



Generate a bioactive natural product library by mining bacterial cytochrome P450 patterns

Xiangyang Liu *

UNT System College of Pharmacy, University of North Texas Health Science Center, Fort Worth, TX 76107, USA

ARTICLE INFO

Article history:

Received 20 January 2016

Accepted 26 January 2016

Available online

Keywords:

Genome mining

Microbial P450

Natural product library

Synthetic biology

ABSTRACT

The increased number of annotated bacterial genomes provides a vast resource for genome mining. Several bacterial natural products with epoxide groups have been identified as pre-mRNA spliceosome inhibitors and antitumor compounds through genome mining. These epoxide-containing natural products feature a common biosynthetic characteristic that cytochrome P450s (CYPs) and its patterns such as epoxidases are employed in the tailoring reactions. The tailoring enzyme patterns are essential to both biological activities and structural diversity of natural products, and can be used for enzyme pattern-based genome mining. Recent development of direct cloning, heterologous expression, manipulation of the biosynthetic pathways and the CRISPR-CAS9 system have provided molecular biology tools to turn on or pull out nascent biosynthetic gene clusters to generate a microbial natural product library. This review focuses on a library of epoxide-containing natural products and their associated CYPs, with the intention to provide strategies on diversifying the structures of CYP-catalyzed bioactive natural products. It is conceivable that a library of diversified bioactive natural products will be created by pattern-based genome mining, direct cloning and heterologous expression as well as the genomic manipulation.

© 2016 The authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

“Discovery consists of seeing what everybody else has seen, and thinking what nobody else has thought.”

— Nobel laureate Albert von Szent-Gyorgyi

1. Introduction

The conception of the natural product library originated from the advocacy for the Biological Resource Center (BRC) via the Organization for the Economic Cooperation and Development (OECD) in 1999.¹ During the past 15 years, the production of a library of microbial natural products has been mainly driven by the development of high-throughput screening methods and the increasing number of reported marine natural products.^{2–4} Broadly speaking, researches on natural product libraries have included the activities of creating a genomic DNA prioritized strain library,⁵ a

metagenome library,^{6,7} a combinational library,⁸ a crude extract library,^{9–11} a fraction library^{12–14} and a pure compound library.¹⁵ Additionally, the bioactive compound library is commercially available from Selleckchem,¹⁶ OTAVA¹⁷ and Cromadex.¹⁸ Strategies for developing a high quality microbial natural product library include diversification and de-replication of strain, natural product structures and high throughput screening methods.^{2,19–23} Even so, the low outcomes from the creation of a crude extract library and a fraction library are far from the expectation in drug discovery. These facts have diminished the interests of big pharma in the natural product library.²⁴

Nature is always full of diversity and natural product biosynthesis is no exception. The “co-linearity” rule, and the diversity and variations in nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs), hybrid NRPS/PKS systems and the beyond have been greatly explored.^{25–31} During last two decades, well over hundreds of natural product biosynthetic pathways and thousands of natural scaffolds such as nonribosomal peptide, polyketide, terpenoid and oligosaccharide have been characterized.^{32,33} Except for the enzymes that catalyze formation of chemical scaffolds, there is a group of enzymes that decorate the functional group density, acting sequentially to transfer the building blocks into complex molecular structures. The tailoring enzymes can be classified into nonoxidative and oxidative enzymes. Nonoxidative enzymes include

* Corresponding author. UNT System College of Pharmacy, University of North Texas Health Science Center, Fort Worth, TX 76107, USA. Tel.: 414-339-1330; fax: 817-735-2603.

E-mail address: shangdongliu@gmail.com.

Peer review under responsibility of KeAi Communications Co., Ltd.

acyltransferases, methyltransferases and glycosyltransferases, while oxidative enzymes encompass those who introduce water-solubilizing oxygen functionalities during maturation and are probably most consequential for introduction of both scaffold and functional group complexity. While considering how to diversify nature's small molecule inventory, the oxygenase enzymes have proven to be most valuable to characterize with the twin goals of prediction of new natural product scaffolds and combinatorial engineering of intermediates.^{33,34} The most prominent of the oxidative enzymes are the heme iron cytochrome P450 (CYPs) monooxygenases found in microbes such as *Streptomyces* species,³⁵ *Bacillus* species³⁶ and Cyanobacteria,³⁷ etc.³⁸

Traditionally, natural product genome mining means “use analysis of DNA sequence data to predict structural elements of new natural products and then use this information to design strategies for rapidly identifying, purifying, and structurally characterizing the compounds”.^{39,40} However, the peculiarities of biosynthetic pathways indicate that textbook “co-linearity” rules cannot be applied to deduce structures from all DNA data.^{41,42} Thanks to the development of gene cluster prediction software such as antiSMASH,⁴³ genome mining has become a quick and inexpensive way to analyze the biosynthetic potential of sequenced microbes. The current genome mining approaches include sequence-based genome mining,^{44,45} bioactivity-guided genome mining,⁴⁶ enzyme-based genome mining,⁴⁷ pattern-based genome mining^{48,49} and genome neighborhood network analysis.^{50,51} Sequence-based genome mining is designed to detect and extract carboxyl (C) – and keto-synthase (KS) – domains from DNA or amino acid sequence data. The high degrees of sequence similarity (E -value $< 10e^{-40}$ for eSNAPD or 85–90% protein sequence identity for NaPDoS) suggest that identified biosynthetic gene clusters are responsible for the biosynthesis of similar natural products resulting from reference gene clusters.^{6,7,44,52} For bioactivity-guided genome mining, the conserved structural motif of bioactive natural products can be used as a reference to perform genome mining. For example, the enediynes PKS (PKSE) is proposed to be involved in the formation of the highly reactive chromophore ring structure (or “warhead”) found in all enediynes. By PKSE based genome mining, the enediynes biosynthetic gene clusters have been identified from sequenced actinomycete genomes.⁵³ FK228 is an antitumor drug with tumor cytotoxicity resulting from the functional disulfide group, which is catalyzed by the FAD-dependent disulfide oxidoreductase.^{54,55} BLAST analysis of FAD-dependent disulfide oxidoreductase has led to the discovery of a homologous gene cluster of thailandepsins in *Burkholderia thailandensis* E264.⁵⁶ Enzyme-based genome mining searches conserved synthase domains against the NCBI database of sequenced bacterial genomes in order to obtain the presumptive enzyme sequences.⁴⁷ The “pattern-based genome mining” refers to the connection of MS/MS fragmentation pattern to the biosynthetic pathways genome mining and de-replication of certain bacterial species. For example, the MS/MS fragmentation pattern of the 827.492 for arenicolide A production was used to identify the uncharacterized gene cluster in *S. pacifica* strains CNQ-748 and CNT-138.⁴⁹ The “genome neighborhood networks (GNN) analysis” is a bioinformatics strategy to predict enzymatic functions on a large scale based on their genomic context. In this case, bioinformatics of PepM and phosphonate GNN were applied for 278 sequenced bacterial genomes and led to the discovery of 19 new phosphonate natural products.⁵⁷ Enediynes GNNs were generated for the virtual screening of the sequenced bacterial genomes resulted in 87 potential enediynes gene clusters from 78 different bacteria strains.⁵⁰ The pattern-based genome mining and the genome neighborhood analysis provide a comprehensive method to identify the biosynthetic gene cluster of sequenced bacterial strains.

With the physiological and medical Nobel Prize awarding to natural product avermectin and artemisinin, the discovery of natural

products has entered a golden age.⁵⁸ Our long term goal is to create a high quality, diversified natural product library. In the postgenomic era, the challenges to generate a microbial natural product library have been switched from the traditional de-replication strategies to issues of how to translate the annotated biosynthetic gene clusters of interest to a bioactive natural product library. The current review will focus on genome mining of CYPs which are involved in the biosynthetic gene clusters of bacterial secondary metabolites, especially those with epoxide functional groups, with an intention to share the considerations to build a diversified natural product library through CYP pattern-based genome mining, direct cloning and heterologous expression, and genome manipulation.

2. Genome mining of CYP-catalyzed bacterial natural products

This section will cover the introduction of CYP, two genome mining methods, and application of genome mining in two genera.

2.1. The importance of microbial CYPs

Microbial natural products catalyzed by CYP biosynthetic pathways have diversified biological activities including antitumor activities, antibacterial activities, antifungal activities, anti-HIV activities, anti-parasitic and anti-cholesterol activities. Natural products such as pladienolides/FD-895,^{59,60} GEX1/herboxidiene,⁶¹ FR901464 (FR)/spliceostatins/thailanstatins⁶² (Fig. 1A) are known to have antitumor activities by targeting pre-mRNA spliceosome. Specifically, pladienolide B and spliceostatin A have been reported to express their antiproliferative activities against tumor cells at the nM range by targeting the splicing factor 3B subunit 1 (SF3B1) with the help of three-membered epoxide groups. Furthermore, CYP-catalyzed natural products with antitumor activities may have different mechanisms. For instance, epothilone A expresses the tubulin-binding activity,⁶³ griseorhodin A is telomerase inhibitor,⁶⁴ tirandamycin is RNA polymerase inhibitor⁶⁵ and epoxomicin is proteasome-inhibitor,⁶⁶ etc (Fig. 1B).

Microbial CYPs are one of the most widely distributed groups of tailoring enzymes to catalyze the formation of the final bioactive natural products.^{31,35,37,67–69} CYPs catalyze the hydroxylation and/or epoxidation reactions in the late stages of biosynthesis after macrolide formation by PKSs.^{70–72} Reactions catalyzed by CYP monooxygenases include hydroxylation of saturated C–H bonds, epoxidation of C=C double bonds, and oxidative decarboxylation (Table 1).¹⁰⁶ For example, The CYP PldB hydroxylates 6-deoxy pladienolide B to pladienolide B in the pladienolide biosynthetic pathway.⁵⁹ The CYP HerG catalyzes the stereospecific hydroxylation at C-18 of herboxidiene.⁶¹ For FR biosynthesis, the CYP Fr9R is not only involved in the hydroxylation at C-4 of FR901464, but is also related to the formation of the hemiketal group at C-1.⁶² CYPs in the biosynthetic pathways of natural products often work together with its patterns such as transferases and epoxidases to catalyze the production of epoxide groups. Moreover, dual-function CYPs have been reported to catalyze sequential epoxidation and hydroxylation of the same substrate.^{65,88,107}

2.2. Bioinformatics of microbial CYPs

The information on three-dimensional structures is essential to understand the molecular basis for substrate recognition and specificity of CYPs. It is universally thought that three elements are essential in all CYPs sequences: (a) the conserved cysteine, which is the fifth ligand to the heme Fe atom and can be represented as FXXGXXXCXG, (b) the EXXR motif forming a charge pair in the K helix (possibly involved in heme binding) and (c) overall CYP fold topology, including stability.¹⁰⁸ However, there are some unusual examples for bacterial CYPs, such as variations at conserved motifs,

Table 1
CYPs involved in the biosynthesis of microbial natural products.

