Indonesian J. Pharm. Vol. 28 No. 1 : 26 – 33
ISSN-p : 2338-9427
DOI: 10.14499/indonesianjpharm28iss1pp26

Research Article

SECONDARY BIOACTIVE METABOLITE GENE CLUSTERS IDENTIFICATION OF ANTICANDIDA-PRODUCING *Streptomyces* Sp. GMR22 ISOLATED FROM WANAGAMA FOREST AS REVEALED BY GENOME MINING APPROACH

Camelia Herdini[1,5]*, Sofia Mubarika[1], Bambang Hariwiyanto[1,5], Nastiti Wijayanti[1], Akira Hosoyama[2], Atsushi Yamazoe[2], Hideaki Nojiri[3], Jaka Widada[4]*

[1]Graduate School of Biotechnology, Universitas Gadjah Mada, Barek Utara, Yogyakarta, Indonesia
[2]Biological Resource Center, Nat Inst of Technology and Evaluation, Nishihara, Shibuya-ku, Tokyo, Japan
[3]Biotechnology Research Center, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
[4]Dept of Agricultural Microbiology, Universitas Gadjah Mada, Bulaksumur, Yogyakarta, Indonesia
[5]Otorhinolaryngology Head and Neck Surgery, Faculty of Medicine, Universitas Gadjah Mada, Sekip, Yogyakarta, Indonesia

Submitted: 09-08-2016
Revised: 08-10-2016
Accepted: 15-12-2016

*Corresponding author
Jaka Widada

Email:
jwidada@ugm.ac.id

ABSTRACT

Streptomyces are a group of Gram-positive bacteria belonging to the Actinobacteria class, which are among the most important bacteria for producing secondary bioactive metabolites such as antibiotics, chemotherapeutics, insecticides and other high-value chemicals. Genome mining of gene clusters that encode the biosynthetic pathways for these metabolites has become a key methodology for novel compound discovery. Recently, we have isolated the *Streptomyces* sp. GMR22 from Cajuput rhizospheric soil at Wanagama Forest, Indonesia. GMR22 produced a secondary metabolite that inhibited *Candida albicans* with IC_{50} of 62.5µg/mL. The objective of this work was to reveal the novel secondary metabolites from GMR22 by genome mining approach. The antiSMASH 3.0 was used to predict gene clusters that encode the biosynthetic pathways of secondary metabolites in the genome of GMR22, and their core chemical structures. The phylogenomic analysis showed that GMR22 was closely related to *Streptomyces bingchenggensis* BCW1, as well as to the large genome size (9.5-12.7Mbp) groups of Streptomyces. AntiSMASH 3.0 analysis revealed that the genome of *Streptomyces* sp. GMR22 harbored at least 63 gene clusters that encode the biosynthetic pathways of secondary metabolites. It was the highest number of gene clusters had been observed among the members of Streptomyces groups, with polyketide synthetase (PKS) was predicted as the major groups of the identified gene cluster products. The results suggested that GMR22 could be a strong potential candidate for secondary bioactive metabolites source.

Key words: *Streptomyces* sp. GMR22, antifungal, genome sequence, antiSMASH

INTRODUCTION

Streptomyces is a representative of a filamentous bacteria responsible for producing variety of valuable secondary metabolites, such as antibiotics, parasiticides, herbicides, and pharmacologically active substances, including antitumor agents and immunosuppressants (Waksman and Woodruff, 1940; Ditsler *et al*., 1992; Euverink, 1995; Newman *et al*., 2000; El Hassan *et al*., 2001; Bentley *et al*., 2002; Bibb, 2005; Quintana *et al.,* 2008; Li *et al*., 2009; Doroghazi *et al*., 2014). Genome sequencing analyses of *Streptomyces avermitilis* (Omura *et al*., 2001), *Streptomyces coelicolor* A3(2) (Bentley *et al*., 2002), *Streptomyces griseus* IFO 13350 (Ohnishi *et al*., 2008), and *Streptomyces bingchenggensis* (Wang *et al*., 2010) revealed that Streptomyces possess an unexpected abundance of natural product biosynthetic gene clusters and thus that they have the potential to make many more compounds than previously thought.

