

Genome-sequenced bacterial collection from sorghum epicuticular wax

Marco E. Mechan-Llontop,¹ John Mullet,^{1,2} Ashley Shade³

AUTHOR AFFILIATIONS See affiliation list on p. 5.

ABSTRACT A collection of 44 isolates isolated from the epicuticular wax of stems of energy sorghum is available at the Great Lakes Bioenergy Researcher Center, Michigan State University, MI, USA. We enriched bacteria with putative plant-beneficial phenotypes and include information on their phenotypic diversity, taxonomy, and whole-genome sequences.

KEYWORDS plant microbiome, host microbiota, genomes, bioenergy, agroecosystems, phyllosphere, exudates, dessication

The heat- and drought-tolerant and high-biomass crop bioenergy sorghum (*Sorghum bicolor* L. Moench) accumulates high levels of epicuticular wax on the outer surface of the cuticle to protect plants against various stresses (1–3). It has been shown that the sorghum epicuticular wax harbors specialized microbiome members with putative plant-beneficial capabilities (4). As a resource for genomic and functional diversity of bacteria associated with phyllosphere exudates, we present a collection of 44 bacterial strains isolated from sorghum epicuticular wax. These strains represent 13 families and include 17 genera with putative plant-beneficial capabilities, including nitrogen fixation, phosphate solubilization, resistance to terpenoids, use of methanol as carbon source, and tolerance to desiccation, based on the culture media in which the strains were originally isolated.

Bacterial strains were isolated from the epicuticular wax of stems of bioenergy sorghum hybrid TX08001 grown at the Texas A&M University Research Farm in College Station, Texas (30°55'5.55" N, 96°43'64.6" W). On 2 September 2020, the fifth and sixth fully elongated stem internodes were destructively collected using razor blades into sterile whirl-pack bags. Within 2 hours of collection, the samples were transported to the laboratory on ice and stored at –80°C. On 8 October 2020, samples were delivered to Michigan State University on dry ice and stored at –80°C until processing. Epicuticular wax was carefully scraped from stems with sterile razor blades and collected into sterile 1.5 mL Eppendorf tubes. Wax from different plants was pooled, and 100 mg was resuspended in 1 mL of sterile water. Resuspended wax was serially diluted from 10^{–1} to 10^{–4}. To capture a diversity of bacteria from the wax, we used a variety of cultivation media (Table 1). To select for anaerobic bacteria, plates were placed in anaerobic jars containing three bags of anaerobic gas generator (Mitsubishi AnaeroPack System). All plates were incubated at 25°C and 37°C for up to 14 days. Isolated colonies were transferred on new plates with the same medium as used for isolation to confirm purity. Bacteria were grown overnight on R2A broth at 28°C with shaking at 200 rpm, and glycerol stock (25% vol/vol) of pure bacteria isolates was stored at –80°C. Prior to performing whole-genome sequencing, we performed full-length 16S rRNA gene sequencing, with universal primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-TACGGTTACCTTGTTACGACTT-3') (5), of each isolate using the Sanger protocol to link

Editor David A. Baltrus, University of Arizona, Tucson, Arizona, USA

Address correspondence to Ashley Shade, shade.ashley@gmail.com.

The authors declare no conflict of interest.

See the funding table on p. 5.

For a companion article on this topic, see <https://doi.org/10.1128/mra.00468-23>.

Received 13 June 2023

Accepted 29 September 2023

Published 1 November 2023

Copyright © 2023 Mechan-Llontop et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

TABLE 1 - Taxonomy, colony phenotype, and genome characteristics of 44 bacterial isolates from sorghum epicuticular wax, as described in this study^a

