

Genomes of diverse Clostridia isolated from a spore forming community in mice that were associated with protection against metabolic syndrome and obesity

Allison M. Weis,¹ Kendra A. Klag,¹ Rickeshia Bell,¹ W. Zac Stephens,¹ June L. Round¹

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT Clostridia are common mammalian gut commensals with emerging roles in human health. Here, we describe 10 Clostridia genomes from a consortium of spore forming bacteria, shown to protect mice from metabolic syndrome. These genomes will provide valuable insight on the beneficial role of spore forming bacteria in the gut.

KEYWORDS Clostridia, spore-former

Spore forming (SF) bacteria are an important part of a healthy microbiome. Loss of SF bacteria has been associated with diseases including obesity and type 2 diabetes (1–3). Recent literature has found that the class Clostridia provides protection from inflammatory bowel disease, metabolic syndrome, infections, and colorectal cancer (4–7). However, this class of bacteria is often fastidious to grow, which has limited the availability of quality genomes to study.

To advance our understanding of SF bacteria, genomes from 10 isolates were sequenced. To isolate SF bacteria, feces from C57BL/6 specific pathogen-free mice were incubated anaerobically with 0.1% cysteine and 3% chloroform at 37°C for 1 hour to kill off vegetative bacteria. Chloroform was removed by bubbling CO₂ through the sample for 30 s. To propagate the enrichment of SF bacteria, the sample was gavaged into a breeder pair of germ-free C57BL/6 mice housed in gnotobiotic conditions and feces from resulting offspring were collected, homogenized, serially diluted, and plated on YCFA anaerobically. Individual colonies were picked, streaked to isolation, and liquid cultures started from individual colonies in YCFA; DNA was extracted using Purelink Microbiome DNA purification kit (Invitrogen). Mouse work was performed under IACUC Protocol 00001562.

Five of the isolates' genomes were hybrid assembled from Illumina NovaSeq paired-end 150 and Oxford Nanopore Technologies (ONT) miniON reads. Illumina libraries were prepared with NEBNext Ultra II FS DNA kit (NEB, E7805S), and reads were adapter and quality-trimmed with cutadapt (v2.10) (8) in the trim_galore (v0.6.6) wrapper using default parameters. ONT libraries were prepared without DNA shearing or size selection using rapid barcoding kit R9.4.1 chemistry (ONT, SQK-RBK004). Reads were basecalled, demultiplexed, adapter, and barcode-trimmed with guppy (v6.0.1_gpu), then quality-filtered with NanoFilt (9) using a minimum average read quality of 10 and a minimum length of 200. Hybrid genomes were assembled with SPAdes v3.15.5 within Unicycler v0.5.0 pipeline "normal mode" and filtered contigs < 200 bp (10, 11). Five genomes were sequenced and assembled with PacBio reads using Flye v2.8.1 (12) with parameters "--plasmids --iterations 2". SMRTbell libraries were prepared without shearing using PacBio Express Template Prep Kit 2.0, pooled, and size selected using Sage Sciences' BluePippin with 0.75% DF Marker S1 High-Pass 6–10 kb v3 run protocol, S1 marker, and a cutoff of 8,000 (BPstart value), and then libraries were bound per SMRT

Editor Vanja Klepac-Ceraj, Department of Biological Sciences, Wellesley College, Wellesley, Massachusetts, USA

Address correspondence to June L. Round, june.round@path.utah.edu.