Previously, we isolated several Streptomyces from rhizospheric soil at Wanagama Forest Indonesia (Nurjasmi *et al*., 2009). *Streptomyces* sp. GMR22 isolated from Cajuput rhizospheric soil produced strong antifungal secondary metabolite and inhibited *Candida albicans* with IC_{50} of 62.5µg/mL (Nurjasmi *et al*., 2009; Alimuddin *et al*., 2010).

Recently, genomic discovery from some Streptomyces can be used as guidance to obtain new bioactive compounds. In addition,

biosynthesis gene clusters mining their genomes becomes a key method to accelerate identification and characterization. The antibiotic and secondary metabolite analysis of the shell (antiSMASH) is a web server and stand-alone software, which combines automatic identification of secondary metabolite gene clusters in the genome sequence analysis of a large collection of algorithms-specific compounds (Medema *et al*., 2011; Weber *et al*., 2015). For the last two years, antiSMASH became a standard tool for analyzing the genomes of bacteria and fungi for their potential of secondary metabolites production (Blin *et al*., 2013). In this paper, we reported the results of the genome mining by using AntiSMASH for analyzing genome sequence of GMR22 and compared to the most closed related *Streptomyces* based on the analysis of whole genome phylogenetic (phylogenomic).

MATERIAL AND METHODS

Analysis of whole genome phylogenetic tree of GMR22

Streptomyces sp. GMR22 genome sequence analysis was done by Next Generation Sequencing (NGS) platforms using 454 pyrosequencing technology (454 GS FLX) and HiSeq1000 (Illumina). Whole-genome phylogenetic analysis (phylogenomic) was done by Neighbour-joining algorithm showed genome relationship of *Streptomyces* sp. GMR22 and 20 complete genome sequences of *Streptomyces*. Phylogeny tree was created using CV tree with a *k*-value of 6 (Xu and Hao, 2009) with *Bacillus subtilis* natto BEST195 as an out-group and visualized with MEGA 6.06 (Tamura *et al*., 2013).

Analysis of Genome Mining of *Streptomyces* Sp. GMR22

Identification, annotation, and analysis of gene cluster involved in the biosynthesis of secondary metabolites and predictions core structure produced done by using antiSMASH 3.0. (Medema *et al*., 2011; Weber *et al*., 2015). Available at HTTP://antism ash. secondarymetabolites.org. Prodigal and run secondary metabolite detection on all possible ORFs were chosen in the Gene finding options parameters.

RESULTS AND DISCUSSION

Phylogenomic of GMR22

Genome sequence analysis results showed that the size of the genome of GMR22 was 11,42 Mbp. Based on the phylogenomic analysis showed that GMR22 isolate closely related to *Streptomyces bingchenggensis* BCW1 and clustered with strains *S. violaceusniger* Tu 4113*, S. rapamycinicus* NRRL 5491*, and S. davawensis* JCM 4913 that have large genome size (9.5-12.7Mbp).

Secondary metabolite biosynthesis gene clusters in *Streptomyces* sp. GMR22

Secondary metabolism of microbes is a rich source of antibiotics, chemotherapy, insecticides and other high value chemicals. Genome mining group biosynthesis pathway genes coding for these metabolites has been a key methodology for the discovery of new compounds. To determine the gene cluster and bioactive secondary metabolites produced by strain GMR22 were then analyzed with AntiSMASH 3.0 program (Table I).

The results indicated that GMR22 harbored at least 63 gene clusters encoding the biosynthetic pathways of secondary metabolites. It was the highest number of gene clusters had been observed among the members of Streptomyces groups. It was also interesting that more than a third of the genes clusters are polyketide syntethase: 21 gene clusters contained of type I of pure polyketide syntethase (T1pks), 3 gene clusters contained of type II of polyketide syntethase (T2pks), 3 gene clusters contained of hybrid type III of polyketide syntethase, and 1 gene cluster of hybrid type I of polyketide syntethase.