P450 enzymes	No. amino acid (AA)	Accession no.	Match in the database ^a				Reference
			CYP name as annotated	Protein identifier	Identity%	AA overlap	
Hydroxylation of saturated C–H bonds							
AmphL	396	AAK73504	CYP107E	AAK73504	100%	396	73
AziB1	401	B4XY99.1	CYP107-like	WP_018771011.1	34%	388	74
ChmPI	407	AAS79447	CYP107B	AAS79447	100%	407	75
EpnK	401	AHB38512	CYP162-like	WP_030872520.1	40%	399	66
EryF	404	Q00441.2	CYP107A	WP_009950397.1	100%	404	76
EryK	397	CYP113A1	CYP113A-like	WP_009950895.1	100%	397	77
HerG	422	AEZ64507	CYP107B-like	WP_016577953.1	48%	404	61
MeiE	459	AAM97314	CYP171A	AAM97314	100%	459	78
MycCl	383	BAC57023	CYP105U	BAC57023	100%	383	79
NcsB3	410	AAM77997	CYP154J	AAM77997	100%	410	80
NysL	394	AAF71769	CYP107E	AAF71769	100%	394	81
OxyD	396	CCD33151	CYP146A	CCD33151	100%	396	82
PikC	416	O87605	CYP107L	AAC64105	100%	416	83
PldB	399	BAH02272	CYP107B-like	BAH02272	100%	399	59
PteC	399	BAC68123	CYP105P1	WP_010981850.1	100%	399	84
TylHI	436	AAD41818	CYP105U	AAD41818	100%	436	85
TylII	417	AAA21341	CYP107-like	AAA21341	100%	417	86
ZbmVIIc	416	ACG60779	CYP185-like	WP_030613068.1	51%	430	78
TxtC	395	AAL36838.1	CYP105A	WP_009073396.1	41%	406	87
Dual function, hydroxylation and epoxidation							
GsfF	414	BAJ16472	CYP105A	BAJ16472	100%	414	88
MycG	397	BAA03672	CYP107B	BAA03672	100%	397	89
Fma-P450	400	AHL19974	_b				90
Fr9R	482	AIC32704	CYP136-like	WP_022984814.1	32%	455	62
PenD	298	ADO85592	–				91
PntD	299	ADO85576	–				92
TamI	413	ADC79647.1	CYP107B-like	ADC79647.1	100%	413	65
TstR	482	AGN11891	CYP136-like	BAP15297	33%	455	93
Catalyze epoxidation							
Asm30	1005	AAM54108	CYP102F	AAM54108	100%	1005	94
ChmPII	401	AAS79446	CYP107B	AAS79446	100%	401	75
EpnI	431	AHB38510	CYP107B-like	WP_030776170.1	48%	412	66
EpoK	419	Q9KIZ4	CYP167A	Q9KIZ4	100%	419	63
EpxC	425	AHB38496	CYP107B-like	WP_020576451.1	42%	406	66
GrhO3	416	AAM33670	CYP105-like	AAM33670	100%	416	64
HedR	409	AAP85338	CYP105-like	AAP85338	100%	409	95
OleP	407	AAA92553	CYP107B	AAA92553	100%	407	96
PimD	397	CAC20932	CYP107E	CAC20932	100%	397	97
PimG	398	CAC20928	CYP105H	CAC20928	100%	398	97
TamI	413	ADC79647	CYP107B-like	ADC79647	100%	413	98
Oxidation of methyl group							
AmphN	399	AAK73509	CYP105H	AAK73509	100%	399	99
NysN	398	AAF71771	CYP105H	AAF71771	100%	398	81
C–C coupling							
DynOrf19	403	ACB47070	CYP107M	WP_015621195.1	51%	398	100
HmtS	397	CBZ42153	CYP113A	WP_030360446.1	52%	398	101
OxyB	398	AAL90878	CYP165B	Q8RN04	100%	398	102
OxyC	406	AAL90879	CYP165B	Q8RN03	100%	406	103
C–N–C coupling							
DynE10	400	ACB47071	CYP107B	WP_029899036.1	53%	400	100
SpcN	390	AGL96571	CYP244A	AGL96571	100%	390	104
Catalyzes the nitration using NO and O ₂							
TxtE	406	ELP66108.1	CYPP450TXTE	ELP66108.1	100%	406	87
Oxidative decarboxylation							
HmtN	419	4E2P_A	CYP113A	4E2P_A	100%	419	101
HmtT	418	4GGV_A	CYP113A	4GGV_A	100%	418	101
Mei-Orf4	398	ADC45514	CYP107-like	ADC45514	100%	398	78
StaP	426	ABI94389	CYP245A	ABI94389	100%	426	105
SpcP	427	AGL96575	CYP245A	AGL96575	100%	427	104

^a CYP names as annotated at website: <https://cyped.biocatnet.de/>.

^b Not found in database.

heme incorporation and topology, substrate binding and functionalities. For example, the amino acid sequence of CYP157C1 contains EQSLW in place of the conserved EXXR.¹⁰⁹ The conserved threonine in CYP107A1 is not present and a hydroxyl group of the substrate 6-deoxyerythronolide B can directly donate a hydrogen bond to the Fe-linked dioxygen for proton transfer.⁷⁶ Different from CYP107A1, CYP158A2 binds two molecules of flaviolin in its active site, and the 2-OH group of flaviolin is responsible for anchoring

the substrate in the active site, while the 5-OH and 7-OH stabilize water molecules which are important for catalysis.¹¹⁰

For the discovery of bacterial CYPs in sequenced genomes, genes encoding the CYP heme binding domain (FXXGXXXCXG) can be screened for the presence of a highly conserved threonine in the putative I-helix, which is proposed to be involved in CYP oxygen activation in most CYPs, and the conserved EXXR motif is located in the K-helix. The sequences of polypeptides containing all three

motifs can be further used as queries for BLAST searches of the GenBank non-redundant protein database (www.ncbi.nlm.nih.gov/BLAST/) to identify their closest homologues in other organisms and tentatively assign the CYP proteins to subfamilies. Briefly, >40% of amino acid sequence identity places a CYP in the same family and >55% places it in the same subfamily.¹¹¹ Besides GenBank, there is a CYPED database which scours Genbank by BLAST-searching and retrieving the CYP sequences to put in their database (<https://cyped.biocatnet.de/>).^{112–115} This database contains the most bacterial CYPs (Table 1) that are involved in the biosynthesis of bioactive natural products (Fig. 1). There is also another online CYP database (<http://drnelson.uthsc.edu/CytochromeP450.html>), which contains 1042 bacterial CYP genes.

2.3. The logic for CYP pattern-based genome mining

There is evidence that CYPs can be used for genome mining for pre-mRNA spliceosome inhibitors. First, all current pre-mRNA spliceosome inhibitors are epoxide-containing natural products (Fig. 1A) whose biosynthesis gene clusters contain genes encoding putative CYP oxygenases or epoxidases. It has been reported that the presence of epoxide group is important in conferring activity to FR analogues.^{116,117} Second, CYPs may not directly be involved in the formation of some epoxide groups,^{70,118} but CYPs are present in the relevant biosynthetic gene clusters (Fig. 2). For example, in a recent metagenome mining report, all six epoxyketone gene clusters contain CYPs,¹¹⁹ indicating that CYP-containing gene clusters may be a rich source for the biosynthesis of epoxide compounds. Third, even antitumor activities of CYP-catalyzed compounds are related to different mechanisms. Small molecule screenings identified that oxaspiro compounds (the farnesyltransferase inhibitor manumycin A analogues) are pre-mRNA splicing inhibitors.¹²⁰ In addition, a number of compounds such as the CYP-related non-epoxide compounds staurosporine (kinase inhibitor), are novel inhibitors of spliceosome assembly.¹²⁰ Fourth, homologous CYPs can be identified through genome mining by a 50–60% similarity except for the 80%–90% rule for PKS and nonribosomal peptide synthetase (NRPS) domain searching. For example, BLAST analysis demonstrated a 50% identity between CYPs PldB and HerG from pladienolide and herboxidiene (both are pre-mRNA spliceosome inhibitors), respectively.¹²¹ In the biosynthesis of FD-891, the CYP GfsF showed 57% identity to the CYP monooxygenase Mflv2418.⁸⁸ Finally, the CYP gene can be used as a signature gene for gene cluster cloning and gene identification.⁵⁹ However, CYP pattern should be considered because CYPs and other tailoring enzymes such as transferases and epoxidases act together on the released PKS or NRPS megasynthases^{68,122,123} (Fig. 3).

2.4. Discovery of thailanstatins from *Burkholderia* sp.

Recently, the genus *Burkholderia* has attracted the attention of several research groups to employ different genome mining strategies for the study of bioactive natural products.¹²⁴ Among those products, FR is a general spliceosome inhibitor discovered in 1992 from *Pseudomonas* sp. No. 2663,^{125–127} now identified as *Burkholderia* sp. FERM BP3421.⁶² There are three oxygenase activities encoded in FR gene cluster: (a) the flavin-dependent monooxygenase (FMO) domain in the last module of *fr9GH*, (b) the CYP encoded by *fr9R*, and (c) the Fe(II)/ α -ketoglutarate-dependent dioxygenase encoded by *fr9P*. The hemiketal FR is mainly biosynthesized through the epoxidation at C3 by the FMO domain of *Fr9GH*, the hydroxylation at C-4 by the *Fr9R*, the hydroxylation at C-1 by *Fr9P* and then decarboxylation.⁶² First, BLAST analysis shows that *Fr9R* has a high identity (82%) to *TstR* in the sequenced genome of *B. thailandensis* MSMB43 (currently being described as “*Burkholderia*

hympydoensis”).^{93,128,129} Second, *Fr9K* is a key 3-hydroxy-3-methylglutaryl-CoA synthase (HCS) homologue-enzyme catalyzing the transfer of $-\text{CH}_2\text{COO}-$ from acyl-S-acyl carrier protein (ACP) to a β -keto thioester polyketide intermediate. *Fr9K* has a high identity (87%) to one of the enzymes (*TstK*) predicted from the genome of *B. thailandensis* MSMB43. Further, a regulator gene expression studies for optimization of production media, and targeted isolation and purification efforts toward diene compounds (UV235 nm), led to the discovery of three compounds called thailanstatins⁹³ and a group of similar or identical compounds called spliceostatins^{130,131} (Fig. 4). Particularly, thailanstatin A contains a carboxyl group, which not only makes it more stable than FR in PBS solution, but also leads to increased activities compared to FR when carboxylic groups were esterified (with enhanced membrane permeability).¹³⁰

2.5. *Kutzneria* species are potential resources of CYP-catalyzed compounds

In the *Actinobacteria* phylum, the genus *Kutzneria* is a minor branch of the *Pseudonocardiaceae* family, currently containing eight species (Fig. 5). Only aculeximycin and kutznerides are known to be produced by *Kutzneria albida* DSM 43870^T and *Kutzneria* sp. 744, respectively.^{132,133} Genome mining of *Kutzneria* species suggests that they could produce compounds similar to the previously reported pre-mRNA spliceosome inhibitors and a group of interesting CYP catalyzed compounds. For example, the pladienolide biosynthetic gene cluster has only four genes (including the *pldB* CYP gene, and the putative epoxidase encoding gene *pldD*) besides the PKS genes.⁵⁹ Inspired by the successful discovery of thailanstatins, pladienolide B CYP monooxygenase-based bioinformatics analysis was conducted and revealed that the *Kutzneria* sp. 744 genome contains the most homologous (78%) enzyme KUTG_06291 (EWM16617) when CYP PldB sequence was used as input for homology search of the GeneBank (Fig. 6). Therefore, the genus *Kutzneria* was postulated to produce compounds similar to pladienolide. However, there are still a lot of gaps in the sequenced genome of *Kutzneria* sp. 744.

Kutzneria albida DSM 43870^T is the only *Kutzneria* strain with a complete sequenced genome. So far, the *K. albida* genome (9.87 Mb) is among the largest actinobacterial genomes sequenced.¹³⁴ Thus, *Kutzneria albida* DSM 43870^T was selected for further phylogenetic analysis of CYPs in parallel with a library of 44 CYPs originate from biosynthetic gene clusters of 32 bioactive natural products.^{37,61–64,73,74,76,78,80–86,88–92,94–100,102–105} AntiSMASH analysis revealed that the *K. albida* DSM 43870^T genome encodes 47 biosynthetic gene clusters which contain at least 13 different CYPs distributed in eight gene clusters (GC4, GC10, GC11, GC18, GC19, GC33, GC34 and GC40)¹³⁴ (Fig. 5, Table S1). Among those gene clusters, only GC40 is known to be responsible for the biosynthesis of aculeximycin. The CYP KALB_6568 is predicted to be responsible for catalyzing the formation of C14 hydroxyl group. However, the compounds biosynthesized by the other CYP-related gene clusters are not known yet. BLAST analysis (Fig. 2) suggests that in the *K. albida* DSM 43870^T genome, GC19 encodes compounds with CYP (KALB_3944, possessing 99% identity to *Mei-Orf4*)-encoded functional groups similar to meilingmycin⁷⁸ and tirandamycin B⁹⁸ (66% identity to *TamI*). As for GC11, the 99% identity between CYP KALB_3411 and *Stap* (in the staurosporine biosynthesis gene cluster¹⁰⁵) suggests that GC11 may be responsible for the biosynthesis of compounds with functional groups formed by similar oxidative decarboxylations. Additionally, antiSMASH analysis indicates that GC11 has a 20% similarity to the biosynthetic gene clusters of the pre-mRNA spliceosome inhibitor staurosporine. Similarly, the 99% identity between CYP KALB_3295 (in GC10) and *Asm30* (CYP in ansamitocin biosynthesis gene cluster⁹⁴) implies that GC10 may produce epoxide compounds similar to ansamitocin. Similar anal-

