Analyzing the biosynthetic potential of antimicrobial-producing actinobacteria originating from Indonesia

Anissa Utami[1,6], **Pamela Apriliana**[2], **Yudi Kusnadi**[3,7], **Dewi S. Zilda**[3], **Zidny Ilmiah**[2], **Puspita Lisdiyanti**[2], **Siswa Setyahadi**[1,4,*], **Agustinus R. Uria**[3,5,*]

[1]Graduate School of Pharmacy, Pancasila University, Jl. Raya Lenteng Agung, Jakarta Selatan 12630, Indonesia
[2]Research Center for Biotechnology, Indonesian Institute of Sciences (LIPI), Jl. LIPI, Cibinong, Bogor, Jawa Barat 16911, Indonesia
[3]Research Center for Marine and Fisheries Product processing and Biotechnology Ministry of Marine Affairs and Fisheries, JL. KS Tubun Petamburan VI, Slipi, Jakarta Pusat 10260, Indonesia
[4]Center of Bio-industrial Technology, Agency for Technology Assessment and Development, Tangerang Selatan, Banten 15314, Indonesia
[5]Faculty of Pharmaceutical Sciences, Hokkaido University, Kita Ward, Sapporo, Hokkaido 060-0812, Japan
[6]Politeknik Hang Tuah Jakarta, Jl. Bendungan Hilir Jakarta Pusat 10210, Indonesia
[7]Life Science Division, ITS Science Indonesia, Jl. Boulevard Artha Gading, Jakarta Utara 14240, Indonesia
*Corresponding author: siswa.setyahadi@bppt.go.id; auria@pharm.hokudai.ac.jp

ABSTRACT We investigated the biosynthetic potential of soil-associated actinobacteria originating from Indonesia, identified as *Streptomyces luridus* and as *Streptomyces luteosporeus*. Antimicrobial assays indicated inhibitory activity by both strains against the pathogen *Pseudomonas aeruginosa*, with *S. luteosporeus* particularly inhibiting the growth of *Bacillus subtilis*. PCR-amplification, cloning, and sequencing of ketosynthase (KS) domains of type I modular polyketide (PKS-I) and adenylation (AD) domains of non-ribosomal peptide synthetase (NRPS) indicated the diversity of KS and AD domains derived from both Indonesian *Streptomyces*. Further phylogenetic analysis showed that KS domains from the subclass cis-AT PKS can be classified as being a part of a loading module or an extension module, along with their predicted substrate specificity. The results suggest that both strains are a potential source of novel biosynthetic pathways. This genetic analysis approach can be used as a fast guide to obtain insight into natural product biosynthetic gene diversity in microorganisms.

KEYWORDS Antimicrobial activity; biosynthetic potential; *Streptomyces luridus*; *Streptomyces luteosporeus*; PKS and NRPS gene diversity

1. Introduction

Phylum Actinobacteria have long been recognized as the major source of compounds representing 45% of all bioactive microbial metabolites discovered so far (Bérdy 2005). Many actinobacterial natural products belong to polyketides and non-ribosomal peptides, two large classes of natural products with various potent biological activities, such as antitumor, antivirus, antibiotics, and immune-suppressants (Cane and Walsh 1999). From the biosynthetic point of view, the huge diversity of most polyketide structures is generated from simple carboxylic acid monomers via a series of enzymatic reactions mediated by polyketide synthases (PKSs) (Risdian et al. 2019; Hertweck 2015). Risdian et al. (2019) stated that among three types of bacterial PKSs known to date (types I, II, and III), type I PKS (PKS-I) in *Streptomyces* has attracted increasing attention due to the involvement of this type in the biosynthesis of many macrocyclic polyketides

(macrolides) with various bioactivities.

Type I PKS system consists of modules, which act in a non-iterative way. Each module harbors a set of catalytic domains that are responsible for one cycle of polyketide elongation and modification. A module minimally consists of three domains: an acyltransferase (AT) domain, an acyl carrier protein (ACP) domain, and a ketoacyl synthase (KS) domain (Hertweck 2009; Rawlings 2001). Each module receives an acyl chain from the upstream module onto KS domain and a suitable extender unit onto ACP through acyltransferase (AT) activity. ACP in this module serves as an anchor for the building block. Subsequent KS-catalyzed addition of the extender unit to the acyl chain results in a polyketide product that can be modified by optional domains such as ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) (Watanabe et al. 2003; Weissman 2015). Due to the highly conserved regions and intermediate specificity of KS domains in PKS systems, KS sequences are ideally used as a basis not only

for PKS classification, but also for predicting structures of intermediates and even final biosynthetic products (Aparicio et al. 1996; Nguyen et al. 2008).

A remarkable structural and functional diversity of non-ribosomal peptides are biosynthetically built from proteinogenic and nonproteinogenic amino acids often along with other unusual acids catalyzed by non-ribosomal peptide synthetases (NRPSs) (Fischbach and Walsh 2006). A typical NRPS module minimally consists of an adenylation (AD) domain, a thiolation (T) domain, and a condensation (C) domain (Fischbach and Walsh 2006). The AD domain responsible for amino acid activation and incorporation is the best-characterized among NRPS domains; and therefore it is suitable to be used as a basis for detecting NRPS systems and predicting the amino acid building blocks (Röttig et al. 2011).

Screening of genes encoding PKS and NRPS systems based on the detection of KS and AD domains have extensively been used to obtain preliminary insights into the biosynthetic potential of various bacterial groups, including the genus Streptomyces (Ayuso-Sacido and Genilloud 2005). The objectives of this work are to analyze the biosynthetic potential of two uncharacterized actinobacterial strains by analyzing KS and AD domains of the PKS-I and NRPS genes. We tested the antimicrobial activity, determined the taxonomic identity of both strains, and predicted their possible environmental distribution based on 16S rRNA analysis. The outcome of these studies become an important basis for genome sequencing to reveal the promising biosynthetic pathways and for chemical analysis to identify secondary metabolites for drug discovery.

