

Genome Sequence of the ϵ -Poly-L-Lysine-Producing Strain *Streptomyces albulus* NK660, Isolated from Soil in Gutian, Fujian Province, China

Yanyan Gu,^a Chao Yang,^a Xiaomeng Wang,^a Weitao Geng,^a Yang Sun,^a Jun Feng,^a Yuanyuan Wang,^a Yufen Quan,^a You Che,^a Chi Zhang,^a Ting Gong,^a Wei Zhang,^a Weixia Gao,^a Zhenqiang Zuo,^a Cunjiang Song,^a Shufang Wang^b

Key Laboratory of Molecular Microbiology and Technology for Ministry of Education, Nankai University, Tianjin, China^a; State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin, China^b

Y.G. and C.Y. contributed equally to this work.

We determined the complete genome sequence of a soil bacterium, *Streptomyces albulus* NK660. It can produce ϵ -poly-L-lysine, which has antimicrobial activity against a spectrum of microorganisms. The genome of *S. albulus* NK660 contains a 9,360,281-bp linear chromosome and a 12,120-bp linear plasmid.

Received 12 May 2014 Accepted 22 May 2014 Published 12 June 2014

Citation Gu Y, Yang C, Wang X, Geng W, Sun Y, Feng J, Wang Y, Quan Y, Che Y, Zhang C, Gong T, Zhang W, Gao W, Zuo Z, Song C, Wang S. 2014. Genome sequence of the ϵ -poly-L-lysine-producing strain *Streptomyces albulus* NK660, isolated from soil in Gutian, Fujian Province, China. *Genome Announc.* 2(3):e00532-14. doi:10.1128/genomeA.00532-14.

Copyright © 2014 Gu et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Cunjiang Song, songcj@nankai.edu.cn, or Shufang Wang, wangshufang@nankai.edu.cn.

Streptomyces spp. are Gram-positive, soil-dwelling, filamentous bacteria, which are responsible for producing a wide variety of economically important natural secondary metabolites used in human and veterinary medicine (1). ϵ -Poly-L-lysine (ϵ -PL), which consists of 21 to 35 L-lysine residues with linkages between the α -carboxyl groups and ϵ -amino groups, is produced by bacteria belonging to the *Streptomycetaceae* family (2). ϵ -PL exhibits antimicrobial activity against a wide spectrum of microorganisms, including Gram-positive and Gram-negative bacteria; it also exhibits antiphage activity. In addition, due to its safety and biodegradability, ϵ -PL has been used as a food preservative in several countries, such as Japan, South Korea, and the United States (3,4).

We report here the genome sequence of the ϵ -PL-producing strain *Streptomyces albulus* NK660, which was isolated from Gutian, Fujian Province, China. The structure of ϵ -PL was determined by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) and gel permeation chromatography (GPC). These results showed that the polymerization degree of the ϵ -PL produced by strain NK660 ranges from 19 to 33 L-lysine monomers, and its molecular mass range is between 2,453 and 4,248 Da (5). The strain is currently maintained at the China General Microbiological Culture Collection Center (CGMCC no. 5392).

The genomic sequence of *S. albulus* NK660 was obtained by the assembly of three data sets, which were generated by the Roche 454 GS-FLX system, a Solexa analyzer (Illumina), and single-molecule real-time sequencing (PacBio). All of the paired reads were assembled using the SMRT Portal version 2.2.0. Annotation was done by merging the results obtained from Prodigal (6), RepeatMasker, tRNAscan-SE 1.21 (7), and RNAmmer 1.2 (8). Gene functional annotation was based on BLASTp with the Nonredundant (NR) databases, Kyoto Encyclopedia of Genes and Genomes

(KEGG) (9), Protein Families (PFAM) database (10), and Clusters of Orthologous Groups (COG) (11).

