Gene Reports 5 (2016) 140–143

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Draft genome sequence of a sponge-derived *Brevibacillus* sp. TP-B0800, a producer of ulbactins with tumor cell migration inhibitory activity

Hisayuki Komaki [a,*], Akira Hosoyama [b], Natsuko Ichikawa [b], Yasuhiro Igarashi [c]

[a] Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292-0818, Japan
[b] NBRC, 2-49-10 Nishihara, Shibuya, Tokyo 151-0066, Japan
[c] Biotechnology Research Center and Department of Biotechnology, Toyama Prefectural University, 5180 Kurokawa, Imizu, Toyama 939-0398, Japan

ARTICLE INFO

Article history:
Received 30 August 2016
Received in revised form 30 September 2016
Accepted 7 October 2016
Available online 10 October 2016

Keywords:
Biosynthesis
Nonribosomal peptide
Secondary metabolite
Thiazoline

ABSTRACT

A sponge-derived *Brevibacillus* sp. TP-B0800 produces new polycyclic thiazoline derivatives designated ulbactins F and G which possess inhibitory activity against tumor cell migration. To elucidate the biosynthetic pathway of ulbactins, we read the genome sequence of this strain. The genome harbors six gene clusters related to nonribosomal peptides, one of which was identified to be responsible for ulbactin biosynthesis. We propose the putative ulbactin-biosynthetic pathway based on bioinformatics analysis. The genome information reported here also suggested potential of this strain to produce edeine, tyrocidin, and unidentified secondary metabolites.

1. Short introduction

Marine microorganisms are currently attracting attention as sources of new secondary metabolites with pharmacologically useful bioactivities (Lam, 2006; Jensen et al., 2005; Fenical and Jensen, 2006). In our search for novel bioactive compounds from marine bacteria, we isolated *Brevibacillus* sp. TP-B0800 from a sponge collected in Iwate, Japan, and discovered structurally unique polycyclic thiazoline derivatives ulbactins F and G from the culture broth as inhibitors of tumor cell migration (Igarashi et al., 2016). These compounds belong to a family of 2-(2-hydroxyphenyl)-2-thiazoline derivatives. Pyochelin, a siderophore produced by *Pseudomonas* (Quadri et al., 1999) and *Streptomyces* (Seipke et al., 2011), is well known as a member of this family. Ulbactins F and G are characterized by their unusual heterocyclic structure in which two thiazolidine rings are fused to construct a nitrogen- and sulfur-containing tricyclic ring system. Despite this unique structural feature, biosynthetic genes of ulbactins have not been reported to date. To identify the biosynthetic gene cluster of ulbactins F and G, we sequenced the genome of a sponge-derived *Brevibacillus* sp. TP-B0800.

2. Data description

Brevibacillus sp. TP-B0800 has been deposited into the NBRC culture collection under the accession number NBRC 110488. The whole genome of *Brevibacillus* sp. TP-B0800 monoisolate was read by using a combined strategy of shotgun sequencing with GS FLX+ (Roche; 79.4 Mb sequences, 12.6-fold coverage) and pair-end sequencing with HiSeq1000 (Illumina; 573.9 Mb sequences, 91.3-fold coverage). These reads were assembled using a Newbler v2.8 software with the default parameters, and subsequently each sequence gap in scaffolds was checked and re-assembled using sequence reads belonging to gap extremes by a GenoFinisher software (Ohtsubo et al., 2012). Branching contigs, one connected to multiple other contigs, were also examined and misassembled linkages were corrected. Consequently, 22 scaffolds and 3 contig sequences of >500 bp each were obtained as the final assembly. The total size of the assembly was 6,280,149 bp, with a G + C content of 47.3%. This whole genome shotgun sequencing project is deposited in DDBJ/EMBL/GenBank under the accession BDFB00000000. Coding sequences were predicted by Prodigal (Hyatt et al., 2010). Nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) gene clusters and their domain organizations were determined in the same manner previously reported (Komaki et al., 2014). Substrates of adenylation (A) domains were predicted by antiSMASH

