsipeptides and potent anti-MRSA agents. Their biosynthetic genes were not reported before Zhang *et al*.'s study (*21*). The genome of *Lysobacter enzymogenes* OH11 contains a large number of biosynthetic gene clusters including several hybrid PKS-NRPS ones. Disrupting one of the gene clusters containing two huge open reading frames, together encoding 12 modules of NRPS, resulted in the HPLC profile change and disappearance of the antibacterial activity of the mutant. Finally, Zhang *et al*. obtained WAP-8294A2 through HPLC- and bioassay-guided fractionation, so to characterize the biosynthetic gene cluster for this type of depsipeptides (*21*).

For the discovery of peptide natural products, an improved MS-guided genome mining method called NPP (natural product peptidogenomics) was invented by iteratively matching de novo tandem MS structures to genomics-based structures. Indeed, NPP enabled the rapid characterization of diverse ribosomal and nonribosomal peptide natural products (*22*). Recently, the MS-guided genome mining method has been evolved to molecular networking, a MS/MS data organizational approach (*23*), which can be integrated with many other methods, *e.g*., bioassays (*24*) and chemical synthesis (*25*). To simultaneously assess large numbers of closely related bacterial strains and their pan-metabolomes, molecular networking should be improved to recognize the MS/MS fragmentation patterns with the availability of biosynthetic gene clusters. This newly invented method identifies molecular families rather than individual molecules thus that enables the discovery of related molecules. Duncan *et al*. invented and applied this method to thirty five *Salinispora* strains, including 30 with draft genome sequences, which successfully resulted in the isolation and characterization of a quinomycin-type depsipeptide, namely retimycin A (*26*).

3. Activating cryptic gene clusters through genetic manipulations

In the genome of actinomycetes, many secondary metabolic gene clusters are silent. If we know how to activate these clusters, we will discover huge number of new enzymes, biological pathways and natural product resources (*27,28*). Microbial secondary metabolism has a very complex transcriptional regulatory system (*29*), probably because: 1. The biosynthetic gene cluster is composed of multiple genes of different origins containing independent regulatory elements. 2. The combined biosynthetic pathway can use different primary metabolites as the starting substances for synthesis, which requires a complex regulatory system. 3. Having more than one metabolic regulation system can help the secondary metabolism process to be sensitive to the external environment changes. In particular, the responses of actinomycetes to environmental factors are more complex than other bacteria. Many regulatory factors are involved in the secondary metabolism regulation of actinomycetes and form a complex multi-level network system. These include high-level multifunctional genes and low-level pathway-specific genes.

Activating cryptic gene clusters can be carried out through manipulating global regulations including: 1. Optimization of culture conditions including the composition of culture media, aeration volume, shape of fermenter, light condition, culture temperature, adding enzyme inhibitors (*30*), elicitors (*31*), rare earth elements (*32*), and co-culture (*33*). 2. Ribosomal engineering, a breeding method of ribosomal mutant, which uses various resistant mutants of microorganisms as screening markers to obtain highly efficient ribosomal mutants with improved secondary metabolite biosynthesis (*34,35*). The methods of activating cryptic gene clusters through manipulating pathway specific regulations include: 1. Overexpression of pathway specific transcriptional factors. Structural genes such as PKS genes and NRPS genes in the secondary metabolic gene cluster of actinomycetes are usually concentrated on 1-2 operons. The wild type promoters can be replaced by gene recombination with constitutive strong promoters such as *PermE*, which bypasses the transcriptional regulations in wild-type strains and activate or partially activate the corresponding gene clusters. 2. Heterologous expression of putative cryptic

gene clusters (*14,36*). The exploitation of secondary metabolic gene clusters requires the continuous development of transgenic cluster techniques among different organisms, including the optimization of super-hosts and the improvement of the methods for constructing/expressing multiple genes in heterologous systems (*37,38*).

4. Transcriptional regulations of secondary metabolisms in actinomycetes

The transcriptional regulations are particularly complex in actinomycetes. Multifunctional global regulators not only regulate the production of multiple secondary metabolites, but also participate in the morphological differentiation of actinomycetes, such as BldA, A-factors, Sigma factor and two-component regulatory systems often cascade into complex regulatory networks. Path-specific regulatory factors are mainly involved in the biosynthesis of specific metabolites, such as members of transcription factor families such as SARP (*Streptomyces* antibiotic regulatory protein) and LuxR. These specific regulatorys are often the only way for various global regulators to regulate specific biosynthetic pathways.

Among the numerous families of transcriptional regulatory proteins of actinomycetes (*39*), TetR (Tetracycline Repressor protein) family plays an important role in coordinating the growth and development of actinomycetes and responses to environmental factors. TetR family members are mainly involved in the regulation of antibiotic biosynthesis, efflux pumps and osmotic stress, and particularly abundant in environmentally vulnerable microorganisms, *e.g*. *Streptomyces coelicolor* A3(2) and *S. avermitilis* MA-4680 have 150 and 116 TetR members, respectively (*40*). TetRs are double domain proteins, one domain binds inducer and receives signals, another domain binds promoter DNA and transduces signals. Only the DNA binding domains are conserved among the TetRs, the elicitor binding domains are not significantly conserved, indicating the fine-tuned responses of actinomycetes to the high diversity of the inducers in the environments. After the unconserved domains binding inducers, the DNA binding domains may be altered allosterically and the ability of binding to DNA be lost, thus releasing the transcriptional repressing effects of TetRs. The TetR-binding inducers, also known as self-regulatory factors, usually are small molecular compounds that initiate secondary metabolism and morphological differentiation at the concentrations of nanomolor, and act as the hormones in eukaryotes.

