AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Metagenome-Assembled Genome Sequences of Five Strains from the *Microtus ochrogaster* (Prairie Vole) Fecal Microbiome

Meghan Donovan,[b] Michael D. J. Lynch,[c,d] Calvin S. Mackey,[a] Grayson N. Platt,[a,b] Brian K. Washburn,[e] Daniel L. Vera,[f]* Darryl J. Trickey,[a] Trevor C. Charles,[c,d] Zuoxin Wang,[b] Kathryn M. Jones[a]

[a]Department of Biological Science, Florida State University, Tallahassee, Florida, USA
[b]Department of Psychology, Florida State University, Tallahassee, Florida, USA
[c]Metagenom Bio, Waterloo, Ontario, Canada
[d]University of Waterloo, Waterloo, Ontario, Canada
[e]Department of Biological Science Core Facilities, Florida State University, Tallahassee, Florida, USA
[f]Department of Biological Science, Center for Genomics and Personalized Medicine, Florida State University, Tallahassee, Florida, USA

Meghan Donovan, Michael D. J. Lynch, Calvin S. Mackey, and Grayson N. Platt made an equal contribution to the manuscript and are ordered alphabetically.

Trevor C. Charles, Zuoxin Wang, and Kathryn M. Jones are co-senior authors.

ABSTRACT The prairie vole (*Microtus ochrogaster*) is an important model for the study of social monogamy and dual parental care of offspring. Characterization of specific host species-microbe strain interactions is critical for understanding the effects of the microbiota on mood and behavior. The five metagenome-assembled genome sequences reported here represent an important step in defining the prairie vole microbiome.

Unusually among rodents, prairie voles form strong mating-induced pair bonds and thus serve as an important model for the study of the neurobiological basis of bonding and associated behaviors (1). Although advances have been made in understanding the neurochemical interactions involved in pair bonding (1), the study of the molecular basis of neuroanatomical responses requires the continued use of a model system. The importance of gut microbes for modulating multiple neurochemical interactions along the "microbiota-gut-brain axis" has been established for humans and mice (2, 3). However, there is considerable variation between mammalian hosts in microbe diversity and metabolism that does not necessarily correlate with host phylogeny, even within a clade such as rodents (4, 5). To date, there have been few studies on the prairie vole microbiome (6, 7). Thus, to facilitate studies of the microbial endocrinology (8) of prairie voles, we have determined the full shotgun metagenome of stool samples from 6 voles and produced an unbinned metagenomic coassembly and 5 metagenome-assembled genomes (MAGs).

Stools were collected from 4 female and 2 male voles (age 3 to 9 months) by temporary isolation of each animal in a bedding-free, sanitized cage. Voles were sexually naive and housed in male/male or female/female cage pairs. All experimental procedures were approved by the Florida State University (FSU) Institutional Animal Care and Use Committee and were in accordance with the U.S. National Institutes of Health Guide for the Care and Use of Laboratory Animals (NIH publication number 80-23).

DNA was prepared from stool samples frozen at −80°C using the MoBio/QIAamp PowerFecal DNA kit, according to the manufacturer's instructions (Qiagen USA). Genomic DNA was sheared using a Covaris E220 focused ultrasonicator. Libraries were prepared using the NEBNext Ultra II DNA library prep kit for Illumina (New England BioLabs, USA), following the manufacturer's protocol. Whole-genome shot-

Citation Donovan M, Lynch MDJ, Mackey CS, Platt GN, Washburn BK, Vera DL, Trickey DJ, Charles TC, Wang Z, Jones KM. 2020. Metagenome-assembled genome sequences of five strains from the *Microtus ochrogaster* (prairie vole) fecal microbiome. Microbiol Resour Announc 9:e01310-19. https://doi.org/10.1128/MRA.01310-19.

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Kathryn M. Jones, kmjones@bio.fsu.edu.

* Present address: Daniel L. Vera, Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.