2. Materials and Methods

2.1. Bacterial strains and cultivation

Two different bacterial strains, designated as ID 04 677 and SDS14LU, used in this work were obtained from Dr. Puspita Lisdiyanti at Research Center for Biotechnology, Indonesian Institute of Sciences (LIPI). Both strains were preserved or deposited with the names A375 and A538 in Indonesian Culture Collection (INaCC) operated by LIPI. The strain ID 04 677 was derived from Purwadadi, West Java, which was identified in this work based on 16S rRNA gene analysis using 27F and 1492R primers (Weisburg et al. 1991). The strain SDS14LU was derived from Mount Tambora Savana, Bima, West Nusa Tenggara, which was previously identified as *Streptomyces thioluteus* (Fawzya et al. 2016). Cultivation of both strains was carried out using International Streptomyces Project-2 (ISP2) solid medium at 25 °C. Test bacteria used for antimicrobial assay included *Escherichia coli* ATCC® 25922, *Bacillus subtilis*, *Staphylococcus aureus* ATCC® 25923, and *Pseudomonas aeruginosa* ATCC® 27853. Media used for antimicrobial testing were Nutrient Broth (NB) for overnight liquid culture, and Mueller Hinton Agar (MHA) for disk-diffusion agar method. *E. coli*

DH5α ATCC® 53868 and pGEM-T Easy (Promega) were used for DNA cloning. All reagents and kits used were in high quality obtained from various manufacturers, such as Promega, Thermo Scientific, and Roche.

2.2. 16S rRNA gene analyses

Genomic DNA was isolated from each strain according to the CTAB method (Piel et al. 2004) with a slight modification and used as the template for the PCR-amplification of nearly complete 16S rDNA. The PCR mixture contained 1x GoTaq Green Master Mix (Promega), 1 µM each 16S primer, and 0.5 µl DNA template in total volume of 50 µl. The PCR program was set up at 30 cycles, consisting of a pre-denaturation at 95 °C for 3 min, denaturation 95 °C for 30 seconds, annealing 55 °C for 1 min, and elongation at 72 °C for 1.5 min, and final elongation 72 °C, 5 min. The target PCR product (1400 bp) was separated on the 1% agarose gel, purified from the gel, and subjected to DNA sequencing using the same primers employed in the PCR amplification. DNA sequencing was conducted using the Big Dye Terminator Cycle sequencing kit on ABI Prism 3700 DNA analyzer (Applied Biosystems, USA). The resulting forward and reverse reads of each 16S rRNA gene sequence were assembled and subjected to the Mega BLAST mode of NCBI (Morgulis et al. 2008) against the preformatted 16S microbial database downloaded from the updated BLAST databases (https://ftp.ncbi.nlm.nih.gov/blast/db/) in Geneious Prime sequence analysis software (ver. 10.2.3). The resulting 16S sequence database created from the maximum BLAST hits of 200 was aligned with the 16S rRNA gene sequences of both ID 04 677 and SDS14LU strains using MUSCLE (Edgar 2004). The alignment result was saved as MEG file and used to build a phylogenetic tree in Molecular Evolutionary Genetic Analysis version 10.0.5 (MEGA X) (Kumar et al. 2018) based on the Neighbor-Joining method (Saitou and Nei 1987). To investigate the microbial habitability, the 16S rRNA gene sequences of both strains were submitted to MetaMetaDB, a database and analytic system for investigating microbial habitability (Yang and Iwasaki 2014). This MetaMetaDB analysis performed a BLAST search against 2,949,852 representative 16S rRNA sequences from 61 diverse environments (update dataset by November 6, 2014) (http://mmdb.aori.u-tokyo.ac.jp/).

2.3. Antimicrobial assays

Streptomyces strains were individually cultivated on ISP2 agar plates at 30 °C for a week. The culture agar in each plate was cut into small pieces in square shape (approximately 1 cm x 1 cm). The agar pieces were transferred into a beaker glass containing 100-200 ml ethyl acetate and mixed well. The beaker glass was closed and let overnight at room temperature. The ethyl acetate extract was transferred into an empty beaker glass and evaporated overnight in a fume hood. The dried extract was dissolved in methanol at the final concentration of 100 mg/ml. The extract dissolved in MeOH (50 µl) was applied onto a sterile paper disc in the clean bench, let it dry

for 15 min, and placed on an agar plate containing a test organism. Four bacterial species as mentioned above were used as test organisms for antimicrobial assays. Each test bacterial strain was cultivated overnight in 5 ml Nutrient Broth (NB). The NB culture of each tested bacterial strain (100 µl) was mixed with 8 ml of semi-solid (half recipe) Mueller-Hinton Agar (MHA), and subsequently poured evenly on the surface of HMA plate (full recipe). The paper discs containing extracts were individually placed on the top of HMA plates containing test bacteria. The HMA plates with paper discs were incubated at 30 °C for 24 h. Zone of inhibition around discs was measured.

2.4. Cloning into *E. coli* and sequencing

KS-encoding fragments were PCR-amplified from the bacterial cell fraction mentioned above using the primer set KSDPQQF (5'-MGNGARGCNNWNSMNATGGAYCCNCARCANMG-3') and KSHGTGR (5'-GGRTCNCCNARNSWNGTNCCNGTNCCRTG-3') (Piel 2002). PCR-amplification of regions coding for AD domain was carried out using the primer pair AG-for (5'-GCSTACSYSATSTACACSTCSGG-3') and AG-rev (5'-SASGTCVCCSGTSCGGTAS-3') (Ayuso-Sacido and Genilloud 2005). The PCR composition and condition are the same as described above, except the annealing temperature was set up at 60.1 °C. The PCR products were separated on 1% agarose gel in the electrophoresis (Power Pac Basic, BioRad). Target fragments of approximately 700 bp were extracted from the gel using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific). The extracted DNA solution was concentrated to 7 µl by evaporator, followed with ligation with pGEM-T Easy (Promega). The ligation product was transformed into electro-competent cells of *E. coli* DH5α using the electroporator Micropulser (BioRad), added with 1 ml LB, and incubated at 37 °C for 1 h in a shaking incubator. The transformed cell suspension was spread on LB plates containing 100 µg/ml ampicillin previously supplemented with 40 µl of 0.1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) and 40 µl of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). After overnight incubation at 37 °C, the resulting white colonies were picked up, transferred to LB agar plates, 96-well microplates, and tubes containing 5-ml LB, followed by overnight incubation. Plasmid constructs were individually recovered from cultures using GeneJet Plasmid Miniprep Kit (Thermo Scientific). The resulting plasmid samples were individually cut with EcoRI to check the presence of inserts, and subsequently digested with BamHI to know the restriction pattern. Recombinant plasmids with unique restriction patters were subjected to DNA sequencing using the primer T7 (5'-GTAATACGACTCACTATAGGG-3') that recognizes the cloning vector.