Presently known self-regulatory factors of actinomycetes include 2,3-disubstituted γ-butyrolactone and its analogues including A-factor, SCB1, 2 and 3, VB-A, B, C, D and E, IM-2 and Factor I, methylenomycin furans (*41*), P1 factor (*42*), peptides including SapB, goadsporin and Factor C, siderophores and certain antibiotics (*43,44*). A-factor was originally found in *S. griseus* (*45,46*). The receptor protein of A-factor is ArpA. When A-factor binds to ArpA, the transcriptional repression of ArpA is removed (*47*). The heterologous expression of *ArpA* always produces inclusion bodies. The CprB in *S. coelicolor* A3(2) is the homologous protein of ArpA, and the crystal structure is obtained. However, the small molecular ligands of CprB have not been found (*48*).

Because self-regulatory factors are similar to the "second messengers" of environmental signals (stimuli), it may be possible to integrate the regulation of secondary metabolisms and morphological differentiations of actinomycetes (integrative regulations). This mechanism is widely used in the regulation of secondary metabolism by actinomycetes, which is consistent with that secondary metabolisms are environment sensitive. Therefore, identification of elicitors and related regulatory mechanisms is another way to activate cryptic gene clusters in actinomycetes (*49*).

5. Transcriptional regulations of ansamycin biosynthesis

The ansamycin class of natural products includes the important anti-tuberculosis drug rifamycin (*50*), and the antitumor drug leads geldanamycins (*51*) and maytansinoids (*52*). These macrolactams arise from the polyketide pathway *via* multidomain modular type I PKSs using 3-amino-5-hydroxybenzoic acid (AHBA) as the starter unit. Based on the structures of AHBA-derived aromatic moieties, ansamycins are divided into benzenic and naphthalenic groups. Each group can be further classified according to the numbers and types of extender units recruited. So far, only about 200 ansamycins have been reported. By analogy to the biosynthesis of other types of type I polyketides such as macrolactones, we speculate that more ansamycin scaffolds, *e.g*. hexaketide, dodecaketide and tridecaketide backbones, are waiting for exploitation. In the existing categories, both of benzenic and naphthalenic types have been identified only in the novel ansamycins divergolides (*53,54*). Moreover, novel ansamycins with unusual extender units have been identified recently (*55*).

Although ansamycin antibiotics are important secondary metabolites of microorganisms, and rifampicin, the derivative of rifamycin, has been used for more than 50 years in clinic, we don't know much about the molecular mechanisms that regulate their biosynthesis. The current literature reports are limited to the regulation mechanism of the biosynthesis of rifamycin, ansamitocin and geldenamycin. *RifO*, *P*, *Q* in rifamycin biosynthetic gene cluster (*rif*) may be involved in the regulation of rifamycin biosynthesis,

and *rifO* may be involved in the regulation of B-factor (*56*). *RifP* encodes a transmembrane protein, a member of the DHA14 Drug:H^+ antiporter family, that pumps antibiotics out of cells using the H^+ gradient produced by respiration. The antisense gene strategy partially silenced the expression of *rifP*, which resulted in the reduction of the production of rifamycin B by 70%, confirming the proton pump function of RifP (*57*). The *rifQ* gene is highly homologous to the transcriptional suppressor gene *actII-ORF1* (*58*). The latter belongs to the TetR family of transcriptional repressor proteins, indicating that RifQ may be involved in the self-regulation of *rifPQ* operon. In addition, the protein encoded by *rif-orf3* may be involved in the negative regulation of rifamycin biosynthesis. The production of rifamycin B in *rif-orf3* knockout mutant increased by about 40% (*59*). Additionally, *gdmR1*, *gdmR2*, *orf19* and *orf20* may be involved in the regulation of geldanamycin biosynthesis, and *orf16* and *orf17* may encode resistant genes (*60*). Gene knockout and complement experiments confirmed that the LuxR family regulators GdmRI and GdmRII positively regulated the transcription of PKS genes, but did not regulate the expression of post-PKS modification genes such as *gdmN* (*61*).

Though the biosynthetic mechanism of ansamitocins has been studied extensively (*62*), much is remained elusive about the regulation of their biosynthesis. However, tremendous effort has been contributed to enhance the production of ansamitocins by optimizations of culture media and fermentation conditions (*63,64*). Indeed, there are ten possible regulatory genes in the ansamitocin biosynthetic gene cluster, which are *asm2*, *asm8*, *asm18*, *asm29*, *asm31*, *asm34*, *asm35*, *asm39*, *asm40* and *asm48*, respectively (*65*). Asm2, Asm29, and Asm34 belong to the TetR-family transcriptional regulators. Knockout of the *asm2* gene resulted in a nine-fold increase in the production of ansamitocin P3 (*66*), showing that Asm2 plays a negative regulatory role in the biosynthesis of ansamitocins. But surprisingly, overexpression of the *asm2* gene also led to 1.3-fold increase in the production of ansamitocin P3 (*67*). Asm31, Asm39, and Asm40 are sigma factor, sigma factor antagonist and anti-sigma antagonist, respectively. Specifically, Asm39 is a regulatory protein having a histidine kinase-like ATPase C-terminal domain. The constitutive overexpression of the *asm39* gene also increased the yield of ansamitocin P3, but the overexpression of either the *asm29* or the *asm34* gene did not significantly affect the yield of ansamitocin P3 (*67*). Asm8 and Asm48 belong to LuxR or LAL (Large ATP-binding regulator of the LuxR) family transcriptional regulators, depletion of the *asm8* gene resulted in the complete loss of ansamitocin production (*65*). Asm18 is homologous to the SARP family proteins. Bioinformatics analysis suggested that Asm18 is a putative transcriptional regulator, but its gene was refractory to amplification from the genomic DNA (gDNA) of *Actinosynnema pretiosum* (*67*).