Isolate ID	Bacterial species	Initial isolation media	Temperature, oxygen availability	Colony phenotype	Sequencing platform	SRA accession number	GenBank assembly	Genome length (bp)	N° contigs	Coverage	GC %	Genome completeness
SORGH_AS_0981	<i>Aeromicrobium</i> sp.	50R2A	37°C, aerobic	Small, white	PacBio	SRP437207	JAVIZD000000000	3,670,284	1	186.7×	25%	98.96%
SORGH_AS_0908	<i>Pseudoxanthomonas winnipegensis</i>	50R2A	37°C, aerobic	Yellow	PacBio	SRP437170	JAUTBB000000000	4,331,500	1	82.7×	69.46%	99.53%
SORGH_AS_0909	<i>Chryseobacterium</i> sp.	50R2A	37°C, aerobic	Yellow	PacBio	SRP437159	JAVIYY000000000	4,186,940	1	137.1×	40.82%	100%
SORGH_AS_0912	<i>Acinetobacter baylyi</i>	50R2A	37°C, aerobic	White, EPS production	PacBio	SRP437185	JAVIZH000000000	3,896,949	3	187.4×	40.23%	99.95%
SORGH_AS_0913	<i>Nocardioideus zeae</i>	50R2A	37°C, aerobic	White, EPS production	PacBio	SRP437161	JAVIYZ000000000	4,248,407	1	90.2×	73.63%	98.19%
SORGH_AS_0919	<i>Microbacterium paludicola</i>	50R2A	37°C, aerobic	Pink	PacBio	SRP437238	JAVIZA000000000	3,351,387	1	196.8×	70.22%	99.24%
SORGH_AS_0974	<i>Agrobacterium larrymoorei</i>	50R2A	37°C, anaerobic	White	PacBio	SRP437176	JAVIZC000000000	5,263,363	3	120.0×	58.10%	98.30%
SORGH_AS_0950	<i>Sphingomonas</i> sp.	Gauze N-1	37°C, aerobic	Orange	PacBio	SRP437227	JAUTAE000000000	4,389,989	3	73.3×	67.10%	99.14%
SORGH_AS_1204	<i>Microbacterium</i> sp.	Gauze N-1	25°C, aerobic	Orange	PacBio	SRP437180	JAUTBD000000000	3,706,564	1	132.3×	70.10%	97.77%
SORGH_AS_1206	<i>Microbacterium arborescens</i>	Gauze N-1	25°C, aerobic	Orange	PacBio	SRP437216	JAUTBE000000000	3,449,815	1	88.6×	70.11%	99.24%
SORGH_AS_1207	<i>Microbacterium trichothecis-nolyticum</i>	Gauze N-1	25°C, aerobic	Orange	PacBio	SRP437150	JAUTBF000000000	3,735,270	1	241.2×	69.92%	97.24%
SORGH_AS_0870	<i>Sphingomonas</i> sp.	Jensen	25°C, aerobic	Orange	PacBio	SRP437151	JAVIYW000000000	4,339,992	2	98.4×	67.06%	99.42%
SORGH_AS_1048	<i>Chryseobacterium</i> sp.	Jensen	37°C, aerobic	Small, orange	PacBio	SRP437220	JAUTAK000000000	4,189,072	1	195.6×	40.80%	99.51%
SORGH_AS_0862	<i>Microbacterium</i> sp.	Jensen	25°C, aerobic	Yellow, EPS production	PacBio	SRP437146	JAUTAY000000000	3,504,869	1	174.0×	69.45%	98.27%
SORGH_AS_0892	<i>Sphingobacterium zeae</i>	KB	25°C, aerobic	Yellow	PacBio	SRP437175	JAUTBA000000000	5,604,819	1	136.4×	40.08%	99.84%
SORGH_AS_0893	<i>Acinetobacter baylyi</i>	KB	25°C, aerobic	White	PacBio	SRP437224	JAVIYX000000000	3,884,413	3	187.0×	40.23%	99.72%
SORGH_AS_0962	<i>Asaia bogorensis</i>	KB	25°C, aerobic	Small, Pink	PacBio	SRP437145	JAVIZB000000000	3,299,069	2	112.5×	59.82%	99.71%
SORGH_AS_0969	<i>Microbacterium</i> sp.	KB	25°C, aerobic	White	PacBio	SRP437218	JAUTAG000000000	3,939,785	1	159.6×	69.65%	98.53%
SORGH_AS_1094	<i>Pantoea dispersa</i>	M9 + 0.4% sucrose	25°C, aerobic	Yellow	PacBio	SRP437173	JAVJAC000000000	4,990,160	3	189.4×	57.57%	99.71%
SORGH_AS_1126	<i>Agrobacterium larrymoorei</i>	M9 + 0.4% sucrose	37°C, aerobic	White	PacBio	SRP437223	JAUTBL000000000	4,882,137	2	142.0×	58.10%	97.93%
SORGH_AS_0906	<i>Klebsiella variicola</i>	M9 nitrogen free, carbon free	37°C, aerobic	White, EPS production	PacBio	SRP437210	JAVJAI000000000	5,683,212	2	135.4×	57.57%	99.85%
SORGH_AS_1173	<i>Klebsiella</i> sp.	M9 nitrogen free, carbon free	37°C, anaerobic	White, EPS production	PacBio	SRP437157	JAVJAD000000000	5,683,290	2	194.8×	57.57%	99.78%
SORGH_AS_1175	<i>Chryseobacterium</i> sp.	M9 nitrogen free, carbon free	37°C, anaerobic	Orange	PacBio	SRP437237	JAVIZF000000000	4,186,971	1	211.3×	40.82%	99.88%
SORGH_AS_0842	<i>Klebsiella variicola</i>	M9 nitrogen free, carbon free	25°C, anaerobic	White, EPS production	PacBio	SRP437149	JAVJAB000000000	5,683,317	2	89.3×	57.57%	99.93%