2.5. Bioinformatic analyses of KS and AD sequences

MIBiG (Minimum Information about a Biosynthetic Gene cluster) database (version 2.0) containing biosynthetic gene clusters of known function was initially downloaded in FASTA format from the Genomic Standards Consortium website (https://mibig.secondarymetabolites.org/) (Kautsar et al. 2020). All of the DNA sequences encoding KS and AD domains isolated in this work were translated into amino acid sequences using the Web-based translation tool Expasy (http://www.expasy.org/tools/dna.html), which was subsequently subjected to BLASTp analysis against the MIBiG database operated in Geneious Prime (ver. 10.2.3). Of the resulting maximal BLAST hits of 50, several domain sequences were selected, downloaded, and aligned with the KS and AD sequences obtained in this work using Clustal Omega program (Sievers and Higgins 2018). A phylogenetic tree was then constructed using the neighbor-joining method (Saitou and Nei 1987) or the Maximum Likelihood method and JTT matrix-based model (Jones et al. 1992) in Mega X (Kumar et al. 2018). Resampling method with boot-strapping values inferred from 1000 replicates was set up on each tree branch to estimate the reliability of phylogenetic tree reconstruction. Based on this phylogenetic tree, we determined the affiliation or classification of KS sequences. The sequence data was also analyzed using some other programs, such as NaPDoS for comparison to a broad set of curated reference genes from well-characterized biosynthetic pathways (Ziemert et al. 2012) and NRPSpredictor2 (Röttig et al. 2011) for predicting the substrate specificity of adenylation domains. The Enzyme Function Initiative – Enzyme Similarity Tool (EFI-EST) SSN analysis (Gerlt et al. 2015) of AD sequences was conducted to generate a protein sequence similarity network.

3. Results and Discussion

3.1. Taxonomic affiliation, antimicrobial activity, and environmental habitability

We evaluated the taxonomic affiliation and antimicrobial activity of two cultivated Streptomyces strains originated from Indonesia. One strain (ID 04 677) isolated from a soil habitat in Purwadadi appeared as small rough-round colonies with white-colored aerial mycelium on the ISP2 solid medium (Supplementary Figure 1a). To identify this strain, 16S rRNA gene analysis was applied. Sequencing and BLASTn analysis of the 16S rRNA gene obtained in this work (Acc. no. KX150801) showed the highest hit (99.85% identity) with that of *S. luridus* YIM 131011 associated with lichen symbiosis (Acc. no. KX502956.1) (Liu et al. 2017). Mega-BLAST analysis against the updated 16S microbial database (https://ftp.ncbi.nlm.nih.gov/blast/db/) exhibited its identity of 99.2% with that of *S. luridus* NBRC 12881 (Acc. no. NR_112280.1) (Labeda et al. 2012). Further phylogenetic analysis revealed its close relationship with *S. luridus*, which was in the same clade as the following characterized Strep-
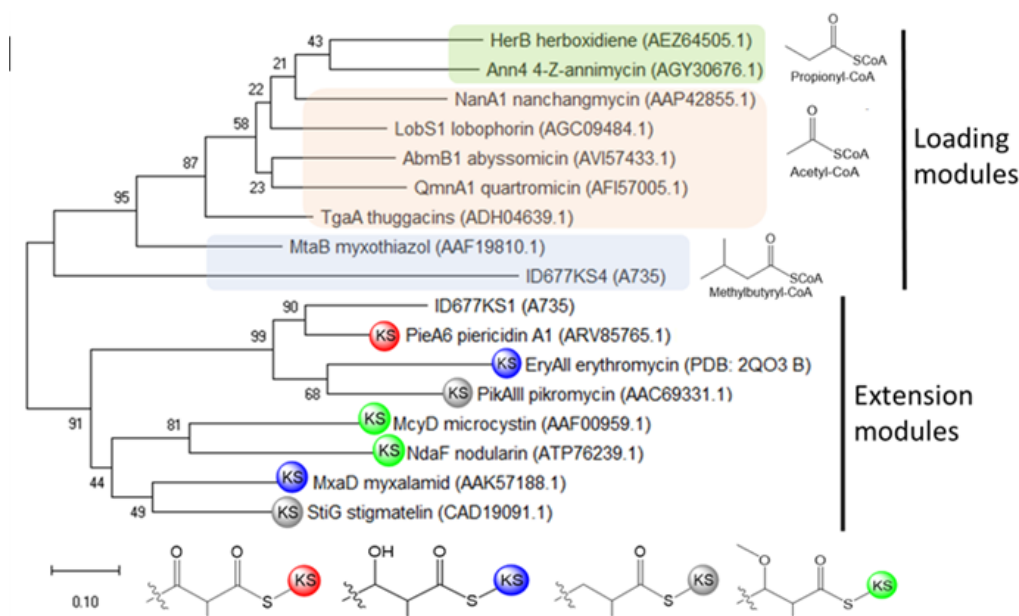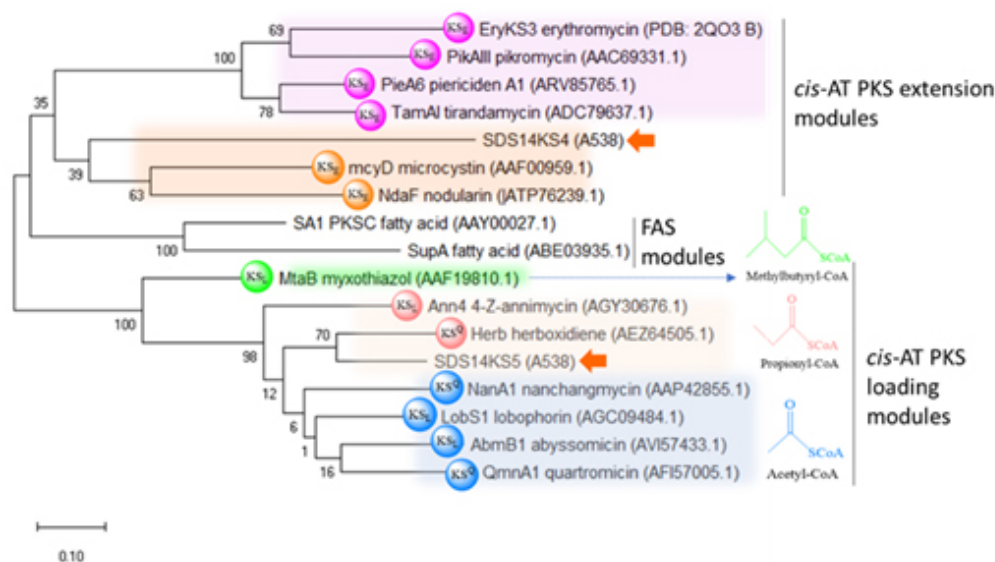
(a)
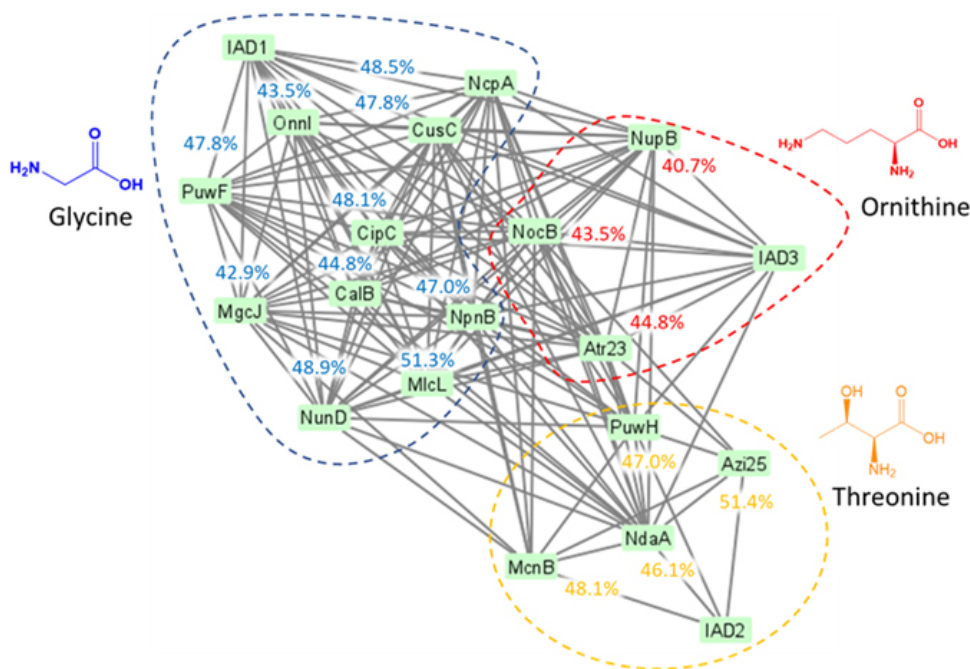
(b)