Very recently, the function of the *asm18* gene was studied through overexpression in the HGF052 strain (*A. pretiosum* ATCC 31565Δ*asm19*). The results of transcription and HPLC analysis indicate that *asm18* gene encodes a positive transcriptional regulator of the *asm* gene cluster. Real-time PCR analysis showed that the overexpression of the *asm18* gene specifically increased the transcription levels of the genes involved in the biosynthesis of the starter unit 3-amino-5-hydroxybenzoic acid (AHBA) (*asm43*), polyketide assembly (*asmA*), post-PKS modification (*asm21*), as well as the transcription levels of the regulatory gene (*asm8*). Asm8 is a pathway specific LAL-type activator in ansamitocin biosynthesis. Previous study revealed its positive regulatory role in ansamitocin biosynthesis (*65*). Further investigation of the chemical constituents of *asm18*-overexpressed HGF052 strain (HGF052+pJTU824-asm18) led to the isolation of seven ansamitocin derivatives (*68*). Interestingly, these ansamitocin derivatives include maytansinol, which is an important component for the production of antibody-maytansinoid conjugates (*69*).

6. Activation of cryptic ansamycin gene clusters

Previously, through PCR screening of the AHBA synthase genes, we obtained dozens of AHBA synthase gene-positive strains. However, the corresponding ansamycins of most strains were refractory to identification under the conventional laboratory experimental conditions, indicating that the gene clusters containing the AHBA synthase gene were not expressed or just expressed at very low levels (*70*). *S.* sp. LZ35 is one of such AHBA-positive strains, and produces two different types of ansamycins, geldanamycins and hygrocins, under the conventional laboratory conditions (*71-73*). Genome sequence analysis of the LZ35 strain revealed the presence of one more ansamycin gene cluster, namely *nam* gene cluster, besides these two ones responsible for the biosynthesis of geldanamycins and hygrocins. The *pks* genes, *namA-E* within the *nam* gene cluster, contain eight modules, implying that the *nam* cluster encodes the biosynthesis of octaketide ansamycins. Bioinformatics analysis of the eight acyltransferase (AT) domains in NamA-E revealed that the possible extender units for the assembly of polyketide chain could be different from those in known octaketide ansamycins, *i.e.* geldanamycins and ansamitocins, and that the products of this gene cluster could contain a novel scaffold. However, no predicted ansamycins were obtained by large scale fermentation under various conditions. Moreover, reverse transcription (RT) PCR analysis indicated that all the genes examined were expressed poorly in the SR101 strain, a geldanamycin-

nonproducing mutant of the strain LZ35 (*73*), indicating that the *nam* gene cluster was cryptic or weakly expressed under the conventional laboratory conditions.

To induce the expression of the *nam* gene cluster, firstly promoter replacement was carried out by introducing the strong constitutive promoter *ermE** at upstream of the *pks* and amide synthase gene operon in the SR101 strain to generate the SR101*namA-ermE** strain. Fermentation and isolation of the SR101*namA-ermE** strain resulted in the identification of a tetraketide, a derivative of SY4B under usual conditions. SY4b is one of the major tetraketides accumulated in Δ*rifF* mutant (rifamycin amide synthase gene deletion) of *Amycolatopsis mediterranei* S699 (*74*). These results indicated that promoter replacement only partially activated the *nam* gene cluster. Further analysis of the *nam* gene cluster identified *nam1* gene encoding a putative positive regulator of the LuxR family proteins. Members of the LuxR family have been previously reported as the activators of polyketide biosynthesis (*75,76*). Constitutive overexpression of the *nam1* gene was carried out by being cloned into the conjugative and integrative vector pJTU824, placed under the control of the *ermE** promoter. The resulting construct was integrated into the chromosome of the SR201 strain (*73*), which is a mutant derived from the SR101 strain with the deletion of hygrocin biosynthetic genes to create the SR201*nam1*OE strain. Fermentation and isolation of the SR201*nam1*OE strain led to the isolation of three novel ansamycins, namely neoansamycins A-C (Figure 2) (*77*), which represents the first successful example of genome mining of ansamycins and illustrates the great potential for exploration of the huge reservoir of cryptic ansamycin biosynthetic gene clusters to obtain novel ansamycins (*70,78*).

Phylogenetic analyses showed that most ansamycin biosynthesis gene clusters are flanked by LuxR-type regulators, *e.g.* geldanamycins (GdmRI, GdmRII) (*61*), hygrocins (Hgc1) (*73*), divegolides (Div8) (*79*) and ansamitocins (Asm8) (*80*). LuxR-type regulators usually contain a DNA-binding helix-turn-helix domain but lack γ-butyrolactone-binding or response regulatory domains. These LuxR-type regulators are activators of ansamycin biosynthesis. However, the mechanism of action of these LuxR-type regulators is still poorly explored. The constitutive overexpression of *nam1* induced the activation of the cryptic ansamycin gene cluster *nam* in *S.* sp. LZ35, further supporting the critical role of LuxR-type regulators in positive regulation of ansamycin biosynthesis. Comparative transcriptional analysis with quantitative RT-PCR (qRT-PCR) showed that five genes *namA*, *namK*, *nam2*, *nam7* and *nam8* involved in the biosynthesis of neoansamycins were upregulated under *nam1*-overexpressing condition in comparison to the control. The upregulation of transcription of the *nam* gene cluster, coupled with the isolation of neoansamycins in the SR201*nam1*OE strain, further confirmed that Nam1 acts as a direct activator of the *nam* gene cluster. The interaction of Nam1 with the promoter regions of these genes was verified by electrophoretic mobility shift assays (EMSA) performed using purified N-terminal truncated Nam1 protein ($Nam1^{DBD}$) and the PCR-generated promoter fragments of these *nam* genes, respectively. $Nam1^{DBD}$ retarded the mobility of all these fragments, suggesting that Nam1 binds specifically to these promoter regions. In order to identify the Nam1 binding site, the promoter region fragments shown to be retarded in EMSAs were studied by DNase I footprinting analysis. Footprinting analyses revealed protected sequences of approximately 20 to 30 nucleotides in target promoters. The comparison of the protected sequences identified the consensus binding sequence and a sequence logo (*81*) that depicts the binding site. The presence of highly conserved binding sites in these *nam* genes suggested a strong relevance of this class of transcriptional regulator in the regulation of *nam* gene expression. Moreover, the knowledge of the LuxR regulators may set the stage for understanding the genetic control of ansamycin antibiotic biosynthesis regulation and provide an effective strategy to discover and improve the yields of these antibiotics.