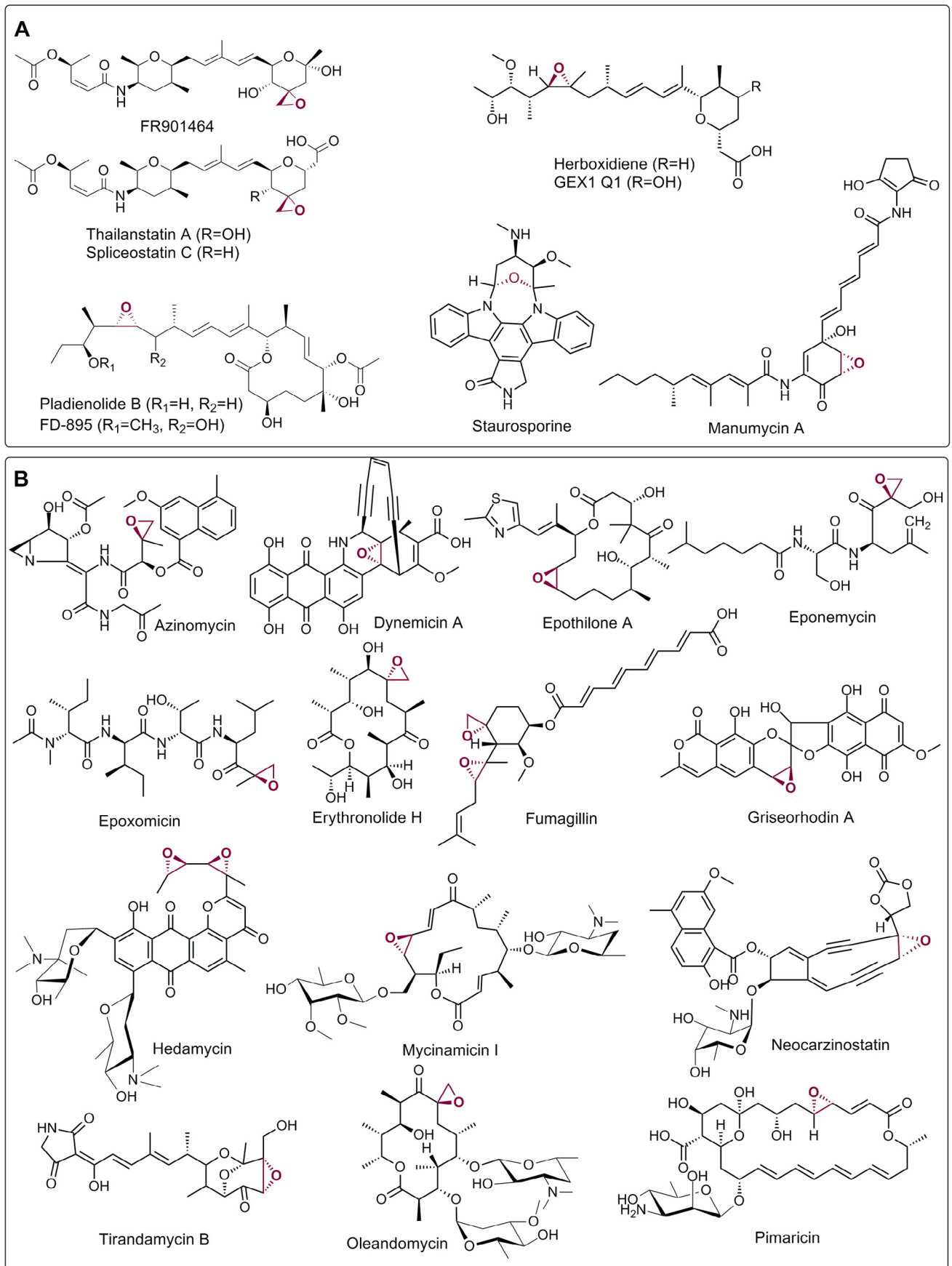


Fig. 1. P450 related microbial natural compounds. Except for staurosporine, all the selected compounds are epoxide-containing compounds. Other epoxide compounds without known biosynthetic pathways, such as antitumor compound trapoxin, are not listed. Panel A: Natural products with known spliceosome inhibitory activities; Panel B: Natural products with other activities such as tubulin-binding compounds epothilone A, telomerase inhibitor griseorhodin A, RNA polymerase inhibitor tirandamycin and proteasome-inhibitor epoxomicin, etc.

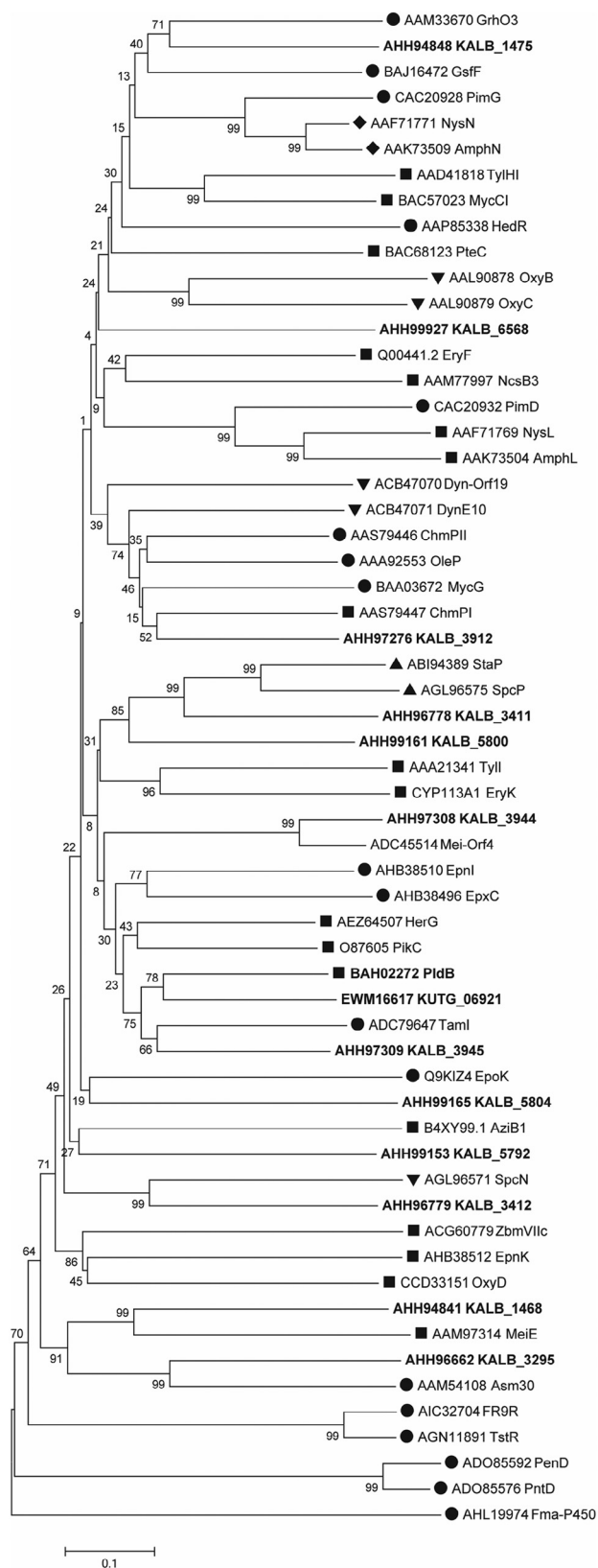


Fig. 2. Neighbor-joining tree of selected P450 enzymes including 13 putative P450 enzymes from *Kutzneria* species. CYPs catalyze epoxide formation (●), hydroxylations (■), oxidation of methyl groups (◆), decarboxylations (▲) and C–C or C–C bond formations (▼). The tree was generated by using MEGA6 using the neighbor joining method. Significant bootstrap values are indicated at the nodes. The scale bar represents 0.1 mutational events per site. PldB, Fr9R, GsfF and StaP associated gene clusters produce pre-mRNA spliceosome inhibitors pladienolid, FR901464, FD-891 and staurosporine, respectively.

ysis suggests that tirandamycin-like compounds may be produced through GC19 based on the 66% identity between KALB_3945 and TamI.^{98,135} Unfortunately, the 27% identity between KALB_5792 (in GC33) and AziB1 (CYP in the azinomycin B biosynthetic gene cluster⁷⁴), and the 19% identity between KALB_5804 (in GC33) and EpoK (CYP in the epothilone biosynthesis gene cluster⁶³) make it unpredictable for the natural products biosynthesized by GC33.

Except for *Kutzneria albida* DSM 43870^T, another top five strains, *Actinoplanes* sp. N902–109 (NC_021191.1), *Micromonospora aurantiaca* ATCC 27029 (CP002162.1), *Streptomyces violaceusniger* Tu 4113 (CP002994.1), *Streptomyces bingchenggensis* BCW-1 (CP002047.1) and *Streptomyces* sp. PVA 94-07 (CM002273.1), were prioritized through CYP pattern-based genome mining of pladienolide biosynthetic pathway. These strains have the potential CYP pattern pathways to produce unknown compounds (Tables S2–S6).

3. Conceivable methods to translate biosynthetic pathways to natural products

This section covers the methods (Fig. 7) or tools that have been used or have the potential to be applied to translate biosynthetic pathways to natural products.

3.1. Direct cloning and heterologous expression of biosynthetic gene clusters

Heterologous expression was primarily used to avoid the lack of genetic tools and regulatory complexity in native hosts via the use of strains that are more amenable to engineering. Direct cloning, genetic engineering and heterologous expression of microbial natural product biosynthesis pathways were reviewed in 2013.¹³⁶ The heterologous expression methods can also be borrowed from those for the discovery of small molecules from metagenomics cosmid library.^{7,137,138} The approaches to cloning targeted gene clusters directly from genomic DNA include RecE-mediated homologous recombination,¹³⁹ oriT-directed capture,¹⁴⁰ transformation-associated recombination (TAR)^{141–143} to facilitate functional expression experiments and phage ϕ T1 integrase-mediated site-specific recombination.¹⁴⁴ Among those, TAR cloning in *Saccharomyces cerevisiae* has been extensively applied to capture and express large biosynthetic gene clusters from environmental DNA samples.^{145–147} Recently, the TAR direct cloning approach was used to clone targeted whole gene cluster from rare actinomycetes and Gram-negative bacteria. By using TAR, the targeted gene cluster was cloned into plasmid pCAP01 by homologous recombination in *Saccharomyces cerevisiae* strain VL6-48 in order to obtain the captured vector, and then the captured vector was transformed into the corresponding host *Streptomyces coelicolor* and *E. coli* for rare actinomycete and Gram-negative bacteria, respectively. An example is the heterologous production of lipopeptide taromycin A (from marine actinomycete *Saccharomonospora* sp. CNQ-490) in *Streptomyces coelicolor* M512¹⁴² and the heterologous production of alterochromide lipopeptides (from *Pseudoalteromonas piscicida* JCM 20779) in *E. coli* BL21 (DE3) utilizing native and *E. coli*-based T7 promoter sequences.¹⁴¹ Since the development of a TAR-based genetic platform allows for heterologous production of lipopeptides in different hosts, efforts have been focused on the development of the high throughput TAR capture method to express the pathway.¹⁴¹

Successful production of the desired products often requires an optimal relationship of timing and flux between primary and secondary cellular metabolism. Besides the use of genetic engineering for the direct cloning of a whole gene cluster, there has been considerable interest in the development of engineered bacterial strains for efficient heterologous production of secondary metabolites.^{148–151} For example, the deletion of a 1.4 Mb segment from the left subtelomeric region of the 9.02 Mb *Streptomyces avermitilis* genome