FIGURE 1 Analysis of KS and AD sequences derived from S. luridus ID 04 677. (a) Phylogenetic analysis using the Maximum Likelihood method and JTT matrix-based model (Saitou and Nei 1987) showing the predictive KS classification based on substrate specificity. (b) Co-occurrence network of AD domains based on substrate specificity: blue for Gly, pink for Orn, and orange for Thr. Edges (lines) indicate that pairwise alignments of nodes. This was conducted by EFI-EST SSN analysis (Gerlt et al. 2015, 2011), which is visualized by Cytoscape (Shannon et al. 2003). Pairwise identity (%) between AD sequences obtained in this study and representaive ADs from other biosynthetic pathways. Note: IAD1 = ID677AD1, IAD2 = ID677AD2, IAD3 = ID677AD3.

tomyces genera: *S. aridus* H9 (98.7%), S. polymachus T258 (98.2%), *S. lavendofoliae* NBRC (99.0%), *S. indoligenes* TRM 43006 (98.7%), *S. roseolilacinus* NRBC 12815 (98.3%), *S. luridus* NBRC 12793 (99.2%), *S. lavendulocolor* NBRC 12881 (99.2%, and *S. gobitricini* NBRC 15419 (99.0%) (Supplementary Figure 1b). The outcome of this 16S rRNA analysis strongly suggested the taxonomic identity of ID 04 677 as *S. luridus*, and therefore

we named it as *Streptomyces luridus* ID 04 677.

Interestingly, the EtOAc extract prepared from this strain exhibited antimicrobial activity against *P. aeruginosa* (Fig. 1C), a Gram-negative pathogenic bacterium known as the leading cause of morbidity and mortality in cystic fibrosis (CF) patients and nosocomial infections (Moradali et al. 2017). To get insight into the habitability of this strain in diverse environmental

FIGURE 2 Phylogenetic analysis of two ketosynthase sequences (SDS14KS4 and SDS14KS5) from *Streptomyces luteosporeus* SDS14LU, which involved a total of 17 KS sequences from cis-AT modules with the specificity of different substrates. The optimal tree was constructed in MEGA X (Kumar et al. 2018) using the Maximum Likelihood method and JTT matrix-based model (Saitou and Nei 1987). Representative types of cis-AT KS domains in loading modules (usually fused with a first extender module) with their starter units are indicated with blue bullets for acetyl-CoA, red bullets for propionyl-CoA, and a green bullet for methylbutyryl-CoA. Representative KS domains from cis-AT loading modules with the substrate specificity for and α-methylated intermadite. The extension units are shown with pink bullets for methylmalonyl-CoA and orange bullets for malonyl-CoA.

types, we submitted the almost complete 16S rRNA gene to MetaMetaDB (http://mmdb.aori.u-tokyo.ac.jp), a database and analytic system for investigating microbial habitability (Yang and Iwasaki 2014). Its 16S rRNA gene sequence showed >97% identity (37 hits) with those associated with root (28.50%), soil (24.38%), rhizosphere (22.44%), ant (21.80%), and marine (2.88%), suggesting the predominant habitability of this strain in soil habitats associated with root and rhizosphere as well as in insects such as ant.