(Continued on next page)

TABLE 1 - Taxonomy, colony phenotype, and genome characteristics of 44 bacterial isolates from sorghum epicuticular wax, as described in this study^a (Continued)

Isolate ID	Bacterial species	Initial isolation media	Temperature, oxygen availability	Colony phenotype	Sequencing platform	SRA accession number	GenBank assembly	Genome length (bp)	N° contigs	Coverage	GC %	Genome completeness
SORGH_AS_0878	<i>Klebsiella variicola</i>	M9 nitrogen free	25°C, aerobic carbon free	White, EPS production	PacBio	SRP437219	JAVIZZ000000000	5,683,197	2	188.5×	57.57%	99.55%
SORGH_AS_0879	<i>Sphingomonas</i> sp.	M9 nitrogen free	25°C, aerobic carbon free	Yellow	PacBio	SRP437133	JAUTBJ000000000	4,326,161	2	185.1×	66.43%	99.47%
SORGH_AS_0885	<i>Nocardioideae zeae</i>	M9 nitrogen free	25°C, aerobic carbon free	Yellow	Illumina	SRP437487	JAVIZJ000000000	4,252,839	27	353.6×	73.44%	99.85%
SORGH_AS_0887	<i>Acinetobacter baylyi</i>	M9 nitrogen free	25°C, aerobic carbon free	White, EPS production	PacBio	SRP437183	JAUTBK000000000	3,884,591	3	186.4×	40.23%	99.86%
SORGH_AS_0888	<i>Microbacterium</i> sp.	MMS methanol	25°C, aerobic	Yellow, EPS production	PacBio	SRP437142	JAUTAZ000000000	3,638,259	1	237.8×	70.61%	98.91%
SORGH_AS_1077	<i>Microbacterium testaceum</i>	R2A	37°C, aerobic	Yellow	PacBio	SRP437132	JAUTBC000000000	3,939,786	1	181.0×	69.65%	98.23%
SORGH_AS_1083	<i>Rhizobium pusense</i>	R2A	37°C, aerobic	White	PacBio	SRP437166	JAVIZE000000000	5,455,244	3	63.6×	59.10%	99.57%
SORGH_AS_0826	<i>Klebsiella</i> sp.	R2A	25°C, anaerobic	White, EPS production	PacBio	SRP437221	JAVJAA000000000	5,685,432	2	136.5×	57.58%	100%
SORGH_AS_0956	<i>Pseudoxanthomonas winnipegensis</i>	R2A + 0.1% caryophyllene	37°C, aerobic	Yellow	PacBio	SRP437156	JAUTAF000000000	4,331,831	1	208.2×	69.46%	98.27%
SORGH_AS_0961	<i>Pseudacidovorax intermedius</i>	R2A + 0.1% caryophyllene	37°C, aerobic	White	PacBio	SRP437178	JAVLUK000000000	4,187,089	1	241.5×	40.82%	98.77%
SORGH_AS_0993	<i>Brevundimonas</i> sp.	R2A + 0.1% caryophyllene	37°C, aerobic	White	PacBio	SRP437129	JAUTAH000000000	2,986,043	1	210.6×	67.51%	99.12%
SORGH_AS_0834	<i>Klebsiella variicola</i>	R2A + 0.1% linalool	25°C, aerobic	White, EPS production	PacBio	SRP437138	JAVJAG000000000	5,683,096	2	163.7×	57.57%	99.64%
SORGH_AS_1014	<i>Xanthomonas sacchari</i>	TSA	37°C, aerobic	Yellow	PacBio	SRP437228	JAUTAI000000000	4,997,627	2	173.7×	69.65%	98.97%
SORGH_AS_1015	<i>Roseomonas cervicalis</i>	TSA	37°C, aerobic	Pink	PacBio	SRP437141	JAUTAJ000000000	4,889,768	5	111.2×	72.05%	98.81%
SORGH_AS_1025	<i>Klebsiella</i> sp.	TSA	37°C, aerobic	White, EPS production	PacBio	SRP437181	JAVIZY000000000	5,684,408	2	172.8×	57.58%	99.70%
SORGH_AS_1064	<i>Chryseobacterium camelliae</i>	TSA	25°C, aerobic	Orange	PacBio	SRP437167	JAUTAL000000000	4,187,050	1	212.0×	40.82%	99.84%
SORGH_AS_1065	<i>Siphonobacter</i> sp.	TSA	25°C, aerobic	Orange	PacBio	SRP437172	JAUTAM000000000	6,150,989	10	125.1×	44.65%	99.70%
SORGH_AS_1067	<i>Nocardioideae zeae</i>	TSA	25°C, aerobic	Yellow	PacBio	SRP43703	JAUTAN000000000	4,380,404	1	164.8×	73.45%	99.22%
SORGH_AS_1070	<i>Klebsiella variicola</i>	TSA	25°C, aerobic	White, EPS production	PacBio	SRP437160	JAVJAH000000000	5,682,994	2	95.0×	57.57%	100%
SORGH_AS_0997	<i>Pseudoxanthomonas</i> sp.	TSA	25°C, aerobic	Yellow	PacBio	SRP437165	JAVIZI000000000	4,331,362	1	171.5×	69.46%	99.09%