Allison M. Weis and Kendra A. Klag contributed equally to this article. The author order was determined by ascending height.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 3 April 2024

Accepted 13 April 2024

Published 20 June 2024

Copyright © 2024 Weis 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 and genome characteristics of spore forming bacteria^a

Isolate ID	GenBank accession	Taxon beyond clostridia order eubacteriales	Class	Closest ref genome	ANI	Sequencing	Contigs	N50	Size	%GC	CDS	SRR	Paired-End read number	PacBio subread number	PacBio subread N50	ONT read number	ONT read N50	ONT flowcell
JLR.KK001	JAYMNO0000	<i>f._Lachnospiraceae;</i>		GCA_9105_74715.1	98.44	Hybrid	31	4,968,565	6,390,410	46%	6,399	SRR28014014, SRR28014007	4,717,052	NA	NA	26,547	7,327	FLO-MIN106
	00000	<i>g._Sporofaciens</i> sp. KK001																
JLR.KK005	JAYMNR0000	<i>f._Candidatus;</i>		GCA_0099_11065.1	96.15	Hybrid	205	197,068	5,083,539	44%	4,949	SRR28014013, SRR28014006	3,632,234	NA	NA	15	4,054	FLO-FLG001
	00000	<i>g._Ventrimonas</i> sp. KK005																
JLR.KK006	JAYMNS0000	<i>f._Lachnospiraceae;</i>		GCA_9105_74255.1	98.86	Hybrid	112	111,685	4,698,239	45%	4,502	SRR28014010, SRR28014005	5,151,028	NA	NA	77	3,513	FLO-FLG001
	00000	<i>g._Candidatus</i> Merdisoma sp. KK006																
JLR.KK011	JAYWSZ0000	<i>f._Lachnospiraceae;</i>		GCA_9105_75725.1	98.57	PacBio	7	2,532,420	5,043,844	48%	4,958	SRR28516891	NA	92,862	11,648	NA	NA	NA
	00000	<i>g._Candidatus</i> Merdisoma sp. KK011																
JLR.KK002	JAYMNP0000	<i>f._Lachnospiraceae</i>		GCF_0004_03845.2	88.4	Hybrid	5	4,212,295	4,276,026	46%	4,102	SRR28014009, SRR28014004	4,880,942	NA	NA	11,608	8,894	FLO-MIN106
	00000	<i>bacterium</i> KK002																
JLR.KK008	CP143548	<i>f._Lachnospiraceae</i>		GCA_9105_85345.1	78.35	PacBio	1	3,188,748	3,188,748	48%	3,078	SRR28516890	NA	39,804	12,282	NA	NA	NA
		<i>bacterium</i> KK008																
JLR.KK009	CP143549	<i>f._Lachnospiraceae</i>		GCA_0004_03315.2	99.31	PacBio	1	5,268,209	5,268,209	47%	5,165	SRR28516889	NA	81,055	12,085	NA	NA	NA
		<i>bacterium</i> KK009																
JLR.KK004	JAYMNQ0000	<i>f._Oscillospiraceae;</i>		GCF_0099_36035.1	96.44	Hybrid	12	2,635,137	3,852,823	54%	4,046	SRR28014008, SRR28014003	3,640,502	NA	NA	15,286	8,990	FLO-MIN106
	00000	<i>g._Acartalibacter</i> sp. KK004																
JLR.KK007	JAYWSY0000	<i>f._Oscillospiraceae;</i>		GCA_9105_84605.1	98.48	PacBio	5	1,637,602	4,295,262	58%	4,458	SRR28516888	NA	109,452	12,756	NA	NA	NA
	00000	<i>g._Lawsonibacter</i> sp. KK007																
JLR.KK010	CP143550	<i>f._Eubacteriales</i>		GCF_0099_36045.1	98.09	PacBio	1	3,076,573	3,076,573	44%	2,874	SRR28516887	NA	260,145	12,915	NA	NA	NA
		Family Xllig_Emergencia sp. KK010																

^aNA = Not applicable.

Link Setup and sequenced on a Sequel II. Assemblies were annotated by NCBI's PGAP v6.6 (13).

All isolates are domain Bacteria, phylum Bacillota, class Clostridia, and order Eubacteriales (see Table 1 for full NCBI-assigned taxonomy). While some of the genomes had published close matching reference genomes, others such as *Lachnospiraceae* KK002 and KK008 were as far away as 88% and 78% best match by average nucleotide identity (ANI), respectively (FastANI v0.1.3 via GTDB-tk toolkit), indicating that they are likely new and undescribed species (14). Each genome's features including GC% content, genome size, and the number of predicted genes are described in Table 1.