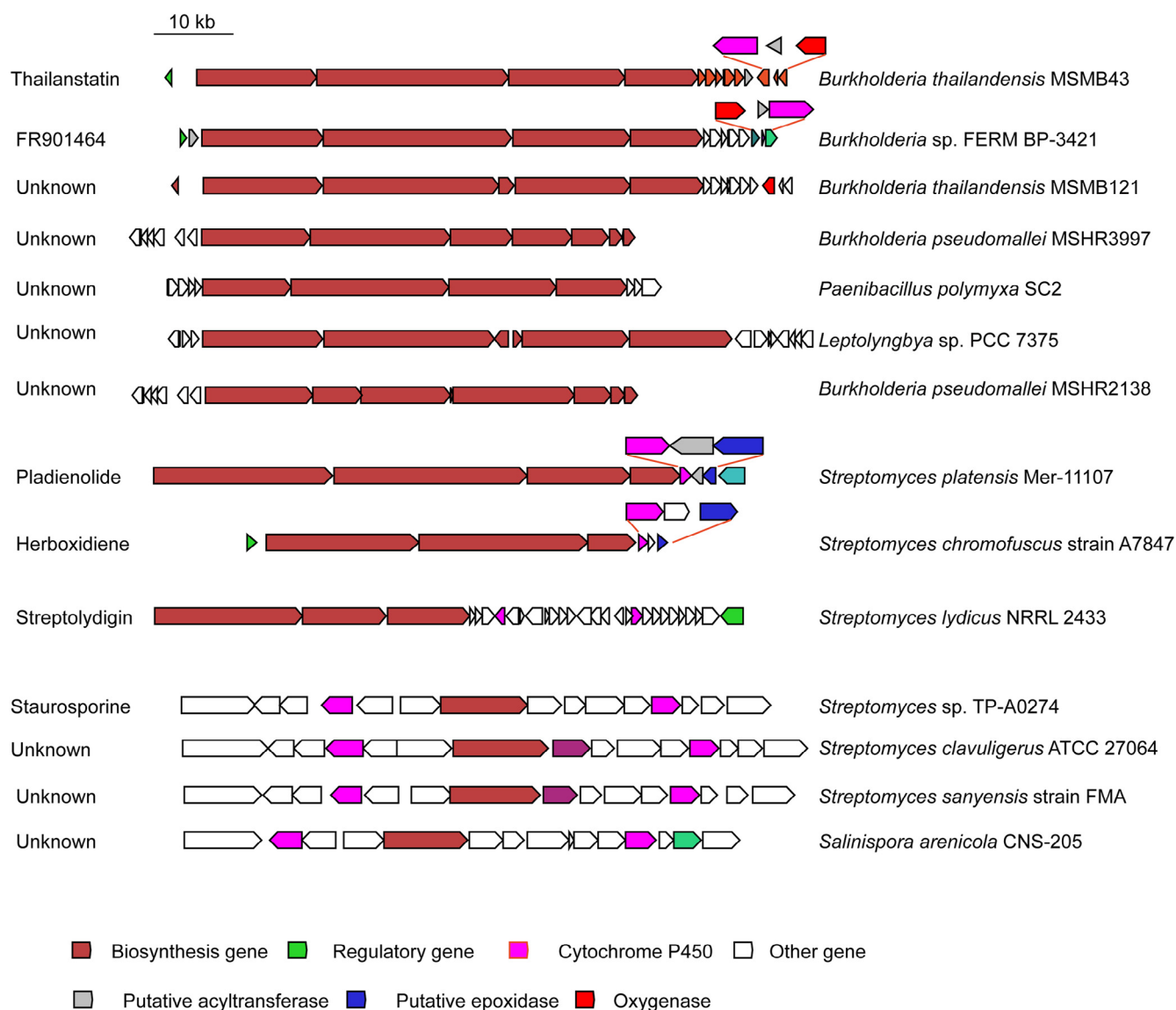


Fig. 3. Pattern-based genome mining of homologous gene clusters of spliceosome inhibitors which contain CYPs. The biosynthetic gene clusters of thailanstatins (Accession no. KJ461964.1) and FR901464 (Accession no. JX307851.1), pladienolide (Accession no. AB435553.1) and herboxidiene (Accession no. JN671974.1) as well as staurosporine (Accession no. AB088119.1) were used for the search of homologous gene clusters by using antiSMASH 3.0.

resulted in the generation of two large-deletion mutants. By using these large-deletion mutants, i.e., *S. avermitilis* SUKA17 or 22, twenty of the entire biosynthetic gene clusters for secondary metabolites, including aminoglycosides, nonribosomal peptides, polyketides and terpenes were successfully expressed.¹⁵² The biosynthetic gene cluster of the polyketide pladienolide has been expressed in a deletion mutant of *Streptomyces avermitilis* with an extra copy of the regulatory gene *pldR* under control of an alternative promoter.¹⁵² It is worthy to note that the engineered hosts are not only useful for the production of exogenous secondary metabolites, but they also facilitate scale-up production and preparation of promising natural products due to their “clean” background.

3.2. Manipulation of the biosynthetic pathways

The recent approaches on the activation and up-regulation of microbial biosynthetic pathways for the discovery of natural products have been reviewed.^{153,154} Except for these methods, there are four types of research reports that may explain how to translate bio-

synthetic gene clusters to novel natural products by using synthetic biology methods. First, optimized bioactive compounds can be produced by engineering the biosynthetic operon and reconstituting the biosynthetic pathway. The group led by Müller and Brönstrup successfully deleted the gene in hydroxymalonyl-CoA biosynthetic operon of the bengamide biosynthetic gene cluster and inserted a promoter on the expression construct pBen32. Heterologous expression of the modified biosynthetic gene cluster led to the discovery of more potent compounds.¹⁵⁵ Second, new compounds with increased activities are produced by inactivation of CYP pattern genes. Bills’ team works on the manipulation of pneumocandin biosynthetic pathway. They focused on the inactivation of three genes (GLP450-1, GLP450-2, and GLOXY1) and generated 13 different pneumocandin analogues that lack one, two, three, or four hydroxyl groups on 4*R*, 5*R*-dihydroxy-ornithine and 3*S*, 4*S*-dihydroxy-homotyrosine of the parent hexapeptide. Among them, seven analogues are previously unreported.¹⁵⁶ Third, natural product analogues can be produced by creating promoter-driven tailoring enzyme constructs. Brady’s group created an *ermE* promoter cas-

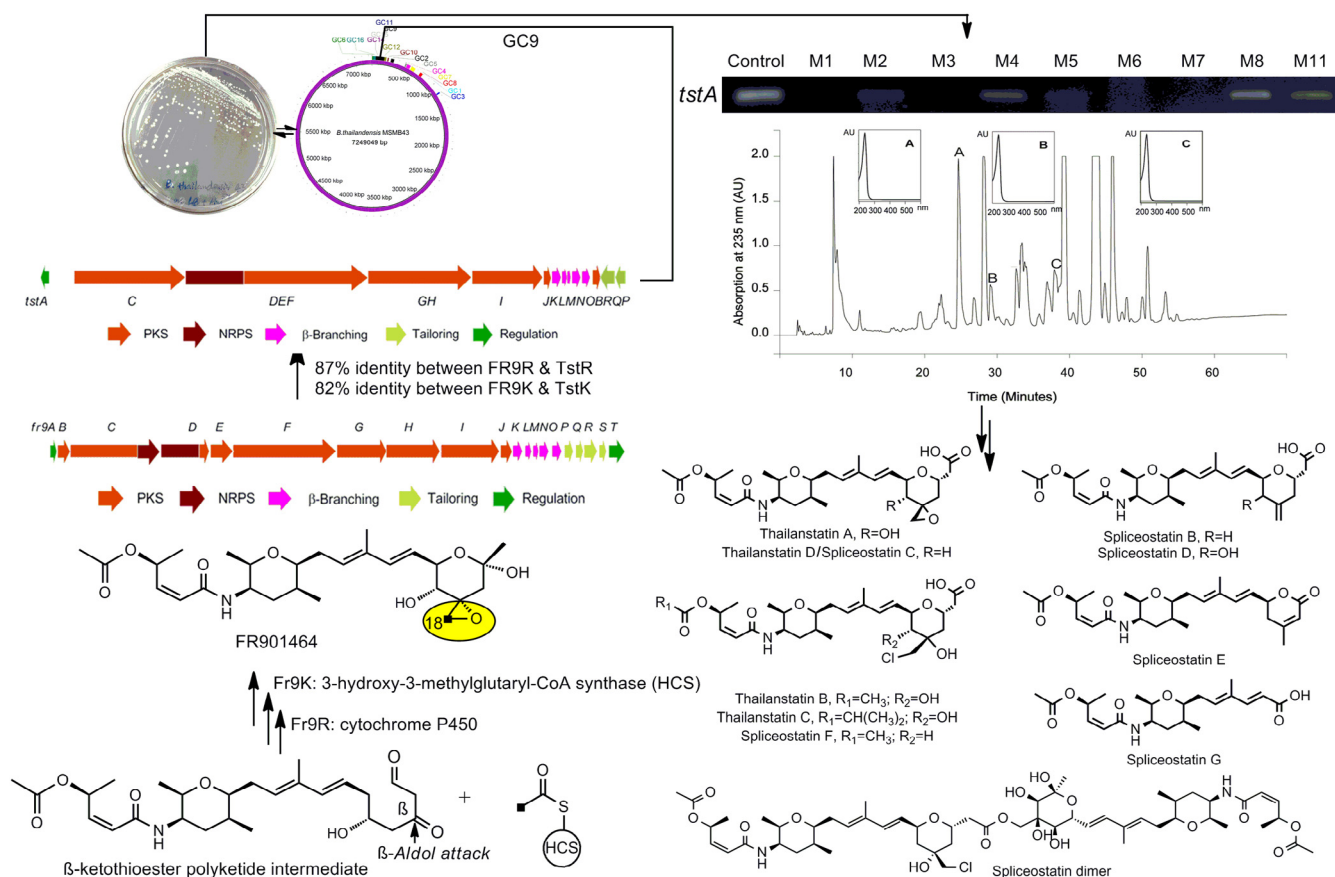


Fig. 4. Genome mining protocol of *Burkholderia thailandensis* MSMB43 for thailanstatins/spliceostatins. Starting from the two key biosynthetic enzymes CYP Fr9R and 3-hydroxy-3-methylglutaryl-CoA synthase (HCS) Fr9K that catalyze the formation of the hemiketal hydroxyl group and the epoxide group of FR901464, the homologous enzymes and biosynthetic gene cluster of thailanstatin were predicted. The Reverse Transcriptional (RT)-PCR help to identify the optimum medium for the expression of regulatory gene *tstA*. The novel compounds were finally obtained through microbial fermentation, natural product isolation and structural elucidation by tracking of the characteristic UV absorption for the diene motif at UV235 nm.

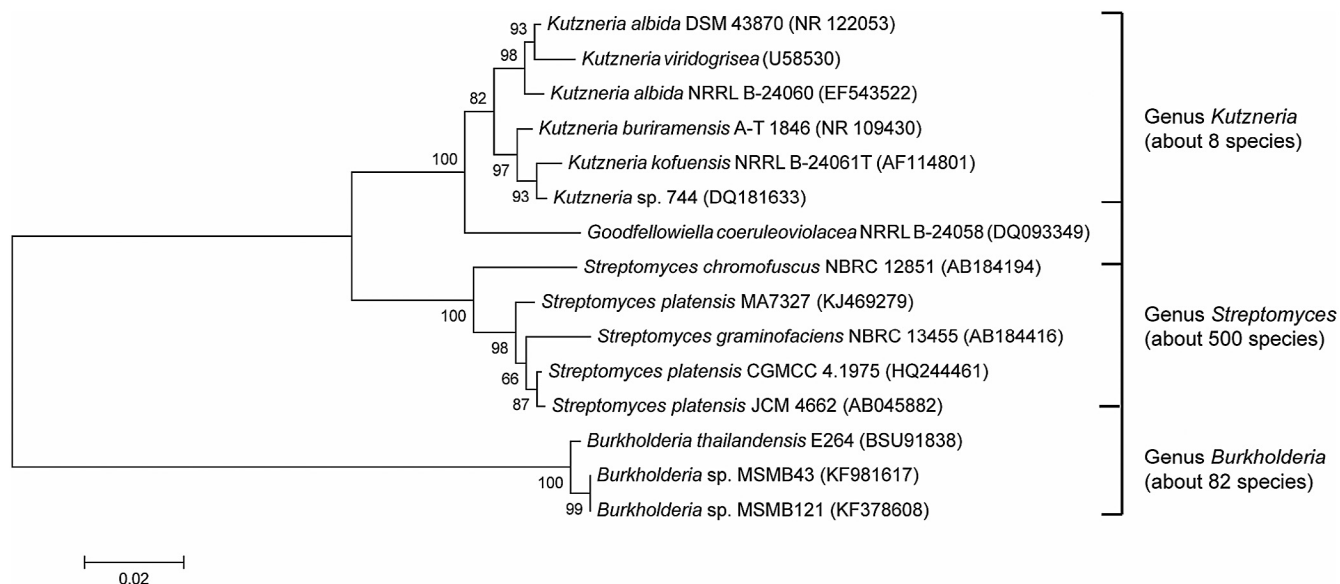


Fig. 5. Phylogenetic tree for taxa of the genus *Kutzneria* and other antitumor compound-producing species (*Streptomyces* and *Burkholderia*). The tree was calculated from complete 16S rRNA gene sequences using the neighbor joining method, illustrating the genus *Kutzneria* position relative to other selected species. All species are type strains except for the strains *Kutzneria* sp. 744, *Burkholderia* sp., MSMB43 and MSMB121. The 16S rRNA sequence of the pladienolide producer *Streptomyces platensis* Mer-11107 is not available, but DNA hybridization suggests that *Streptomyces platensis* Mer-11107 has 87% similarity to the type strain CGMCC4.1975. Percentages at nodes represent levels of bootstrap support from 1000 resampled datasets; Scale bar, 0.02 nucleotide substitutions per site.