Compared to the closest strain in phylogenomic tree (Figure 1), namely *Streptomyces bingchenggensis* BCW-1 (Wang *et al*., 2010), some secondary metabolites have some similarities. AntiSMASH analysis revealed that *Streptomyces* sp. GMR22 harboring at least 63 bioactive secondary metabolites gene clusters (12 NRPS and 28 PKS), while BCW-1 only has 53 bioactive secondary metabolites gene clusters (16 NRPS and 20 PKS). At least there are 9 of bioactive secondary metabolites produced by both strain, they were ectoin, hopene, actinomycin,, skyllamycin, geldanamycin, desferrioxamine B, echosides, meridamycin and spores pigments.

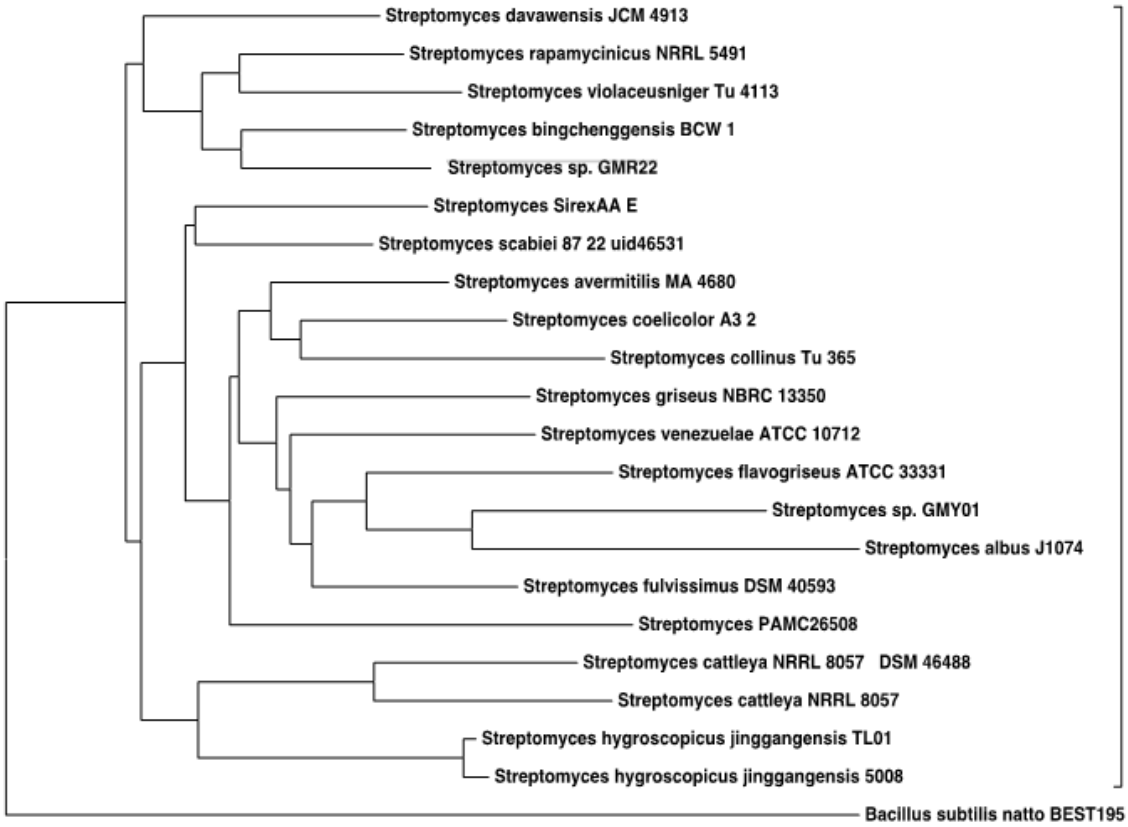

Figure 1. Analysis of whole-genome phylogenetic tree with Neighbour-joining algorithm showed the relationship of *Streptomyces* sp. GMR22 and other 20 completed genomic sequences of *Streptomyces*. Phylogeny tree was created using CVtree with a *k*-value of 6 (Xu and Hao, 2009) with *Bacillus subtilis* natto BEST195 as an out-group and was visualized with MEGA 6.06 (Tamura *et al.,* 2013).

