2017

Genome sequence of Christensenella minuta DSM 22607T

Bruce A. Rosa
Washington University School of Medicine in St. Louis

Kymberlie Hallsworth-Pepin
Washington University School of Medicine in St. Louis

John Martin
Washington University School of Medicine in St. Louis

Aye Wollam
Washington University School of Medicine in St. Louis

Makedonka Mitreva
Washington University School of Medicine in St. Louis

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Genome Sequence of *Christensenella minuta* DSM 22607^T

Bruce A. Rosa,[a] Kymberlie Hallsworth-Pepin,[a] John Martin,[a] Aye Wollam,[a] Makedonka Mitreva[a,b]

The McDonnell Genome Institute, Washington University School of Medicine, St. Louis, Missouri, USA[a]; Department of Medicine, Washington University School of Medicine, St. Louis, Missouri, USA[b]

ABSTRACT Obesity influences and is influenced by the human gut microbiome. Here, we present the genome of *Christensenella minuta*, a highly heritable bacterial species which has been found to be strongly associated with obesity through an unknown biological mechanism. This novel genome provides a valuable resource for future obesity therapeutic studies.

Received 28 October 2016 **Accepted** 1 November 2016 **Published** 12 January 2017

Citation Rosa BA, Hallsworth-Pepin K, Martin J, Wollam A, Mitreva M. 2017. Genome sequence of *Christensenella minuta* DSM 22607^T. Genome Announc 5:e01451-16. https://doi.org/10.1128/genomeA.01451-16.

Address correspondence to Makedonka Mitreva, mmitreva@wustl.edu.

The gut microbiome is acquired naturally from birth, after which it is affected by host genetics, environmental factors, and diet (1). Due to its widespread and significant effects on human health, modulation of the gut microbiome is an emerging therapeutic paradigm, particularly for obesity, diabetes, and inflammatory bowel diseases (2).

Through complex interactions with host metabolism, the gut microbiome both influences and is influenced by the obesity phenotype (3–5). For example, *Firmicutes* bacteria are more abundant at the expense of *Bacteroides* bacteria, both as a result of a high fat/high sugar diet (6) and as a consequence of host genetic obesity due to leptin deficiency (7). However, microbiome-targeted therapeutic efforts for obesity have been hampered by a lack of understanding of the interactions between host genetics and the microbiome, as well as the complexity and diversity of the microbiome (2). Recently, *Christensenella minuta* (the first described member of the *Christensenellaceae* family [8]) was described as being extremely highly heritable and as promoting a lean host phenotype through an unknown biological mechanism, a finding which was experimentally verified using transplantation techniques in germ-free mice (3).

In order to facilitate further research into identifying the biological mechanisms underlying the important role this organism plays in obesity prevention, here we describe the first genome sequence of *Christensenella minuta* DSM 22607^T.

Genomic DNA was obtained from the DSMZ repository and sequenced using the Illumina HiSeq 2000, to a depth of 115×. *De novo* assembly of the genome was conducted using the One Button Velvet assembly pipeline version 1.1.06 (9). Gene annotation was performed using both *ab initio* and evidence-based (BLAST) predictions. Coding sequences were predicted using GeneMark and Glimmer3 (10, 11). Intergenic regions not identified by GeneMark and Glimmer3 were searched by BLAST in NCBI's nonredundant bacterial (NR) database. The best prediction for each open reading frame was selected by evaluating all predictions against the best evidence (NR and Pfam [12]) and resolving overlaps between adjacent coding genes. tRNA genes were determined using tRNAscan-SE (13) and noncoding RNA genes by RNAmmer (14) and Rfam (15). We performed a screen for core genes, as defined by the HMP project (16), on many of the assemblies to test for completeness of the genome. Metabolic pathways and subcellular localization were predicted using KEGG (17) and psortB (18), respectively, and functional domains were evaluated using Interproscan (19) (used to infer gene ontology [GO] terms [20, 21]) and Pfam (22). A total of 2,487 genes (79.7% of all genes) had some predicted functional annotation, with 66.1% being assigned to

at least one of 949 unique GO term annotations, 74.4% containing at least one of 2,822 unique InterPro domains, and 69.5% containing at least one of 1,412 unique Pfam domains.