Abbreviations: A, adenylation; ACP, acyl carrier protein; AMT, aminotransferase; AT, acyltransferase; C, condensation; CoL, CoA ligase; Cy, cyclization; E, epimerization; KS, ketosynthase; KR, ketoreductase; MT, methyltransferase; NBRC, Biological Resource Center, National Institute of Technology and Evaluation; NRPS, nonribosomal peptide synthetase; pk, polyketide unit; PKS, polyketide synthase; Sal, salicylic acid; T, thiolation; TE, thioesterase; X, unidentified amino acid; Y, starter molecule loaded by CoL.

* Corresponding author at: Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292-0818, Japan.
E-mail addresses: komaki-hisayuki@nite.go.jp (H. Komaki), hosoyama-akira@nite.go.jp (A. Hosoyama), ichikawa-natsuko@nite.go.jp (N. Ichikawa), yas@pu-toyama.ac.jp (Y. Igarashi).

http://dx.doi.org/10.1016/j.genrep.2016.10.006
2452-0144/

(Weber et al., 2015). BLASTP searches were conducted against the NCBI nr database.

The genome harbors three NRPS gene clusters and three hybrid PKS/NRPS gene clusters as shown in Table 1. BLAST search results suggested that NRPSs (01-668 to 01-666) and PKSs/NRPSs (01-357 to 01-366) encoded in scaffold00001 are responsible for biosyntheses of tyrocidin and edeine (Westman et al., 2013), respectively. However, products of the remaining four clusters are not speculated because these clusters do not show high similarity to clusters whose products are identified. Then, we predicted chemical backbones that the four clusters may synthesize, according to the collinearity rule of modular PKS/NRPS pathways (Fischbach and Walsh, 2006) and the prediction of A domain substrates. The biosynthetic pathway of pyochelin, a 2-(2-hydroxyphenyl)-2-thiazoline compound related to ulbactins, has been analyzed: pyochelin is synthesized by the condensation of salicylic acid and cysteines via an NRPS pathway (Quadri et al., 1999). Because the NRPS gene cluster in scaffold00004 (04-337 to 04-333) contains three cyclization (Cy) domains for thiazoline formation from cysteines and also its four A domains are predicted to incorporate salicylic acid (or dihydroxybenzoate) and three cysteines, which is in good accordance

Table 1
NRPSs and PKSs encoded in each gene cluster of *Brevibacillus* sp. TP-B0800.