Received 22 October 2019
Accepted 30 November 2019
Published 9 January 2020

TABLE 1 Characteristics of MAGs and vole stool coassemblies

	Data for:													
	MAG						Stool sample from vole no. (sex):							
Characteristic	MAG13	MAG21	MAG84	MAG226	MAG234	Coassembled_Vole1-6_2017[a]	1 (female)	2 (female)	3 (female)	4 (female)	4 (female) (resequencing)	5 (male)	6 (male)	6 (male) (resequencing)
Taxon	*Bacteroidales*	*Bacteroidales*	*Bacteroidales*	*Clostridia*	*Firmicutes*	NA	Prairie vole	Prairie vole	Prairie vole	Prairie vole	Prairie vole	Prairie vole	Prairie vole	Prairie vole
Genome size or no. of bases (Mb)[b]	2.1	2.42	2.67	3.11	2.99	269,886,203[c]	16.5	13.3	11.1	13.7	25.2	10.4	12.1	33.2
G+C content (%)	40.89	48.39	49.28	38.34	41.89	ND	47	46.8	47.2	47.7	47.7	46.6	46.6	46.7
No. of contigs	213	147	244	348	318	1,551,628								
N_{50} (bp)	15,620	26,297	15,553	17,329	15,004	960								
No. of genes identified	1,934	2,092	2,450	3,216	2,957	ND								
No. of tRNAs	34	36	38	44	43	ND								
Completion (%)[d]	90.65	90.65	92.09	89.21	94.96	NA								
Estimated redundancy, single-copy core genes (%)[d]	2.16	0.72	2.16	0	0	NA								
Estimated contamination (%)[e]	1.26	1.15	5.51	1.68	3.38	NA								
Strain heterogeneity (%)[e]	50	50	88	66.67	76.92	NA								
GenBank accession no. for sample	WFMC00000000	WFMB00000000	WFMA00000000	WFLZ00000000	WFLY00000000	WFLX000000000								
Platform							Illumina paired end	Illumina paired end	Illumina paired end	Illumina paired end	Illumina paired end	Illumina paired end	Illumina paired end	Illumina paired end
Library name							v1_NEB6_GCCAAT_May2017	v2_NEB1_ATCACG_May2017	v3_NEB2_CGATGT_May2017	v4_NEB3_TTAGGC_May2017	vole4_NEB3_TS1_TTAGGC_Oct2017_deep	v5_NEB4_TGACCA_May2017	v6_NEB5_ACAGTG_May2017	v6_NEB5_ACAGTG_May2017
SRA accession no. for raw reads							SRR9122728	SRR9129758	SRR9130027	SRR9130028	SRR9130365	SRR9130099	SRR9130108	SRR9335148
SRA accession no. for expt							SRX5896710	SRX5903730	SRX5903999	SRX5904000	SRX5904189	SRX5904051	SRX5904060	SRX6101484

[a] ND, no data; NA, not applicable.
[b] Genome size is provided for MAGs, and the number of bases is provided for vole stool coassemblies.
[c] Total number of merged reads.
[d] From Anvi'o v.5.5.0 (15).
[e] From CheckM v.1.0.18 (17).

gun sequencing of libraries (average fragment size, 765 bp) was performed on an Illumina HiSeq 2500 instrument in the FSU College of Medicine Translational Science Laboratory using paired-end 250-base sequence reads. The total numbers of reads were 66,059,128 (vole 1), 53,198,216 (vole 2), 44,221,130 (vole 3), 54,740,608 (vole 4), 41,707,738 (vole 5), and 48,354,762 (vole 6). Additional sequencing performed on the HiSeq instrument with paired-end 200-base sequence reads generated 125,776,266 (vole 4) and 173,617,670 (vole 6) reads. Read quality control was performed using standard pnnl-atlas v.1.0.35 (9) filtering. The coassembly of all 8 sequence runs and the binning was managed using SqueezeMeta v.1.1.2 (10). Coassembly was performed on reads merged before assembly using Megahit v.1.1.2 (11) (see Table 1 for coassembly details). Data were binned as contigs after coassembly. Binning was performed with MaxBin v.2.2.6 (12) (producing 235 bins) and with metabat2 v.2.12.1 (13) (producing 38 bins). Bins were subsequently processed using DAS Tool v.1.1.1 (14), producing 77 bins. Five bins with high percent completion and low percent contamination were chosen for immediate refinement into MAGs using Anvi'o v.5.5.0 (15), according to an online tutorial (http://merenlab.org/data/refining-espinoza-mags/) (16). Quality was assessed with CheckM v.1.0.18 (17). Default parameters were used for all software, unless otherwise specified.

