AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Streptomyces* sp. Strain RKCA744, Isolated from the Arauca River, Colombia

Christopher Cartmell,[a] Alyssa L. Grunwald,[c] Carolina Arango,[b,d] Noelle J. Duncan,[c] Bradley A. Haltli,[b,c] Luis E. Diaz,[d] Russell G. Kerr[a,b,c]

[a]Department of Chemistry, University of Prince Edward Island, Charlottetown, PEI, Canada
[b]Department of Biomedical Sciences, Atlantic Veterinary College, Charlottetown, PEI, Canada
[c]Nautilus Biosciences Croda, Charlottetown, PEI, Canada
[d]Universidad de La Sabana, Chia, Colombia

ABSTRACT *Streptomyces* sp. strain RKCA744 was isolated from sediment collected from the Arauca River, Colombia.

In continuing bioprospecting efforts to discover new bioactive natural products, sediment (ca. 400 g) was collected from the banks of the Arauca River located at 7.07263889°N, 70.82741667°W, at a water depth of 50 cm. The bacterium was isolated by dilution plating as described previously (1). BLASTN (v. 2.13.0) analysis of the full-length 16S rRNA gene sequence extracted from the genome sequence of *Streptomyces* sp. strain RKCA744 (1,527 bp) indicated that the closest match in the GenBank 16S rRNA gene sequence database was *Streptomyces hygroscopicus* (100% identity; accession number NR_043379.1) (2, 3). To further interrogate the biosynthetic potential of RKCA744 biosynthetic gene clusters (BGCs) within the genome of this organism, we obtained a draft whole-genome sequence.

Streptomyces sp. strain RKCA744 was cultured in Difco ISP2 broth at 30°C with shaking at 200 rpm. DNA was isolated using the DNeasy UltraClean microbial kit (Qiagen) following the manufacturer's protocol, and 16 μg of high-quality genomic DNA was obtained. Library preparation and Pacific Biosciences RS II DNA sequencing were performed by McGill University and the Génome Québec Innovation Centre using a standard PacBio protocol for sheared large-insert (20-kb) libraries (4) and a PacBio RS II single-molecule real-time (SMRT) platform with P4-C2 chemistry (Pacific Biosciences, USA). Sequencing was performed using one SMRT cell, which generated a total of 221,223 raw subreads (423,176,179 bp) with an average length of 1,913 bp and a GC content of 72.1%. The raw reads were filtered for quality using the SMRT Analysis software (version 2.3.0) (5) with a minimum subread length of 500, a minimum polymerase read quality of 0.75, and a minimum polymerase read length of 100. *De novo* genome assembly of the filtered reads was performed by the sequencing facility using the Hierarchical Genome Assembly Process (HGAP2)/Quiver protocol (5). Assembly resulted in the generation of 392 contigs with a total assembly length of 9,759,387 bp. The longest contig was 164,905 bp, the N_{50} was 40,528 bp, and sequencing coverage of the assembled genome was 43×.

The *Streptomyces* sp. RKCA744 draft genome was annotated using the Prokaryotic Genome Annotation Pipeline (6) yielding 7,860 predicted protein-coding sequences and 61 tRNAs (3 rRNA operons) from all contigs.

AntiSMASH (v. 6.1.1) was used to identify natural product BGCs. A total of 62 putative BGCs were identified and included putative siderophore (3), ectoine (2), butyrolactone (1), arylpolyene (1), homoserine lactone (1), terpene (6), nonribosomal peptide (NRP; 10), polyketide (PK; 24), and hybrid NRP/PK (8) BGCs. Several BGCs exhibited a

Editor J. Cameron Thrash, University of Southern California

Address correspondence to Russell G. Kerr, rkerr@upei.ca.

The authors declare no conflict of interest.

Received 10 August 2022
Accepted 15 September 2022
Published 26 September 2022

high percentage of similarity to several known gene clusters including isorenieratene (85% similarity) (7), ectoine (100% similarity) (8), ochronotic pigment (75% similarity), hopene (76% similarity) (9), pristinol (100% similarity) (10), desferrioxamine B (100% similarity) (11), and geosmin (100% similarity) (12).

Data availability. This whole-genome sequencing project has been deposited at GenBank under the accession number JAMYEP000000000. Raw sequencing data sets, including the longest filtered subreads, filtered subreads, and circular consensus sequences (CCSs), have been registered in the NCBI SRA database under the accession number SRP389029 with the BioProject accession number PRJNA853303.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Council of Canada (NSERC), the Canada Research Chairs Program, the Atlantic Canada Opportunities Agency, the Canada Foundation for Innovation, Nautilus Biosciences Croda, and the Jean-Louis Lévesque Foundation.

We also acknowledge McGill University and the Génome Québec Innovation Centre in Canada for PacBio RS II sequencing services.

REFERENCES

1. Arango C, Acosta-Gonzalez A, Parra-Giraldo CM, Sánchez-Quitian ZA, Kerr R, Díaz LE. 2018. Characterization of actinobacterial communities from Arauca River sediments (Colombia) reveals antimicrobial potential presented in low abundant isolates. Open Microbiol J 12:181–194. https://doi.org/10.2174/1874285801812010181.
2. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214. https://doi.org/10.1089/10665270050081478.
3. Baranasic D, Gacesa R, Starcevic A, Zucko J, Blazic M, Horvat M, Gjuracic K, Fujs S, Hranueli D, Kosec G, Cullum J, Petkovic H. 2013. Draft genome sequence of *Streptomyces rapamycinicus* strain NRRL 5491, the producer of the immunosuppressant rapamycin. Genome Announc 1:e00581-13. https://doi.org/10.1128/genomeA.00581-13.
4. Pacific Biosciences. 2022. Procedure checklist—preparing gDNA libraries using the SMRTbell express template preparation kit v2.0. Pacific Biosciences, Menlo Park, CA. https://www.pacb.com/wp-content/uploads/Procedure-Checklist-Preparing-gDNA-Libraries-Using-the-SMRTbell-Express-Template-Preparation-Kit-2.0.pdf.
5. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
6. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
7. Tsumaki T, Yamaguchi M, Tsumaki T. 1954. Pigments of marine animals. I. On the pigments of a sponge, Ilenierа Japonica. Nippon Kagaku Zasshi 75:297–300. https://doi.org/10.1246/nikkashi1948.75.297.
8. Peters P, Galinski EA, Trüper HG. 1990. The biosynthesis of ectoine. FEMS Microbiol Lett 71:157–162. https://doi.org/10.1111/j.1574-6968.1990.tb03815.x.
9. Hoshino T, Sato T. 2002. Squalene–hopene cyclase: catalytic mechanism and substrate recognition. Chem Commun (Camb) (4):291–301. https://doi.org/10.1039/b108995c.
10. Klapschinski TA, Rabe P, Dickschat JS. 2016. Pristinol, a sesquiterpene alcohol with an unusual skeleton from *Streptomyces pristinaespiralis*. Angew Chem Int Ed Engl 55:10141–10144. https://doi.org/10.1002/anie.201605425.
11. Codd R, Richardson-Sanchez T, Telfer TJ, Gotsbacher MP. 2018. Advances in the chemical biology of desferrioxamine B. ACS Chem Biol 13:11–25. https://doi.org/10.1021/acschembio.7b00851.
12. Izaguirre G, Taylor WD. 2004. A guide to geosmin- and MIB-producing cyanobacteria in the United States. Water Sci Technol 49:19–24. https://doi.org/10.2166/wst.2004.0524.