Table 1. Cluster of genes and bioactive secondary metabolites of *Streptomyces* sp. GMR22 based on the analysis of genome sequences with AntiSMASH 3.0 program cluster 1-14.

Cluster	**Type**	**Size (bp)**	**Most silmilar known cluster (%)**	**MIB***
Cluster 1	T1pks	40541	Ansamitosin (21)	BGC0000020_c1
Cluster 2	Nrps	53307	Actinomycin (7)	BGC0000296_c1
Cluster 3	T1pks	1553	-	
Cluster 4	T1pks	26277	Hygrocin (74)	BGC0000075_c1
Cluster 5	Lantipeptide	23294	Reveromycin (9)	BGC0000135_c1
Cluster 6	T1pks-Other ks	123792	Bafilomycin (50)	BGC0000957_c1
Cluster 7	Bacteriocin-Lantipeptide	28221	-	
Cluster 8	T1pks	47645	Ansatrienin (7)	BGC0000957_c1
Cluster 9	T2pks	42504	Lysolipin (23)	BGC0000242_c1
Cluster 10	Terpene	21101	-	
Cluster 11	Nrps-Lantipeptide	111619	Hormaomycin (8)	BGC0000374_c1
Cluster 12	T1pks	45292	-	
Cluster 13	Lassopeptide	22386	-	
Cluster 14	Siderophore	13368	-	

Table II. Cluster of genes and bioactive secondary metabolites of *Streptomyces* sp. GMR22 based on the analysis of genome sequences with AntiSMASH 3.0 program cluster 15-55

Cluster	Type	Size (bp)	Most silmilar known cluster (%)	MIB*
Cluster 15	Butyrolactone	10990	-	
Cluster 16	Nrps	88213	Laspartomycin (53)	BGC0000379_c1
Cluster 17	T1pks	7156	Chalcomycin (17)	BGC0000035_c1
Cluster 18	T1pks	12100	ECO-02301 (32)	BGC0000052_c1
Cluster 19	Terpene	18680	Hopene (53)	BGC0000663_c1
Cluster 20	Ectoine	10404	Ectoine (75)	BGC0000853_c1
Cluster 21	T1pks	9684	-	
Cluster 22	T1pks	79794	Elaiophylin (87)	BGC0000053_c1
Cluster 23	Hserlactone	20756	Daptomycin (3)	BGC0000336_c1
Cluster 24	Butyrolactone	10936	Meilingmycin (2)	BGC0000093_c1
Cluster 25	Nrps-T3pks	100064	A47934 (44)	BGC0000290_c1
Cluster 26	Nrps	39174	-	
Cluster 27	Nrps	2918	-	
Cluster 28	T1pks	33977	Geldanamycin (69)	BGC0000066_c1
Cluster 29	T2pks	42492	Medermycin (50)	BGC0000245_c1
Cluster 30	T1pks	63580	ECO-02301 (53)	BGC0000052_c1
Cluster 31	Lassopeptide	22646	Polyxypeptin (8)	BGC0001036_c1
Cluster 32	Terpene	21074	Merochlorin (4)	BGC0001083_c1
Cluster 33	Siderophore	13771	-	
Cluster 34	Terpene	25477	Salinilactam (8)	BGC0000109_c1
Cluster 35	T1pks	114377	Nemadectin (50)	BGC0000109_c1
Cluster 36	Arylpolyne-Ladderane	42384	Skyllamycin (18)	BGC0000429_c1
Cluster 37	Nrps	44011	Ochronotic (75)	BGC0000918_c1
Cluster 38	Siderophore	12058	-	
Cluster 39	Bacteriocin	11341	-	
Cluster 40	T1pks	62548	Medermycin (13)	BGC0000245_c1
Cluster 41	Siderophore	11788	DesferrioxamineB (100)	BGC0000941_c1
Cluster 42	T3pks-Nrps	57880	Pheganomycin (38)	BGC0001148_c1
Cluster 43	Other	25323	Echosides (52)	BGC0000340_c1
Cluster 44	Laddarane-Arylpolyene-Nrps	104183	Skyllamycin (46)	BGC0000429_c1
Cluster 45	Nrps	34119	Meridamycin (18)	BGC0001011_c1
Cluster 46	Nrps	78699	Leinamycin (4)	BGC0001101_c1
Cluster 47	Terpene	22238	Brasilicardin (36)	BGC0000632_c1
Cluster 48	Terpene	20909	-	
Cluster 49	Other	42739	Echosides (11)	BGC0000340_c1
Cluster 50	Lantipeptide	24154	-	
Cluster 51	T1pks	44652	Ambruticin (14)	BGC0000014_c1
Cluster 52	T1pks	63597	Elaiophylin (45)	BGC0000053_c1
Cluster 53	T1pks	77675	Nigericin (94)	BGC0000114_c1
Cluster 54	T1pks	126001	Indanomycin (21)	BGC0000079_c1
Cluster 55	Terpene-T3pks-Nrps	34999	Lobophorin (20)	BGC0001183_c1