^aSRA is the sequence read archive of the U.S. National Institutes of Health.

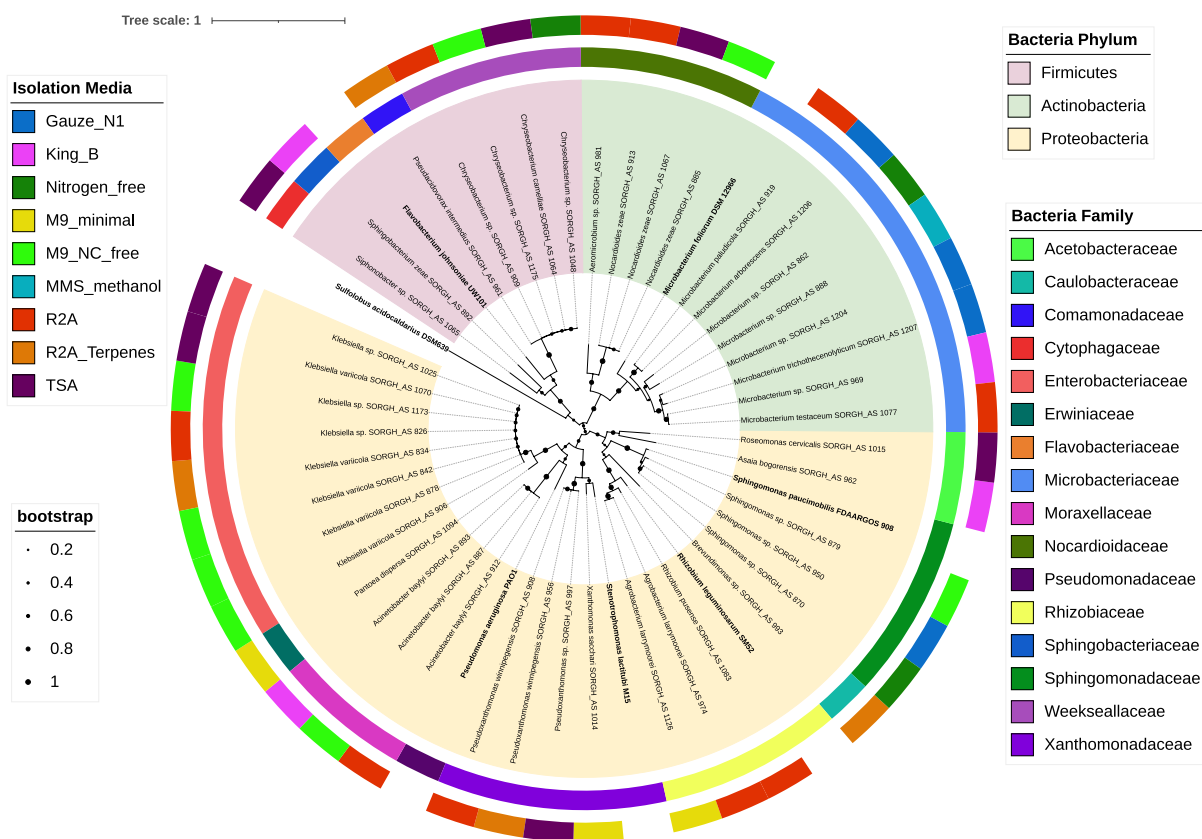


FIG 1 Phylogenetic diversity of the bacterial collection cultivated from sorghum epicuticular wax. The archaea *Sulfolobus acidocaldarius* DSM 639 (ASM2847236v1) was included as an outgroup. The bacterial reference strains *Pseudomonas aeruginosa* PA01 (ASM676v1), *Microbacterium foliorum* DSM 12966 (ASM95641v1), *Stenotrophomonas lactitubi* M15 (ASM280351v1), *Flavobacterium johnsoniae* UW101 (ASM1664v1), *Sphingomonas paucimobilis* FDAARGOS_908 (ASM1602709v1), and *Rhizobium leguminosarum* SM52 (ASM430655v1) were also included. The species tree was inferred from unrooted orthogroup gene trees using STAG and then rooted using the STRIDE algorithms implemented in OrthoFinder. Species tree was annotated with iTOLs.

them to corresponding 16S rRNA gene amplicon dynamics in a field microbiome study that was executed at the same field site at Texas A&M University Research Farm (4).

Individual isolates or the full cryopreserved collections are available. Cultures can be regrown using standard microbiological procedures to transfer a sterile inoculation loop of freezer stock onto the original isolation medium and incubation time (Table 1).

The high molecular weight genomic DNA of each wax bacterial isolate was extracted by using a phenol/chloroform extraction protocol (6). The genomic DNA of each strain was submitted to the Joint Genome Institute, a U.S. Department of Energy User Facility, for library preparation