AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Pseudomonas* sp. Strain CES, Containing the Entire Alkylxanthine Gene Cluster for Caffeine Breakdown

Ryan M. Summers,[a] Jonathan Shao,[b] Meredith B. Mock,[a] Chi Li Yu,[c] Fernando E. Vega[d]

[a]Chemical and Biological Engineering, The University of Alabama, Tuscaloosa, Alabama, USA
[b]U.S. Department of Agriculture, Agricultural Research Service, Northeast Area, Beltsville, Maryland, USA
[c]University of Iowa Carver College of Medicine, Proteomics Facility, Iowa City, Iowa, USA
[d]Sustainable Perennial Crops Laboratory, U.S. Department of Agriculture, Agricultural Research Service, Beltsville, Maryland, USA

ABSTRACT *Pseudomonas* strain CES was isolated from caffeine-enriched soil and found to possess the *N*-demethylation pathway for caffeine breakdown. We report the nucleotide sequence of the draft genome with 5,827,822 bp, 62.6% G+C content, and 5,427 protein-coding regions.

Various bacteria present in the soil, in water, or in the alimentary canal of insects have been shown to catabolize caffeine (1,3,7-trimethylxanthine) using *N*-demethylation or C-8-oxidative pathways (1–4). The bacterium *Pseudomonas* sp. strain CES was isolated from coffee-enriched soil through a caffeine-enrichment procedure as described previously (5, 6). Prior to isolation, coffee grounds were amended to the soil for 10 days. The strain was initially characterized as a *Pseudomonas* sp. based on 16S rRNA sequencing and fatty acid methyl ester (FAME) analysis (5) and can grow on up to 8 g/liter caffeine, which is over three times higher than the concentration used for most caffeine-degrading bacteria (2, 3).

Here, we report the draft genome sequence of *Pseudomonas* sp. CES, which uses the *N*-demethylation pathway for caffeine breakdown (6). Genomic DNA extraction and library preparation were performed at the Beijing Genomics Institute (BGI). DNA was extracted from CES cells using the BGI standard operating procedure based on the cetyltrimethylammonium bromide procedure (7), and 200- to 400-bp fragments were recovered after fragmentation by sonication. Fragments were end repaired, ligated to adaptors, amplified by PCR, heat denatured, and circularized with splint oligonucleotides. This final library was sequenced using next-generation sequencing on the Illumina HiSeq BGISEQ-500 platform, resulting in 177,377,152 total reads (88,688,576 paired-end reads) of 90 to 100 bp. Low-quality reads (50% Q < 20) and adapter sequences were identified and removed using SOAPnuke (8), and cleaned reads were assembled *de novo* using the IDBA-UD assembly program (9) using the parameters mink 33 and maxk 59. This analysis resulted in 3,143,938 paired-end reads and 28,542,398 single-end reads assembled into 198 contigs. Reads were mapped back to the assembled contigs for validation, and the N_{50} contig length was 71,749 nucleotides. The 198 contigs of 5,827,822 bp have an overall base composition of 62.6% G+C; the average coverage per base position was 553×. The assembled 198 contigs were run through Prodigal version 2.5.3 GeneFinder (10) and annotated using BLASTP against the RefSeq database of bacterial proteins downloaded from NCBI. The programs tRNAscan-SE version 1.4 (11) and RNAmmer version 1.2 (12) were used to predict regions encoding tRNAs and rRNAs. Default parameters were used for all software unless otherwise specified.

The chromosome has 5,427 protein-coding regions, 16S and 23S rRNA genes,

Citation Summers RM, Shao J, Mock MB, Yu CL, Vega FE. 2020. Draft genome sequence of *Pseudomonas* sp. strain CES, containing the entire alkylxanthine gene cluster for caffeine breakdown. Microbiol Resour Announc 9:e00484-20. https://doi.org/10.1128/MRA.00484-20.

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to Ryan M. Summers, rmsummers@eng.ua.edu.

Received 8 May 2020
Accepted 22 June 2020
Published 9 July 2020

TABLE 1 Genes in the Alx gene cluster for *Pseudomonas* sp. CES, *Pseudomonas putida* CBB5, and *Pseudomonas* sp. NCIM 5235

		CBB5		NCIM 5235		*E. coli* MG1655	
Gene	CES GenBank accession no.	GenBank accession no.	Nucleotide identity (%)	GenBank accession no.	Nucleotide identity (%)	GenBank accession no.	Nucleotide identity (%)
ndmA	KAF4561170	KKX58007	90.31	AWV66915	99.72		
ndmB	KAF4561167	KKX58004	90.14	AWV66912	98.03		
ndmC	KAF4561164	KKX58001	87.68	AWV66909	98.59		
ndmD	KAF4561163	KKX58000	80.78	AWV66908	96.26		
ndmE	KAF4561162	KKX57999	84.68	AWV66906	98.20		
cafR	KAF4561168	KKX58005	90.45	AWV66913	100		
cafT[a]	KAF4561173	KKX58009	64.13				
cafP	KAF4561166	KKX58003	87.25	AWV66911	96.42		
frmA	KAF4561172	KKX58008	92.7	AWV66917	99.19	AAC73459	73.17
frmB	KAF4561161	KKX57998	75.35	AWV66907	96.84	AAC73458	48.01
VOC[b]	KAF4561171			AWV66916	99.21		

