Korean Journal of Microbiology (2018) Vol. 54, No. 4, pp. 448-452
DOI https://doi.org/10.7845/kjm.2018.8069

pISSN 0440-2413
eISSN 2383-9902

The complete genome sequence of a marine sponge-associated bacteria, *Bacillus safensis* KCTC 12796BP, which produces the anti-allergic compounds

Nguyen Phan Kieu Hanh†, Soo Hee Kim†, Geum Jin Kim, Hyukjae Choi*, and Doo Hyun Nam*

College of Pharmacy, Yeungnam University, Gyeongsan 38541, Republic of Korea

해양 해면체로부터 분리한 세균으로 항알러지성 물질을 생산하는 *Bacillus safensis* KCTC 12796BP의 유전체 해독

한 응엔 판 기우† · 김수희† · 김금진 · 최혁재* · 남두현*

영남대학교 약학대학

(Received September 17, 2018; Revised October 16, 2018; Accepted October 16, 2018)

The full genome sequence of *Bacillus safensis* KCTC 12796BP which had been isolated from the marine sponge in the seawater of Jeju Island, was determined by Pac-Bio next-generation sequencing system. A circular chromosome in the length of 3,935,874 bp was obtained in addition to a circular form of plasmid having 36,690 bp. The G + C content of chromosome was 41.4%, and that of plasmid was 37.3%. The number of deduced CDSs in the chromosome was 3,980, whereas 36 CDS regions were determined in a plasmid. Among the deduced CDSs in chromosome, 81 tRNA genes and 24 rRNA genes in addition to one tmRNA were allocated. More than 30 CDSs for sporulation, 16 CDSs for spore coat, and 20 CDSs for germination were also assigned in the chromosome. Several genes for capsular polysaccharide biosynthesis and for flagella biosynthesis and chemotaxis in addition to genes for osmotic tolerance through glycine-choline betaine pathway were also identified. Above all, the biosynthetic gene cluster for anti-allergic compounds seongsanamides were found among two non-ribosomal peptide synthetase (NRPS) gene clusters for secondary metabolites.

† These authors contributed equally to this work.
***For correspondence.** (H. Choi) E-mail: h5choi@yu.ac.kr;
Tel.: +82-53-810-2824; Fax: +82-53-810-4654/
(D.H. Nam) E-mail: dhnam@ynu.ac.kr;
Tel.: +82-53-810-2825; Fax: +82-53-810-4654

Keywords: *Bacillus safensis*, genome sequence, marine bacteria, non-ribosomal peptide synthetase, seongsanamide

Due to the limitation and difficulty in finding novel natural products from terrestrial biological sources, natural product chemists are starting to search novel biological active compounds from marine biological sources. Ocean covering almost 70% of earth surface possesses tremendous biological and chemical diversity. Based on the diverse oceanic environment, the exploitation of marine natural products having biologically functional activities led the emergence of some important pharmaceutical candidates (Lindequist, 2016; Romano *et al*., 2017).

Particularly, the marine sponges living in benthic environment has been attracted an interests from scientists, because those are good hosts to many microorganisms constituting up to 40–60% of its total biomass (Santos-Gandelman *et al*., 2014). The interaction of marine sponges and diverse microorganisms belonging to different phyla provides the rich source of novel functional biological molecules (Paul *et al*., 2011; Choi and Oh, 2015; Agrawal *et al*., 2016). In addition to marine sponges itself, the symbiotic bacteria on marine sponges has been also

explored as a source of marine active metabolites (Thomas *et al*., 2012; Choi and Oh, 2015; Bibi *et al*., 2017), and more than 5,000 natural compounds has been found from sponges and their associated microorganisms (Santos-Gandelman *et al*., 2014).

Bacillus safensis KCTC 12796BP was isolated from a marine sponge in the seawater in front of Seongsan-ri, Jeju Island, Korea, in the process of exploiting novel marine metabolites from ocean environment. Through the metabolite analysis of this strain by LC-Mass spectrometer, novel compounds belonging to a non-ribosomal lipopeptide group were identified, one of which showed the anti-allergic activity (Kim *et al*., 2018). In order to confirm the exact chemical structure of those compounds in this family, the full genome sequencing of the strain was attempted, as likely as the case of other marine sponge-associated *Bacillus* species including *Bacillus atrophaeus* producing algaecide bacillamide (Liu *et al*., 2016) and three *Bacillus* species producing antibacterial PE8-15 (van Zyl *et al*., 2012).