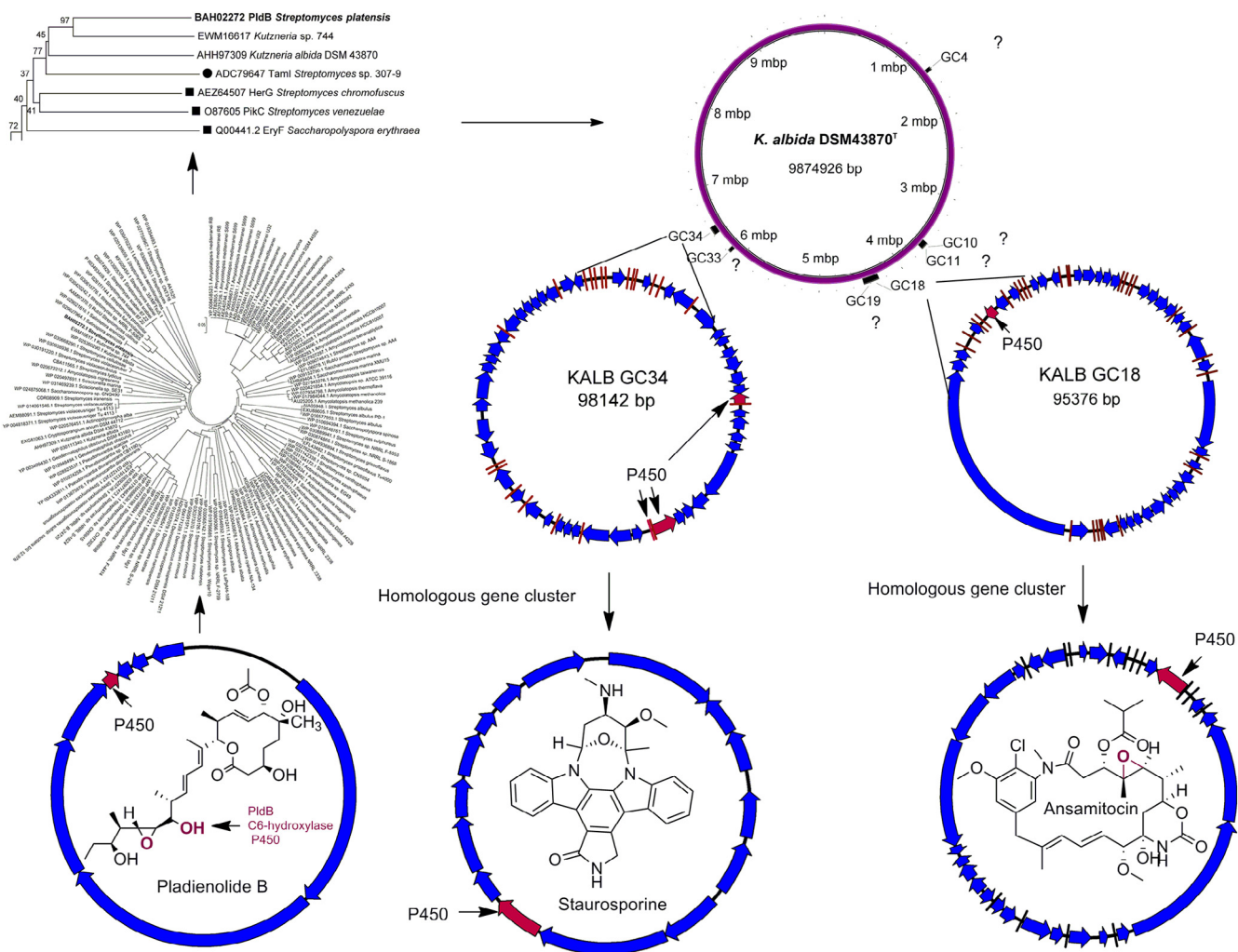


Fig. 6. Strategy for the discovery of novel compounds from *Kutzneria* species. By starting with the structure of pladienolide B, and moving through phylogenetic analysis of the streptomycetes-related cytochrome P450s, *Kutzneria* species were selected for further analysis of P450-related polyketides. The genome sequencing of the closest strain of *Kutzneria* sp. 744 was finished on February, 2014 without detailed annotation and there are still a lot of gaps waiting to be closed.

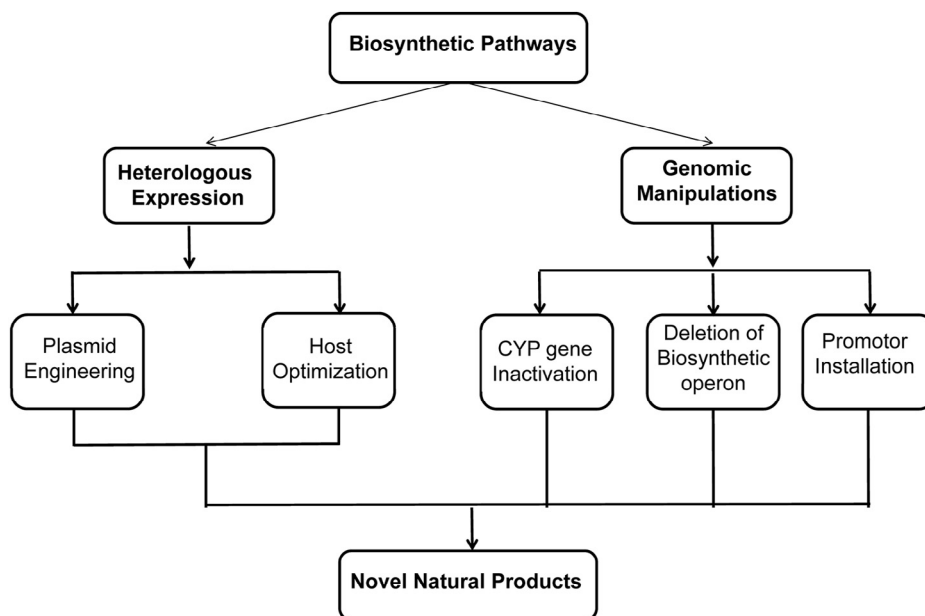


Fig. 7. Schematic representation of methods to translate biosynthetic pathways to novel natural products.

sette, which was introduced to the upstream of the first ORF of the biosynthetic gene cluster of interest. Further introduction of the cosmid containing the constitutively expressed tailoring operon into *S. toyocaensis*: Δ Stal resulted in the production of three new glycopeptides.⁴⁴ Fourth, stable compounds with increased titer are produced by engineering the CYP pathway. Researchers in Pfizer engineered a pAE-PF29 vector which enabled the overexpression of Fr9R encoded by CYP *fr9R* in *Burkholderia* sp. FERM BP-3421, and led to the enhanced production of stable thailanstatin A from spliceostatin C.¹⁵⁷

3.3. Synthetic biology tool kits

Except for the four examples of genetic manipulations, synthetic biology tools to edit bacterial genomes have been reviewed by other researchers.¹⁵⁸ The Clustered Regularly Interspaced Short Palindromic Repeats and Cas proteins (CRISPR-Cas) systems are composed of a powerful and broadly applicable set of tools to manipulate *Streptomyces* genomes. There are three research groups that worked on the gene deletion of *Streptomyces* species by using different CRISPR-Cas toolkits. The group led by Zhao developed a temperature-sensitive pCRISPomyces system, which is applicable to the genome editing of *Streptomyces lividans* 66, *Streptomyces viridochromogenes* DSM 40736 and *Streptomyces albus* J1074.¹⁵⁹ Lee and Webber's group focused on the gene deletion efficiency of actinorhodin biosynthetic gene cluster in *Streptomyces coelicolor* A3 by an engineered CRISPR-Cas system with improved efficiency.¹⁶⁰ Sun's team worked on the discovery of a novel CRISPR-Cas system in combination with the counterselection system CodA(sm), and the D314A mutant of cytosine deaminase, to delete the actinorhodin biosynthetic gene cluster in *Streptomyces coelicolor* M145 genome.¹⁶¹

4. The advantages and challenges of the above mentioned strategies

The advantages and challenges to generate a bioactive natural product library by CYP pattern-based genome mining as compared to the traditional strategies such as microbial fermentation with different media and chromatographic fractionation for different polarities can be summarized as follows.

4.1. Advantages

- CYP is a class of most important tailoring enzymes that can be explored and exploited for the structural diversification of bioactive natural products.
- Genome mining is an efficient method. The combination of PCR screening and genome mining will prioritize the cryptic gene clusters of interest for annotation and heterologous expression. This is more efficient when compared to the traditional conception of one strain many active compounds (OSMAC).
- Genome mining can be used for large scale and high throughput method to accelerate drug discovery, as exemplified by a four-year case study on genome mining of 10,000 actinomycetes for 19 novel phosphonic acid natural products.⁵⁷
- Both heterologous expression and activation of cryptic pathways can generate novel compounds that have unique biological activities but are difficult to access using synthetic chemistry.
- CRISPR-cas9 enabled genome editing tools are very efficient. First, compared to the 5% (2/38) efficiency in the traditional in-frame gene deletion method,⁶² the efficiency of the CRISPR-Cas9 systems can be 100% without unwanted side- or off-target effects.¹⁶⁰ Second, CRISPR-Cas9 systems can overcome the employment of the counter selectable marker for the selection of double crossover mutants. Third, CRISPR-Cas9 systems can avoid the scar sequence left at the target site and realize mul-

multiple gene deletions. Fourth, CRISPR/Cas system provides new modularity, which can target the site of interest by inserting of a short spacer into a CRISPR array/sgRNA construct. The insertion can be achieved with high throughput using modern DNA assembly techniques. CRISPR/Cas9 genome editing can shorten the two-month period of previous gene deletion method to the current 1–2 week time course.¹⁶¹

4.2. Main challenges

- How to prioritize the characterization of orphan biosynthetic gene clusters, and how to rapidly connect genes to biosynthesized small molecules will come with increase in DNA-sequenced data.¹⁶²
- Pathway manipulations in native hosts are subjected to the genetic complexity of the bacteria. For example, researchers have developed a procedure which takes about 10 days for the chromosomal knock out of Gram-negative bacteria *B. pseudomallei* by using the *phe5-gat* cassette-assisted by plasmid pKaKa2.¹⁶³ However, this method is not applicable to the gene knockout of another strain called *B. gladioli*, which requires another pCR4Blunt-TOPO-based vector for the recombination assisted by plasmid pKD46.¹⁶⁴
- Up to now, Cas9 targeting is still limited by protospacer adjacent motif (PAM) sequences. Also, targeting efficiency is site-dependent.
- Although the development of genetic manipulation tools could greatly enhance the chances toward discovery and production of natural products, a major challenge in the process of microbial genome mining is to produce compounds in high titers.¹⁶⁵
- Bacterial genomes are publically available but the accessibility of strains of interest is limited due to the regulations of international or domestic properties. Decreasing DNA synthesis costs and advances in DNA assembly could help to solve the issue with limited material access. A larger fraction of strains that are isolated in research labs worldwide will be the future challenge.¹⁶⁶

5. Conclusion

The current review provides strategies on the discovery of a group of epoxide-containing natural products (Fig. 8), highlights (1) CYP and its pattern enzyme-based genome mining as a guidance for the generation of a diversified natural product library, (2) direct cloning and heterologous expression of biosynthetic pathways and genomic manipulation methods and tools will translate the selected biosynthetic gene clusters to the bioactive natural product library. It is worthy to note that tailoring reactions play important roles on diversifications of scaffolds such as polyketide, peptide and hybrid polyketide-peptide backbones, and often tailoring enzyme patterns such as CYPs, ligases, Cyclases, ketoreductases, transferases, and oxygenases can be used for genome mining. This advanced genome mining will avoid the de-replications of biosynthetic gene clusters to quickly identify the annotated biosynthetic gene clusters from the vast pool of sequenced bacterial genomes. It is predicted that we will be sequencing genomes for pennies as nearly as 2020 (<https://youtu.be/j88APStUcp4>). Synthetic biology tools pioneered by different researchers will be continuously developed toward a high-throughput potential due to the increasing numbers of sequenced bacterial genomes. This review also points out the potential mechanisms and diversity of CYPs in the microbial biosynthesis of natural product antitumor agents. The *Kutzneria* strains, as rare actinobacterial species, can be explored and exploited for CYP catalyzed compounds, and can be used as a rich resource for the diversification of microbial CYPs.