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number LSZW00000000. The version described in this paper is the first version, LSZW01000000.

ACKNOWLEDGMENTS

This work was funded by the National Institutes of Health (NIH), Human Microbiome Project 5U54HG004968 (M.M.). We declare no conflicts of interest.

REFERENCES

1. Rodríguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, Avershina E, Rudi K, Narbad A, Jenmalm MC, Marchesi JR, Collado MC. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. Microb Ecol Health Dis 26:26050. https://doi.org/10.3402/mehd.v26.26050.
2. Rajpal DK, Brown JR. 2013. Modulating the human gut microbiome as an emerging therapeutic paradigm. Sci Prog 96:224–236. https://doi.org/10.3184/003685013X13691404141587.
3. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT, Spector TD, Clark AG, Ley RE. 2014. Human genetics shape the gut microbiome. Cell 159:789–799. https://doi.org/10.1016/j.cell.2014.09.053.
4. Karlsson FH, Tremaroli V, Nookaew I, Bergström G, Behre CJ, Fagerberg B, Nielsen J, Bäckhed F. 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature 498:99–103. https://doi.org/10.1038/nature12198.
5. Musso G, Gambino R, Cassader M. 2011. Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. Annu Rev Med 62:361–380. https://doi.org/10.1146/annurev-med-012510-175505.
6. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A 102:11070–11075. https://doi.org/10.1073/pnas.0504978102.
7. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. 2008. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. Cell Host Microbe 3:213–223. https://doi.org/10.1016/j.chom.2008.02.015.
8. Morotomi M, Nagai F, Watanabe Y. 2012. Description of *Christensenella minuta* gen. nov., sp. nov., isolated from human faeces, which forms a distinct branch in the order *Clostridiales*, and proposal of *Christensenellaceae* fam. nov. Int J Syst Evol Microbiol 62:144–149. https://doi.org/10.1099/ijs.0.026989-0.
9. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
10. Borodovsky M, Mills R, Besemer J, Lomsadze A. 2003. Prokaryotic gene prediction using GeneMark and GeneMark.hmm. Curr Protoc Bioinformatics Chapter 4, Unit 4.5. https://doi.org/10.1002/0471250953.bi0405s01.
11. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
12. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A. 2008. The Pfam protein families database. Nucleic Acids Res 36:D281–D288. https://doi.org/10.1093/nar/gkm960.
13. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.0955.
14. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
15. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res 33:D121–D124. https://doi.org/10.1093/nar/gki081.
16. Human Microbiome Project Consortium. 2012. A framework for human microbiome research. Nature 486:215–221. https://doi.org/10.1038/nature11209.
17. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res 32:D277–D280. https://doi.org/10.1093/nar/gkh063.
18. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–1615. https://doi.org/10.1093/bioinformatics/btq249.
19. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. Nucleic Acids Res 33:W116–W120. https://doi.org/10.1093/nar/gki442.
20. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. Nat Genet 25:25–29. https://doi.org/10.1038/75556.
21. Gene Ontology Consortium, Blake JA, Dolan M, Drabkin H, Hill DP, Li N, Sitnikov D, Bridges S, Burgess S, Buza T, McCarthy F, Peddinti D, Pillai L, Carbon S, Dietze H, Ireland A, Lewis SE, Mungall CJ, Gaudet P, Chrisholm RL, Fey P, Kibbe WA, Basu S, Siegele DA, McIntosh BK, Renfro DP, Zweifel AE, Hu JC, Brown NH, Tweedie S, Alam-Faruque Y, Apweiler R, Auchinchloss A, Axelsen K, Bely B, Blatter M-, Bonilla C, Bouguerleret L, Boutet E, Breuza L, et al. 2013. Gene ontology annotations and resources. Nucleic Acids Res 41:D530–D535. https://doi.org/10.1093/nar/gks1050.
22. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. 2014. Pfam: the protein families database. Nucleic Acids Res 42:D222–D230. https://doi.org/10.1093/nar/gkt1223.