The genome sequence of *B. safensis* KCTC 12796BP was obtained by PacBio_10K yielding 242,134 reads; > 381.4-fold coverage with mean subread length of 8,443 bp. Total 3 linear contigs were firstly produced by *de novo* assembly using FALCON software (Pacific Biosciences). The large sequence of contig 1 was assumed to be a chromosome, and the contig 2 to be a plasmid. It was found that contig 3 was mistakenly assembled by connecting two reverse part sequences of contig 1. Because the *Bacillus* genomes are generally circular, the gene amplification was attempted using probes corresponding 5'-end region of contig 1 and 3'-end region of contig 3 reverse sequence. The gap (120 bp) between both ends of contig 1 and contig 3 were filled up to make finally a complete circular chromosome. Contig 2 was also confirmed by gene amplification using probes corresponding 5'-end and 3'-end regions, and one base missing was detected between both ends, finally to make a complete circular plasmid.

The genome sequences of this strain are 97.42% identical by average nucleotide identity (ANI) to the type genome of *B. safensis* FO-36b (BioProject PRJNA270528), with 86.2% coverage of the genome.

The complete sequence of chromosome of *B. safensis* is consisted of 3,935,874 bp with 41.4% G + C content, and the complete sequence of plasmid is composed of 36,690 bp with 37.3% G + C content (Table 1). Through open reading frames (ORFs) analysis using a prokka pipeline, 3,905 CDSs and 75 pseudogenes were found with 81 tRNA genes, 24 rRNA genes, and one tmRNA gene in a chromosome. In contrast 37 CDSs with one pseudogene were assigned in a plasmid.

The COG (cluster of orthrologous genes) analysis of the annotated CDSs in the chromosome by eggNOG pipeline (Jensen *et al*., 2008) showed that 16.7% of total was involved in information storage and processing, 21.8% participated in cellular processes, 29.6% encoded the proteins for cell metabolism, and 31.9% poorly characterized.

The most distinguishable feature of genera *Bacillus* is the formation of endospore. In the chromosome, more than 30 genes involved in the sporulation process (BSL056_11545~BSL056_11570, BSL056_12030~BSL056_12065, BSL056_13350~BSL056_13355, BSL056_14540~BSL056_14545, BSL056_14620~BSL056_14625, and so on) and 16 genes encoding spore coat protein (BSL056_05790~BSL056_05835, BSL056_18760~BSL056_18770, and so on) were identified. In addition more than 20 genes related to spore germination (BSL056_16065~BSL056_16070, BSL056_16595~BSL056_16605, BSL056_18005~BSL056_18015, and so on) were also confirmed (Eijlander *et al*., 2013).

Same as other *Bacillus* species, the genes for capsule biosynthesis by poly-γ-glutamate synthase including *capABC* (BSL056_17990~BSL056_18000) in addition to *capD* (BSL056_10680) were found with its positive regulator *acpB* (BSL056_04010) in the chromosome (Candela *et al*., 2005).

The osmotolerance of this strain in the saline environment

Table 1. Genome Feature of *Bacillus safensis* KCTC 12796BP

DNA	Bases (bp)	G + C content (%)	CDS	Pseudogene	tRNA	rRNA	tmRNA
chromosome	3,935,874	41.4	3,905	75	81	24	1
plasmid	36,690	37.3	36	1	0	0	0
Total	3,972,564	41.3	3,941	76	81	24	1

seems to be provided by three transport systems (Kappes *et al*., 1996). The glycine-betaine transport system conferred by *opuA* operon (BSL056_01645~BSL056_01655) and *opuD* gene (BSL056_14970), choline transport system by *opuB* operon (BSL056_20205~BSL056_20210), glycine betaine/carnitine/choline transport system by *opuC* operon (BSL056_16980~BSL056_16990) were denoted in the chromosome.

The main *fla-che* operon (BSL056_07875~BSL056_08030) for flagella biosynthesis and chemotaxis was located in the chromosome (Guttenplan *et al*., 2013). Additional two operons (BSL056_17645~BSL056_17695 and BSL056_18135~BSL056_18140) and *mot* operon (BSL056_06575~BSL056_06580) for flagella motor proteins were also found. Those gene products provide the bacterial motility of this strain.

