



# Feces Metagenomes and Metagenome-Assembled Genome Sequences from Two Separate Dogs (*Canis lupus familiaris*) with Multiple Diarrheal Episodes

 Tshepiso Pleasure Ateba,<sup>a</sup> Kazeem Adekunle Alayande,<sup>b</sup> Mulunda Mwanza<sup>a</sup>

<sup>a</sup>Department of Animal Health, Faculty of Natural and Agricultural Sciences, Mafikeng Campus, North West University, Mmabatho, Mafikeng, South Africa

<sup>b</sup>Department of Microbiology, Faculty of Natural and Agricultural Sciences, Mafikeng Campus, North West University, Mmabatho, Mafikeng, South Africa

**ABSTRACT** This study reports two feces metagenomes (D84 and D85) and six metagenome-assembled genomes (MAGs). The assembled MAGs include *Pseudomonas* sp. strain N1D84 and *Acinetobacter* sp. strain N2D84 from D84 and *Enterococcus* sp. strain N4D85, *Enterococcus* sp. strain N5D85, *Lactobacillus* sp. strain N6D85, and *Leuconostoc* sp. strain N7D85 from D85. *Acinetobacter* sp. N2D84 was identified as a human pathogen with a probability of 92%.

**G**astroenteritis is a common disease in dogs caused by a number of microorganisms, including bacteria and viruses. Its symptoms vary from diarrhea and abdominal pain to acute hemorrhagic syndrome, and it can be fatal (1).

Fecal samples were separately collected from two diarrhea-affected dogs (*Canis lupus familiaris*) that were housed by the Society for the Prevention of Cruelty to Animals (SPCA) in Mafikeng, North West Province, South Africa. The samples were collected directly from the rectum with sterile gloves and immediately placed in sterile containers. Metagenomic DNA was extracted from 150 mg of the fecal sample using a Quick-DNA fecal/soil microbe miniprep kit (Zymo Research Corp., USA) following the manufacturer's instruction. The library was prepared with a Nextera DNA Flex library prep kit using Nextera DNA CD index adapters (96 indexes plated) and sequenced on an Illumina NovaSeq instrument. A total of 12,512,020 reads were generated from sample D84 and 19,928,158 from sample D85, both with 1 × 150-bp paired-end read lengths, while the depth of sequencing for D84 was 156× and that for D85 was 153×.

The metagenomic data were filtered for adapter regions and low-quality reads using Trimmomatic v0.36 (2). The quality reads were taxonomically classified using Kaiju v1.7.2 (3) and GOTTCHA2 v2.1.6 and thereafter assembled using metaSPAdes v3.13.0 (4) and MEGAHIT v1.2.9 (5) for samples D84 and D85, respectively. Each assembly was binned into metagenome-assembled genomes (MAGs) using MaxBin2 v2.2.4 (6). The quality and completeness of each MAG was assessed using CheckM v1.0.18 (7), and redundant bins were excluded from further analysis. The taxonomic assignments were obtained for the MAGs based on the genome taxonomy database using GTDB-Tk v1.1.0 (8).

The MAGs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.12 (9). The extent of pathogenicity and acquired drug-resistant genes were determined for the MAGs using PathogenFinder v1.1 (10) and ResFinder v4.0, respectively (11). Most of the software was accessed through the KBase workspace service v0.11.1 (12), and default parameters were used for all the software employed in the analysis.

The assembly size of the feces metagenome from sample D84 is 10,293,073 bp, and it contains 575 contigs with a G+C content of 52.23%, an  $N_{50}$  value of 52,139 bp, and an  $L_{50}$  value of 37. The entire assembly, which contains 7,693 coding sequences, was

**Citation** Ateba TP, Alayande KA, Mwanza M. 2020. Feces metagenomes and metagenome-assembled genome sequences from two separate dogs (*Canis lupus familiaris*) with multiple diarrheal episodes. Microbiol Resour Announc 9:e01065-20. <https://doi.org/10.1128/MRA.01065-20>.

**Editor** David Rasko, University of Maryland School of Medicine

**Copyright** © 2020 Ateba 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/).

Address correspondence to Tshepiso Pleasure Ateba, 16100581@nwu.ac.za.