Table 3. Cluster of genes and bioactive secondary metabolites of *Streptomyces* sp. GMR22 based on the analysis of genome sequences with AntiSMASH 3.0 program cluster 56-63.

Cluster	Type	Size (bp)	Most silmilar known cluster (%)	MIB*
Cluster 56	T2pks	42516	Spore-pigment (75)	BGC0000271_c1
Cluster 57	Terpene	21191	2-methylisoborneol (100)	BGC0000658_c1
Cluster 58	Otherks	41029	Galbonolides (20)	BGC0000065_c1
Cluster 59	Other	44029	-	
Cluster 60	T1pks	71853	Meridamycin (55)	BGC0001011_c1
Cluster 61	T1pks	45286	Hygrocin (38)	BGC0000075_c1
Cluster 62	T1pks	62304	Oligomycin (38)	BGC0000117_c1
Cluster 63	T1pks	46363	Salinomycin (8)	BGC0000144_ct

*The Minimum Information on Biosynthetic Geneclusters (Genecluster repository)

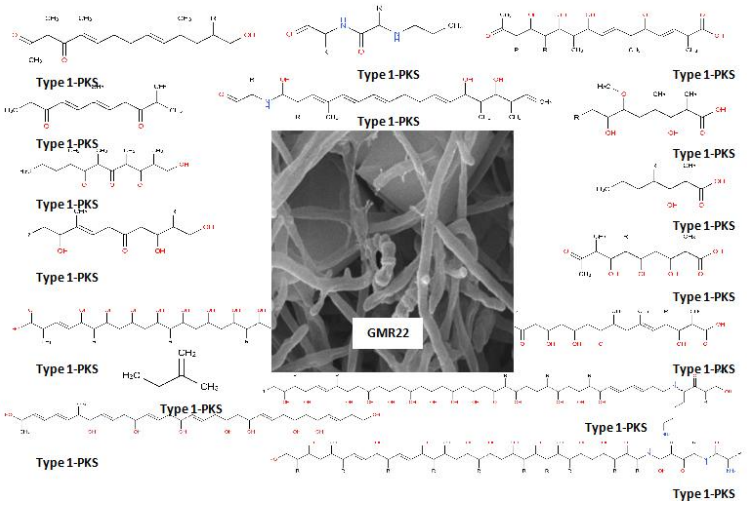

Figure 2. The chemical structure of the type 1-PKS predictable core group of secondary metabolites *Streptomyces* sp. GMR22 based on analysis of genome mining