Two non-ribosomal peptide synthetase (NRPS) gene clusters (BSL056_01920~BSL056_01930, BSL056_13660~BSL056_13665) in addition to two non-ribosomal peptide synthetase-polyketide synthase (NRPS-PKS) gene clusters (BSL056_01935~BSL056_01940, BSL056_03345~ BSL056_03365) probably involved in the secondary metabolism were found in the chromosome. The substrate specificity of each adenylation domain of NRPS genes based on the amino acid residues lining on the substrate binding pocket (Stachelhaus *et al*., 1999) was predicted using NRPSsp web server (Prieto *et al*., 2011) and NRPSpredictor2 web (Röttig *et al*., 2011) server. Based on the analysis of adenylation domains of each module (Table 2), one NRPS gene cluster (BSL056_01920~BSL056_01930) is proven to be involved in the biosynthesis of surfactin, and the other NRPS gene cluster (BSL056_13660~BSL056_13665) involved in the biosynthesis of seongsanamide group.

Marine *Bacillus* species produce versatile secondary metabolites including lipopeptides (Mondol *et al.*, 2013). Both surfactin and seongsanamides produced by this strain are belonging to lipopeptide family. Even though surfactin is a well-known representative lipopeptide, seongsanamide family (seongsanamides A~E) is a novel category of compounds isolated from the culture broth of *B. safensis* KCTC 12796BP (Kim *et al*., 2018). *In vitro* analysis of biological activity revealed that seongsanamide A among this group showed the most potent anti-allergic activity. The gene organization of adenylation domains of NRPSs as well as the location of epimerase domain is well matched with its chemical configuration with its stereospecificity (Fig. 1). The intramolecular linkage between two tyrosine residues in this compound might be mediated by cytochrome P450 encoded

Table 2. Amino acid residues lining the binding pocket of each adenylation domain in NRPS gene clusters found in the chromosome of *Bacillus safensis* KCTC 12796BP

Gene	Module	235	236	239	278	299	301	322	330	331	517	Predicted amino acid specificity
BSL056_01920~BSL056_01925 (Surfactin biosynthetic gene cluster)												
BSL056_01920	Module 1	D	A	K	D	L	G	V	V	D	K	L-Glu
	Module 2	D	A	F	M	L	G	M	I	F	K	L-Leu
	Module 3	D	A	W	F	L	G	N	V	V	K	D-Leu
BSL056_01925	Module 4	D	A	L	F	F	G	V	D	I	K	L-Val
	Module 5	D	L	T	K	V	G	H	I	G	K	L-Asp
	Module 6	D	A	W	F	L	G	N	V	V	K	D-Leu
BSL056_01930	Module 7	D	G	F	F	L	G	V	V	F	K	L-Leu
BSL056_13660~BSL056_13665 (Seongsanamide biosynthetic gene cluster)												
BSL056_13665	Module 1	D	A	W	F	L	G	H	V	V	K	L-Leu
	Module 2	D	L	F	N	N	A	L	T	Y	K	D-Ala
	Module 3	D	A	S	T	I	A	A	V	C	K	D-Tyr
	Module 4	D	F	W	N	I	G	M	V	H	K	L-Thr
	Module 5	D	A	W	F	L	G	H	V	V	K	D-Leu
BSL056_13660	Module 6	D	A	W	F	L	G	H	V	V	K	L-Leu
	Module 7	D	G	F	F	L	G	V	V	F	K	L-Ile
	Module 8	D	A	S	T	I	A	A	V	C	K	L-Tyr