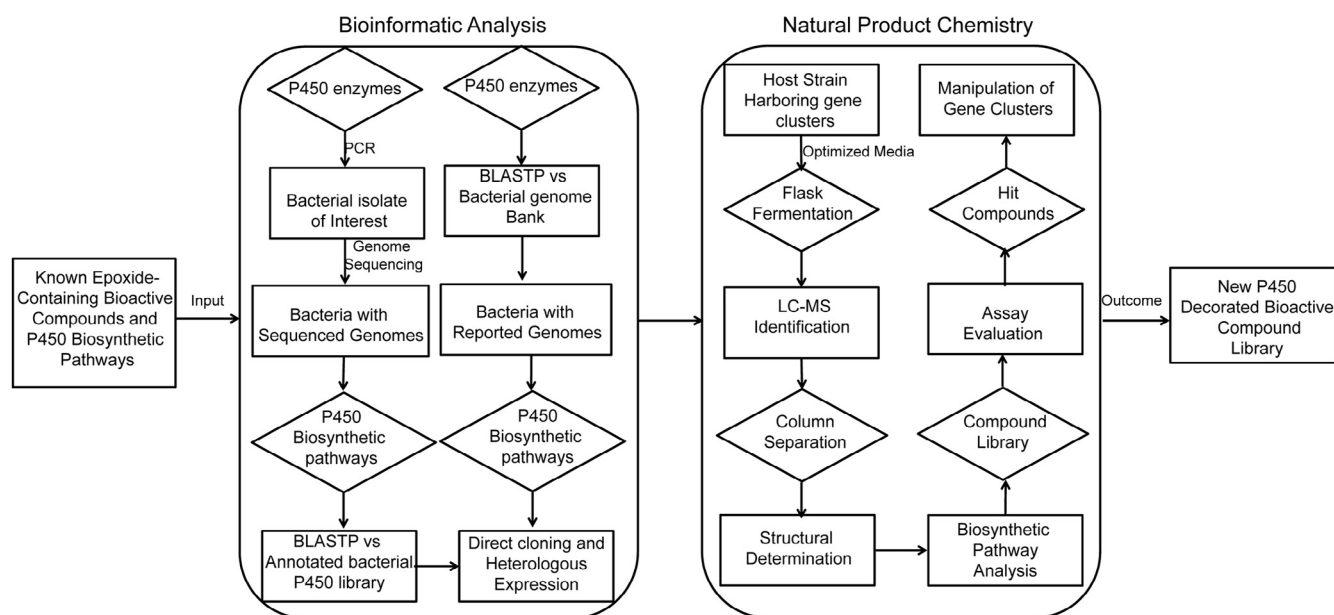


Fig. 8. The proposed protocol for the development of a natural product library in postgenomic stage.

Acknowledgement

Prof. Lixin Zhang in Institute of Microbiology, Chinese Academy of Sciences is greatly appreciated for the invitation for this review. Prof. Arnold Demain from Drew University, Dr. Jeremy Barnes from Alcon, a Novartis Company, Prof. Chandrakant Kokare from Sinhgad Institute of Pharmacy, and Prof. Shengying Li from Qingdao Institute of Bioenergy and Bioprocess Technology are acknowledged for the critical linguistic comments. Prof. David R. Nelson from University of Tennessee is thanked for the providing the guidance on the microbial P450 database. The author has been supported by a US National Institute of Health/National Cancer Institute grant (R01 CA152212) awarded to Prof. Yi-Qiang Cheng in the UNT System College of Pharmacy, University of North Texas Health Center in Fort Worth, Texas, USA.

Appendix: Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.synbio.2016.01.007.

References

- OECD. Biological resource centres. Underpinning the future of life science and biotechnology. <http://www.SourceOECD.org>.
- Liu X, Ashforth E, Ren B, Song F, Dai H, Liu M, et al. Bioprospecting microbial natural product libraries from the marine environment for drug discovery. *J Antibiot (Tokyo)* 2010;**63**:415–22.
- Knight V, Sanglier JJ, DiTullio D, Braccili S, Bonner P, Waters J, et al. Diversifying microbial natural products for drug discovery. *Appl Microbiol Biotechnol* 2003;**62**:446–58.
- Zhang L, An R, Wang J, Sun N, Zhang S, Hu J, et al. Exploring novel bioactive compounds from marine microbes. *Curr Opin Microbiol* 2005;**8**:276–81.
- Hindra, Huang T, Yang D, Rudolf JD, Xie P, Xie G, et al. Strain prioritization for natural product discovery by a high-throughput real-time PCR method. *J Nat Prod* 2014;**77**:2296–303.
- Reddy BV, Milshteyn A, Charlop-Powers Z, Brady SF. eSNaPD: a versatile, web-based bioinformatics platform for surveying and mining natural product biosynthetic diversity from metagenomes. *Chem Biol* 2014;**21**:1023–33.
- Charlop-Powers Z, Milshteyn A, Brady SF. Metagenomic small molecule discovery methods. *Curr Opin Microbiol* 2014;**19**:70–5.
- Bixby KA, Thomas MG, Shen B. Combinatorial biosynthesis of anticancer natural products. In: Cragg GM, Kingdon DG, Newman DJ, editors. *Anticancer agents from natural products*. Boca Raton, FL: Taylor & Francis Group; 2012. p. 671–98.
- Singh SB, Jayasuriya H, Ondeyka JG, Herath KB, Zhang C, Zink DL, et al. Isolation, structure, and absolute stereochemistry of platensimycin, a broad spectrum antibiotic discovered using an antisense differential sensitivity strategy. *J Am Chem Soc* 2006;**128**:11916–20.
- Wang J, Kodali S, Lee SH, Galgoci A, Painter R, Dorso K, et al. Discovery of platencin, a dual FabF and FabH inhibitor with in vivo antibiotic properties. *Proc Natl Acad Sci U S A* 2007;**104**:7612–6.
- Bugni TS, Richards B, Bhoite L, Cimbora D, Harper MK, Ireland CM. Marine natural product libraries for high-throughput screening and rapid drug discovery. *J Nat Prod* 2008;**71**:1095–8.
- Hu Y, Potts MB, Colosimo D, Herrera-Herrera ML, Legako AG, Yousufuddin M, et al. Discoipyrroles A-D: isolation, structure determination, and synthesis of potent migration inhibitors from *Bacillus humanensis*. *J Am Chem Soc* 2013;**135**:13387–92.
- Du L, King JB, Morrow BH, Shen JK, Miller AN, Cichewicz RH. Diarylcyclopentendione metabolite obtained from a *Preussia typharum* isolate procured using an unconventional cultivation approach. *J Nat Prod* 2012;**75**:1819–23.
- Du L, Robles AJ, King JB, Powell DR, Miller AN, Mooberry SL, et al. Crowdsourcing natural products discovery to access uncharted dimensions of fungal metabolite diversity. *Angew Chem Int Ed Engl* 2014;**53**:804–9.
- Quinn RJ, Carroll AR, Pham NB, Baron P, Palframan ME, Suraweera L, et al. Developing a drug-like natural product library. *J Nat Prod* 2008;**71**:464–8.
- Kim MK, Kim S, Park SJ, Lee H, Kim TS, Jung SY, et al. Abstract 1702: High-throughput cell-based screening of drug library identifies albendazole as a sensitizer with combination of bortezomib for treatment multiple myeloma. *Cancer Res* 2014;**74**:1702. <http://www.selleckchem.com/>.
- Starosyla SA, Volynets GP, Lukashov SS, Gorbatiuk OB, Golub AG, Bdzhola VG, et al. Identification of apoptosis signal-regulating kinase 1 (ASK1) inhibitors among the derivatives of benzothiazol-2-yl-3-hydroxy-5-phenyl-1,5-dihydropyrrol-2-one. *Bioorg Med Chem* 2015;**23**:2489–97. <http://www.otavachemicals.com>.
- Venkataraman SK, Moores A, Stone A, Hurst A, Mikell JR, Moraes RM, et al. Fractional extraction of plant biomass: generation of botanical extract libraries. *Planta Med* 2009;**75**:PA36. <https://chromadex.com/Default.aspx>.
- Camp D, Davis RA, Campitelli M, Ebdon J, Quinn RJ. Drug-like properties: guiding principles for the design of natural product libraries. *J Nat Prod* 2012;**75**:72–81.
- Gaudencio SP, Pereira F. Dereplication: racing to speed up the natural products discovery process. *Nat Prod Rep* 2015;**32**(6):779–810.
- Fredenhagen A, Derrien C, Gassmann E, An MS. MS library on an ion-trap instrument for efficient dereplication of natural products. Different fragmentation patterns for $[M + H]^+$ and $[M + Na]^+$ ions. *J Nat Prod* 2005;**68**:385–91.
- Yang JY, Sanchez LM, Rath CM, Liu XT, Boudreau PD, Bruns N, et al. Molecular networking as a dereplication strategy. *J Nat Prod* 2013;**76**:1686–99.
- Bian J, Song F, Zhang L. Strategies on the construction of high-quality microbial natural product library – a review. *Acta Microbiol Sin* 2008;**48**:1132–7.
- Carter GT. Natural products and Pharma 2011: strategic changes spur new opportunities. *Nat Prod Rep* 2011;**28**:1783–9.
- Walsh CT. Combinatorial biosynthesis of antibiotics: challenges and opportunities. *ChemBiochem* 2002;**3**:125–34.