and sequencing with the Pacific Biosciences (PacBio) platform (7). DNA was sheared to 10 kb using g-TUBE columns (Covaris) and subjected to library preparation using the SMRTbell Express Template Prep Kit 2.0 and sequenced on the PacBio Sequel platform. Samples that failed DNA quality control required by the PacBio pipeline were instead sequenced using the NovaSeq S4 Illumina platform (8). High-fidelity PacBio Circular Consensus Sequencing (CCS) reads >5 kb were assembled with Flye 2.8.3 (9) using default settings. Raw Illumina sequences were quality filtered using BBTools (10). Artifact-filtered and normalized Illumina reads were assembled with SPAdes v3.15.3 (11) (–phred-offset 33 –cov-cutoff auto –t 16 m 64 –careful –k 25,55,95), and contigs were discarded if the length was <1 kb (BBTools reformat.sh: minlength = 1,000 ow = t). Genome completeness and contamination were estimated with CheckM v1.2.2 (12). The assembly was annotated using Prokka 1.14.6 (13), and the phylogenetic tree was inferred from using Orthofinder 2.5.5 (14) and edited with iTOLs v.6.5.8 (15) (Fig. 1).

ACKNOWLEDGMENTS

This work was supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409. Support for field research was provided by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Awards DE-SC0018409 and DE-FC02-07ER64494), by the National Science Foundation Long-term Ecological Research Program (DEB 1637653) at the Kellogg Biological Station, and by Michigan State University AgBioResearch. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231.

