Microbiology Resource Announcements

| Applied and Industrial Microbiology | Announcement

Draft genome sequences of six bacterial strains isolated from Cannabis rhizosphere soil

Tyson Bookout,[1] Kira L. Goff,[1,2] Jeff Gauthier,[3] Roger C. Levesque,[3] Shawn Lewenza[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Although bacterial isolates from Cannabis flowers were reported and sequenced, few from its rhizosphere have been characterized. Here we report the draft genomes of six bacterial strains isolated from Cannabis rhizosphere soil samples. These sequences may shed light on plant-microbe interactions in the Cannabis rhizosphere at the molecular level.

KEYWORDS *Pseudomonas*, *Cannabis sativa*, rhizosphere, soil microbiology

The Cannabis plant has been cultivated for its nutritional value, durable fibers, medicinal use, and recreation for thousands of years (1). The ratio of cannabidiol (CBD) and tetrahydrocannabinol (THC) is important for both medicinal and recreational use, and the yields of these metabolites could be impacted by the microbial communities associated with the cannabis plant (2). Although bacterial isolates have been recovered and sequenced from Cannabis flowers (3), few have been reported from the Cannabis rhizosphere. Such bacterial isolates may have applications in biotechnology or agriculture.

Here we present the draft genome sequences of six bacterial isolates (Table 1) obtained from the rhizosphere of cultivated Cannabis plants (Anandia Labs; Vancouver, Canada). One gram of soil, recovered from Cannabis roots, was suspended in 10 mL of double-distilled water. The soil particulates were sedimented by centrifugation and the soil water extract spread-plated on *Pseudomonas* isolation agar (PIA, BD Difco). After incubation at 37°C, single colonies were isolated by three rounds of single-colony isolation with streaking, again on PIA agar.

Brain-heart infusion (BHI) agar (BD Difco) was used to culture bacterial isolates at room temperature (~25°C) for 48 h. Single colonies were suspended in 2 mL BHI broth and grown to an OD_{600} of 1.0. Cells were pelleted by centrifugation (10,000× *g* for 20 min at 4°C) and genomic DNA was extracted using a DNeasy Blood & Tissue kit (QIAGEN) according to the recommended protocol for Gram-negative bacteria. DNA purity was confirmed using a NanoDrop 2000 spectrophotometer (Thermo Fisher) and quantified using a Qubit dsDNA BR kit (Thermo Fisher). The G-tube fragmentation spin-column protocol (Covaris) was used to shear 1 µg genomic DNA into 400–700 bp fragments. These were then sequenced in paired ends (2 × 300 bp) on a MiSeq apparatus (Illumina) using the KAPA HyperPrep kit (Roche Sequencing Solutions) with Illumina TruSeq 3-PE adapters and the manufacturer's recommended protocol for bacterial WGS sequencing. Read quality was visualized with FastQC (version 0.11.8 (4)) and quality filtering/adapter removal was performed with TRIMMOMATIC (version 0.39–2 (5)), keeping sequences below Q15 quality score over a minimum read length of 36 bp. Trimmed Illumina reads were assembled with the A5 pipeline (version 20150522 (6)). Annotation was done *via* the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 6.4 (7). Taxonomy was assigned to draft genomes with whole-genome *in silico* digital DNA-DNA hybridization

Editor Irene L. G. Newton, Indiana University, Bloomington, Bloomington, Indiana, USA

Address correspondence to Shawn Lewenza, slewenza@athabascau.ca.

Tyson Bookout and Kira L. Goff are joint first authors. Author order was determined by who initiated the project.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 19 July 2023
Accepted 25 September 2023

Published 31 October 2023

TABLE 1 Main assembly metrics and number of features for each assembled genome[a]

Isolate ID	Organism	TYGS dDDH d4 (% and 95% CI)	Raw coverage (x)	Genome size (Mbp)	# Scaffolds	Scaffold N50 (bp)	Completeness	GC content (%)	# CDS	# tRNAs
SPPC 2814	*Pseudomonas sp.* CAN1	41.5 (39.0–44.0)[b]	33	6.81	99	638,296	100.0	65.7	6,221	116
SPPC 2815	*Pseudomonas citronellolis* CAN5	89.2 (85.9–91.9)	40	7.11	64	936,389	100.0	67.5	6,291	114
SPPC 2816	*Aeromonas caviae* CAN6	84.4 (81.7–86.8)	25	4.77	114	228,986	100.0	61.0	4,438	163
SPPC 2817	*Serratia bockelmannii* CAN8	87.8 (84.3–90.6)	27	5.23	55	1,162,636	99.7	59.4	4,953	136
SPPC 2818	*Pseudomonas rhodesiae* CAN13	80.3 (77.4–82.9)	44	5.91	43	898,063	99.7	60.3	5,414	105
SPPC 2819	*Variovorax sp.* CAN15	50.4 (47.8–53.0)[b]	30	6.92	193	401,761	100.0	67.2	6,602	98

