

Complete Genome Sequence of the Novel Leech Symbiont *Mucinivorans hirudinis* M3^T

Michael C. Nelson, Lindsey Bomar,* Joerg Graf

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut, USA

* Present address: Lindsey Bomar, The Forsyth Institute, Department of Microbiology, Cambridge, Massachusetts, USA.

***Mucinivorans hirudinis* M3^T was isolated from the digestive tract of the medicinal leech, *Hirudo verbana*, and is the type species of a new genus within the *Rikenellaceae*. Here, we report the complete annotated genome sequence of this bacterium.**

Received 18 December 2014 Accepted 22 December 2014 Published 5 February 2015

Citation Nelson MC, Bomar L, Graf J. 2015. Complete genome sequence of the novel leech symbiont *Mucinivorans hirudinis* M3^T. *Genome Announc* 3(1):e01530-14. doi:10.1128/genomeA.01530-14.

Copyright © 2015 Nelson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Joerg Graf, joerg.graf@uconn.edu.

Bacteria belonging to the phylum *Bacteroidetes* are commonly found within the alimentary tracts of humans and animals and have previously been shown to play numerous roles in both host development and regulation of the gut microbiome (1–5). The *Rikenellaceae* is a family within the *Bacteroidetes* for which only four genera are currently described, and few full genome sequences are available. Previous work in our lab identified and isolated a novel bacterium present in the crop of medicinal leech *Hirudo verbana* that represents a new genus and species within the *Rikenellaceae* (6, 7). Here, we report the complete annotated genome sequence of the type strain isolate for this bacterium, known as *Mucinivorans hirudinis* M3^T (8).

An initial Illumina mate-pair library was sequenced as 100 bp paired-ends on a single lane of a HiSeq 2000. The reads were trimmed to 36 bp and assembled using CLC Genomics Workbench (version 4.0), yielding a 2.7-Mbp draft genome consisting of 332 contigs larger than 500 bp ($N_{50} = 21,974$ bp). A preliminary analysis of the draft genome after annotation by RAST identified numerous repetitive and transposable elements and phage sequences, which hindered further assembly using available short-read methods.

Because of this complexity, Pacific Bioscience's single-molecule real-time (SMRT) sequencing technology was chosen as an optimal sequencing strategy. Subsequently, the full genome was sequenced from two libraries, one size selected greater than 7 kb and the other not, using two SMRT cells for each library on the PacBio RS II. The genome was assembled using the Hierarchical Genome Assembly Process (HGAP) assembler (9) yielding a single contig approximately 3.15 Mbp in length, which was manually edited to circularize the overlapping ends of the genome. The final genome sequence is 3,149,227 bp and has a G+C content of 44.9%. Self dot plot analysis using Gepard (10) confirmed the presence of several large (>1.5 kbp) sequence repeats that had hindered the initial assembly.

Functional annotation by RAST combined with manual annotation and curation identified 3,015 open reading frames (ORFs), 39 tRNAs, and 2 complete ribosomal operons. As initially observed in the draft genome, more than 200 ORFs were annotated

as encoding phage and/or mobile element proteins. As expected for a member of the *Bacteroidetes*, the M3^T genome encodes for a number of carbohydrate metabolism genes, with many of these organized in polysaccharide utilization loci (PULs) similar to those present in other *Bacteroidetes*. Other notable features of the M3^T genomes are a genes involved in cobalamin biosynthesis, which may be utilized to provide the leech host with vitamin B₁₂ which is deficient in vertebrate blood (11, 12). More than 1,200 ORFs (~40%) were annotated as hypothetical proteins, a result not surprising for the annotation of the genome of a new genus with few annotated genomes for other genera within the *Rikenellaceae*. The availability of a new, fully sequenced and annotated genome for a novel member of the *Bacteroidetes* should allow for more thorough analysis of members of this important group of bacteria.

Nucleotide sequence accession numbers. Raw PacBio read data and the annotated genome were deposited in EMBL/GenBank/DDBJ under the study accession no. PRJEB5062. The annotated genome sequence is available under the accession no. [HG934468](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/seqview.fcgi?acc=HG934468).

ACKNOWLEDGMENTS

We thank M. Maltz for her initial isolation of the M3^T strain.

PacBio sequencing and HGAP assembly were provided by Pacific Biosciences as part of their 2013 ASM SMRT Sequencing Solutions Grant Contest. This research was supported by NIH grant R01 GM095390 (to Joerg Graf, Peter Visscher, and Hilary G. Morrison).

REFERENCES

- An D, Oh SF, Olszak T, Neves JF, Avci FY, Erturk-Hasdemir D, Lu X, Zeissig S, Blumberg RS, Kasper DL. 2014. Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* 156:123–133. <http://dx.doi.org/10.1016/j.cell.2013.11.042>.
- Arora T, Sharma R. 2011. Fermentation potential of the gut microbiome: implications for energy homeostasis and weight management. *Nutr Rev* 69:99–106. <http://dx.doi.org/10.1111/j.1753-4887.2010.00365.x>.
- Degnan PH, Barry NA, Mok KC, Taga ME, Goodman AL. 2014. Human gut microbes use multiple transporters to distinguish vitamin B₁₂ analogs and compete in the gut. *Cell Host Microbe* 15:47–57. <http://dx.doi.org/10.1016/j.chom.2013.12.007>.

4. Hooper LV, Gordon JL. 2001. Commensal host-bacterial relationships in the gut. *Science* 292:1115–1118. <http://dx.doi.org/10.1126/science.1058709>.
5. Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, Nakanishi Y, Uetake C, Kato K, Kato T, Takahashi M, Fukuda NN, Murakami S, Miyauchi E, Hino S, Atarashi K, Onawa S, Fujimura Y, Lockett T, Clarke JM, Topping DL, Tomita M, Hori S, Ohara O, Morita T, Koseki H, Kikuchi J, Honda K, Hase K, Ohno H. 2013. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504:446–450. <http://dx.doi.org/10.1038/nature12721>.
6. Worthen PL, Gode CJ, Graf J. 2006. Culture-independent characterization of the digestive-tract microbiota of the medicinal leech reveals a tripartite symbiosis. *Appl Environ Microbiol* 72:4775–4781. <http://dx.doi.org/10.1128/AEM.00356-06>.
7. Bomar L, Maltz M, Colston S, Graf J. 2011. Directed culturing of microorganisms using metatranscriptomics. *MBio* 2: <http://dx.doi.org/10.1128/mBio.00012-11>.
8. Nelson MC, Bomar L, Maltz M, Graf J. 6 January 2015. *Mucinivorans hirudinis* gen. nov., sp. nov., an anaerobic, mucin-degrading bacterium isolated from the digestive tract of the medicinal leech, *Hirudo verbana*. *Int J Syst Evol Microbiol*. <http://dx.doi.org/10.1099/ijss.0.000052>.
9. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <http://dx.doi.org/10.1038/nmeth.2474>.
10. Krumsiek J, Arnold R, Rattei T. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23:1026–1028. <http://dx.doi.org/10.1093/bioinformatics/btm039>.
11. Nelson MC, Graf J. 2012. Bacterial symbioses of the medicinal leech *Hirudo verbana*. *Gut Microbes* 3:322–331. <http://dx.doi.org/10.4161/gmic.20227>.
12. Graf J. 2002. The effect of symbionts on the physiology of *Hirudo medicinalis*, the medicinal leech. *Invertebr Reprod Dev* 41:269–275. <http://dx.doi.org/10.1080/07924259.2002.9652760>.