This data were generated for JGI Proposal #508590.

The authors declare no conflict of interest.

AUTHOR AFFILIATIONS

¹Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, Michigan, USA

²Department of Biochemistry & Biophysics, Texas A&M University, College Station, Texas, USA

³Centre National de la Recherche Scientifique, Laboratoire d'Ecologie Microbienne (UMR CNRS 5557, UMR INRAE 1418, VetAgro Sup), Universite Claude Bernard Lyon 1, Villeurbanne, France

AUTHOR ORCIDs

Marco E. Mechan-Llontop  <http://orcid.org/0000-0001-5644-4768>

Ashley Shade  <http://orcid.org/0000-0002-7189-3067>

FUNDING

Funder	Grant(s)	Author(s)
U.S. Department of Energy (DOE)	DE-SC0018409	Marco E. Mechan-Llontop John Mullet Ashley Shade
U.S. Department of Energy (DOE)	DE-FC02-07ER64494	Marco E. Mechan-Llontop John Mullet Ashley Shade
National Science Foundation (NSF)	1832042	Marco E. Mechan-Llontop John Mullet Ashley Shade
U.S. Department of Agriculture (USDA)	Michigan State AgBioResearch	Ashley Shade

AUTHOR CONTRIBUTIONS

Marco E. Mechan-Llontop, Data curation, Funding acquisition, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing | John Mullet, Funding acquisition, Resources, Writing – review and editing | Ashley Shade, Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing – review and editing

DATA AVAILABILITY

Raw sequencing data are available at the Joint Genome Institute's GOLD database under Study ID [Gs0157305](https://doi.org/10.5584/JGI.GOLD.StudyID.Gs0157305), as well as with the National Institutes of Health Sequence Read Archive (see Table 1). Genome sequencing assembly data have been deposited in the GenBank with accession numbers listed in Table 1. The bacterial collection is available at The Great Lakes Bioenergy Research Center at Michigan State University, USA (gblbrc@msu.edu).

REFERENCES

- Xue D, Zhang X, Lu X, Chen G, Chen Z-H. 2017. Molecular and evolutionary mechanisms of cuticular wax for plant drought tolerance. *Front Plant Sci* 8:621. <https://doi.org/10.3389/fpls.2017.00621>
- Steinmüller D, Tevini M. 1985. Action of ultraviolet radiation (UV-B) upon cuticular waxes in some crop plants. *Planta* 164:557–564. <https://doi.org/10.1007/BF00395975>
- Wang X, Kong L, Zhi P, Chang C. 2020. Update on cuticular wax biosynthesis and its roles in plant disease resistance. *Int J Mol Sci* 21:5514. <https://doi.org/10.3390/ijms21155514>
- Mechan-Llontop ME, Mullet J, Shade A. 2023. Phyllosphere exudates select for distinct microbiome members in sorghum epicuticular wax and aerial root mucilage. *Phytobiomes Journal* 7:184–197. <https://doi.org/10.1094/PBIOMES-08-22-0046-FI>
- Miller CS, Handley KM, Wrighton KC, Frischkorn KR, Thomas BC, Banfield JF. 2013. Short-read assembly of full-length 16S amplicons reveals bacterial diversity in subsurface sediments. *PLoS One* 8:e56018. <https://doi.org/10.1371/journal.pone.0056018>
- Kutchma AJ, Roberts MA, Knaebel DB, Crawford DL. 1998. Small-scale isolation of genomic DNA from *Streptomyces* mycelia or spores. *Biotechniques* 24:452–456. <https://doi.org/10.2144/98243st05>
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Veceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138. <https://doi.org/10.1126/science.1162986>
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438. <https://doi.org/10.1517/14622416.5.4.433>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Bushnell B. 2014. Bbtools software package (version 39.01). <https://sourceforge.net/projects/bbmap/>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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>
- 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>
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- Emms DM, Kelly S. 2019. Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20:238. <https://doi.org/10.1186/s13059-019-1832-y>
- Letunic I, Bork P. 2021. Interactive tree of life (iTOL) V5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>