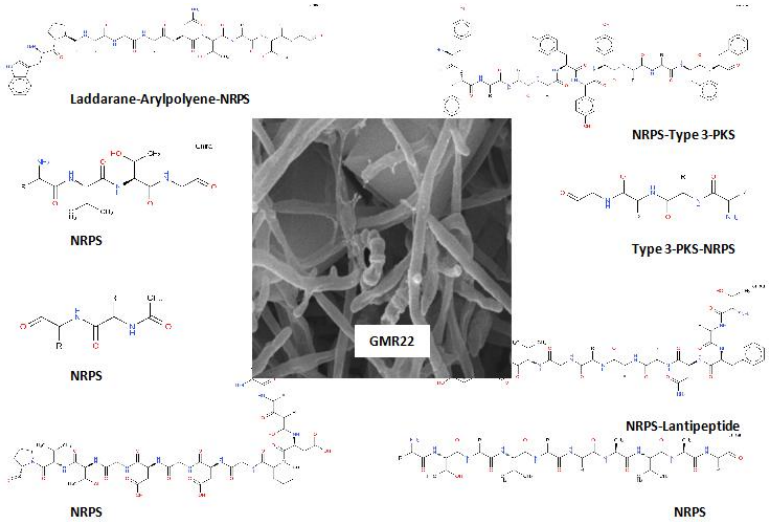

Figure 3. The chemical structure of the predictable core of NRPS group and NRPS hybrid of secondary metabolites *Streptomyces* sp. GMR22 based on analysis of genome mining

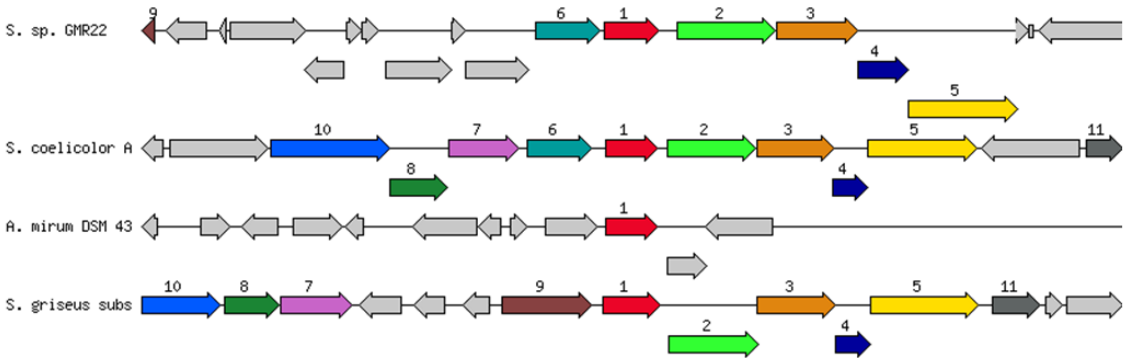

Figure 4. The cluster of genes Streptomyces sp. GMR22 involved in Desferrioxamine B production. S. sp GMR22: *Streptomyces sp.* GMR22; S. coelicolor: *Streptomyces coelicolor*A3(2); A. mirum DSM 43*: Actinosynnema mirum* DSM 43827; S. griseus: *Streptomyces griseus subs sp.* The number listed in the gene are as follows: 1. Hypothetical protein associated with desferrioxamine E biosynthesis; 2. Siderophore biosynthesis L-2,4-diaminobutyrate decarboxylase (DesA); 3. Siderophore biosynthesis protein, monooxygenase (DesB); 4. Siderophoresynthetase small component, acetyltransferase (DesC); 5. Siderophoresynthetase component, ligase (DesD); 6. Putative Desferrioxamine E transporter.

Here are some predictable chemical structures of secondary metabolites *Streptomyces* sp. GMR22 (Figures 2 and 3). The analysis was based on the genetic organization of secondary metabolites.