[a]dDDH d4: proportion of sequence identity matched between draft genome and the most similar reference genome. Raw coverage: average number of reads covering each nucleotide position in the assembly. Completeness: an estimate calculated by CheckM, based on the presence of core genes, and the consistency of nucleotide usage statistics across scaffolds. CDS: coding DNA sequences. tRNAs: transfer RNA genes.
[b]Potential new species (dDDH with closest Type Strain Genome Server [TYGS] reference < 70%).

(dDDH) *via* the TYGS (8). Genome completeness was evaluated with CheckM (9). All software was run with default parameters unless otherwise specified.

Our results include two potential new species with dDDH d4 below 70% (Table 1). Further genomic analysis of these strains will facilitate a greater understanding of the molecular mechanisms of plant-microbe interactions in the growth of the Cannabis plant and the potential impact of secondary metabolite production.

ACKNOWLEDGMENTS

The authors wish to thank Jonathan Page for providing Cannabis plant soil samples, as well as Brian Boyle and the IBIS Genomics Platform for Illumina library preparation and sequencing.

This study was funded with an ARF grant from Athabasca University (S.L.). R.C. Levesque is funded by the Canadian Institutes of Health Research, Genome Canada, Genome Quebec, and Ontario Genomics.

The authors declare no conflict of interest. The supplier of Cannabis soil samples used in this study (Anandia Labs) had no influence on funding, experimental design, or interpretation of the results.

AUTHOR AFFILIATIONS

[1]Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Alberta, Canada
[2]Faculty of Science and Technology, Athabasca University, Athabasca, Alberta, Canada
[3]Institute of Integrative and Systems Biology, Laval University, Quebec, Canada

AUTHOR ORCIDs

Jeff Gauthier http://orcid.org/0000-0002-2523-2698
Shawn Lewenza http://orcid.org/0000-0002-7289-9911

FUNDING

Funder	Grant(s)	Author(s)
Athabasca University (AU)		Shawn Lewenza
Gouvernement du Canada | Canadian Institutes of Health Research (IRSC)		Roger C. Levesque

Funder	Grant(s)	Author(s)
Genome Canada (GC)		Roger C. Levesque

AUTHOR CONTRIBUTIONS

Tyson Bookout, Data curation, Formal analysis, Investigation, Writing – original draft | Kira L. Goff, Data curation, Formal analysis, Investigation, Writing – original draft | Jeff Gauthier, Methodology, Software, Validation, Writing – review and editing | Roger C. Levesque, Funding acquisition, Supervision, Visualization, Writing – review and editing | Shawn Lewenza, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

These draft genome sequences have been deposited in DDBJ/ENA/GenBank BioProject accession PRJNA971652. Draft genome accession numbers are as follows: JASMRU000000000, JASMRV000000000, JASMRW000000000, JASMRX000000000, JASMRY000000000, and JASMRZ000000000, and raw reads were, respectively, deposited in the Sequence Read Archive under accessions: SRX20708416, SRX20708417, SRX20708418, SRX20708419, SRX20708420, and SRX20708421.

REFERENCES

1. Bonini SA, Premoli M, Tambaro S, Kumar A, Maccarinelli G, Memo M, Mastinu AJ. 2018. *Cannabis sativa*: a comprehensive ethnopharmacological review of a medicinal plant with a long history. J Ethnopharmacol 227:300–315. https://doi.org/10.1016/j.jep.2018.09.004
2. Ahmed B, Hijri M. 2021. Potential impacts of soil microbiota manipulation on secondary metabolites production in cannabis. J Cannabis Res 3:25. https://doi.org/10.1186/s42238-021-00082-0
3. McKernan K, Helbert Y, Kane L, Houde N, Zhang L, McLaughlin S. 2021. Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools. F1000Res 10:624. https://doi.org/10.12688/-f1000research.53467.2
4. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
6. Tritt A, Eisen JA, Facciotti MT, Darling AE. 2012. An integrated pipeline for de Novo assembly of microbial genomes. PLoS ONE 7:e42304. https://doi.org/10.1371/journal.pone.0042304
7. 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
8. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10:2182. https://doi.org/10.1038/s41467-019-10210-3
9. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. Checkm: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114