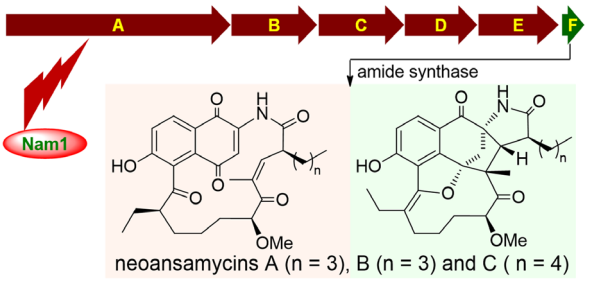


Figure 2. The first example of obtaining new ansamycins through genome mining of transcriptionally activating cryptic gene clusters.

Additionally, another AHBA synthase gene-positive strain *S.* sp. XZQH13 contains a cryptic ansamycin biosynthetic gene cluster *ast* as well. The constitutive overexpression of the *astG1* gene, encoding a Large-ATP-binding regulator of the LuxR family, induced the expression of the *ast* gene cluster including these key biosynthetic genes *astB4*, *astD1* and *astF1*. This led to the isolation of two ansatrienins hydroxymycotrienin A and thiazinotrienomycin G, further supporting that LuxR-type regulators are pathway-specific positive regulators for the biosynthesis of ansamycins (*82*). Moreover, overexpression of the LuxR-type regulator gene *div8* increases the production of divergolides in the *S.* sp. W112 strain (*79*), which further facilitates the isolation of five new ansamycins, namely divergolides O-S (*83*).

7. Perspectives

With the aid of whole genome sequences of thousands

of bacterial genomes and metagenomes, the fast development of genome mining has dramastically facilitated the automated discovery of promising new natural products, however, mostly ignored their roles in the environments. Though this development has been avoid of duplications in the discovery of natural products incurred by bioactivity-guided fractionation, it is insufficient in function predictions. This gap can be amended by adding bio-filters during genome mining (*21*). Specifically, microbial libraries for producers of both glycopeptide and ansamycin antibacterial compounds were enriched by screening the self-resistance mechanism of antibiotic producers. Together with screenings for biosynthetic genes resulted in the discovery of a new glycopeptide antibiotic, pekiskomycin (*84*). The self-resistance guided genome mining approach can be also applied in the discovery of inhibitors of target proteins beyond microorganisms. Panter *et al*. utilized the pentapeptide repeat protein (PRP)-mediated self-resistance mechanism against topoisomerase inhibitors to search the genome of *Pyxidicoccus fallax*. A gene cluster was identified adjacent to a predicted PRP (co-clustering), however, no corresponding compounds were reported before their study. Activation in the native host and heterologous expression of this gene cluster resulted in the isolation of pyxidicyclines, a group of new in inhibitors of *E. coli* DNA topoisomerase IV and human DNA topoisomerase I (*85*). Similarly, the fungal sesquiterpenoid aspterric acid, a potent herbicide, was obtained by the approach of co-clustering self-resistance gene and biosynthetic gene cluster as well (*86*). Indeed, the self-resistance mechanisms of microorganisms against natural products are diverse (*50,87-89*) thus that the self-resistance-directed genome mining would be fruitful in the focused discovery of novel bioactive natural products (*90,91*).

Some natural products are produced by microorganisms evolved from natural selection for acquisition of improved defence against competing deleterious organisms. The production of these defensive secondary metabolites is inducible because they originally serve as responses to environmental challenges. Therefore, the combination of chemical ecology and genome mining can be an integrative approach to the discovery of novel bioactive natural products (*92*). Much more ambitious, Medema and Fischbach suggested to use globe-wide metagenomes to footprint microbial ecosystems, and then to target these for exhaustive single-cell sequencing, which would facilitate our understanding the ecological roles, and even further the natural roles, of natural products. Ideally, computational analysis can predict the action targets of natural products with the availability of ecological distribution and evolution of biosynthetic gene clusters (*93*).

Besides ecology, evolution is another critical issue of natural products that should be considered during genome mining. Natural products are produced by secondary metabolism. Secondary metabolism is the response and adaptation mechanisms of organisms to environmental biotic and abiotic stresses, and natural products are the executors of these mechanisms. Based on the evolutionary theory of natural selection, biological traits are derived from the selection of variations in a population. It is beneficial to improve the fitness of organisms to their environments. The mutations that reduce fitness are directly harmful or irreparable, and will disappear from the population. Some of the new properties of variations have no profit or loss, so they belong to neutral variations. Applying these ideas to natural products, it is generally believed that the cost will increase in producing new natural products. If these costs are accompanied by benefits, organisms will be more adaptable. Moreover, mutations of biosynthetic gene clusters are beneficial to fitness but are not lethal to organisms, which exactly fits the nearly neutral theory that the adaptation may be due not to strong selection of rare variants with large effects, but to weak selection of common variants (*94-96*). Thus, computational analysis of the evolution of the biosynthetic gene clusters in sequences and organizations could help in simulating the evolution of pathways to facilitate genome mining (*97*).

