




Genome Sequences of 42 Bacteria Isolated from *Sorghum bicolor* Roots

Dale A. Pelletier,^a Zhou Li,^b Tse-Yuan S. Lu,^a Li Zhang,^c Zhenbin Hu,^d Geoffrey P. Morris,^d Tijana Glavina del Rio,^e Dongyu Wang,^c Jin-Gui Chen,^a  Chongle Pan^{c,f}

^aBiosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

^bGenome Science and Technology, University of Tennessee, Knoxville, Tennessee, USA

^cDepartment of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA

^dDepartment of Agronomy, Kansas State University, Manhattan, Kansas, USA

^eDOE Joint Genome Institute, Berkeley, California, USA

^fSchool of Computer Science, University of Oklahoma, Norman, Oklahoma, USA

ABSTRACT Forty-two bacterial strains were isolated from root samples of *Sorghum bicolor*. The strains spanned 17 genera, including *Dechloromonas*, *Duganella*, *Dyella*, *Flavobacterium*, *Herbaspirillum*, *Lutibacter*, *Mucilaginibacter*, *Novosphingobium*, *Paraburkholderia*, *Pedobacter*, *Pleomorphomonas*, *Rhizobacter*, *Rhizobium*, *Rhizomicrobium*, *Rugamonas*, *Variovorax*, and *Xanthobacter*. Their whole-genome sequences revealed diverse metabolic processes, including biological nitrogen fixation, in sorghum root microbiota.

Sorghum bicolor is a C4 flowering plant in the grass family Poaceae and the tribe Andropogoneae. Owing to its high tolerance to heat and drought, sorghum is especially important for agriculture in semiarid lands. Root microbiota may play important roles in the growth of sorghum in such environments (1, 2). Here, we report the isolation and whole-genome sequences of 42 bacterial strains in the sorghum root microbiota.

Natural soils in the root regions of actively growing sorghum plants were collected from a sorghum field site (latitude, 39.1392778; longitude, -96.6322763) at Kansas State University (Manhattan, KS). The natural soils were mixed with perlite and used to grow sorghum from seeds for 8 weeks in a greenhouse at Oak Ridge National Laboratory (Oak Ridge, TN). Sorghum roots were harvested for the isolation of bacteria. Root tissue was washed with sterile water to remove loosely adherent soil. Fine roots (<2 mm) were excised, macerated in 10 ml of MgSO₄ (10 mM), and serially diluted onto nitrogen-free minimal salt agar medium with sucrose, glucose, or succinate as the sole carbon source (Table 1). Cultures were isolated through three rounds of restreaking onto the aforementioned medium at 28°C aerobically or anaerobically in GasPak jars with CO₂ generator pouches (BD BBL). We isolated 27 strains under anaerobic conditions and 15 strains under aerobic conditions (Table 1). Initial taxonomy was assigned using Sanger sequencing of the 16S rRNA amplicon with primers 27F and 1492R (3), and then a BLAST search of the NCBI nonredundant/nucleotide database was performed with a cutoff value of >90% identity across a minimum of 1,300 nucleotides (4, 5).

Single colonies were subsequently inoculated into Reasoner's 2A (R2A) liquid medium and grown for 2 days at 30°C with shaking, and cells were pelleted and stored at -80°C until they were processed for DNA extraction (6). DNA samples from all bacterial cultures were prepared utilizing a Qiagen DNeasy kit according to the manufacturer's instructions. The whole-genome sequences were generated at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) utilizing Illumina technology (7). Illumina

Citation Pelletier DA, Li Z, Lu T-YS, Zhang L, Hu Z, Morris GP, Glavina del Rio T, Wang D, Chen J-G, Pan C. 2020. Genome sequences of 42 bacteria isolated from *Sorghum bicolor* roots. *Microbiol Resour Announc* 9:e00736-20. <https://doi.org/10.1128/MRA.00736-20>.

Editor David A. Baltus, University of Arizona
This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.
Address correspondence to Jin-Gui Chen, chenj@ornl.gov, or Chongle Pan, cpan@ou.edu.