From figure 2, it can be seen that mostly the secondary metabolites of GMR22 dominated by type I-PKS. Polyketide is a large family of natural products found in bacteria, fungi, and plants. In addition, many clinically important drugs are included in this class, such as tetracycline, daunorubicin, erythromycin, rapamycin, and lovastatin. These compounds were biosynthesized from the precursor acyl-CoA by polyketide synthetase (PKS). Recent literature on polyketide biosynthesis suggests that polyketide synthases have much greater diversity in both mechanism and structure than the current type I, II and III paradigms. These examples serve as an inspiration for searching novel polyketide synthases to give new insights into polyketide biosynthesis and to provide new opportunities for combinatorial biosynthesis (Shen, 2003).

Secondary metabolites were also found in GMR22 with NRPS structure and hybrid NRPS with PKSs or others (Figure 3). However, NRPS structure is less compared to the class of PKSs compounds. Two secondary metabolites that encoded by type III of PKS were found in GMR22 as well. Type III of PKS is commonly found in fungi and plants. In general, type III of PKSs are involved in the biosynthesis of several lipid compounds and variety of secondary metabolite has some interesting characteristics, which does not belong to type III of PKS plants. In addition, many compounds produced by type III of PKS bacteria have significant biological functions in the interests of the pharmaceutical (Katsuyama and Ohnishi, 2012).

Secondary metabolite biosynthesis gene clusters involved in anticancer production

Based on the genome mining with antiSMASH 3.0, GMR22 has gene cluster for Desferrioxamine B which is 100% homologous with the most similar known cluster. Figure 4 shows a comparison of the gene cluster Desferrioxamine B GMR22 with Desferrioxamine B gene cluster of *Streptomyces coelicolor*A3(2) and *Streptomyces griseus sub sp. griseus.*

Desferrioxamine B (MIB BGC0000941_c1) is a highly conserved gene cluster in *Streptomyces* genera (Figure 4). From this Figure, it can be seen that both genes that encode enzymes, as well as location and sequence of genes for GMR22, were similar to the gene cluster in *Streptomyces coelicolor* A3(2). The gene cluster comprises a set of 4 main enzymes (2 to 5) with essential functions for the siderophore biosynthesis: the gene

encoding siderophore biosynthesis of L-2,4-diaminobutyrate decarboxylase DesA (2), siderophore biosynthesis protein, mono-oxygenase DesB (3), siderophoresynthetase small component, and acetyltransferase DesC (4), and siderophoresynthetase component ligase DesD (5). There are 2 addition enzymes that may be also involved, namely: a protein associated with Desferrioxamine E biosynthesis (1) and Desferrioxamine E transporter (6). In GMR22 mostly the size of Desferrioxamine genes are longer than Desferrioxamine genes in *Streptomyces coelicolor* A3(2). Functional analysis comparison between of gene cluster from GMR22 and strain A3(2) would be an interesting study to prove whether the activities are the same. Desferrioxamine B is an iron chelator, which has a variety of clinical applications for patients with iron overload in terms of improving the quality of life and overall survival. In addition, Desferrioxamin B in the experiment showed potential anticancer in the cells of colon cancer (Salis *et al*., 2014).

CONCLUSION

Based on the analysis of phylogeny relationship on the whole genome sequence showed that GMR22 was closed related with *Streptomyces bingchenggensis* BCW1 and clustered with strains of Streptomyces which have large genome size (9.5-12.7Mbp). AntiSMASH analysis revealed that *Streptomyces* sp. GMR22 harbored at least 63 gene clusters that encode the biosynthetic pathways of secondary metabolites. It was the highest number of gene clusters had been observed among the members of Streptomyces groups, with PKS was predicted as the major groups of the identified gene cluster products. Further studies on *Streptomyces* sp. GMR22 will provide more insights into natural product biosynthesis potential of related Streptomyces.

REFERENCES

Alimuddin AW., Widada J., Mustofa NR., 2010. An actinomycete producing anticandida isolated from cajuput rhizosphere: partial identification of isolates and amplification of PKS-I genes. *Indonesian J Biotechnol*, *15*, pp.1-8.

Bentley SD., Chater KF., Cerdeno-Tarraga, AM., Challis GL., Thomson NR., *et al.,* 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3 (2). *Nature*, *417*(6885), pp.141-147.