[a] The NCIM 5235 gene cluster sequencing did not extend to include *cafT*.
[b] The CBB5 gene cluster does not contain a VOC gene.

and 29 tRNA genes. The putative *N*-demethylase genes *ndmABCDE*, required for *N*-demethylation of caffeine to xanthine, were identified in the genomic sequence with a BLAST search of the encoded protein using BLASTP (13) against NCBI's nonredundant protein database. The caffeine-degrading gene cluster in CES is similar to the alkylxanthine (Alx) gene clusters in *Pseudomonas putida* CBB5 and *Pseudomonas* sp. NCIM 5235 (Table 1); Clustal Omega analysis of the three cluster nucleotide sequences revealed that the CES cluster has identities of 71.92% and 91.65% to the clusters from CBB5 and NCIM 5235, respectively. The CES Alx gene cluster contains putative transcriptional regulators *cafT* and *cafR*, a putative caffeine permease, *cafP*, and homologs to *Escherichia coli frmAB* genes, which encode enzymes for formaldehyde metabolism (14–16), but differs from that of CBB5 due to the presence of a vicinal oxygen chelate (VOC) family protein upstream of the *ndmA* gene. This VOC gene was found in strain NCIM 5235, which can also survive at high levels of caffeine (17) and may be one reason that CES can grow above 2.5 g/liter caffeine. The CES draft genome could help elucidate the molecular evolution of *Pseudomonas* species and may reveal additional genes or regulatory mechanisms to enable bacteria to grow at higher caffeine concentrations.

Data availability. This genome was deposited in GenBank under the BioProject number PRJNA613486, BioSample accession number SAMN14432513, and SRA accession number SRP259409. This whole-genome shotgun project was deposited at DDBJ/ENA/GenBank under the accession number JAAXLL000000000. The version described in this paper is version JAAXLL010000000.

REFERENCES

1. Dash SS, Gummadi SN. 2006. Catabolic pathways and biotechnological applications of microbial caffeine degradation. Biotechnol Lett 28:1993–2002. https://doi.org/10.1007/s10529-006-9196-2.
2. Yu CL, Kale Y, Gopishetty S, Louie TM, Subramanian M. 2008. A novel caffeine dehydrogenase in *Pseudomonas* sp. strain CBB1 oxidizes caffeine to trimethyluric acid. J Bacteriol 190:772–776. https://doi.org/10.1128/JB.01390-07.
3. Yu CL, Louie TM, Summers R, Kale Y, Gopishetty S, Subramanian M. 2009. Two distinct pathways for metabolism of theophylline and caffeine are coexpressed in *Pseudomonas putida* CBB5. J Bacteriol 191:4624–4632. https://doi.org/10.1128/JB.00409-09.
4. Ceja-Navarro JA, Vega FE, Karaoz U, Hao Z, Jenkins S, Lim HC, Kosina P, Infante F, Northen TR, Brodie EL. 2015. Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. Nat Commun 6:7618. https://doi.org/10.1038/ncomms8618.
5. Yu CL, Kale Y, Gopishetty S, Subramanian M. 2007. Bacterial degradation of caffeine via demethylation and oxidation, abstr Q-023. *In* 107th general meeting of the American Society for Microbiology, Toronto, Canada. ASM Press, Washington, DC.
6. Yu CL, Summers RM, Li Y, Mohanty SK, Subramanian M, Pope RM. 2015. Rapid identification and quantitative validation of a caffeine-degrading pathway in *Pseudomonas* sp. CES. J Proteome Res 14:95–106. https://doi.org/10.1021/pr500751w.
7. Wilson K. 2001. Preparation of genomic DNA from bacteria. Curr Protoc Mol Biol 56:2.4.1–2.4.5. https://doi.org/10.1002/0471142727.mb0204s56.
8. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience 7:1–6. https://doi.org/10.1093/gigascience/gix120.
9. Peng Y, Leung HCM, Yiu SM, Chin F. 2012. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174.
10. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.
11. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.

12. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
13. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
14. Quandt EM, Summers RM, Subramanian MV, Barrick JE. 2015. Draft genome sequence of the bacterium *Pseudomonas putida* CBB5, which can utilize caffeine as a sole carbon and nitrogen source. Genome Announc 3:e00640-15. https://doi.org/10.1128/genomeA.00640-15.
15. Summers RM, Louie TM, Yu C-L, Gakhar L, Louie KC, Subramanian M. 2012. Novel, highly specific *N*-demethylases enable bacteria to live on caffeine and related purine alkaloids. J Bacteriol 194:2041–2049. https://doi.org/10.1128/JB.06637-11.
16. Summers RM, Seffernick JL, Quandt EM, Yu CL, Barrick JE, Subramanian MV. 2013. Caffeine junkie: an unprecedented glutathione *S*-transferase-dependent oxygenase required for caffeine degradation by *Pseudomonas putida* CBB5. J Bacteriol 195:3933–3939. https://doi.org/10.1128/JB.00585-13.
17. Retnadhas S, Gummadi SN. 2018. Identification and characterization of oxidoreductase component (NdmD) of methylxanthine oxygenase system in *Pseudomonas* sp. NCIM 5235. Appl Microbiol Biotechnol 102:7913–7926. https://doi.org/10.1007/s00253-018-9224-x.