Received 25 June 2020

Accepted 17 August 2020

Published 10 September 2020

TABLE 1 Isolation and whole-genome sequencing of 42 bacterial strains from *Sorghum bicolor* roots

Genome name	Order	GenBank accession no.	SRA accession no.	No. of reads ($\times 10^6$)	Genome size (kbp)	No. of scaffolds	No. of genes	GC content (%)	M_{50} (kbp)	Genome coverage (X)	Encodes nitrogenase	Incubation conditions	Isolation medium
<i>Dechloromonas</i> sp. SG708	Rhodocyclales	JAAZQM0000000000	SRR10150408	33.7	4,539	27	4,444	61	462	330	Yes	Anaerobic	Sucrose
<i>Duganella</i> sp. SG902	Burkholderiales	JABMJK0000000000	SRR10150426	38.8	6,806	47	6,084	65	416	219	No	Anaerobic	Glucose
<i>Dyella</i> sp. SG562	Xanthomonadales	JAAOY5000000000	SRR10150377	29.1	5,263	31	4,477	67	319	285	No	Aerobic	Sucrose
<i>Dyella</i> sp. SG609	Xanthomonadales	JAAZQJ0000000000	SRR10150400	40.2	5,127	27	4,472	68	316	292	No	Aerobic	Sucrose
<i>Flavobacterium</i> sp. SG815	Flavobacteriales	JABBHH0000000000	SRR10150419	15.1	3,997	16	3,702	37	357	375	No	Anaerobic	Succinate
<i>Herbaspirillum</i> sp. SG826	Burkholderiales	JAAOYU0000000000	SRR10150418	18.9	5,474	16	5,019	63	612	274	No	Anaerobic	Succinate
<i>Lutibacter</i> sp. SG786	Flavobacteriales	JAAOYX0000000000	SRR10150417	34.7	4,379	4	3,962	65	932	342	No	Anaerobic	Sucrose
<i>Mucilagibacter</i> sp. SG538B	Sphingobacteriales	JABMJJ0000000000	SRR10150382	39.5	7,151	23	6,188	42	152	208	No	Aerobic	Sucrose
<i>Mucilagibacter</i> sp. SG564	Sphingobacteriales	JABGNA0000000000	SRR10150380	41.4	6,829	38	5,942	42	494	220	No	Aerobic	Sucrose
<i>Mucilagibacter</i> sp. SG486	Sphingobacteriales	JABGBY0000000000	SRR10150311	34.8	4,764	35	4,392	65	431	315	No	Aerobic	Sucrose
<i>Novosphingobium</i> sp. SG619	Sphingomonadales	JABGBZ0000000000	SRR10150403	42.2	4,803	37	4,440	65	431	312	No	Aerobic	Sucrose
<i>Novosphingobium</i> sp. SG707	Sphingomonadales	JAAZQL0000000000	SRR10150409	29.4	5,642	41	5,177	63	565	266	No	Anaerobic	Sucrose
<i>Novosphingobium</i> sp. SG718	Sphingomonadales	JABGCA0000000000	SRR10150412	15.8	4,670	21	4,199	65	2,582	321	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG720	Sphingomonadales	JAAZQN0000000000	SRR10150410	30.8	5,171	24	4,675	65	614	290	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG726	Sphingomonadales	JAAZQO000000000	SRR10150413	35.3	4,666	42	4,276	65	420	321	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG751A	Sphingomonadales	JABGCB0000000000	SRR10150416	27.3	5,730	75	5,280	62	249	262	Yes	Anaerobic	Sucrose
<i>Novosphingobium</i> sp. SG754	Sphingomonadales	JABGCC0000000000	SRR10150423	37.9	4,893	30	4,432	65	634	306	No	Anaerobic	Sucrose
<i>Novosphingobium</i> sp. SG774	Sphingomonadales	JABBBH0000000000	SRR10150414	36.8	4,785	43	4,323	65	599	313	No	Anaerobic	Succinate
<i>Novosphingobium</i> sp. SG781	Sphingomonadales	JABBBH0000000000	SRR10150422	44.7	4,786	43	4,313	65	599	313	No	Anaerobic	Succinate