The phylogeny-guided genome mining represents an approach on the basis of evolutionary biology to the discovery of natural products. This approach was established on the assumption that "a single biosynthetic gene that must have co-evolved with its respective biosynthetic gene cluster could be used as a phylogenetic marker that might represent an evolutionary path of its entire biosynthetic gene cluster" (*98*). This approach provides a very straightforward means to screen a large number of microbial genomes or metagenomes attributed to the use of a single gene as a molecular marker in search for new biosynthetic gene clusters of interest, which has already resulted in some fruitful findings (*98*). In particular, with the advance of natural product phylogeny, *e.g*. the Bayesian analysis of biosynthetic modules and/or domains rather than the sequence of a single gene (*99*), novel algorithms-promoted integrative pan-genome mining approach will play an important role in both the mining and the synthetic biology of natural products (*100*).

8. Conclusion

In conclusion, genome mining has brought revolutionary influence on the discovery of novel natural products and the characterization of biosynthetic pathways. The future genome mining would evolve to an integrative approach by the involvement of multiple disciplines to bioactivity-targeted mining of novel natural products.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (81530091), and and Program for Changjiang Scholars and Innovative Research Team in University (IRT_17R68).

References

1. Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. Nat Rev Drug Discov. 2015; 14:111-129.
2. Ikeda H. Natural products discovery from micro-organisms in the post-genome era. Biosci Biotechnol Biochem. 2017; 81:13-22.
3. Winter JM, Behnken S, Hertweck C. Genomics-inspired discovery of natural products. Curr Opin Chem Biol. 2011; 15:22-31.
4. Mukherjee S, Stamatis D, Bertsch J, Ovchinnikova G, Verezemska O, Isbandi M, Thomas AD, Ali R, Sharma K, Kyrpides NC, Reddy TB. Genomes OnLine Database (GOLD) v.6: Data updates and feature enhancements. Nucleic Acids Res. 2017; 45:D446-D456.
5. Choi S-S, Katsuyama Y, Bai L, Deng Z, Ohnishi Y, Kim E-S. Genome engineering for microbial natural product discovery. Curr Opin Microbiol. 2018; 45:53-60.
6. Malpartida F, Hopwood DA. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature. 1984; 309:462-464.
7. Tremblay N, Hill P, Conway KR, Boddy CN. The use of ClusterMine360 for the analysis of polyketide and nonribosomal peptide biosynthetic pathways. In: Nonribosomal peptide and polyketide biosynthesis: Methods and Protocols (Evans BS, ed.). Springer New York, New York, NY, 2016; pp. 233-252.
8. Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick WH. The genomisotopic approach: A systematic method to isolate products of orphan biosynthetic gene clusters. Chem Biol. 2007; 14:53-63.
9. Robbel L, Knappe TA, Linne U, Xie X, Marahiel MA. Erythrochelin − a hydroxamate-type siderophore predicted from the genome of *Saccharopolyspora erythraea*. FEBS J. 2010; 277:663-676.
10. Peano C, Bicciato S, Corti G, Ferrari F, Rizzi E, Bonnal RJ, Bordoni R, Albertini A, Bernardi LR, Donadio S, De Bellis G. Complete gene expression profiling of *Saccharopolyspora erythraea* using GeneChip DNA microarrays. Microb Cell Fact. 2007; 6:37.
11. Lautru S, Deeth RJ, Bailey LM, Challis GL. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat Chem Biol. 2005; 1:265-269.
12. Scherlach K, Hertweck C. Discovery of aspoquinolones A-D, prenylated quinoline-2-one alkaloids from *Aspergillus nidulans*, motivated by genome mining. Org Biomol Chem. 2006; 4:3517-3520.
13. Udwary DW, Zeigler L, Asolkar RN, Singan V, Lapidus A, Fenical W, Jensen PR, Moore BS. Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. Proc Natl Acad Sci U S A. 2007; 104:10376-10381.
14. Nguyen T, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, Taudien S, Platzer M, Hertweck C, Piel J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nat Biotechnol. 2008; 26:225-233.
15. Ju K-S, Gao J, Doroghazi JR, *et al*. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. Proc Natl Acad Sci U S A. 2015; 112:12175-12180.
16. Owen JG, Reddy BVB, Ternei MA, Charlop-Powers Z, Calle PY, Kim JH, Brady SF. Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. Proc Natl Acad Sci U S A. 2013; 110:11797-11802.
17. Chen LY, Wang XQ, Wang YM, Geng X, Xu XN, Su C, Yang YL, Tang YJ, Bai FW, Zhao XQ. Genome mining of *Streptomyces xinghaiensis* NRRL B-24674(T) for the discovery of the gene cluster involved in anticomplement activities and detection of novel xiamycin analogs. Appl Microbiol Biotechnol. 2018; 102:9549-9562.
18. Krug D, Müller R. Secondary metabolomics: The impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. Nat Prod Rep. 2014; 31:768-783.
19. Surup F, Viehrig K, Rachid S, Plaza A, Maurer CK, Hartmann RW, Muller R. Crocadepsins-depsipeptides from the myxobacterium *Chondromyces crocatus* found by a genome mining approach. ACS Chem Biol. 2018; 13:267-272.
20. Doroghazi JR, Albright JC, Goering AW, Ju K-S, Haines RR, Tchalukov KA, Labeda DP, Kelleher NL, Metcalf WW. A roadmap for natural product discovery based on large-scale genomics and metabolomics. Nat Chem Biol. 2014; 10:963.
21. Zhang W, Li Y, Qian G, Wang Y, Chen H, Li Y, Liu F, Shen Y, Du L. Identification and characterization of the anti-methicillin-resistant *Staphylococcus aureus* WAP-8294A2 biosynthetic gene cluster from *Lysobacter enzymogenes* OH11. Antimicrob Agents Chemother. 2011; 55:5581-5589.
22. Kersten RD, Yang Y-L, Xu Y, Cimermancic P, Nam S-J, Fenical W, Fischbach MA, Moore BS, Dorrestein PC. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. Nat Chem Biol. 2011; 7:794-802.
23. Quinn RA, Nothias LF, Vining O, Meehan M, Esquenazi E, Dorrestein PC. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. Trends Pharmacol Sci. 2017; 38:143-154.
24. Nothias LF, Nothias-Esposito M, da Silva R, Wang M, Protsyuk I, Zhang Z, Sarvepalli A, Leyssen P, Touboul D, Costa J, Paolini J, Alexandrov T, Litaudon M, Dorrestein PC. Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. J Nat Prod. 2018; 81:758-767.
25. Bonneau N, Chen G, Lachkar D, Boufridi A, Gallard JF, Retailleau P, Petek S, Debitus C, Evanno L, Beniddir MA, Poupon E. An unprecedented blue chromophore found in nature using a "chemistry first" and molecular networking approach: Discovery of dactylocyanines A-H. Chemistry. 2017; 23:14454-14461.
26. Duncan K, Crüsemann M, Lechner A, Sarkar A, Li J, Ziemert N, Wang M, Bandeira N, Moore B, Dorrestein P, Jensen P. Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene

clusters and their products from *Salinispora* species. Chem Biol. 2015; 22:460-471.
27. Chiang YM, Chang SL, Oakley BR, Wang CCC. Recent advances in awakening silent biosynthetic gene clusters and linking orphan clusters to natural products in microorganisms. Curr Opin Chem Biol. 2011; 15:137-143.
28. Niu G. Genomics-driven natural product discovery in actinomycetes. Trends Biotechnol. 2018; 36:238-241.
29. Ishihama A. Prokaryotic genome regulation: Multifactor promoters, multitarget regulators and hierarchic networks. FEMS Microbiol Rev. 2010; 34:628-645.
30. Bode HB, Bethe B, Hofs R, Zeeck A. Big effects from small changes: Possible ways to explore nature's chemical diversity. Chembiochem. 2002; 3:619-627.
31. Yoon V, Nodwell JR. Activating secondary metabolism with stress and chemicals. J Ind Microbiol Biotechnol. 2014; 41:415-424.
32. Ochi K, Tanaka Y, Tojo S. Activating the expression of bacterial cryptic genes by *rpoB* mutations in RNA polymerase or by rare earth elements. J Ind Microbiol Biotechnol. 2014; 41:403-414.
33. Lu C, Shen Y. Harnessing the potential of chemical defenses from antimicrobial activities. Bioessays. 2004; 26:808-813.
34. Hosaka T, Ohnishi-Kameyama M, Muramatsu H, Murakami K, Tsurumi Y, Kodani S, Yoshida M, Fujie A, Ochi K. Antibacterial discovery in actinomycete strains with mutations in RNA polymerase or ribosomal protein S12. Nat Biotechnol. 2009; 27:462-464.
35. Ochi K, Okamoto S, Tozawa Y, Inaoka T, Hosaka T, Xu J, Kurosawa K. Ribosome engineering and secondary metabolite production. Adv Appl Microbiol. 2004; 56:155-184.
36. Hrvatin S, Piel J. Rapid isolation of rare clones from highly complex DNA libraries by PCR analysis of liquid gel pools. J Microbiol Methods. 2007; 68:434-436.
37. Bian X, Tang B, Yu Y, *et al*. Heterologous production and yield improvement of epothilones in *Burkholderiales* strain DSM 7029. ACS Chem Biol. 2017; 12:1805-1812.
38. Wang X, Zhou H, Chen H, *et al*. Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in *Burkholderiales* species. Proc Natl Acad Sci U S A. 2018; 115:E4255-E4263.
39. Bentley SD, Chater KF, Cerdeno-Tarraga AM, *et al*. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature. 2002; 417:141-147.
40. Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R. The TetR family of transcriptional repressors. Microbiol Mol Biol Rev. 2005; 69:326-356.
41. Corre C, Song L, O'Rourke S, Chater KF, Challis GL. 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. Proc Natl Acad Sci U S A. 2008; 105:17510-17515.
42. Recio E, Colinas A, Rumbero A, Aparicio JF, Martin JF. PI factor, a novel type quorum-sensing inducer elicits pimaricin production in *Streptomyces natalensis*. J Biol Chem. 2004; 279:41586-41593.
43. Willey JM, Gaskell AA. Morphogenetic signaling molecules of the streptomycetes. Chem Rev. 2011; 111:174-187.
44. Mlot C. Antibiotics in nature: Beyond biological warfare. Science. 2009; 324:1637-1639.
45. Khokhlov AS, Tovarova I, Borisova LN, Pliner SA, Shevchenko LN, Kornitskaia E, Ivkina NS, Rapoport IA. A-factor responsible for the biosynthesis of streptomycin by a mutant strain of *Actinomyces streptomycini*. Dokl Akad Nauk SSSR. 1967; 177:232-235.
46. Kato JY, Funa N, Watanabe H, Ohnishi Y, Horinouchi S. Biosynthesis of gamma-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. Proc Natl Acad Sci U S A. 2007; 104:2378-2383.
47. Nishida H, Ohnishi Y, Beppu T, Horinouchi S. Evolution of gamma-butyrolactone synthases and receptors in *Streptomyces*. Environ Microbiol. 2007; 9:1986-1994.
48. Natsume R, Ohnishi Y, Senda T, Horinouchi S. Crystal structure of a gamma-butyrolactone autoregulator receptor protein in *Streptomyces coelicolor* A3(2). J Mol Biol. 2004; 336:409-419.
49. Osbourn A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. Trends Gene. 2010; 26:449-457.
50. Floss HG, Yu TW. Rifamycin-mode of action, resistance, and biosynthesis. Chem Rev. 2005; 105:621-632.
51. Fukuyo Y, Hunt CR, Horikoshi N. Geldanamycin and its anti-cancer activities. Cancer Lett. 2010; 290:24-35.
52. Cassady JM, Chan KK, Floss HG, Leistner E. Recent developments in the maytansinoid antitumor agents. Chem Pharm Bull. 2004; 52:1-26.
53. Ding L, Maier A, Fiebig HH, Gorls H, Lin WH, Peschel G, Hertweck C. Divergolides A-D from a mangrove endophyte reveal an unparalleled plasticity in ansa-macrolide biosynthesis. Angew Chem Int Ed Engl. 2011; 50:1630-1634.
54. Xu Z, Baunach M, Ding L, Peng H, Franke J, Hertweck C. Biosynthetic code for divergolide assembly in a bacterial mangrove endophyte. Chembiochem. 2014; 15:1274-1279.
55. Wilson MC, Nam SJ, Gulder TA, Kauffman CA, Jensen PR, Fenical W, Moore BS. Structure and biosynthesis of the marine streptomycete ansamycin ansalactam A and its distinctive branched chain polyketide extender unit. J Am Chem Soc. 2011; 133:1971-1977.
56. August PR, Tang L, Yoon YJ, Ning S, Muller R, Yu TW, Taylor M, Hoffmann D, Kim CG, Zhang X, Hutchinson CR, Floss HG. Biosynthesis of the ansamycin antibiotic rifamycin: Deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. Chem Biol. 1998; 5:69-79.
57. Absalon AE, Fernandez FJ, Olivares PX, Barrios-Gonzalez J, Campos C, Mejia A. RifP; a membrane protein involved in rifamycin export in *Amycolatopsis mediterranei*. Biotechnol Lett. 2007; 29:951-958.
58. Fernández-Moreno MA, Caballero J, Hopwood DA, Malpartida F. The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. Cell. 1991; 66:769-780.
59. Xu J, Wan E, Kim CJ, Floss HG, Mahmud T. Identification of tailoring genes involved in the modification of the polyketide backbone of rifamycin B by *Amycolatopsis mediterranei* S699. Microbiology. 2005; 151:2515-2528.
60. Rascher A, Hu Z, Viswanathan N, Schirmer A, Reid R, Nierman WC, Lewis M, Hutchinson CR. Cloning and characterization of a gene cluster for geldanamycin

