Complete Genome Sequences of *Streptomyces albus* Strain INA 01303

D. V. Poshvina, T. A. Efimenko, M. V. Demiankova, A. A. Glukhova, T. D. Ivankova, O. V. Efremenkova, V. S. Sadykova, A. S. Balkin, A. S. Vasilchenko

Laboratory of Antimicrobial Resistance, Institute of Environmental and Agricultural Biology (X-BIO), Tyumen State University, Tyumen, Russia

Gause Institute of New Antibiotics, Moscow, Russia

Laboratory of Biomedical Technologies, Institute for Cellular and Intracellular Symbiosis of the Ural Branch of the Russian Academy of Sciences, Orenburg, Russia

**ABSTRACT**

Here, we report the complete genome sequence of *Streptomyces albus* strain INA 01303, which was isolated from the Salt Lake Tambukan (Russia). The genome consists of a linear 6,840,896-nucleotide chromosome. This strain is predicted to produce a range of novel secondary metabolites with antibiotic activity.

The genus *Streptomyces*, a unique subgroup of actinomycete bacteria, is best known for its ability to produce amounts of bioactive secondary metabolites with antimicrobial, antiparasitic, antitumoral, and immunosuppressant activity (1, 2).

*Streptomyces albus* strain INA 01303 was isolated from the littoral zone of the bottom sediments of Lake Tambukan (Stavropol Krai, Russia). The strain was isolated using the standard method of sowing soil suspensions on dense nutrient medium organic agar number 2 Gause (glucose, 1.0 g/L; peptone, 0.5 g/L; tryptone, 0.3 g/L; NaCl, 0.5 g/L; agar, 2.0 g/L; tap water, pH 7.2 to 7.4) (3). To prepare soil suspensions, 100 mg of soil was placed in 10 mL of sterile water and shaken for 10 min, then the suspension was diluted in ratios of 1:1,000 and 1:10,000 and sown on the medium in petri dishes. The culture was incubated at a temperature of 28°C for 14 days. The overnight culture of *S. albus* strain INA 01303, which was grown on the liquid number 2 Gause medium, was transferred for DNA isolation.

Genomic DNA was isolated by the phenol-chloroform method. Briefly, 350 μL of Tris-salt buffer (100 mM/L Tris-HCl, 20 mM/L EDTA, and 750 mM/L NaCl; pH 8.0) and 50 μL of lysozyme solution (10 mg/mL) was added to 50 μL of sediment of the studied culture and incubated for 30 min at 37°C. After that step, a 10% SDS solution was added to the mixture to a final concentration of 1%, 4.0 μL of a proteinase K solution (10 mg/mL), and 1.0 μL RNase I (10 mg/mL) and incubated for 30 min at 60°C. After extraction with a mixture of phenol-chloroform-isooamyl alcohol (25:24:1, vol/vol/vol) and subsequent extraction with chloroform-isooamyl alcohol (24:1, vol/vol), DNA was precipitated from the aqueous phase in 3 volumes of absolute ethanol with the addition of 10 M ammonium acetate (1:10, vol/vol) at −20°C overnight. After centrifugation and double washing with 70% ethanol, the DNA was dried and dissolved in 30 μL of deionized water. DNA concentration was estimated using Qubit 4.0 fluorimeter (Thermo Fisher Scientific, Germany). DNA quality was evaluated using a Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific). Nanopore sequencing libraries were generated using the Native barcoding kit SQK-NBD112.24 (Oxford Nanopore Technologies, UK) protocol and sequenced on a SpotON flow cell vR9 (catalog number FLO-MIN106) using a MinION Mk1 device for 48 h in the Center of Shared Scientific Equipment “Microorganisms Persistence” of the Institute of Cellular and Intracellular Symbiosis of the Ural Branch of the Russian Academy of Sciences. The obtained Nanopore long reads were trimmed for minimal length (1,000 bp) and quality (q = 8) with Nanofilt (4). The total number of raw reads was 597,783 (957 Mb) with an N₅₀ value of 6,234 bp. Reads were basecalled using Guppy basecaller v5.0.13 in «fast» mode.
The de novo genome assembly of the isolate was generated by Nanopore sequencing data using Flye v2.8.3 (5), and the assembly was polished with two rounds of Racon v1.4.12 (6). The resulting assembly contigs were analyzed using the QUAST software v5.0.2 (7) to assess assembly quality, as well as using BUSCO v4.1.1 (8) to evaluate assembly completeness. Genome annotation was performed locally using Prokka v1.14.6 (9). The prediction of the specialized metabolites biosynthetic gene clusters (BGCs) was performed with antiSMASH 6.0 using strict detection (10). All mentioned software has been used with default parameters unless otherwise specified.

The genome of Streptomyces albus strain INA 01303 consists of 6,840,896 bp with 85-fold overall coverage and average GC content of 73%. In the genome of strain INA 01303, a total of 7,893 protein-coding genes were annotated. The genome annotation contained 85 tRNAs, 21 rRNAs (5S, 16S, and 23S), 71 miscellaneous RNAs (miscRNAs), and 1 transfer-messenger RNA (tmRNA). Genome annotation revealed a number of biosynthetic gene clusters (BGCs). In total, 17 regions were predicted, covering 4 polyketide synthase (PKS), 3 nonribosomal peptide synthetase (NRPS), 4 PKS/NRPS, 4 terpene, 1 siderophore, and 1 ectoine biosynthetic gene clusters.

**Data availability.** The assembled genome sequence of Streptomyces albus strain INA 01303 was deposited in GenBank under accession number CP102119.1. Nanopore raw reads have been deposited in the NCBI Sequence Read Archive under accession number SRR20718915 (BioProject number PRJNA863464 and BioSample number SAMN30035760).

**ACKNOWLEDGMENTS**

This study was supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of the Federal Scientific and Technical Program for the Development of Genetic Technologies for 2019 to 2027 (agreement number 075-15-2021-1345, unique identifier RF—193021X0012).

**REFERENCES**

1. Kong D, Wang X, Nie J, Niu G. 2019. Regulation of antibiotic production by signaling molecules in Streptomyces. Front Microbiol 10:2927. https://doi.org/10.3389/fmicb.2019.02927.
2. Jones SE, Elliot MA. 2017. Streptomyces exploration: competition, volatile communication and new bacterial behaviours. Trends Microbiol 25:522–531. https://doi.org/10.1016/j.tim.2017.02.001.
3. Gauze GF, Preobrazhenskaya TP, Terekhova LP, Maksimova TS. 1983. Opredelitel’ aktinomitsetov. Nauka, Moscow, Russia.
4. Coster WDe, D’Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.
5. Kolmogorov M, Yuan J, Lin Y, Pevzner P. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
6. Vašer R, Sovič I, Nagarajan N, Šikotic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
8. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
9. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
10. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335.