genomeA
Journals.ASM.org

Draft Genome Sequence of *Streptomyces* sp. Strain CT34, Isolated from a Ghanaian Soil Sample

Yin Zhai,[a] Bin Cheng,[b] Juan Hu,[b] Kwaku Kyeremeh,[c] Xiaoling Wang,[d] Marcel Jaspars,[d] Hai Deng,[d] Zi-Xin Deng,[a] Kui Hong[a]

Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan, China[a]; Beijing Genomics Institution, Shenzhen, China[b]; Department of Chemistry, University of Ghana, Accra, Ghana[c]; Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Aberdeen, United Kingdom[d]

Presented here is a draft genome sequence of *Streptomyces* sp. strain CT34, which produces a novel ribosomally synthesized and posttranslationally modified peptide (RiPP). Analysis of the deduced open reading frame set identified the putative RiPP biosynthesis gene cluster, as well as other secondary metabolite gene clusters.

Received 13 December 2014 **Accepted** 22 December 2014 **Published** 5 February 2015

Citation Zhai Y, Cheng B, Hu J, Kyeremeh K, Wang X, Jaspars M, Deng H, Deng Z-X, Hong K. 2015. Draft genome sequence of *Streptomyces* sp. strain CT34, isolated from a Ghanaian soil sample. Genome Announc 3(1):e01508-14. doi:10.1128/genomeA.01508-14.

Address correspondence to Kui Hong, kuihong31@whu.edu.cn.

The newly emerged ribosomally synthesized and posttranslationally modified peptides (RiPPs) form a very large group of important natural products with vast structure diversity (1). A new RiPP with a molecular weight of 3.6 kDa was isolated from Ghanaian *Streptomyces* sp. strain CT34. The structure of the peptide is quite different from the current ribosomal peptides, and it appears to fall into the linaridin class (2).

Genomic DNA of strain CT34 was extracted via the method of Pospiech and Neumann (3) and submitted for whole-genome sequencing by Beijing Genomics Institution (BGI). Two types of 500-bp and 6-Kb sequence libraries were prepared and sequenced by the Illumina HiSeq2000. A total of 24,826,204 raw reads were obtained with two libraries of 904 Mbp and 451 Mbp after deletion of the low-quality and adapter-containing reads (4). The resulting reads were assembled into 16 scaffolds and 310 contigs using SOAPdenovo (5). The total length of the assembly was 8,066,430 bp, with a genome coverage of 99.85% and an average GC content of approximately 71.39%.

Glimmer software (6, 7) was used to predict the genes of *Streptomyces* sp. CT34. The genome possesses 7,781 genes, with an average length of 875 bp. The predicted total gene length was 6,809,991 bp, which makes up 84.42% of the genome. Based on the analysis of the antiSMASH (8) program, 30 gene clusters were revealed for secondary metabolites biosynthesis, including 3 siderophores, 4 terpenes, 1 mixed nonribosomal peptide synthetase (NRPS)/polyketide synthase (T1-PKS), 1 mixed oligosaccharide/terpene, 1mixed PKS (T4-PKS)/PKS (T1-PKS), 1 mixed lantipeptide/PKS (T1-PKS), 3 PKSs (2 T2-PKSs, 1 T3-PKS), 2 lantipeptides, 3 NRPSs, 3 bacteriocins, 3 butyrolactones, 1 ectoine, and 4 unspecified clusters. A putative gene cluster related to the biosynthesis of the new linaridin RiPP was revealed with a length of 12,108 bp, including 10 open reading frames (ORFs) of catalytic and auxiliary functions and 1 ORF encoding the prepeptide. Further analysis of the novel peptide is under way.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number JSFP00000000. The version described in this paper is version JSFP01000000.

ACKNOWLEDGMENTS

This work was supported by the China "973" program (2012CB721001), the "863" Program (2012AA092201), the National Natural Science Foundation of China (31170467), and the EU FP7 project PharmaSea (312184). K.K., M.J., and H.D. thank the Royal Society–Leverhulme Trust Africa for the financial support (award AA090088) that enabled the sampling of sediments and subsequent isolation of this unique Ghanaian strain.

REFERENCES

1. **McIntosh JA, Donia MS, Schmidt EW.** 2009. Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. Nat Prod Rep **26:**537–559. http://dx.doi.org/10.1039/b714132g.
2. **Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJ, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süssmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenze SC, Willey JM, van der Donk WA.** 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Nat Prod Rep **30:**108–160. http://dx.doi.org/10.1039/c2np20085f.
3. **Pospiech A, Neumann B.** 1995. A versatile quick-prep of genomic DNA

from Gram-positive bacteria. Trends Genet **11:**217–218. http://dx.doi.org/10.1016/S0168-9525(00)89052-6.
4. **Kelley DR, Schatz MC, Salzberg SL.** 2010. Quake: quality-aware detection and correction of sequencing errors. Genome Biol **11:**R116. http://dx.doi.org/10.1186/gb-2010-11-11-r116.
5. **Li R, Li Y, Kristiansen K, Wang J.** 2008. SOAP: short oligonucleotide alignment program. BioInformatics **24:**713–714. http://dx.doi.org/10.1093/bioinformatics/btn025.
6. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res **27:**4636–4641. http://dx.doi.org/10.1093/nar/27.23.4636.
7. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. BioInformatics **23:**673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.
8. **Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T.** 2013. AntiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res **41:**W204–W212. http://dx.doi.org/10.1093/nar/gkt449.