production in *Streptomyces hygroscopicus* NRRL 3602. FEMS Microbiol Lett. 2003; 218:223-230.
61. He W, Lei J, Liu Y, Wang Y. The LuxR family members GdmRI and GdmRII are positive regulators of geldanamycin biosynthesis in *Streptomyces hygroscopicus* 17997. Arch Microbiol. 2008; 189:501-510.
62. Li Y, Zhao P, Kang Q, Ma J, Bai L, Deng Z. Dual carbamoylations on the polyketide and glycosyl moiety by Asm21 result in extended ansamitocin biosynthesis. Chem Biol. 2011; 18:1571-1580.
63. Gao Y, Fan Y, Nambou K, Wei L, Liu Z, Imanaka T, Hua Q. Enhancement of ansamitocin P-3 production in *Actinosynnema pretiosum* by a synergistic effect of glycerol and glucose. J Ind Microbiol Biotechnol. 2014; 41:143-152.
64. Lin J, Bai L, Deng Z, Zhong JJ. Effect of ammonium in medium on ansamitocin P-3 production by *Actinosynnema pretiosum*. Biotechnol Bioproc Eng. 2010; 15:119-125.
65. Yu TW, Bai L, Clade D, Hoffmann D, Toelzer S, Trinh KQ, Xu J, Moss SJ, Leistner E, Floss HG. The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from *Actinosynnema pretiosum*. Proc Natl Acad Sci U S A. 2002; 99:7968-7973.
66. Bandi S, Kim Y, Chang YK, Shang GD, Yu TW, Floss HG. Construction of asm2 deletion mutant of *Actinosynnema pretiosum* and medium optimization for ansamitocin P-3 production using statistical approach. J Ind Microbiol Biotechnol. 2006; 16:1338-1346.
67. Ng D, Chin HK, Wong VV. Constitutive overexpression of asm2 and asm39 increases AP-3 production in the actinomycete *Actinosynnema pretiosum*. J Ind Microbiol Biotechnol. 2009; 36:1345-1351.
68. Li S, Lu C, Chang X, Shen Y. Constitutive overexpression of asm18 increases the production and diversity of maytansinoids in *Actinosynnema pretiosum*. Appl Microbiol Biotechnol. 2016; 100:2641-2649.
69. Amiri-Kordestani L, Blumenthal GM, Xu QC, *et al*. FDA approval: Ado-trastuzumab emtansine for the treatment of patients with HER2-positive metastatic breast cancer. Clin Cancer Res. 2014; 20:4436-4441.
70. Wang H, Chen Y, Ge L, *et al*. PCR screening reveals considerable unexploited biosynthetic potential of ansamycins and a mysterious family of AHBA-containing natural products in actinomycetes. J Appl Microbiol. 2013; 115:77-85.
71. Lu C, Li Y, Deng J, Li S, Wang H, Shen Y. Hygrocins C-G, cytotoxic naphthoquinone ansamycins from *gdmAI*-disrupted *Streptomyces* sp. LZ35. J Nat Prod. 2013; 76:2175-2179.
72. Shi N, Wang H, Lu C, Liu Z, Shen Y. Ansamycins produced by *Streptomyces* sp. LZ35. Chin Pharm J. 2011; 46:1317-1320.
73. Li S, Wang H, Li Y, Deng J, Lu C, Shen Y, Shen Y. Biosynthesis of hygrocins, antitumor naphthoquinone ansamycins produced by *Streptomyces* sp. LZ35. Chembiochem. 2014; 15:94-102.
74. Yu TW, Shen Y, Doi-Katayama Y, Tang L, Park C, Moore BS, Richard Hutchinson C, Floss HG. Direct evidence that the rifamycin polyketide synthase assembles polyketide chains processively. Proc Natl Acad Sci U S A. 1999; 96:9051-9056.
75. Laureti L, Song L, Huang S, Corre C, Leblond P, Challis GL, Aigle B. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. Proc Natl Acad Sci U S A. 2011; 108:6258-6263.
76. Kuscer E, Coates N, Challis I, Gregory M, Wilkinson B, Sheridan R, Petkovic H. Roles of rapH and rapG in positive regulation of rapamycin biosynthesis in *Streptomyces hygroscopicus*. J Bacteriol. 2007; 189:4756-4763.
77. Li S, Li Y, Lu C, Zhang J, Zhu J, Wang H, Shen Y. Activating a cryptic ansamycin biosynthetic gene cluster to produce three new naphthalenic octaketide ansamycins with *n*-pentyl and n-butyl side chains. Org Lett. 2015; 17:3706-3709.
78. Huitu Z, Linzhuan W, Aiming L, Guizhi S, Feng H, Qiuping L, Yuzhen W, Huanzhang X, Qunjie G, Yiguang W. PCR screening of 3-amino-5-hydroxybenzoic acid synthase gene leads to identification of ansamycins and AHBA-related antibiotic producers in actinomycetes. J Appl Microbiol. 2009; 106:755-763.
79. Li SR, Zhao GS, Sun MW, He HG, Wang HX, Li YY, Lu CH, Shen YM. Identification and characterization of the biosynthetic gene cluster of divergolides from *Streptomyces* sp. W112. Gene. 2014; 544:93-99.
80. Pan W, Kang Q, Wang L, Bai L, Deng Z. Asm8, a specific LAL-type activator of 3-amino-5-hydroxybenzoate biosynthesis in ansamitocin production. Sci China Life Sci. 2013; 56:601-608.
81. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Genome Res. 2004; 14:1188-1190.
82. Xie C, Deng JJ, Wang HX. Identification of AstG1, a LAL family regulator that positively controls ansatrienins production in *Streptomyces* sp. XZQH13. Curr Microbiol. 2015; 70:859-864.
83. Zhao G, Li S, Guo Z, Sun M, Lu C. Overexpression of div8 increases the production and diversity of divergolides in *Streptomyces* sp. W112. RSC Advances. 2015; 5:98209-98214.
84. Thaker MN, Wang W, Spanogiannopoulos P, Waglechner N, King AM, Medina R, Wright GD. Identifying producers of antibacterial compounds by screening for antibiotic resistance. Nat Biotechnol. 2013; 31:922-927.
85. Panter F, Krug D, Baumann S, Muller R. Self-resistance guided genome mining uncovers new topoisomerase inhibitors from myxobacteria. Chem Sci. 2018; 9:4898-4908.
86. Yan Y, Liu Q, Zang X, Yuan S, Bat-Erdene U, Nguyen C, Gan J, Zhou J, Jacobsen SE, Tang Y. Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. Nature. 2018; 559:415-418.
87. Gomi S, Ikeda D, Nakamura H, Naganawa H, Yamashita F, Hotta K, Kondo S, Okami Y, Umezawa H, Iitaka Y. Isolation and structure of a new antibiotic, indolizomycin, produced by a strain SK2-52 obtained by interspecies fusion treatment. J Antibiot. 1984; 37:1491-1494.
88. Hopwood DA. How do antibiotic-producing bacteria ensure their self-resistance before antibiotic biosynthesis incapacitates them? Mol Microbiol. 2007; 63:937-940.
89. Xu H, Huang W, He QL, Zhao ZX, Zhang F, Wang R, Kang J, Tang GL. Self-resistance to an antitumor antibiotic: A DNA glycosylase triggers the base-excision repair system in yatakemycin biosynthesis. Angew Chem Int Ed Engl. 2012; 51:10532-10536.
90. Ziemert N, Alanjary M, Weber T. The evolution of genome mining in microbes – a review. Nat Prod Rep. 2016; 33:988-1005.