Anvi'o estimated all but one of the MAGs at >90% completeness (see Table 1). Recovery of rRNA genes was poor, which is not unusual for MAGs due to the difficulty of assembling these sequences (18). However, tRNAscan-SE (19) detected 34 to 44 tRNA genes in all of the MAGs, with predicted anticodons for 16 to 20 amino acids. These data will be extremely useful in studies of metabolic functions in the vole microbiome and for comparison with other rodent models.

Data availability. The MAG sequences and associated experiment and run data have been deposited in GenBank under BioProject accession number PRJNA449069 and the GenBank and SRA accession numbers given in Table 1.

ACKNOWLEDGMENTS

This work was partially supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, award number 2014-67013-21579 to K.M.J., and by National Institutes of Health grants MH058616-R01 and MH108527-R01 to Z.W. M.D. was supported by NIH program training grant T32 MH093311.

REFERENCES

1. Lieberwirth C, Wang ZX. 2016. The neurobiology of pair bond formation, bond disruption, and social buffering. Curr Opin Neurobiol 40:8–13. https://doi.org/10.1016/j.conb.2016.05.006.
2. Cryan JF, O'Riordan KJ, Cowan CSM, Sandhu KV, Bastiaanssen TFS, Boehme M, Codagnone MG, Cussotto S, Fulling C, Golubeva AV, Guzzetta KE, Jaggar M, Long-Smith CM, Lyte JM, Martin JA, Molinero-Perez A, Moloney G, Morelli E, Morillas E, O'Connor R, Cruz-Pereira JS, Peterson VL, Rea K, Ritz NL, Sherwin E, Spichak S, Teichman EM, van de Wouw M, Ventura-Silva AP, Wallace-Fitzsimons SE, Hyland N, Clarke G, Dinan TG. 2019. The microbiota-gut-brain axis. Physiol Rev 99:1877–2013. https://doi.org/10.1152/physrev.00018.2018.
3. Johnson KV, Foster KR. 2018. Why does the microbiome affect behaviour? Nat Rev Microbiol 16:647–655. https://doi.org/10.1038/s41579-018-0014-3.
4. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008. Evolution of mammals and their gut microbes. Science 320:1647–1651. https://doi.org/10.1126/science.1155725.
5. Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L, Henrissat B, Knight R, Gordon JI. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332:970–974. https://doi.org/10.1126/science.1198719.
6. Curtis JT, Assefa S, Francis A, Kohler GA. 2018. Fecal microbiota in the female prairie vole (*Microtus ochrogaster*). PLoS One 13:e0190648. https://doi.org/10.1371/journal.pone.0190648.
7. Assefa S, Ahles K, Bigelow S, Curtis JT, Kohler GA. 2015. Lactobacilli with probiotic potential in the prairie vole (*Microtus ochrogaster*). Gut Pathog 7:35. https://doi.org/10.1186/s13099-015-0082-0.
8. Sandrini S, Aldriwesh M, Alruways M, Freestone P. 2015. Microbial endocrinology: host-bacteria communication within the gut microbiome. J Endocrinol 225:R21–R34. https://doi.org/10.1530/JOE-14-0615.
9. White RA III, Brown J, Colby S, Overall CC, Lee J-Y, Zucker J, Glaesemann KR, Jansson C, Jansson JK. 2017. ATLAS (Automatic Tool for Local Assembly Structures)—a comprehensive infrastructure for assembly, annotation, and genomic binning of metagenomic and metatranscriptomic data. PeerJ Prepr 5:e2843v1. https://doi.org/10.7287/peerj.preprints.2843v1.
10. Tamames J, Puente-Sánchez F. 2018. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. Front Microbiol 9:3349. https://doi.org/10.3389/fmicb.2018.03349.
11. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
12. Wu YW, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638.
13. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359.

14. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843. https://doi.org/10.1038/s41564-018-0171-1.
15. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319. https://doi.org/10.7717/peerj.1319.
16. Shaiber A, Eren AM, Shaiber A, Eren AM. 2019. Composite metagenome-assembled genomes reduce the quality of public genome repositories. mBio 10:e00725-19. https://doi.org/10.1128/mBio.00725-19.
17. 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.
18. Yuan C, Lei J, Cole J, Sun Y. 2015. Reconstructing 16S rRNA genes in metagenomic data. Bioinformatics 31:i35–i43. https://doi.org/10.1093/bioinformatics/btv231.
19. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.