Bibb MJ., 2005. Regulation of secondary metabolism in streptomycetes. *Current Opinion in Microbiology*, 8, pp.208-215.

Blin K., Medema MH., Kazempour D., Fischbach MA., Breitling R., *et al.,* 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic acids research*, p.gkt449.

Distler J., Mansouri K., Mayer G., Stockmann M., Piepersberg W., 1992. Streptomycin biosynthesis and its regulation in Streptomycetes. *Gene*, *115*(1), pp.105-111.

Doroghazi JR., Albright JC., Goering AW., Ju KS., Haines RR., Tchalukov KA., *et al.,* 2014. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nature chemical biology*, *10*(11), pp.963-968.

El Hassan AM., Fahal AH., Ahmed AO., Ismail A., Veress, B., 2001. The immunopathology of actinomycetoma lesions caused by Streptomyces somaliensis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *95*(1), pp.89-92.

Euverink GJW., 1995. Aromatic amino acid biosynthesis in actinomycetes. *Groningen: University of Groningen*.

Katsuyama Y., Ohnishi Y., 2012. Type III polyketide synthases in microorganisms. *Methods Enzymol*, *515*, pp.359-377.

Li MH., Ung PM., Zajkowski J., Garneau-Tsodikova S. Sherman DH., 2009. Automated genome mining for natural products. *BMC bioinformatics*, *10*(1), p.1.

Medema MH., Blin K., Cimermancic P., de Jager V., Zakrzewski P., Fischbach MA., *et al.,* 2011. antiSMASH: rapid identification, annotation, and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids research*, *39*(suppl 2), pp.W339-W346.

Newman DJ., Cragg GM., and Snader KM., 2000. The influence of natural products upon drug discovery. *Natural product reports*, *17*(3), pp.215-234.

Nurjasmi R., Widada J., Ngadiman N., 2009. Diversity of Actinomycetes at Several Forest Types in Wanagama I Yogyakarta and Their Potency as a Producer of Antifungal Compound. *Indonesian Journal of Biotechnology*, *14*(2).

Ohnishi Y., Ishikawa J., Hara H., Suzuki H., Ikenoya M., *et al.,* 2008. Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350. *Journal of bacteriology*, *190*(11), pp.4050-4060.

Ōmura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., *et al.,* 2001. Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites. *Proceedings of the National Academy of Sciences*, *98*(21), pp.12215-12220.

Quintana ET., Wierzbicka K., Mackiewicz P., Osman A., Fahal, AH., *et al.,* 2008. Streptomyces sudanensis sp. nov., a new pathogen isolated from patients with actinomycetoma. *Antonie Van Leeuwenhoek*, *93*(3), pp.305-313.

Salis O., Bedir A., Kilinc V., Alacam H., Gulten S., Okuyucu A., 2014. The anticancer effects of desferrioxamine on human breast adenocarcinoma and hepatocellular carcinoma cells. *Cancer Biomarkers*, *14*(6), pp.419-426.

Shen B., 2003. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current opinion in chemical biology*, *7*(2), pp.285-295.

Tamura K., Stecher G., Peterson D., Filipski, A., Kumar S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, *30*(12), pp.2725-2729.

Waksman SA., Woodruff HB., 1940. Bacteriostatic and Bactericidal Substances Produced by a Soil Actinomyces. *Experimental Biology and Medicine*, *45*(2), pp.609-614.

Weber T., Blin K., Duddela S., Krug D., Kim HU., *et al.,* 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic acids research*, *43*(W1), pp.W237-W243.

Wang XJ., Yan YJ., Zhang B., An J., Wang JJ., *et al.* 2010. Genome sequence of the milbemycin-producing bacterium Streptomyces bingchenggensis. *Journal of bacteriology*, *192*(17), pp.4526-4527.

Xu Z., Hao B., 2009. CVTree update: a newly designed phylogenetic study platform using composition vectors and whole genomes. *Nucleic acids research*, *37*(suppl 2), pp.W174-W178.