91. Machado H, Tuttle RN, Jensen PR. Omics-based natural product discovery and the lexicon of genome mining. Curr Opin Microbiol. 2017; 39:136-142.
92. Vizcaino MI, Guo X, Crawford JM. Merging chemical ecology with bacterial genome mining for secondary metabolite discovery. J Ind Microbiol Biotechnol. 2014; 41:285-299.
93. Medema MH, Fischbach MA. Computational approaches to natural product discovery. Nat Chem Biol. 2015; 11:639-648.
94. Hurst LD. Fundamental concepts in genetics: Genetics and the understanding of selection. Nat Rev Genet. 2009; 10:83-93.
95. Kimura M. The neutral theory of molecular evolution and the world view of the neutralists. Genome. 1989; 31:24-31.
96. Valenzuela CY. Foundational errors in the Neutral and Nearly-Neutral theories of evolution in relation to the Synthetic Theory: Is a new evolutionary paradigm necessary? Biol Res. 2013; 46:101-119.
97. Rudolf JD, Yan X, Shen B. Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery. J Ind Microbiol Biotechnol. 2016; 43:261-276.
98. Kang HS. Phylogeny-guided (meta)genome mining approach for the targeted discovery of new microbial natural products. J Ind Microbiol Biotechnol. 2017; 44:285-293.
99. Schmitt I, Barker FK. Phylogenetic methods in natural product research. Nat Prod Rep. 2009; 26:1585-1602.
100. van der Lee TAJ, Medema MH. Computational strategies for genome-based natural product discovery and engineering in fungi. Fungal Genet Biol. 2016; 89:29-36.

(Received November 1, 2018; Revised December 25, 2018; Accepted December 28, 2018)