The second strain (SDS14LU) derived from Mount Tambora Savana appeared as large rough-round colonies with light red-colored aerial mycelium observed for 7 days on the ISP2 solid medium (Supplementary Figure 2a). Its partial 16S rRNA gene was previously sequenced and deposited with the accession number KT593926.1 (Fawzya et al. 2016), and Fawzya et al. (2016) suggested its identity as Streptomyces thioluteus (Fawzya et al. 2016). However, more recent BLASTn analysis indicated its highest identity with other *Streptomyces* species, such as *S. cinnamoneus* (Labeda et al. 2012) (99.32%), followed with that of *S. griseocarneus* (99.25%), *S. luteosporous* (99.24%), and *S. lilacinus* (99.17%). Mega-BLAST search against the updated 16S microbial database (from characterized microbial strains) exhibited the highest identity with that of *S. luteosporeus* NBRC 14657 (98.4%), *S. aureoversilis* NBRC 13021 (98.7%), *S. rectiverticillatus* NBRC 13079 (98.2%). Further phylogenetic analysis showed that its taxonomic affiliation was closer to *S. luteosporeus* than to other two *Streptomyces* species mentioned above (Supplementary Figure 2b). This suggests the identity of SDS14LU strain as *Streptomyces luteosporeus*. An-

timicrobial assay of the EtOAc extract from this strain showed inhibitory activity against *B. subtilis* and *P. aeruginosa*, two representatives of Gram-positive and Gram-negative food-borne bacteria (Supplementary Figure 2c). MetaMetaDB (Yang and Iwasaki 2014) search indicated its high identity (>97%, 28 hits) with those present in root (26.60%), rhizosphere (22.34%), soil (21.49%), and ant (18.08%) (Supplementary Figure 2d), suggesting that this *S. luteosporeus* can be found or may be adapted in habitats associated with soil and insects.

3.2. Secondary metabolite biosynthetic potential

Using the degenerate primers for detecting PKS (Piel 2002) genes, we PCR-amplified and cloned KS-encoding fragments into *E. coli*. Subsequent sequencing and analyzing some clones with correct-sized inserts indicated that four amplicons belonged to KS domains characterized by the presence of typical KS conserved motifs (e.g., MD-PQQR) and two of the conserved catalytic triad (C-H-H) (Robbins et al. 2016) (Supplementary Figure 3). They retain the active sites CSG/SSL and HGTGT, suggesting its ability to catalyze translocation of carbon chain and decarboxylative condensation (Robbins et al. 2016).

One of the two KS amplicons from *S. luridus* ID 04 677 (ID677KS1) exhibited high identity (73%-74%) with those in cis-AT PKS modules, exemplified by PieA6 in the biosynthesis of piericiden A1 (Engl et al. 2018) (Supplementary Figure 4). Interestingly, the sequence alignment between ID677KS1 and other KS domains from the same cis-AT PKS clade indicated highly conserved amino acid residues that could be responsible for its specificity determinant, as has been shown for EryKS3 (erythromycin

PKS), PikAIII (pikromycin PKS), MlsA2KS (mycolactone PKS) (Murphy et al. 2016) (Supplementary Figure 4). Based on NaPDoS analysis (Ziemert et al. 2012), the amplicon ID677KS1 displayed the highest similarity (87% identity) with KS in cis-AT PKS catalyzing lymphostin biosynthesis (Aotani et al. 1997) (Supplementery Table 1). The second KS (ID677KS4) from *S. luridus* ID 04 677 was closely related to the first KS in the extension module McyD (Nishizawa et al. 1999) (Supplementary Table 1). This sequence analysis strongly suggests the classification of both KS amplicons into cis-AT PKS type.

Further phylogenetic analysis against several KS domains from some representative biosynthetic pathways retrieved from MIBiG database (Kautsar et al. 2020) classified ID677KS1 into the extension module clade and ID677KS4 into the loading module type (Figure 1a). ID677KS1 was phylogenetically closer to PieA6-KS2 (74% pairwise idenity), a cis-AT KS domain that accepts an acyl intermediate containing α-methyl and β-keto moiety (Liu et al. 2012), suggesting that they may share a similar substrate specificity (Figure 1a). ID677KS4 phylogenetically belonged to the same clade as KS domains in the loading modules. It was closer to the KS from MtaB with the substrate specificity for methyl butyryl-CoA (Silakowski et al. 1999).

Using the degenerate primers for detecting NRPS (Ayuso-Sacido and Genilloud 2005) genes, we obtained three different AD sequences (ID677AD1, ID677AD2, ID677AD3) from *S. luridus* ID 04 677. The three amplicons were characterized with the presence of the core motifs A3 to A7 (e.g., GPTE), including residues predicted to be responsible for substrate specificity (Konz and Marahiel 1999) (Sypplementary Figure 5). BLASTp analysis of ID677AD1 against MIBiG database (Kautsar et al. 2020) showed 48.5% identity with AD in NcpA (4-methylproline NRPS) that activates glycine (Supplementary Table 2). NRPSpredictor2 (Röttig et al. 2011) predicted its substrate specificity for glycine. The other domain ID677AD2 shared 49.8% identity with AD1 in NdaA (nodularin NRPS) (Jokela et al. 2017) from *Nostoc sp.* with the predicted substrate specificity for threonine (Röttig et al. 2011). The third domain ID677AD3 showed 46.2% identity with AD in NocB (nocardicin A NRPS) (Gunsior et al. 2004) from *Nocardia uniformis*. It was predicted to confer specificity for ornithine (match: 50%) based on NRPSpredictor2 (Röttig et al. 2011) (Table S2). This was supported by EFI-EST SSN analysis (Gerlt et al. 2015) showing the co-occurrence network of AD sequences with the same substrate specificity (Figure 1b).