<i>Novosphingobium</i> sp. SG788	Sphingomonadales	JABBBH0000000000	SRR10150415	36.1	4,620	41	4,185	65	247	324	No	Anaerobic	Sucrose
<i>Novosphingobium</i> sp. SG900	Sphingomonadales	JABBBH0000000000	SRR10150424	36.0	5,642	44	5,176	63	354	266	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG906	Sphingomonadales	JABBBH0000000000	SRR10150444	32.5	4,613	27	4,180	65	386	325	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG916	Sphingomonadales	JABBBH0000000000	SRR10150554	23.0	5,454	21	4,923	65	728	275	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG917	Sphingomonadales	JABCNE0000000000	SRR10150552	20.7	5,454	22	4,901	65	728	275	No	Anaerobic	Glucose
<i>Novosphingobium</i> sp. SG919	Sphingomonadales	JABCNG0000000000	SRR10150556	44.5	5,454	21	4,899	65	728	275	No	Anaerobic	Glucose
<i>Paraburkholderia</i> sp. SG903	Burkholderiales	JABGCD0000000000	SRR10150427	49.4	7,348	25	6,759	62	778	204	No	Anaerobic	Glucose
<i>Pedobacter</i> sp. SG908	Sphingobacteriales	JAAOY0000000000	SRR10150445	27.6	6,000	5	5,204	38	1,543	250	No	Anaerobic	Glucose
<i>Pedobacter</i> sp. SG918	Sphingobacteriales	JABCNF0000000000	SRR10150555	27.0	5,831	6	5,051	38	1,392	257	No	Anaerobic	Glucose
<i>Pleomorphomonas</i> sp. SG524	Rhizobiales	JAAOYR0000000000	SRR10150307	37.4	5,024	18	4,805	63	649	298	Yes	Aerobic	Sucrose
<i>Pleomorphomonas</i> sp. SG806	Rhizobiales	JAAOYT0000000000	SRR10150421	41.5	5,024	18	4,807	63	649	298	Yes	Aerobic	Sucrose
<i>Rhizobacter</i> sp. SG490	Burkholderiales	JABBBH0000000000	SRR10150312	46.2	6,725	36	6,216	69	384	223	No	Aerobic	Sucrose
<i>Rhizobacter</i> sp. SG703	Burkholderiales	JAAZQK0000000000	SRR10150401	41.3	6,510	29	5,947	69	468	230	No	Anaerobic	Sucrose
<i>Rhizobium</i> sp. SG526	Rhizobiales	JABBBH0000000000	SRR10150310	34.3	7,014	45	6,810	59	441	214	No	Aerobic	Sucrose
<i>Rhizobium</i> sp. SG570	Rhizobiales	JAAZQJ0000000000	SRR10150379	44.4	7,548	58	7,357	59	328	199	No	Aerobic	Sucrose
<i>Rhizobium</i> sp. SG741	Rhizobiales	JAAZQP0000000000	SRR10150425	57.7	7,224	39	7,029	59	442	208	No	Anaerobic	Succinate
<i>Rhizomicrobium</i> sp. SG733	Micropales	JABGCE0000000000	SRR10150411	28.1	4,588	15	4,292	59	1,159	327	Yes	Anaerobic	Succinate
<i>Rugomonas</i> sp. SG757	Pseudomonadales	JABMJJ0000000000	SRR10150420	46.8	7,005	24	6,262	64	717	212	No	Anaerobic	Sucrose
<i>Variovorax</i> sp. SG517	Burkholderiales	JABMJG0000000000	SRR10150306	60.2	7,423	31	7,024	67	611	200	No	Aerobic	Sucrose
<i>Variovorax</i> sp. SG533	Burkholderiales	JABMJH0000000000	SRR10150381	53.3	6,922	20	6,484	66	963	214	No	Aerobic	Sucrose
<i>Xanthobacter</i> sp. SG518	Rhizobiales	JAAZQG0000000000	SRR10150309	35.6	5,578	21	5,275	67	528	269	Yes	Aerobic	Sucrose
<i>Xanthobacter</i> sp. SG563	Rhizobiales	JAAZQH0000000000	SRR10150378	37.0	5,578	21	5,273	67	528	269	Yes	Aerobic	Sucrose
<i>Xanthobacter</i> sp. SG618	Rhizobiales	JABBBH0000000000	SRR10150402	24.7	5,205	15	4,877	68	536	288	Yes	Aerobic	Sucrose