Gene cluster	ORF (scaffold-orf no.)	Size (aa)	Domain organization	Closest homolog (definition [*origin*], accession no., % of similarity/identity)	Predicted product (or its backbone)
NRPS	01–668	1092	A(phe)/T/E	tyrocidine non-ribosomal peptide synthetase TycA [*Brevibacillus.* sp. BC25], WP_007721721, 96/97	Tyrocidin
	01–667	3589	C/A(pro)/T-C/A(phe)/T-C/A(phe)/T/E	tyrocidine non-ribosomal peptide synthetase TycB [*B. brevis*], WP_017248471, 99/99	
	01–666	6487	C/A(asn)/T-C/A(gln)/T-C/A(tyr)/T-C/A(val)/T-C/A(orn)/T-C/A(leu)/T-TE	tyrocidine non-ribosomal peptide synthetase TycC [*B. brevis*], WP_017248472, 99/99	
NRPS	11–63	2371	A(glu)/T-C/A(ser)/T/C	non-ribosomal peptide synthetase [*B. brevis*], WP_026043332, 99/99	Glu-Ser-Val-Val-Ser-X-Orn
	11–64	3188	A(val)/T-C/A(val)/T-C/A(ser)/T/C	non-ribosomal peptide synthetase [*B. brevis*], WP_026043333, 98/98	
	11–66	2920	C/A/T-C/A(orn)/T/E	non-ribosomal peptide synthetase [*B. brevis*], WP_017251707, 98/99	
NRPS	04–337	986	A(sal)-salycilate synthase	AMP-dependent synthetase [*B. brevis*], WP_017252055, 99/99	Sal-Cys-Cys-Cys (ulbactin)
	04–336	3024	T-Cy/A(cys)/MT/T-Cy/A(cys)/MT/T	non-ribosomal peptide synthetase [*B. brevis*], WP_017252054, 98/98	
	04–333	1601	Cy/A(cys)/T-TE	pyochelin synthetase F [*B. brevis*], WP_017252051, 96/96	
Hybrid PKS/NRPS*	01–357	1501	C/A(asn)/T/E	non-ribosomal peptide synthetase [*B. brevis*], WP_017249075, 99/99	Edeine
	01–359	2171	C/A/T/C-C	non-ribosomal peptide synthetase [*B. brevis*], WP_017249073, 99/99	
	01–361	1096	A/T-C	non-ribosomal peptide synthetase [*B. brevis*], WP_017249071, 99/99	
	01–362	1327	A/T-C	non-ribosomal peptide synthetase [*B. brevis*], WP_017249070, 99/99	
	01–363	895	A/T	non-ribosomal peptide synthetase [*B. brevis*], WP_017249069, 99/99	
	01–364	2884	KS/MT/KR/ACP-C/A(gly)/T	non-ribosomal peptide synthetase [*B. brevis*], WP_017249068, 98/98	
	01–365	442	AT	malonyl CoA-acyl carrier protein transacylase [*B. brevis*], WP_026043124, 99/99	
	01–366	1121	KS/ACP	polyketide synthase [*B. brevis*], WP_017249065, 99/99	
Hybrid PKS/NRPS	10–85	1100	AT/AT	malonyl CoA-ACP transacylase [*B. brevis*], WP_017250287, 98/98	Large polyketide containing Ser and X
	10–82	2740	ACP-KS/MT/ACP-KS/DH/KR	polyketide synthase [*B. brevis*], WP_017250284, 97/98	
	10–81	3976	MT/ACP-KS/DH/ACP-KS/DH/KR/ACP-KS/ACP	polyketide synthase [*B. brevis*], WP_017250283, 98/98	
	10–80	7167	C/A/T-KS/KR/ACP-KS/DH/KR/ACP-KS/DH/KR/ACP-KS/KR/ACP-KS	non-ribosomal peptide synthetase [*B. brevis*], WP_017250282, 97/98	
	10–79	3741	DH/ACP-KS/DH/KR/ACP-KS/DH/KR/ACP	polyketide synthase [*B. brevis*], WP_017250281, 97/98	
	10–78	2452	MT/ACP-KS/MT/ACP-KS/ACP	polyketide synthase [*B. brevis*], WP_017250280, 98/98	
	10–77	2003	C/A(ser)/MT/T/C	non-ribosomal peptide synthetase [*B. brevis*], WP_017250279, 98/98	
hybrid PKS/NRPS	01–304	5186	CoL/T-KS/ACP-AMT/C/T-C/A(orn)/T-C/A/T/E	non-ribosomal peptide synthetase [*B. brevis*], WP_017249129, 98/99	Y-pk-Orn-X-Phe
	01–305	1556	C/A(phe)/T-C	non-ribosomal peptide synthetase [*B. brevis*], WP_017249128, 97/98	

A, adenylation; ACP, acyl carrier protein; AMT, aminotransferase; AT, acyltransferase; C, condensation; CoL, CoA ligase; Cy, cyclization; E, epimerization; KS, ketosynthase; KR, ketoreductase; MT, methyltransferase; Sal, salicylic acid; T, thiolation; TE, thioesterase; X, unidentified amino acid; Y, starter molecule loaded by CoL; pk, polyketide unit. Substrates of A domains, predicted by antiSMASH (Weber et al., 2015), are shown in parentheses. *Each protein also showed over 92% identity to EdeP, N, L, K, J, I, H and G (Westman et al., 2013), respectively.