BLAST search showed that two different KS sequences from the strain *S. luteosporeus* SDS14LU (SDS14KS4 and SDS14KS5) were similar with many KS domains from type I PKS system (Supplementary Figure 6). The first KS sequence (SDS14KS4) shared the identity of 41% to 53% with those in the cis-AT modules, such as McyD (microcystin PKS) (Tillett et al. 2000). The other KS (SDS14KS5) showed the high identity of 74-76% with

those in the cis-AT modules, such as PikAI_Q9ZGI51KSB (pikromycin PKS) (Xue et al. 1998) (Supplementary Table 3). This sequence analysis suggests that both SDS14KS4 and SDS14KS5 belong to cis-AT PKS system.

Phylogenetic analysis of both KS sequences against those from several characterized biosynthetic pathways divided them into two distinct cis-AT PKS clades, which were separated from the FAS clade (Figure 2). A close look at the multiple alignment suggests the presence of conserved residues in cis-AT KS of loading modules that are different from the counterpart residues in cis-AT KS of extension modules (Supplementary Figure 6). SDS14KS4 was particularly grouped in the same clade as KS domains from cis-AT chain-extension modules, such as McyD (Tillett et al. 2000). We propose that a cis-AT chain-extension module harboring SDS14KS4 may recruit malonyl-CoA through the cognate AT domain and load it onto the ACP domain. The ACP may deliver this extender unit to SDS14KS4 for a decarboxylative condensation reaction with the installed intermediate (Weissman 2015).

In contrast to SDS14KS4, SDS14KS5 was classified into a different cis-AT KS group associated with loading modules, which is usually fused to the first extender module (Buntin et al. 2010) (Figure 2). In particular, SDS14KS5 was closely related to the KS in HerB, a loading module in herboxidiene biosynthesis that accepts the starter unit propionate (Shao et al. 2012). One of the unique conserved residues in SDS14KS5 is a glutamine in place of a cysteine in the motif CSSL, which is highly conserved in some KS domains in loading modules. These so called KSQ domains have no condensation activity but can decarboxylate dicarboxylic acid starters. In the KSQ-type loading module harboring this KS, it is assumed that the AT domain might select propionyl-CoA as a starter substrate and transfers it onto the ACP domain. Subsequent KSQ-catalyzed decarboxylation generates a propionate unit, which is then transferred to the downstream extension module for polyketide chain elongation (Miyanaga 2017).

3.3. Accession numbers

The partial 16S rRNA gene sequence of the strain ID04-677 obtained in this work was deposited in the GenBank with Accession Number of KX150801. KS and AD sequences isolated from ID 04 677 obtained in this work was deposited in the GenBank database under the following accession numbers: MT676434, MT676435, MT676436, MT676437, and MT676438. KS sequence derived from SDS14LU in this work was deposited under the accession numbers MT676439.

3.4. Discussion

In this work, we analyzed biosynthetic potential of two *Streptomyces* strains derived from geographically different locations in Indonesia, namely Puwardadi (a region in West java) and Mount Tambora (an active stratovolcano Savanna located in the northern part of Sumbawa, West Nusa Tenggara). Phylogenetic analysis indicated the iden-

tity of the Puwardadi strain (ID 04 677) as *S. luridus* and the Tambora strain (SDS14LU) as *S. luteosporeus*. To the best of our knowledge, both species remain poorly investigated. Based on MetaMetaDB search (Yang and Iwasaki 2014), their 16S rRNA gene sequences shared >97% identity with thoses found predominantly in root, rhizosphere, soil, and ant, suggesting that they live not only as free-living soil bacteria, but also as symbiosis with plant roots and insects.

The antimicrobial assays indicated that both strains exhibited antibacterial activity towards the pathogenic bacterium *P. aeruginosa*. Since this opportunistic pathogen is resistant to a wide range of antibiotics (Moradali et al. 2017), both strains promise biosynthetic potential to produce multi-drug resistant antimicrobial compounds. The ability of them to produce antimicrobial metabolites combined with their predominant distribution in root, rhizosphere, soil, and ant may indicate their ecological role as biocontrols to protect the plant and ant hosts from microbial pathogens (van der Meij et al. 2017). This bioecological information provides a strategic direction to explore them from unique habitat types located in geographically different regions for drug discovery program.

Previous studies indicated that *S. luridus* was a producer of compounds with bioemulsification potential (Lamilla et al. 2018). *S. luteosporeus* is known to produce thiolutin and indolmycin (Celmer and Solomons 1955). Although some compounds have been described for both strains, such previous studies indicated that their biosynthetic capacity remain poorly investigated, especially in terms of polyketide and non-ribosomal peptide biosynthetic pathways. Therefore, identification of PKS and NRPS based on KS and AD sequence analysis described in this work is necessary to reveal the biosynthetic potential of both strains.

Our genetic analysis indicated that two KS domains obtained from *S. luridus* ID 04 677 are phylogenetically derived from different cis-AT PKS module types: the first one (ID677KS1) belongs to an extension module with the predictive specificity for an α-Me and β-keto intermediate, and the second one (ID677KS1) is a part of a loading module with the predictive specificity for methylbutyryl-CoA. Phylogenetic analysis of two KS domains derived from *S. luteosporeus* SDS14LU suggests that one of them (SDS14KS4) belongs to a cis-AT chain extension module with the predictive specificity for an α-Me and β-Ome intermediate. The other KS (SDS14KS5) seems to be a part of a KSQ-type loading module that consists of KSQ, AT, and ACP (Miyanaga 2017) with the predicted specificity for propionyl-CoA. Three different AD domains of NRPS system were cloned from *S. luridus* ID 04 677. Their predicted specificity are Gly, Thr, and Orn (Röttig et al. 2011). To the best of our knowledge, no PKS and NRPS pathways harboring such KS and AD features have been reported so far from both strains.