300-bp-fragment plate-based DNA libraries were constructed and sequenced using NovaSeq XP v1 reagent kits with an S4 flow cell, following a 2×150 -bp indexed run recipe. BBDuk in BBTools v38.44 (8) was used to remove contaminants, to trim reads that contained adapter sequences, and to remove reads containing ≥ 1 N bases or having a length of ≤ 51 bp or $\leq 33\%$ of the full read length. BBSplit in BBTools v38.44 was used to remove reads that mapped to masked human, cat, dog, or mouse references at 93% identity. The parameter settings were `rqc.filter2.sh rna=f minlength=51 phix=t trimfragadapter=t maxns=1 mlf=0.33 removehuman=t removedog=t removecat=t removemouse=t khist=t removemicrobes=t sketch kapa=t clumpify=t barcodefilter=f trimpolyg=5`. Artifact-filtered and normalized Illumina reads were assembled using SPAdes v3.12.0 (9) with the following parameters: `phred-offset 33, cov cutoff auto, t 16, m 64, careful, and k 25,55,95`. Scaffolds were discarded if the length was < 1 kbp (BBTools `reformat.sh, minlength`). Genome annotation was performed using the standard IMG Annotation Pipeline v5.0.1 (10).

The taxonomy, isolation conditions, assembly statistics, and annotation results are shown in Table 1 for all 42 isolates. The genomes were sequenced at a 273-fold average coverage depth and assembled with an average N_{50} value of 663 kbp and an average total length of 5.6 Mbp. Eight of the genomes contain nitrogenase gene clusters (*nifHDK*) for biological nitrogen fixation. These genomes will aid in investigations of plant-microbe interactions with sorghum (11).

Data availability. The 42 whole-genome sequences have been deposited in the NCBI GenBank and SRA databases, and their accession numbers are listed in Table 1.

ACKNOWLEDGMENTS

Sample collection, strain isolation, and processing were supported by the Laboratory-Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. DOE under contract DE-AC05-00OR22725. The sequencing data were generated under JGI Community Science Program proposal 503207. The work conducted by the U.S. DOE JGI, a DOE Office of Science User Facility, is supported under contract DE-AC02-05CH11231.

REFERENCES

- Hartman K, Tringe SG. 2019. Interactions between plants and soil shaping the root microbiome under abiotic stress. *Biochem J* 476:2705–2724. <https://doi.org/10.1042/BCJ20180615>.
- Kuramae EE, Derksen S, Schlemper TR, Dimitrov MR, Costa OY, da Silveira AP. 2020. Sorghum growth promotion by *Paraburkholderia tropica* and *Herbaspirillum frisingense*: putative mechanisms revealed by genomics and metagenomics. *Microorganisms* 8:725. <https://doi.org/10.3390/microorganisms8050725>.
- Utturkar SM, Cude WN, Robeson MS, Yang ZK, Klingeman DM, Land ML, Allman SL, Lu T-Y, Brown SD, Schadt CW, Podar M, Doktycz MJ, Pelletier DA. 2016. Enrichment of root endophytic bacteria from *Populus deltoides* and single-cell-genomics analysis. *Appl Environ Microbiol* 82:5698–5708. <https://doi.org/10.1128/AEM.01285-16>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Claesson MJ, Wang Q, O'Sullivan O, Greene-Diniz R, Cole JR, Ross RP, O'Toole PW. 2010. Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Res* 38:e200. <https://doi.org/10.1093/nar/gkq873>.
- Reasoner DJ, Geldreich E. 1985. A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol* 49:1–7. <https://doi.org/10.1128/AEM.49.1.1-7.1985>.
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438. <https://doi.org/10.1517/14622416.5.4.433>.
- Bushnell B. 2016. BBTools. <https://jgi.doe.gov/data-and-tools/bbtools>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen I-MA, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v. 4). *Stand Genomic Sci* 10:86. <https://doi.org/10.1186/s40793-015-0077-y>.
- Li Z, Yao Q, Guo X, Crits-Christoph A, Mayes MA, Hervey WJ, Lebeis SL, Banfield JF, Hurst GB, Hettich RL, Pan C. 2019. Genome-resolved proteomic stable isotope probing of soil microbial communities using $^{13}\text{C}_2$ and ^{13}C -methanol. *Front Microbiol* 10:2706. <https://doi.org/10.3389/fmicb.2019.02706>.