The genome of *S. albulus* NK660, with a G+C content of 72.32%, comprises a 9,360,281-bp linear chromosome and a 12,120-bp linear plasmid. The chromosome contains 7 rRNA operons, 70 tRNA genes, and 8,190 protein-coding genes (CDSs), holding about 85.32% of the whole genome. Among the 8,190 open reading frames (ORFs), 4,978 (60.78%) have a clear function, 573 (7.00%) encode products with high similarity to putative proteins, and the remaining 2,639 (32.22%) have no match in the protein database of the National Center for Biotechnology Information (NCBI).

Yamanaka et al. (12) clarified the catalytic mechanisms of ϵ -PL synthetase (Pls). Pls is a membrane protein with six transmembrane domains surrounding three tandem soluble domains. These tandem domains iteratively catalyzed L-lysine polymerization using free L-lysine polymer as the receptor and T-domain-bound L-lysine as the donor, directly yielding chains of different lengths. However, additional details on the regulation of ϵ -PL biosynthesis need further investigation. Therefore, this draft genome sequence will allow for developing a deeper understanding of its regulatory mechanism, thereby allowing us to construct an ϵ -PL overproducer for large-scale industrial production.

Nucleotide sequence accession numbers. The genome sequence of *S. albulus* NK660 has been deposited in GenBank, where the linear chromosome accession no. is CP007574 and the linear plasmid sequence accession no. is CP007575.

ACKNOWLEDGMENTS

We thank Lei Wang and Lu Feng in the genomic sequencing platform of the TEDA School of Biological Sciences and Biotechnology and Tianjin Key Laboratory of Microbial Functional Genomics for their support.

This work was supported by the National Key Basic Research Program

of China (973 program) grant 2012CB725204, the National High Technology Research and Development Program of China (863 program) grant 2012AA021505, the Natural Science Foundation of China grants 31070039, 31170030, 31300032, and 51073081, the Project of Tianjin, China (grants 13JCZDJC27800, 13JCYBJC24900, and 13JCQNJC09700), and the Ph.D. Candidate Research Innovation Fund of Nankai University.

REFERENCES

- Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, Yamashita A, Hattori M, Horinouchi S. 2008. Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J. Bacteriol.* **190**:4050–4060. <http://dx.doi.org/10.1128/JB.00204-08>.
- Hamano Y, Yoshida T, Kito M, Nakamori S, Nagasawa T, Takagi H. 2006. Biological function of the *pld* gene product that degrades ϵ -poly-L-lysine in *Streptomyces albulus*. *Appl. Microbiol. Biotechnol.* **72**:173–181. <http://dx.doi.org/10.1007/s00253-006-0396-4>.
- Hiraki J. 2000. ϵ -Polylysine, its development and utilization. *Fine Chem.* **29**:18–25.
- Yoshida T, Nagasawa T. 2003. Epsilon-poly-L-lysine: microbial production, biodegradation and application potential. *Appl. Microbiol. Biotechnol.* **62**:21–26. <http://dx.doi.org/10.1007/s00253-003-1312-9>.
- Geng WT, Yang C, Gu YY, Liu RH, Guo WB, Wang XM, Song CJ, Wang SF. 2014. Cloning of ϵ -poly-L-lysine (ϵ -PL) synthetase gene from a newly isolated ϵ -PL-producing *Streptomyces albulus* NK660 and its heterologous expression in *Streptomyces lividans*. *J. Microbiol. Biotechnol.* **7**:155–164. <http://dx.doi.org/10.1111/1751-7915.12108>.
- Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**: 955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
- Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. 2010. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**:355–360. <http://dx.doi.org/10.1093/nar/gkp896>.
- Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD. 2012. The Pfam protein families database. *Nucleic Acids Res.* **40**:D290–D301. <http://dx.doi.org/10.1093/nar/gkr1065>.
- Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**:33–36. <http://dx.doi.org/10.1093/nar/28.1.33>.
- Yamanaka K, Maruyama C, Takagi H, Hamano Y. 2008. ϵ -poly-L-lysine dispersity is controlled by a highly unusual nonribosomal peptide synthetase. *Nat. Chem. Biol.* **4**:766–772. <http://dx.doi.org/10.1038/nchembio.125>.