Through genetic and bioinformatic analyses, we have shown the biosynthetic potential of both strains. Further genome sequencing and chemical analyses are necessary to reveal their entire biosynthetic pathways responsible for polyketide and non-ribosomal peptide production. The genetic analysis approach described in this work can be used as a simple guide not only to obtain rapid insight into natural product biosynthetic gene diversity in microorganisms (both cultivated and uncultivated), but also to prevent the re-discovery of polyketide and peptide biosynthetic pathways described previously.

4. Conclusions

Bacterial strain (ID 04-677 and SDS14LU) was screened from Purwadadi and Mount Tambora. ID 04-677 was identity as Streptomyces luridus and SDS14LU was identity as Streptomyces thioluteus. One of the two KS amplicons from S. luridus ID 04 677 exhibited high identity (73%-74%) with those in cis-AT PKS modules. Two different KS sequences from the strain S. luteosporeus SDS14LU were similar with many KS domains from type I PKS system. The antimicrobial assays indicated that both strains exhibited antibacterial activity towards the pathogenic bacterium P. aeruginosa. Genetic analysis is one of technique for obtain natural biosynthetic gene diversity in microorganisms.

Authors' contributions

ARU, DSZ, SS, PL designed the study. SS and ARU performed study supervision. AU, PA, DSZ, YK, ZI carried out the laboratory work PA, YK, ZI BY analyzed the data. ARU, SS, DSZ, PL, AU wrote the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The author declare that they have no competing interest.

References

Aotani Y, Nagata H, Yoshida M. 1997. Lymphostin (LK6-A), a novel immunosuppressant from *Streptomyces* sp. KY11783: Structural elucidation. J Antibiot. 50(7):543–545. doi:10.7164/antibiotics.50.543.

Aparicio JF, Molnár I, Schwecke T, König A, Haydock SF, Khaw LE, Staunton J, Leadlay PF. 1996. Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: Analysis of the enzymatic domains in the modular polyketide synthase. Gene. 169(1):9–16. doi:10.1016/0378-1119(95)00800-4.

Ayuso-Sacido A, Genilloud O. 2005. New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: Detection and distribution of these biosynthetic gene sequences in major taxonomic groups. Microb Ecol. 49(1):10–24. doi:10.1007/s00248-004-0249-6.

Bérdy J. 2005. Bioactive microbial metabolites: A personal view. J Antibiot. 58(1):1–26. doi:10.1038/ja.2005.1.

Buntin K, Irschik H, Weissman KJ, Luxenburger E, Blöcker H, Müller R. 2010. Biosynthesis of Thuggacins in Myxobacteria: Comparative Cluster Analysis Reveals Basis for Natural Product Structural Diversity. Chem Biol. 17(4):342–356. doi:10.1016/j.chembiol.2010.02.013.

Cane DE, Walsh CT. 1999. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem Biol. 6(12). doi:10.1016/S1074-5521(00)80001-0.

Celmer WD, Solomons IA. 1955. The Structures of Thiolutin and Aureothricin, Antibiotics Containing a Unique Pyrrolinonodithiole Nucleus. J Am Chem. Soc. 77(10):2861–2865. doi:10.1021/ja01615a058.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797. doi:10.1093/nar/gkh340.

Engl T, Kroiss J, Kai M, Nechitaylo TY, Svatoš A, Kaltenpoth M. 2018. Evolutionary stability of antibiotic protection in a defensive symbiosis. Proc Natl Acad Sci USA. 115(9):E2020–E2029. doi:10.1073/pnas.1719797115.

Fawzya YN, Zilda DS, Chaniago S, Prestisia HN, Lisdiyanti P, Khasanah N. 2016. Screening of Indonesian *Streptomyces* sp. capable of secreting transglutaminase (MTGase) and optimization of MTGase production using different growth media. Squalen Bulletin of Marine and Fisheries Postharvest and Biotechnology 11(1):13–21.

Fischbach MA, Walsh CT. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms. Chem Rev. 106(8):3468–3496. doi:10.1021/cr0503097.

Gerlt JA, Allen KN, Almo SC, Armstrong RN, Babbitt PC, Cronan JE, Dunaway-Mariano D, Imker HJ, Jacobson MP, Minor W, et al. 2011. The enzyme function initiative. Biochemistry. 50(46):9950–9962. doi:10.1021/bi201312u.

Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL. 2015. Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta, Proteins Proteomics. 1854(8):1019–1037. doi:10.1016/j.bbapap.2015.04.015.

Gunsior M, Breazeale SD, Lind AJ, Ravel J, Janc JW, Townsend CA. 2004. The biosynthetic gene cluster for a monocyclic β-lactam antibiotic, nocardicin A. Chem Biol. 11(7):927–938. doi:10.1016/j.chembiol.2004.04.012.

Hertweck C. 2009. The biosynthetic logic of polyketide diversity. Angewandte Chemie - International Edition 48(26):4688–4716. doi:10.1002/anie.200806121.

Hertweck C. 2015. Decoding and reprogramming complex polyketide assembly lines: Prospects for synthetic biology. Trends Biochem Sci. 40(4):189–199. doi:10.1016/j.tibs.2015.02.001.

Jokela J, Heinilä LM, Shishido TK, Wahlsten M, Fewer DP, Fiore MF, Wang H, Haapaniemi E, Permi P, Sivonen K. 2017. Production of high amounts of hepatotoxin nodularin and new protease inhibitors pseudospumigins by the brazilian benthic *Nostoc sp*. CENA543. Front Microbiol. 8(OCT). doi:10.3389/fmicb.2017.01963.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8(3):275–282. doi:10.1093/bioinformatics/8.3.275.

Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, Van Der Hooft JJ, Van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, et al. 2020. MIBiG 2.0: A repository for biosynthetic gene clusters of known function. Nucleic Acids Res. 48(D1):D454–D458. doi:10.1093/nar/gkz882.

Konz D, Marahiel MA. 1999. How do peptide synthetases generate structural diversity? Chem. Biol. 6(2). doi:10.1016/S1074-5521(99)80002-7.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol. Evol. 35(6):1547–1549. doi:10.1093/molbev/msy096.