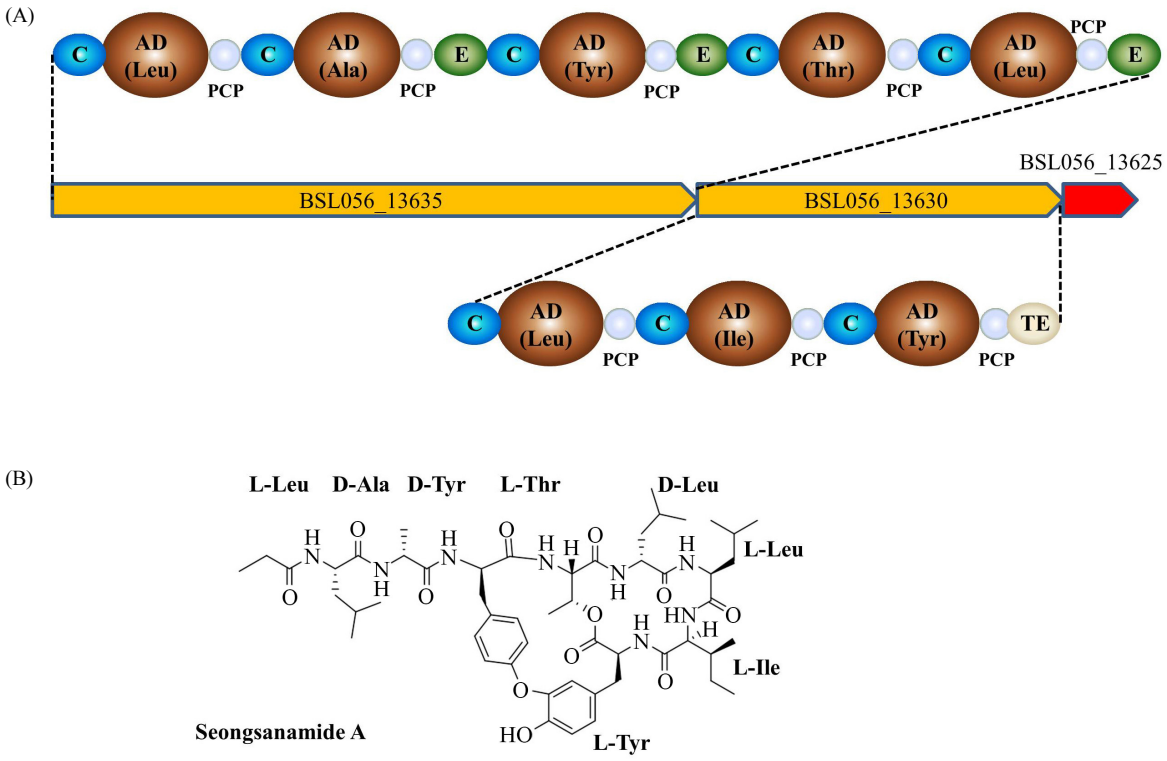

Fig. 1. The seongsanamide biosynthetic gene cluster of *Bacillus safensis* KCTC 12796BP. (A) The architecture and the organization of non-ribosomal peptide synthase domains in seongsanamide biosynthetic gene cluster. C, condensation domain; AD, adenylation domain; E, epimerase domain; PCP, peptidyl carrier protein domain; TE, thioesterase domain. (B) Chemical structure of seongsanamide A possessing an anti-allergic activity.

by adjacent gene (BSL056_13655).

Nucleotide sequence accession number

The complete genome sequences of *B. safensis* KCTC 12796BP has been deposited under the accession number of CP018197.1 (chromosome) and CP018198.1 (plasmid) of BioProject: PRJNA353573 at DDBJ/EMBL/GenBank.

적 요

제주도 성산리 앞 바다 속 해면체로부터 분리한 *Bacillus safensis* KCTC 12796BP의 유전체를 분석하였다. 그 결과 3,935,874 bp의 환형 염색체와 36,690 bp의 plasmid 염기 서열을 확인하였다. 염색체는 G + C 함량이 41.4%로 75개의 위유전자를 포함한 3,980개의 코딩 서열을, plasmid는 G + C 함량이 37.3%로 36개의 코딩 서열을 포함하고 있었다. 염색체 코딩 서열 중에는 81개의 tRNA 유전자, 24개 rRNA 유전자와 1개의 tmRNA 유전자가 있었다. 또한 포자 생성에 필요한 30개의 유전자, 포자피를 지령하는 16개의 유전자, 그리고 발아에 필요한 20개의 유전자도 발견되었다. 이외에 협막 다당체 생합성에 필요한 유전자와 편모 생합성 및 주화성에 필요한 유전자, 그리고 염 내성에 필요한 glycine-choline betaine 수송체에 관한 유전자도 존재하였다. 무엇보다도 항알러지활성을 보이는 이차대사산물 seongsanamide의 생합성을 지령하는 비리보좀성 펩타이드 합성효소 유전자를 확인할 수 있었다.