26. Reeves CD. The enzymology of combinatorial biosynthesis. *Crit Rev Biotechnol* 2003;**23**:95–147.
27. Sieber SA, Marahiel MA. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* 2005;**105**:715–38.
28. Menzella HG, Reeves CD. Combinatorial biosynthesis for drug development. *Curr Opin Microbiol* 2007;**10**:238–45.
29. Mootz HD, Schwarzer D, Marahiel MA. Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBiochem* 2002;**3**:490–504.
30. Shen B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* 2003;**7**:285–95.
31. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* 2006;**106**:3468–96.
32. Wilkinson B, Micklefield J. Mining and engineering natural-product biosynthetic pathways. *Nat Chem Biol* 2007;**3**:379–86.
33. Walsh CT. A chemocentric view of the natural product inventory. *Nat Chem Biol* 2015;**11**:620–4.
34. Skinnider MA, Dejong CA, Rees PN, Johnston CW, Li H, Webster AL, et al. Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res* 2015;**43**:9645–62.
35. Urlacher VB, Girhard M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol* 2012;**30**:26–36.
36. Whitehouse CJ, Bell SG, Wong LL. P450(BM3) (CYP102A1): connecting the dots. *Chem Soc Rev* 2012;**41**:1218–60.
37. Robert FO, Pandhal J, Wright PC. Exploiting cyanobacterial P450 pathways. *Curr Opin Microbiol* 2010;**13**:301–6.
38. Kelly SL, Kelly DE. Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philos Trans R Soc Lond B Biol Sci* 2013;**368**:20120476.
39. Wiseman A. Enzyme induction in microbial organisms. *Basic Life Sci* 1975;**6**:1–26.
40. Challis GL, Ravel J, Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol Lett* 2000;**187**:111–4.
41. Moss SJ, Martin CJ, Wilkinson B. Loss of co-linearity by modular polyketide synthases: a mechanism for the evolution of chemical diversity. *Nat Prod Rep* 2004;**21**:575–93.
42. Teta R, Gurgui M, Helfrich EJ, Kunne S, Schneider A, Van Echten-Deckert G, et al. Genome mining reveals *trans*-AT polyketide synthase directed antibiotic biosynthesis in the bacterial phylum bacteroidetes. *ChemBiochem* 2010;**11**:2506–12.
43. Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* 2015;**43**(W1):W237–43.
44. Owen JG, Reddy BV, Ternei MA, Charlop-Powers Z, Calle PY, Kim JH, et al. 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**:11797–802.
45. Gontang EA, Gaudencio SP, Fenical W, Jensen PR. Sequence-based analysis of secondary-metabolite biosynthesis in marine actinobacteria. *Appl Environ Microbiol* 2010;**76**:2487–99.
46. Kersten RD, Lane AL, Nett M, Richter TK, Duggan BM, Dorrestein PC, et al. Bioactivity-guided genome mining reveals the lomaiviticin biosynthetic gene cluster in *Salinispora tropica*. *ChemBiochem* 2013;**14**:955–62.
47. Cane DE, Ikeda H. Exploration and mining of the bacterial terpenome. *Acc Chem Res* 2012;**45**:463–72.
48. Johnston CW, Skinnider MA, Wyatt MA, Li X, Ranieri MR, Yang L, et al. An automated genomes-to-natural products platform (GNP) for the discovery of modular natural products. *Nat Commun* 2015;**6**:8421.
49. Duncan KR, Crusemann M, Lechner A, Sarkar A, Li J, Ziemert N, et al. Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species. *Chem Biol* 2015;**22**:460–71.
50. 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* 2015;doi:10.1007/s10295-015-1671-0.
51. Zhao S, Sakai A, Zhang X, Vetting MW, Kumar R, Hillerich B, et al. Prediction and characterization of enzymatic activities guided by sequence similarity and genome neighborhood networks. *eLife* 2014;**3**.
52. Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR. The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS ONE* 2012;**7**:e34064.
53. Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, et al. A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat Biotechnol* 2003;**21**:187–90.
54. Wang C, Wesener SR, Zhang H, Cheng YQ. An FAD-dependent pyridine nucleotide-disulfide oxidoreductase is involved in disulfide bond formation in FK228 anticancer depsipeptide. *Chem Biol* 2009;**16**:585–93.
55. Li J, Wang C, Zhang ZM, Cheng YQ, Zhou J. The structural basis of an NADP(+)-independent dithiol oxidase in FK228 biosynthesis. *Sci Rep* 2014;**4**:4145.
56. Wang C, Henkes LM, Doughty LB, He M, Wang D, Meyer-Almes FJ, et al. Thailandepsins: bacterial products with potent histone deacetylase inhibitory activities and broad-spectrum antiproliferative activities. *J Nat Prod* 2011;**74**:2031–8.
57. Ju KS, Gao J, Doroghazi JR, Wang KK, Thibodeaux CJ, Li S, 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**:12175–80.
58. Shen B. A new golden age of natural products drug discovery. *Cell* 2015;**163**:1297–300.
59. Machida K, Arisawa A, Takeda S, Tsuchida T, Aritoku Y, Yoshida M, et al. Organization of the biosynthetic gene cluster for the polyketide antitumor macrolide, pladienolide, in *Streptomyces platensis* Mer-11107. *Biosci Biotechnol Biochem* 2008;**72**:2946–52.
60. Seki-Asano M, Okazaki T, Yamagishi M, Sakai N, Takayama Y, Hanada K, et al. Isolation and characterization of a new 12-membered macrolide FD-895. *J Antibiot (Tokyo)* 1994;**47**:1395–401.
61. Yu D, Xu F, Shao L, Zhan J. A specific cytochrome P450 hydroxylase in herboxidiene biosynthesis. *Bioorg Med Chem Lett* 2014;**24**:4511–4.
62. Eustaquio AS, Janso JE, Ratnayake AS, O'Donnell CJ, Koehn FE. Spliceostatin hemiketal biosynthesis in *Burkholderia* spp. is catalyzed by an iron/alpha-ketoglutarate-dependent dioxygenase. *Proc Natl Acad Sci U S A* 2014;**111**:E3376–85.
63. Nagano S, Li H, Shimizu H, Nishida C, Ogura H, Ortiz de Montellano PR, et al. Crystal structures of epothilone D-bound, epothilone B-bound, and substrate-free forms of cytochrome P450epoK. *J Biol Chem* 2003;**278**:44886–93.
64. Li A, Piel J. A gene cluster from a marine *Streptomyces* encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A. *Chem Biol* 2002;**9**:1017–26.
65. Carlson JC, Li S, Gunatilleke SS, Anzai Y, Burr DA, Podust LM, et al. Tirandamycin biosynthesis is mediated by co-dependent oxidative enzymes. *Nat Chem* 2011;**3**:628–33.
66. Schorn M, Zettler J, Noel JP, Dorrestein PC, Moore BS, Kaysser L. Genetic basis for the biosynthesis of the pharmacologically important class of epoxyketone proteasome inhibitors. *ACS Chem Biol* 2014;**9**:301–9.
67. Ortiz de Montellano PR. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chem Rev* 2010;**110**:932–48.
68. Lin S, Huang T, Shen B. Tailoring enzymes acting on carrier protein-tethered substrates in natural product biosynthesis. *Methods Enzymol* 2012;**516**:321–43.
69. Walsh CT. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* 2004;**303**:1805–10.
70. Thibodeaux CJ, Chang WC, Liu HW. Enzymatic chemistry of cyclopropane, epoxide, and aziridine biosynthesis. *Chem Rev* 2012;**112**:1681–709.
71. Podust LM, Sherman DH. Diversity of P450 enzymes in the biosynthesis of natural products. *Nat Prod Rep* 2012;**29**:1251–66.
72. Ma J, Wang Z, Huang H, Luo M, Zuo D, Wang B, et al. Biosynthesis of himastatin: assembly line and characterization of three cytochrome P450 enzymes involved in the post-tailoring oxidative steps. *Angew Chem Int Ed Engl* 2011;**50**:7797–802.
73. Caffrey P, Lynch S, Flood E, Finnan S, O'Leary M. Amphoterin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem Biol* 2001;**8**:713–23.
74. Zhao Q, He Q, Ding W, Tang M, Kang Q, Yu Y, et al. Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem Biol* 2008;**15**:693–705.
75. Ward SL, Hu Z, Schirmer A, Reid R, Revill WP, Reeves CD, et al. Chalcomycin biosynthesis gene cluster from *Streptomyces bikiniensis*: novel features of an unusual ketolide produced through expression of the *chm* polyketide synthase in *Streptomyces fradiae*. *Antimicrob Agents Chemother* 2004;**48**:4703–12.
76. Nagano S, Cupp-Vickery JR, Poulos TL. Crystal structures of the ferrous dioxygen complex of wild-type cytochrome P450eryF and its mutants, A245S and A245T: investigation of the proton transfer system in P450eryF. *J Biol Chem* 2005;**280**:22102–7.
77. Stassi D, Donadio S, Staver MJ, Katz L. Identification of a *Saccharopolyspora erythraea* gene required for the final hydroxylation step in erythromycin biosynthesis. *J Bacteriol* 1993;**175**:182–9.
78. He Y, Sun Y, Liu T, Zhou X, Bai L, Deng Z. Cloning of separate meilingmycin biosynthesis gene clusters by use of acyltransferase-ketoreductase didomain PCR amplification. *Appl Environ Microbiol* 2010;**76**:3283–92.
79. Anzai Y, Saito N, Tanaka M, Kinoshita K, Koyama Y, Kato F. Organization of the biosynthetic gene cluster for the polyketide macrolide mycinamicin in *Micromonospora griseorubida*. *FEMS Microbiol Lett* 2003;**218**:135–41.
80. Liu W, Nonaka K, Nie L, Zhang J, Christenson SD, Bae J, et al. The neocarzinostatin biosynthetic gene cluster from *Streptomyces carzinostaticus* ATCC 15944 involving two iterative type I polyketide synthases. *Chem Biol* 2005;**12**:293–302.
81. Brautaset T, Sekurova ON, Sletta H, Ellingsen TE, StrLm AR, Valla S, et al. Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem Biol* 2000;**7**:395–403.
82. Cryle MJ, Meinhart A, Schlichting I. Structural characterization of OxyD, a cytochrome P450 involved in beta-hydroxytyrosine formation in vancomycin biosynthesis. *J Biol Chem* 2010;**285**:24562–74.
83. Sherman DH, Li S, Yermalitskaya LV, Kim Y, Smith JA, Waterman MR, et al. The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*. *J Biol Chem* 2006;**281**:26289–97.
84. Ikeda H, Kazuo SY, Omura S. Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J Ind Microbiol Biotechnol* 2014;**41**:233–50.
85. Bate N, Cundliffe E. The mycinose-biosynthetic genes of *Streptomyces fradiae*, producer of tylosin. *J Ind Microbiol Biotechnol* 1999;**23**:118–22.