**Received** 16 September 2020

**Accepted** 25 October 2020

**Published** 25 November 2020

**TABLE 1** Features and accession numbers of the metagenome-assembled genomes from the feces metagenomes<sup>a</sup>

Physical sample	MAG identity	Genome size (bp)	Total no. of CDS <sup>b</sup>	No. of contigs	$N_{50}$ (bp)	G + C content (%)	CMP <sup>c</sup> (%)	CNT <sup>d</sup> (%)	GenBank accession no.
D84	<i>Pseudomonas</i> sp. NID84	6,374,370	5,798	168	98,074	59.00	99.68	4.06	JACRYQ000000000
	<i>Acinetobacter</i> sp. N2D84	3,755,466	3,729	373	11,956	41.17	94.72	10.15	JACRYP000000000
D85	<i>Enterococcus</i> sp. N4D85	3,185,421	3,438	678	5,355	37.72	76.84	21.84	JACUTJ000000000
	<i>Enterococcus</i> sp. N5D85	4,273,788	4,421	616	9,391	39.29	75.78	33.33	JACUTK000000000
	<i>Lactobacillus</i> sp. N6D85	3,558,646	3,945	592	8,543	34.23	67.24	14.37	JACUTL000000000
	<i>Leuconostoc</i> sp. N7D85	2,401,161	2,582	445	5,665	39.19	87.69	38.16	JACUTM000000000

<sup>a</sup>The values for completeness and contamination of each MAG were determined using CheckM v1.0.18. The genome sizes were determined using v1-KBaseGenomeAnnotations.Assembly-5.0 and NCBI Prokaryotic Genome Annotation Pipeline (PGAP).

<sup>b</sup>CDS, coding sequences.

<sup>c</sup>CMP, completeness.

<sup>d</sup>CNT, contamination.

binned into two MAGs, which were identified as *Pseudomonas* sp. strain NID84 and *Acinetobacter* sp. strain N2D84 (Table 1). As for sample D85, the assembly size is 16,848,868 bp, distributed in 2,968 contigs, and it has an average G+C content of 37.35%. The  $N_{50}$  value of the assembly is 7,085 bp, while the  $L_{50}$  value is 570, and the entire metagenome contains 8,163 predicted genes. Four different MAGs were extracted from the metagenome and identified as *Enterococcus* sp. strain N4D85, *Enterococcus* sp. strain N5D85, *Lactobacillus* sp. strain N6D85, and *Leuconostoc* sp. strain N7D85 (Table 1). Not all the contigs contained the MAGs in both samples, and the unbinned contigs were discarded. *Acinetobacter* sp. N2D84 was identified as a human pathogen with a probability of 92% and 13 matched pathogenic families.

Ethical clearance for the study was approved by the Research Ethics Committee of the North West University, South Africa (NWU-00160-14-A9).

**Data availability.** All data were deposited under the GenBank BioProject number [PRJNA655841](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA655841). These whole-genome shotgun projects have been deposited in DDBJ/ENA/GenBank under the accession numbers [JACRYN000000000](https://www.ncbi.nlm.nih.gov/nuccore/JACRYN000000000) and [JACRYO000000000](https://www.ncbi.nlm.nih.gov/nuccore/JACRYO000000000). The versions described in this paper are the first versions, [JACRYN010000000](https://www.ncbi.nlm.nih.gov/nuccore/JACRYN010000000) and [JACRYO010000000](https://www.ncbi.nlm.nih.gov/nuccore/JACRYO010000000). The SRA accession numbers for the raw reads are [SRX8905271](https://www.ncbi.nlm.nih.gov/sra/SRX8905271) and [SRX8949278](https://www.ncbi.nlm.nih.gov/sra/SRX8949278) for samples D84 and D85, respectively.

## ACKNOWLEDGMENTS

We acknowledge the Molecular Research (MRDNA) Laboratory (Shallowater, TX, USA) for providing sequencing services.

This work was supported in part by the Health and Welfare Sector Education and Training Authority (HWSETA) Postgraduate Bursary, South Africa, as well as by an emerging research grant from the Faculty of Natural and Agricultural Sciences, North West University, awarded to T.P.A.

## REFERENCES

- Rainha K, Ferreira RF, Trindade CNR, Carneiro LG, Penna B, Endres BT, Begum K, Alam MJ, Garey KW, Domingues Regina Maria CP, Ferreira EO. 2019. Characterization of *Clostridioides difficile* ribotypes in domestic dogs in Rio de Janeiro, Brazil. *Anaerobe* 58:22–29. <https://doi.org/10.1016/j.anaerobe.2019.06.007>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
- Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 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>.
- Wu Y-W, 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>.
- 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>.
- Chaumeil P-A, 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>.
- 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>.
- Cosentino S, Larsen MV, Aarestrup FM, Lund O. 2013. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. *PLoS One* 8:e77302. <https://doi.org/10.1371/journal.pone.0077302>.
- Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A,

- Allesoe RL, Rebelo AR, Florensa AR, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wiczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* dkaa345. <https://doi.org/10.1093/jac/dkaa345>.
12. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.