

Genome Sequence of *Saccharopolyspora erythraea* D, a Hyperproducer of Erythromycin

Wei-Bing Liu, Wen-Bang Yu, Shu-Hong Gao, Bang-Ce Ye

Laboratory of Biosystems and Microanalysis, State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, China

***Saccharopolyspora erythraea* is a Gram-positive bacterium that can produce antibiotics. However, this microorganism must often be genetically improved for higher production before it can be used in an industrial setting. Here, we report the whole-genome sequence of the industrial hyperproducer strong mutator *Saccharopolyspora erythraea* strain D.**

Received 12 August 2013 Accepted 21 August 2013 Published 19 September 2013

Citation Liu W-B, Yu W-B, Gao S-H, Ye B-C. 2013. Genome sequence of *Saccharopolyspora erythraea* D, a hyperproducer of erythromycin. *Genome Announc.* 1(5):e00718-13. doi:10.1128/genomeA.00718-13.

Copyright © 2013 Liu et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Bang-Ce Ye, bcye@ecust.edu.cn.

Saccharopolyspora erythraea is a soil dwelling, Gram-positive, GC-rich bacterium originally identified as *Streptomyces erythraeus*, but it was later assigned to the genus *Saccharopolyspora* (1). *S. erythraea*, as an excellent producer, produces a large number of secondary metabolites with biological activities, including some used as antibiotics. However, this microorganism must often be genetically improved for higher production before it can be used in an industrial setting. Historically, strain improvement has been empirically carried out by multiple rounds of random mutagenesis and screening (2). More recently, these rational strain improvement strategies benefit from the support of genomic, transcriptomic, proteomic, and metabolomic technologies (3–6). Since 2007, when the complete genome sequence of *S. erythraea* strain NRRL 23338 was completed (7), many researchers analyzed the difference between NRRL 23338 and mutants at the genome and transcriptome level using microarray and comparative genomics technologies, as well as other methods (8–12). The whole-genome sequencing of *S. erythraea* will help scientists to better understand the regulation of the erythromycin biosynthesis pathway and its relationship to other genes, which will in turn allow for more sophisticated strategies for improving the production of erythromycin. According to the Genomes OnLine Database (GOLD) (13), genome sequencing of the industrial hyperproducer strain of this organism has not been done yet, and there is only one completed and one draft sequence available for *S. erythraea*, specifically, that of the model strain *S. erythraea* NRRL 23338. Here, we report the whole-genome sequence of the industrial hyperproducer strong mutator *S. erythraea* D. The analysis of the sequence may help us to gain further insight into the overall genetic variation in *S. erythraea* and to understand the mechanism of its high production of erythromycin in order to investigate strategies for further improving the strain.

Genome sequencing was performed using the Roche GS Junior technology. Roche 454-GS Junior generated 167,320 reads (average length, ~519 nucleotides) that provided ~10× coverage of the entire genome. Several assembly procedures were applied, and manual editing was performed with the 454-GS Junior data. The genome was annotated using the IGS Annotation Engine (14), and

the NCBI Prokaryotic Genomes Annotation Pipeline (PGAP) 2.0 was employed for submission (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/).

The final assembly contains 384 contigs, with an N_{50} contig length of 37,346 bp; the largest contigs assembled measure 152,774 bp. The genome of *S. erythraea* D is composed of 79,374,806 bp, with an average G+C content of 71.23%. Compared with NRRL 23338, 304 high-confidence mutations were found, in which 100 nonsense mutations are located in a coding region, 141 sense mutations are located in a coding region, and there is one small fragment insertion/deletion, 59 mutations, and three small fragment insertions/deletions located in a noncoding region.

A more-detailed comparative analysis of this genome with those of other mutant and NRRL 23338 strains will provide further insight into the specific properties of this high producer of erythromycin.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. [AVCN00000000](https://www.ncbi.nlm.nih.gov/nuccore/AVCN00000000). The version described in this paper is the first version, [AVCN01000000](https://www.ncbi.nlm.nih.gov/nuccore/AVCN01000000).