86. Baltz RH, Seno ET. Properties of *Streptomyces fradiae* mutants blocked in biosynthesis of the macrolide antibiotic tylosin. *Antimicrob Agents Chemother* 1981;**20**:214–25.
87. Barry SM, Kers JA, Johnson EG, Song L, Aston PR, Patel B, et al. Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat Chem Biol* 2012;**8**:814–6.
88. Kudo F, Motegi A, Mizoue K, Eguchi T. Cloning and characterization of the biosynthetic gene cluster of 16-membered macrolide antibiotic FD-891: involvement of a dual functional cytochrome P450 monooxygenase catalyzing epoxidation and hydroxylation. *Chembiochem* 2010;**11**:1574–82.
89. Zhang W, Liu Y, Yan J, Cao S, Bai F, Yang Y, et al. New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *J Am Chem Soc* 2014;**136**:3640–6.
90. Lin HC, Tsunematsu Y, Dhingra S, Xu W, Fukutomi M, Chooi YH, et al. Generation of complexity in fungal terpene biosynthesis: discovery of a multifunctional cytochrome P450 in the fumagillin pathway. *J Am Chem Soc* 2014;**136**:4426–36.
91. Zhu D, Seo MJ, Ikeda H, Cane DE. Genome mining in *Streptomyces*. Discovery of an unprecedented P450-catalyzed oxidative rearrangement that is the final step in the biosynthesis of pentalenolactone. *J Am Chem Soc* 2011;**133**:2128–31.
92. Seo MJ, Zhu D, Endo S, Ikeda H, Cane DE. Genome mining in *Streptomyces*. Elucidation of the role of Baeyer-Villiger monooxygenases and non-heme iron-dependent dehydrogenase/oxygenases in the final steps of the biosynthesis of pentalenolactone and neopentalenolactone. *Biochemistry* 2011;**50**:1739–54.
93. Liu X, Biswas S, Berg MG, Antapli CM, Xie F, Wang Q, et al. Genomics-guided discovery of thailanstatins A, B, and C As pre-mRNA splicing inhibitors and antiproliferative agents from *Burkholderia thailandensis* MSMB43. *J Nat Prod* 2013;**76**:685–93.
94. Yu TW, Bai L, Clade D, Hoffmann D, Toelzer S, Trinh KQ, et al. The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from *Actinosynnema pretiosum*. *Proc Natl Acad Sci U S A* 2002;**99**:7968–73.
95. Billign T, Hyun CG, Williams JS, Czisny AM, Thorson JS. The hedamycin locus implicates a novel aromatic PKS priming mechanism. *Chem Biol* 2004;**11**:959–69.
96. Rodriguez AM, Olano C, Mendez C, Hutchinson CR, Salas JA. A cytochrome P450-like gene possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*. *FEMS Microbiol Lett* 1995;**127**:117–20.
97. Aparicio JF, Fouces R, Mendes MV, Olivera N, Martin JF. A complex multienzyme system encoded by five polyketide synthase genes is involved in the biosynthesis of the 26-membered polyene macrolide pimaricin in *Streptomyces natalensis*. *Chem Biol* 2000;**7**:895–905.
98. Carlson JC, Fortman JL, Anzai Y, Li S, Burr DA, Sherman DH. Identification of the tirandamycin biosynthetic gene cluster from *Streptomyces* sp. 307-9. *Chembiochem* 2010;**11**:564–72.
99. Carmody M, Murphy B, Byrne B, Power P, Rai D, Rawlings B, et al. Biosynthesis of amphoterin derivatives lacking exocyclic carboxyl groups. *J Biol Chem* 2005;**280**:34420–6.
100. Gao Q, Thorson JS. The biosynthetic genes encoding for the production of the dynamic enediyne core in *Micromonospora chersina* ATCC53710. *FEMS Microbiol Lett* 2008;**282**:105–14.
101. Zhang H, Chen J, Wang H, Xie Y, Ju J, Yan Y. Structural analysis of HmtT and HmtN involved in the tailoring steps of himastatin biosynthesis. *FEBS Lett* 2013;**587**:1675–80.
102. Zerbe K, Pylypenko O, Vitali F, Zhang W, Rousset S, Heck M, et al. Crystal structure of OxyB, a cytochrome P450 implicated in an oxidative phenol coupling reaction during vancomycin biosynthesis. *J Biol Chem* 2002;**277**:47476–85.
103. Pylypenko O, Vitali F, Zerbe K, Robinson JA, Schlichting I. Crystal structure of OxyC, a cytochrome P450 implicated in an oxidative C-C coupling reaction during vancomycin biosynthesis. *J Biol Chem* 2003;**278**:46727–33.
104. Li T, Du Y, Cui Q, Zhang J, Zhu W, Hong K, et al. Cloning, characterization and heterologous expression of the indolocarbazole biosynthetic gene cluster from marine-derived *Streptomyces sanyensis* FMA. *Mar Drugs* 2013;**11**:466–88.
105. Howard-Jones AR, Walsh CT. Staurosporine and rebeccamycin aglycones are assembled by the oxidative action of StaP, StaC, and RebC on chromopyrrolic acid. *J Am Chem Soc* 2006;**128**:12289–98.
106. Meunier B, de Visser SP, Shaik S. Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. *Chem Rev* 2004;**104**:3947–80.
107. Anzai Y, Li S, Chaulagain MR, Kinoshita K, Kato F, Montgomery J, et al. Functional analysis of MycC1 and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics. *Chem Biol* 2008;**15**:950–9.
108. Lamb DC, Waterman MR. Unusual properties of the cytochrome P450 superfamily. *Philos Trans R Soc Lond B Biol Sci* 2013;**368**:20120434.
109. Rupasinghe S, Schuler MA, Kagawa N, Yuan H, Lei L, Zhao B, et al. The cytochrome P450 gene family CYP157 does not contain EXXR in the K-helix reducing the absolute conserved P450 residues to a single cysteine. *FEBS Lett* 2006;**580**:6338–42.
110. Zhao B, Guengerich FP, Bellamine A, Lamb DC, Izumikawa M, Lei L, et al. Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of *Streptomyces coelicolor* A3(2) cytochrome P450 158A2. *J Biol Chem* 2005;**280**:11599–607.
111. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, et al. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 1996;**6**:1–42.
112. Gricman L, Vogel C, Pleiss J. Conservation analysis of class-specific positions in cytochrome P450 monooxygenases: functional and structural relevance. *Proteins* 2014;**82**:491–504.
113. Sirim D, Widmann M, Wagner F, Pleiss J. Prediction and analysis of the modular structure of cytochrome P450 monooxygenases. *BMC Struct Biol* 2010;**10**:34.
114. Sirim D, Wagner F, Lisitsa A, Pleiss J. The cytochrome P450 engineering database: integration of biochemical properties. *BMC Biochem* 2009;**10**:27.
115. Fischer M, Knoll M, Sirim D, Wagner F, Funke S, Pleiss J. The cytochrome P450 engineering database: a navigation and prediction tool for the cytochrome P450 protein family. *Bioinformatics* 2007;**23**:2015–7.
116. Thompson CF, Jamison TF, Jacobsen EN. FR901464: total synthesis, proof of structure, and evaluation of synthetic analogues. *J Am Chem Soc* 2001;**123**:9974–83.
117. Motoyoshi H, Horigome M, Ishigami K, Yoshida T, Horinouchi S, Yoshida M, et al. Structure-activity relationship for FR901464: a versatile method for the conversion and preparation of biologically active biotinylated probes. *Biosci Biotechnol Biochem* 2004;**68**:2178–82.
118. Patrizia D, Girolamo C. Biosynthesis of heterocycles: from isolation to gene cluster, 2015.
119. Owen JG, Charlop-Powers Z, Smith AG, Ternei MA, Calle PY, Reddy BV, et al. Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proc Natl Acad Sci U S A* 2015;**112**:4221–6.
120. Aukema KG, Chohan KK, Plourde GL, Reimer KB, Rader SD. Small molecule inhibitors of yeast pre-mRNA splicing. *ACS Chem Biol* 2009;**4**:759–68.
121. Shao L, Zi J, Zeng J, Zhan J. Identification of the herboxidiene biosynthetic gene cluster in *Streptomyces chromofuscus* ATCC 49982. *Appl Environ Microbiol* 2012;**78**:2034–8.
122. Chen H, Walsh CT. Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chem Biol* 2001;**8**:301–12.
123. Walsh CT, Fischbach MA. Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc* 2010;**132**:2469–93.
124. Liu X, Cheng YQ. Genome-guided discovery of diverse natural products from *Burkholderia* sp. *J Ind Microbiol Biotechnol* 2014;**41**:275–84.
125. Nakajima H, Hori Y, Terano H, Okuhara M, Manda T, Matsumoto S, et al. New antitumor substances, FR901463, FR901464 and FR901465. II. Activities against experimental tumors in mice and mechanism of action. *J Antibiot (Tokyo)* 1996;**49**:1204–11.
126. Nakajima H, Sato B, Fujita T, Takase S, Terano H, Okuhara M. New antitumor substances, FR901463, FR901464 and FR901465. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities. *J Antibiot (Tokyo)* 1996;**49**:1196–203.
127. Nakajima H, Takase S, Terano H, Tanaka H. New antitumor substances, FR901463, FR901464 and FR901465. III. Structures of FR901463, FR901464 and FR901465. *J Antibiot (Tokyo)* 1997;**50**:96–9.
128. Gee JE, Glass MB, Novak RT, Gal D, Mayo MJ, Steigerwalt AG, et al. Recovery of a *Burkholderia thailandensis*-like isolate from an Australian water source. *BMC Microbiol* 2008;**8**:54.
129. Zhuo Y, Liu L, Wang Q, Liu X, Ren B, Liu M, et al. Revised genome sequence of *Burkholderia thailandensis* MSMB43 with improved annotation. *J Bacteriol* 2012;**194**:4749–50.
130. He H, Ratnayake AS, Janso JE, He M, Yang HY, Loganzo F, et al. Cytotoxic spliceostatin B from *Burkholderia* sp. and their semisynthetic analogues. *J Nat Prod* 2014;**77**:1864–70.
131. Liu X, Biswas S, Tang GL, Cheng YQ. Isolation and characterization of spliceostatin B, a new analogue of FR901464, from *Pseudomonas* sp. No. 2663. *J Antibiot (Tokyo)* 2013;**66**:555–8.
132. Broberg A, Menkis A, Vasiliauskas R. Kutznerides 1–4, depsipeptides from the actinomycete *Kutzneria* sp. 744 inhabiting mycorrhizal roots of *Picea abies* seedlings. *J Nat Prod* 2006;**69**:97–102.
133. Pohanka A, Menkis A, Levenfors J, Broberg A. Low-abundance kutznerides from *Kutzneria* sp. 744. *J Nat Prod* 2006;**69**:1776–81.
134. Rebets Y, Tokovenko B, Lushchik I, Ruckert C, Zaburannyi N, Bechthold A, et al. Complete genome sequence of producer of the glycopeptide antibiotic Aculeximycin *Kutzneria albida* DSM 43870T, a representative of minor genus of *Pseudonocardiaceae*. *BMC Genomics* 2014;**15**:885.
135. Mo X, Wang Z, Wang B, Ma J, Huang H, Tian X, et al. Cloning and characterization of the biosynthetic gene cluster of the bacterial RNA polymerase inhibitor tirandamycin from marine-derived *Streptomyces* sp. SCSIO1666. *Biochem Biophys Res Commun* 2011;**406**:341–7.
136. Ongley SE, Bian X, Neilan BA, Muller R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat Prod Rep* 2013;**30**:1121–38.
137. Brady SF. Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. *Nat Protoc* 2007;**2**:1297–305.
138. Milshcheyn A, Schneider JS, Brady SF. Mining the metabiome: identifying novel natural products from microbial communities. *Chem Biol* 2014;**21**:1211–23.
139. Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat Biotechnol* 2012;**30**:440–6.
140. Kvitko BH, McMillan IA, Schweizer HP. An improved method for oriT-directed cloning and functionalization of large bacterial genomic regions. *Appl Environ Microbiol* 2013;**79**:4869–78.
141. Ross AC, Gulland LE, Dorrestein PC, Moore BS. Targeted capture and heterologous expression of the *Pseudomonas* alterochromide gene cluster in *Escherichia*

- coli* represents a promising natural product exploratory platform. *ACS Synth Biol* 2015;**4**:414–20.
142. 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. *Proc Natl Acad Sci U S A* 2014;**111**:1957–62.
 143. Kouprina N, Larionov V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat Protoc* 2008;**3**:371–7.
 144. Du D, Wang L, Tian Y, Liu H, Tan H, Niu G. Genome engineering and direct cloning of antibiotic gene clusters via phage varphiBT1 integrase-mediated site-specific recombination in *Streptomyces*. *Sci Rep* 2015;**5**:8740.
 145. Feng Z, Kallifidas D, Brady SF. Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. *Proc Natl Acad Sci U S A* 2011;**108**:12629–34.
 146. Feng Z, Kim JH, Brady SF. Fluostatin produced by the heterologous expression of a TAR reassembled environmental DNA derived type II PKS gene cluster. *J Am Chem Soc* 2010;**132**:11902–3.
 147. Kim JH, Feng Z, Bauer JD, Kallifidas D, Calle PY, Brady SF. Cloning large natural product gene clusters from the environment: piecing environmental DNA gene clusters back together with TAR. *Biopolymers* 2010;**93**:833–44.
 148. Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, et al. Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl Environ Microbiol* 2004;**70**:2452–63.
 149. Feng Z, Wang L, Rajski SR, Xu Z, Coeffet-LeGal MF, Shen B. Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts. *Bioorg Med Chem* 2009;**17**:2147–53.
 150. Pfeifer B, Hu Z, Licari P, Khosla C. Process and metabolic strategies for improved production of *Escherichia coli*-derived 6-deoxyerythronolide B. *Appl Environ Microbiol* 2002;**68**:3287–92.
 151. Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci U S A* 2010;**107**:2646–51.
 152. Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozone I, Izumikawa M, et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth Biol* 2013;**2**:384–96.
 153. Rutledge PJ, Challis GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* 2015;**13**:509–23.
 154. Zarins-Tutt JS, Barberi TT, Gao H, Mearns-Spragg A, Zhang L, Newman DJ, et al. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep* 2015;**33**:54–72.
 155. Wenzel SC, Hoffmann H, Zhang J, Debussche L, Haag-Richter S, Kurz M, et al. Production of the bengamide class of marine natural products in *Myxobacteria*: biosynthesis and structure-activity relationships. *Angew Chem Int Ed Engl* 2015;**54**(51):15560–4.
 156. Li Y, Chen L, Yue Q, Liu X, An Z, Bills GF. Genetic manipulation of the pneumocandin biosynthetic pathway for generation of analogues and evaluation of their antifungal activity. *ACS Chem Biol* 2015;**10**:1702–10.
 157. Eustaquio AS, Chang LP, Steele GL, O'Donnell CJ, Koehn FE. Biosynthetic engineering and fermentation media development leads to gram-scale production of spliceostatin natural products in *Burkholderia* sp. *Metab Eng* 2015;**33**:67–75.
 158. Kim E, Moore BS, Yoon YJ. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat Chem Biol* 2015;**11**:649–59.
 159. Cobb RE, Wang Y, Zhao H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth Biol* 2015;**4**(6):723–8.
 160. Tong Y, Charusanti P, Zhang L, Weber T, Lee SY. CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS Synth Biol* 2015;**4**(9):1020–9.
 161. Zeng H, Wen S, Xu W, He Z, Zhai G, Liu Y, et al. Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl Microbiol Biotechnol* 2015;**99**:10575–85.
 162. Tang X, Li J, Millan-Aguinaga N, Zhang JJ, O'Neill EC, Ugalde JA, et al. Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chem Biol* 2015;**10**:2841–9.
 163. Kang Y, Norris MH, Wilcox BA, Tuanyok A, Keim PS, Hoang TT. Knockout and pullout recombineering for naturally transformable *Burkholderia thailandensis* and *Burkholderia pseudomallei*. *Nat Protoc* 2011;**6**:1085–104.
 164. Moebius N, Ross C, Scherlach K, Rohm B, Roth M, Hertweck C. Biosynthesis of the respiratory toxin bongkrekic acid in the pathogenic bacterium *Burkholderia gladioli*. *Chem Biol* 2012;**19**:1164–74.
 165. Lee SY, Kim HU, Park JH, Park JM, Kim TY. Metabolic engineering of microorganisms: general strategies and drug production. *Drug Discov Today* 2009;**14**:78–88.
 166. Overmann J. Significance and future role of microbial resource centers. *Syst Appl Microbiol* 2015;**38**:258–65.