ACKNOWLEDGMENTS

We thank the Round Lab for help with mouse work and bacteria isolation.

This work was funded by R01AT011423-03, a W.M. Keck Award, and a Burrough's Welcome grant to J.L.R. K.A.K. was funded by NIH NRSA 30DK127846. A.M.W. was funded by NIH NCI NRSA F32CA243501.

AUTHOR AFFILIATION

¹Department of Pathology, Division of Microbiology and Immunology, University of Utah School of Medicine, Salt Lake City, Utah, USA

AUTHOR ORCIDs

Allison M. Weis  <http://orcid.org/0000-0001-9991-1661>

Kendra A. Klag  <http://orcid.org/0000-0001-5237-8628>

W. Zac Stephens  <http://orcid.org/0000-0002-8072-9023>

June L. Round  <http://orcid.org/0000-0002-7158-9874>

FUNDING

Funder	Grant(s)	Author(s)
HHS NIH National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)	R01AT011423	Allison M. Weis Kendra A. Klag Rickeshia Bell W. Zac Stephens

DATA AVAILABILITY

All sequences are available through NCBI Bioproject [PRJNA1061597](#). GenBank and SRA accession numbers listed in Table 1. Isolates are available from the Round Lab.

REFERENCES

- Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, et al. 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490:55–60. <https://doi.org/10.1038/nature11450>
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JL. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027–1031. <https://doi.org/10.1038/nature05414>
- Thingholm LB, Rühlemann MC, Koch M, Fuqua B, Laucke G, Boehm R, Bang C, Franzosa EA, Hübenthal M, Rahnavard A, Frost F, Lloyd-Price J, Schirmer M, Lusi AJ, Vulpe CD, Lerch MM, Homuth G, Kacprowski T, Schmidt CO, Nöthlings U, Karlsen TH, Lieb W, Laudes M, Franke A, Huttenhower C. 2019. Obese individuals with and without type 2 diabetes show different gut microbial functional capacity and composition. *Cell Host Microbe* 26:252–264. <https://doi.org/10.1016/j.chom.2019.07.004>
- Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, Fukuda S, Saito T, Narushima S, Hase K, Kim S, Fritz JV, Wilmes P, Ueha S, Matsushima K, Ohno H, Olle B, Sakaguchi S, Taniguchi T, Morita H, Hattori M, Honda K. 2013. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* 500:232–236. <https://doi.org/10.1038/nature12331>
- Zhang X, Yu D, Wu D, Gao X, Shao F, Zhao M, Wang J, Ma J, Wang W, Qin X, Chen Y, Xia P, Wang S. 2023. Tissue-resident Lachnospiraceae family bacteria protect against colorectal carcinogenesis by promoting tumor immune surveillance. *Cell Host Microbe* 31:418–432. <https://doi.org/10.1016/j.chom.2023.01.013>
- Kim Y-G, Sakamoto K, Seo S-U, Pickard JM, Gilliland MG III, Pudlo NA, Hoostal M, Li X, Wang TD, Feehley T, Stefka AT, Schmidt TM, Martens EC, Fukuda S, Inohara N, Nagler CR, Núñez G. 2017. Neonatal acquisition of Clostridia species protects against colonization by bacterial pathogens. *Science* 356:315–319. <https://doi.org/10.1126/science.aag2029>
- Petersen C, Bell R, Klag KA, Lee S-H, Soto R, Ghazaryan A, Buhrke K, Ekiz HA, Ost KS, Boudina S, O'Connell RM, Cox JE, Villanueva CJ, Stephens WZ, Round JL. 2019. T cell-mediated regulation of the microbiota protects

- against obesity. *Science* 365:eaat9351. <https://doi.org/10.1126/science.aat9351>
8. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j* 17:10. <https://doi.org/10.14806/ej.17.1.200>
 9. De Coster W, D'Hert S, Schultz DT, Cruys M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
 10. 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>
 11. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
 12. 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>
 13. 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>
 14. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>