ACKNOWLEDGMENTS

This study was supported by the Shanghai Postdoctoral Scientific Program, National Natural Science Foundation of China (NSFC) general projects grant 21276079, and the Fundamental Research Funds for the Central Universities.

We are grateful to Roche Diagnostics (Shanghai) Limited for assistance with the genome sequencing.

REFERENCES

1. Labeda DP. 1987. Transfer of the type strain of *Streptomyces erythraeus* (Waksman 1923) Waksman and Henrici 1948 to the genus *Saccharopolyspora* Lacey and Goodfellow 1975 as *Saccharopolyspora erythraea* sp. nov., and designation of a neotype strain for *Streptomyces erythraeus*. *Int. J. Syst. Bacteriol.* 37:19–22.
2. Adrio JL, Demain AL. 2006. Genetic improvement of processes yielding microbial products. *FEMS Microbiol. Rev.* 30:187–214.
3. Lee SY, Lee DY, Kim TY. 2005. Systems biotechnology for strain improvement. *Trends Biotechnol.* 23:349–358.

4. van der Werf MJ, Jellema RH, Hankemeier T. 2005. Microbial metabolomics: replacing trial-and-error by the unbiased selection and ranking of targets. *J. Ind. Microbiol. Biotechnol.* 32:234–252.
5. Stafford DE, Stephanopoulos G. 2001. Metabolic engineering as an integrating platform for strain development. *Curr. Opin. Microbiol.* 4:336–340.
6. Donadio S, Monciardini P, Sosio M. 2007. Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Nat. Prod. Rep.* 24:1073–1109.
7. Oliynyk M, Samborskyy M, Lester JB, Mironenko T, Scott N, Dickens S, Haydock SF, Leadlay PF. 2007. Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat. Biotechnol.* 25:447–453.
8. Peano C, Talà A, Corti G, Pasanisi D, Durante M, Mita G, Biccato S, De Bellis G, Alifano P. 2012. Comparative genomics and transcriptional profiles of *Saccharopolyspora erythraea* NRRL 2338 and a classically improved erythromycin over-producing strain. *Microb. Cell Fact.* 11:32. doi:10.1186/1475-2859-11-32.
9. Peano C, Biccato S, Corti G, Ferrari F, Rizzi E, Bonnal RJ, Bordoni R, Albertini A, Bernardi LR, Donadio S, De Bellis G. 2007. Complete gene expression profiling of *Saccharopolyspora erythraea* using GeneChip DNA microarrays. *Microb. Cell Fact.* 6:37. doi:10.1186/1475-2859-6-37.
10. Marcellin E, Mercer TR, Licon-Cassani C, Palfreyman RW, Dinger ME, Steen JA, Mattick JS, Nielsen LK. 2013. *Saccharopolyspora erythraea*'s genome is organised in high-order transcriptional regions mediated by targeted degradation at the metabolic switch. *BMC Genomics* 14:15. doi:10.1186/1471-2164-14-15.
11. Carata E, Peano C, Tredici SM, Ferrari F, Talà A, Corti G, Biccato S, De Bellis G, Alifano P. 2009. Phenotypes and gene expression profiles of *Saccharopolyspora erythraea* rifampicin-resistant (*rif*) mutants affected in erythromycin production. *Microb. Cell Fact.* 8:18. doi:10.1186/1475-2859-8-18.
12. Li YY, Chang X, Yu WB, Li H, Ye ZQ, Yu H, Liu BH, Zhang Y, Zhang SL, Ye BC, Li YX. 2013. Systems perspectives on erythromycin biosynthesis by comparative genomic and transcriptomic analyses of *S. erythraea* E3 and NRRL23338 strains. *BMC Genomics* 14:523. doi:10.1186/1471-2164-14-523.
13. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. 2010. The Genomes OnLine Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 38:D346–D354. doi:10.1093/nar/gkp848.
14. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. *Stand. Genomic Sci.* 4:244–251.