Complete Genome Sequence of *Akkermansia glycaniphila* Strain Pyt^T, a Mucin-Degrading Specialist of the Reticulated Python Gut

Janneke P. Ouwerkerk,[a] Jasper J. Koehorst,[b] Peter J. Schaap,[b] Jarmo Ritari,[d] Lars Paulin,[c] Clara Belzer,[a] Willem M. de Vos[a,d,e]

Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands[a]; Systems and Synthetic Biology, Wageningen University, Wageningen, The Netherlands[b]; Institute of Biotechnology, University of Helsinki, Helsinki, Finland[c]; Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland[d]; RPU Immunobiology, Department of Bacteriology & Immunology, University of Helsinki, Helsinki, Finland[e]

ABSTRACT *Akkermansia glycaniphila* is a novel *Akkermansia* species that was isolated from the intestine of the reticulated python and shares the capacity to degrade mucin with the human strain *Akkermansia muciniphila* Muc^T. Here, we report the complete genome sequence of strain Pyt^T of 3,074,121 bp. The genomic analysis reveals genes for mucin degradation and aerobic respiration.

The ability to grow on mucin as a sole carbon and nitrogen source is a distinctive feature of the human isolate *Akkermansia muciniphila* Muc^T (1). *A. glycaniphila* strain Pyt^T is the second *Akkermansia* species described to be mucolytic in pure culture (2). Based on 16S rRNA gene sequences, the *Akkermansia* genus has recently been associated with the intestine or feces from a wide range of mammals and nonmammals, such as birds, fish, and reptiles, such as the Burmese python (3, 4). Here, we describe the complete genome sequence of *Akkermansia glycaniphila* strain Pyt^T, an anaerobic mucin-degrading specialist isolated from the intestine of the reticulated python (*Malayopython reticulatus*).

Total DNA of strain Pyt^T was extracted using the MasterPure Gram-positive DNA purification kit (Epicentre). Single-molecule sequencing was performed using a PacBio RSII instrument at the Institute of Biotechnology (University of Helsinki, Finland). Assembly was performed with PacBio SMRT Analysis pipeline version 2.2 and the HGAP protocol (5). Default settings were used, except for the following adaptations: minimum subread length, 500; minimum polymerase read length quality, 0.80; minimum seed read length, 7,000; split target into chunks, 1; alignment candidate per chunk, 24; genome size, 3,000,000; target coverage, 30; overlapper error rate, 0.06; overlapper mini length, 40; and overlapper k-mer, 14. Annotation was carried out with an in-house pipeline consisting of Prodigal version 2.5 for prediction of protein-coding DNA sequences (CDSs) (6), InterProScan 5RC7 for protein annotation (7), tRNAscan-SE version 1.3.1 for prediction of tRNAs (8), and RNAmmer version 1.2 for prediction of rRNAs (9). Additional protein function predictions were derived via BLAST identifications against the UniRef50 (10) and Swiss-Prot (11) databases (download August 2013). Subsequently, the annotation was further enhanced by adding EC numbers via PRIAM version 2013-03-06 (12). Noncoding RNAs were identified using rfam_scan.pl version 1.04 on release 11.0 of the RFAM database (13).

The Pyt^T genome is composed of a single chromosome of 3,074,121 bp, with a G+C content of 57.6%. It contains 2,532 CDSs, all 21 tRNA genes, and three complete rRNA operons. For 72% (1,811) of the coding sequences, a function could be predicted. Analysis of the genome revealed the presence of numerous mucin-degrading enzymes,

of which many are predicted to be secreted: 54 glycoside hydrolases, one glycosyl hydrolase, seven sialidases, and three sulfatases. The Pyt^T genome is predicted to encode a cytochrome *bd* ubiquinol oxidase, indicating the potential for aerobic respiration. This bacterium might use this in the oxic-anoxic interface of its probable habitat: the intestinal mucin layer, as recently has been experimentally verified for *A. muciniphila* Muc^T (14).

The availability of this genome will enhance the understanding of metabolic and physiological properties of members of the genus *Akkermansia* present in the intestinal tract of *Animalia* in general.

Accession number(s). The complete genome sequence of *Akkermansia glycaniphila* strain Pyt^T was deposited at EMBL-EBI under the accession number LT629973.

REFERENCES

1. Derrien M, Vaughan EE, Plugge CM, de Vos WM. 2004. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. Int J Syst Evol Microbiol 54:1469–1476. https://doi.org/10.1099/ijs.0.02873-0.

2. Ouwerkerk JP, Aalvink S, Belzer C, de Vos WM. 5 August 2016. *Akkermansia glycaniphila* sp. nov.: an anaerobic mucin-degrading bacterium isolated from reticulated python faeces. Int J Syst Evol Microbiol. https://doi.org/10.1099/ijsem.0.001399.

3. Belzer C, de Vos WM. 2012. Microbes inside—from diversity to function: the case of *Akkermansia*. ISME J 6:1449–1458. https://doi.org/10.1038/ismej.2012.6.

4. Costello EK, Gordon JI, Secor SM, Knight R. 2010. Postprandial remodeling of the gut microbiota in Burmese pythons. ISME J 4:1375–1385. https://doi.org/10.1038/ismej.2010.71.

5. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.

6. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.

7. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY. 2012. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res 40:D306–D312. https://doi.org/10.1093/nar/gkr948.

8. 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.

9. 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.

10. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics 23:1282–1288. https://doi.org/10.1093/bioinformatics/btm098.

11. UniProt Consortium. 2014. Activities at the universal protein resource (UniProt). Nucleic Acids Res 42:D191–D198. https://doi.org/10.1093/nar/gkt1140.

12. Claudel-Renard C, Chevalet C, Faraut T, Kahn D. 2003. Enzyme-specific profiles for genome annotation: PRIAM. Nucleic Acids Res 31:6633–6639. https://doi.org/10.1093/nar/gkg847.

13. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, Eddy SR, Gardner PP, Bateman A. 2013. Rfam 11.0: 10 years of RNA families. Nucleic Acids Res 41:D226–D232. https://doi.org/10.1093/nar/gks1005.

14. Ouwerkerk JP, van der Ark KC, Davids M, Claassens NJ, Robert Finestra T, de Vos WM, Belzer C. 23 September 2016. Adaptation of *Akkermansia muciniphila* to the oxic-anoxic interface of the mucus layer. Appl Environ Microbiol. https://doi.org/10.1128/AEM.01641-16.