Labeda DP, Goodfellow M, Brown R, Ward AC, Lanoot B, Vanncanneyt M, Swings J, Kim SB, Liu Z, Chun J, et al. 2012. Phylogenetic study of the species within the family Streptomycetaceae. In: Antonie Van Leeuwenhoek., volume 101. p. 73–104. doi:10.1007/s10482-011-9656-0.

Lamilla C, Braga D, Castro R, Guimarães C, de Castilho LV, Freire DM, Barrientos L. 2018. *Streptomyces luridus* So3.2 from Antarctic soil as a novel producer of compounds with bioemulsification potential. PLoS ONE. 13(4). doi:10.1371/journal.pone.0196054.

Liu C, Jiang Y, Wang X, Chen D, Chen X, Wang L, Han L, Huang X, Jiang C. 2017. Diversity, Antimicrobial Activity, and Biosynthetic Potential of Cultivable Actinomycetes Associated with Lichen Symbiosis. Microb Ecol. 74(3):570–584. doi:10.1007/s00248-017-0972-4.

Liu Q, Yao F, Chooi YH, Kang Q, Xu W, Li Y,

Shao Y, Shi Y, Deng Z, Tang Y, You D. 2012. Elucidation of piericidin A1 biosynthetic locus revealed a thioesterase-dependent mechanism of α-pyridone ring formation. Chem Biol. 19(2):243–253. doi:10.1016/j.chembiol.2011.12.018.

Miyanaga A. 2017. Structure and function of polyketide biosynthetic enzymes: various strategies for production of structurally diverse polyketides. doi:10.1080/09168451.2017.1391687.

Moradali MF, Ghods S, Rehm BH. 2017. *Pseudomonas aeruginosa* lifestyle: A paradigm for adaptation, survival, and persistence. Front Cell Infect Microbiol. 7(FEB). doi:10.3389/fcimb.2017.00039.

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. 2008. Erratum: Database indexing for production MegaBLAST searches. Bioinformatics 24(24):2942. doi:10.1093/bioinformatics/btn554.

Murphy AC, Hong H, Vance S, Broadhurst RW, Leadlay PF. 2016. Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. Chem Commun. 52(54):8373–8376. doi:10.1039/c6cc03501a.

Nguyen TA, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, Taudien S, Platzer M, Hertweck C, Piel J. 2008. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nat Biotechnol. 26(2):225–233. doi:10.1038/nbt1379.

Nishizawa T, Asayama M, Fujii K, Harada KI, Shirai M. 1999. Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp. J Biochem. 126(3):520–529. doi:10.1093/oxfordjournals.jbchem.a022481.

Piel J. 2002. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus beetles*. Proc Natl Acad Sci USA. 99(22):14002–14007. doi:10.1073/pnas.222481399.

Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. 2004. Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. Proc Natl Acad Sci USA. 101(46):16222–16227. doi:10.1073/pnas.0405976101.

Rawlings BJ. 2001. Type I polyketide biosynthesis in bacteria (Part A -Erythromycin biosynthesis). Nat Prod Rep. 18(2):190–227. doi:10.1039/b009329g.

Risdian C, Mozef T, Wink J. 2019. Biosynthesis of polyketides in *Streptomyces*. Microorganisms. 7(5). doi:10.3390/microorganisms7050124.

Robbins T, Kapilivsky J, Cane DE, Khosla C. 2016. Roles of Conserved Active Site Residues in the Ketosynthase Domain of an Assembly Line Polyketide Synthase. Biochemistry. 55(32):4476–4484. doi:10.1021/acs.biochem.6b00639.

Röttig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O. 2011. NRPSpredictor2 - A web server for predicting NRPS adenylation domain speci-

ficity. Nucleic Acids Research 39(SUPPL. 2). doi:10.1093/nar/gkr323.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4(4):406–425. doi:10.1093/oxfordjournals.molbev.a040454.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 13(11):2498–2504. doi:10.1101/gr.1239303.

Shao L, Zi J, Zeng J, Zhan J. 2012. Identification of the herboxidiene biosynthetic gene cluster in *Streptomyces chromofuscus* ATCC 49982. Appl Environ. Microbiol. 78(6):2034–2038. doi:10.1128/AEM.06904-11.

Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. Protein Science 27(1):135–145. doi:10.1002/pro.3290.

Silakowski B, Schairer HU, Ehret H, Kunze B, Weinig S, Nordsiek G, Brandt P, Blöcker H, Höfle G, Beyer S, Müller R. 1999. New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. J Biol Chem. 274(52):37391–37399. doi:10.1074/jbc.274.52.37391.

Tillett D, Dittmann E, Erhard M, Von Döhren H, Börner T, Neilan BA. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: An integrated peptide-polyketide synthetase system. Chem Biol. 7(10):753–764. doi:10.1016/S1074-5521(00)00021-1.

van der Meij A, Worsley SF, Hutchings MI, van Wezel GP. 2017. Chemical ecology of antibiotic production by actinomycetes. doi:10.1093/femsre/fux005.

Watanabe K, Wang CC, Boddy CN, Cane DE, Khosla C. 2003. Understanding Substrate Specificity of Polyketide Synthase Modules by Generating Hybrid Multimodular Synthases. J Biol Chem. 278(43):42020–42026. doi:10.1074/jbc.M305339200.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol. 173(2):697–703. doi:10.1128/jb.173.2.697-703.1991.

Weissman KJ. 2015. Uncovering the structures of modular polyketide synthases. Natural Product Reports 32(3):436–453. doi:10.1039/c4np00098f.

Xue Y, Zhao L, Liu HW, Sherman DH. 1998. A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: Architecture of metabolic diversity. Proc Natl Acad Sci US. A. 95(21):12111–12116. doi:10.1073/pnas.95.21.12111.

Yang CC, Iwasaki W. 2014. MetaMetaDB: A database and analytic system for investigating microbial habitability. PLoS ONE. 9(1). doi:10.1371/journal.pone.0087126.

Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen

PR. 2012. The natural product domain seeker NaP-
DoS: A phylogeny based bioinformatic tool to clas-
sify secondary metabolite gene diversity. PLoS ONE.
7(3). doi:10.1371/journal.pone.0034064.