Acknowledgements

This work was kindly supported by National Research Fund of Korea (NRF-2017R1A2B4006110).

References

Agrawal S, Adholeya A, and Deshmukh SK. 2016. The pharmacological potential of non-ribosomal peptides from marine sponge and tunicates. *Front. Pharmacol*. **7**, 333.

Bibi F, Faheem M, Azhar EI, Yasir M, Alvi SA, Kamal MA, Ullah I,

and Nasser MI. 2017. Bacteria from marine sponges: A source of new drugs. *Curr. Drug Metab*. **18**, 11–15.

Candela T, Mock M, and Fouet A. 2005. CapE, a 47-amino-acid peptide, is necessary for *Bacillus anthracis* polyglutamate capsule synthesis. *J. Bacteriol.* **187**, 7765–7772.

Choi H and Oh DC. 2015. Considerations of the chemical biology of microbial natural products provide an effective drug discovery strategy. *Arch. Pharm. Res.* **38**, 1591–1605.

Eijlander RT, de Jong A, Krawczyk AO, Holsappel S, and Kuipers OP. 2013. SporeWeb: an interactive journey through the complete sporulation cycle of *Bacillus subtilis*. *Nucleic Acids Res*. **42**, D685–D691.

Guttenplan SB, Shaw S, and Kearns DB. 2013. The cell biology of peritrichous flagella in *Bacillus subtilis. Mol. Microbiol*. **87**, 211–229.

Jensen LJ, Julien P, Kuhn M, von Mering C, Muller J, Doerks T, and Bork P. 2008. eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res*. **36**, D250–D254.

Kappes RM, Kempf B, and Bremer E. 1996. Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD. *J. Bacteriol*. **178**, 5071–5079.

Kim GJ, Li X, Kim SH, Yang I, Hahn D, Chin J, Nam SJ, Nam JW, Kang H, Nam DH, *et al*. 2018. Seongsanamides A-D, bicyclic peptides isolated from *Bacillus safensis* L056 with their anti-allergic activities. *Org. Lett*. doi: 10.1021/acs.orglett.8b03293

Lindequist U. 2016. Marine-derived pharmaceuticals - challenges and opportunities. *Biomol. Ther*. **24**, 561–571.

Liu F, Sun W, Su F, Zhou K, and Li Z. 2012. Draft genome sequence of the sponge-associated strain *Bacillus atrophaeus* C89, a potential producer of marine drugs. *J. Bacteriol.* **194**, 4454.

Mondol MAM, Shin HJ, and Islam MT. 2013. Diversity of secondary metabolites from marine *Bacillus* species: chemistry and biological activity. *Mar. Drugs* **11**, 2846–2872.

Paul VJ, Ritson-Williams R, and Sharp K. 2011. Marine chemical ecology in benthic environments. *Nat. Prod. Rep*. **28**, 345–387.

Prieto C, Garcia-Estrada C, Lorenzana D, and Martin JF. 2011. NRPSsp: Non-ribosomal peptide synthase substrate predictor. *Bioinformatics* **28**, 426–427.

Romano G, Costantini M, Sansone C, Lauritano C, Ruocco N, and Ianora A. 2017. Marine microorganisms as a promising and sustainable source of bioactive molecules. *Mar. Environ. Res*. **128**, 58–69.

Röttig M, Medema MH, Blin K, Weber T, Rausch C, and Kohlbacher O. 2011. NRPSpredictor2 – a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res*. **39**, W362–W367.

Santos-Gandelman JF, Giambiagi-deMarval M, Oelemann WM, and Laport MS. 2014. Biotechnological potential of sponge-associated bacteria. *Curr. Pharm. Biotechnol*. **15**, 143–155.

Stachelhaus T, Mootz HD, and Marahiel M. 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol*. **6**, 493–505.

Thomas TRA, Kavlekar DP, and LokaBharathi PA. 2012. Marine drugs from sponge-microbe association--a review. *Mar. Drugs* **8**, 1417–1468.

van Zyl LJ, Matobole R, Biteghe FAN, Klein T, Kirby B, and Trindade M. 2016. Draft genome sequences of three *Bacillus* species from South African marine sponges. *Genome Announc*. **4**, e00143-16.