



Complete Genome Sequence and Annotation for *Turicibacter sanguinis* MOL361^T (DSM 14220)

Joel J. Maki,^{a,b,c} Daniel W. Nielsen,^{a,c} Torey Looft^a

^aFood Safety and Enteric Pathogens Research Unit, National Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Ames, Iowa, USA

^bInterdepartmental Microbiology Graduate Program, Iowa State University, Ames, Iowa, USA

^cOak Ridge Institute for Science and Education, ARS Research Participation Program, Oak Ridge, Tennessee, USA

ABSTRACT *Turicibacter sanguinis* MOL361 (DSM 14220) is the reference and type strain for the *Turicibacter* genus, commonly found in the intestinal tract of animal species. Long-read sequencing was performed on this strain to complement publicly available Illumina HiSeq-based data, producing a complete annotated genome sequence.

Turicibacter is a genus of the *Firmicutes* phylum, commonly identified in the intestine of animals (1–8). *Turicibacter* species are capable of interaction with host-derived compounds, suggesting that they contribute to physiological processes (9–13). Here, we describe the complete genome sequence of *Turicibacter sanguinis* MOL361^T.

Turicibacter sanguinis MOL361 is the type strain for the type species of *Turicibacter* and was sequenced in 2019 as part of the Department of Energy (DOE) Joint Genome Institute project Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase IV: the One Thousand Microbial Genomes (KMG IV). The incomplete draft genome was uploaded to NCBI (assembly accession number [GCA_004338625.1](https://www.ncbi.nlm.nih.gov/assembly/GCA_004338625.1)). To complete the *T. sanguinis* genome, long-read sequence data were generated for assembly with existing short-read data. *T. sanguinis* DSM 14220 was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ) and grown anaerobically at 37°C on brain heart infusion (BHI) broth prior to sequencing. Genomic DNA was isolated using the PureLink genomic DNA extraction minikit (Invitrogen, Waltham, MA) according to the manufacturer's instructions. The extracted DNA was assessed using gel electrophoresis, a Qubit fluorimeter (double-stranded DNA [dsDNA] broad-range [BR] kit; Life Technologies, Waltham, MA), and a NanoDrop instrument (Thermo Fisher Scientific, Waltham, MA) prior to sequencing. The genomic library for Nanopore sequencing was prepared with the SQK-RBK004 rapid barcoding kit (Oxford Nanopore, Oxford, UK) according to the manufacturer's instructions.

Default parameters were used for all software unless otherwise noted. A FLO-MIN106 (R9.4.1) flow cell was used on a MinION device (Oxford Nanopore) for sequencing of the *T. sanguinis* MOL361^T genomic DNA. The flow cell was run for 48 h, and reads were quality (Q) scored (Q ≥ 7), demultiplexed, and trimmed with Guppy v3.1.5 (14). A total of 66,759 MinION reads (genomic coverage, 30×) with a mean length of 2,632 bp (N_{50} , 4,949 bp) were generated. Paired-end Illumina HiSeq reads (2 × 150 bp; genomic coverage, 240×) were downloaded from GenBank's Sequence Read Archive (SRA) (accession number [PRJNA500327](https://www.ncbi.nlm.nih.gov/sra/PRJNA500327)) using the SRA Toolkit (15). Illumina and Nanopore reads were assembled into 2 contigs with Unicycler v0.4.7 in "bold" mode (16). Geneious Prime v2019.1.1 (Biomatters, Ltd., Auckland, New Zealand) was used to map the Illumina HiSeq and MinION reads to both contigs using the "map to reference" command with "medium sensitivity" settings. The smaller (1,400-bp) contig was present in the Illumina reads but not in the Nanopore reads. A BLASTn search of the NCBI

Citation Maki JJ, Nielsen DW, Looft T. 2020. Complete genome sequence and annotation for *Turicibacter sanguinis* MOL361^T (DSM 14220). *Microbiol Resour Announc* 9:e00475-20. <https://doi.org/10.1128/MRA.00475-20>.

Editor Frank J. Stewart, Georgia Institute of Technology

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Torey Looft, torey.looft@usda.gov.

Received 12 May 2020

Accepted 24 May 2020

Published 18 June 2020

nucleotide database identified the small contig as similar (>97% identity) to DNA sequences originating from several ungulates, a *Clostridium botulinum* genome, an *Actinoalloteichus* sp. genome, and a *Babesia bigemina* genome. This observation, and the Geneious mapping, suggested that this second contig was a contaminant, and it was removed from the final assembly.

The remaining circular contig was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.11 (17). The annotated chromosome for *Turicibacter sanguinis* MOL36^T was 2,999,687 bp long with 34.4% G+C content. *Turicibacter sanguinis* MOL36^T has 2,752 protein-coding sequences, with 43 pseudogenes and 121 tRNA genes.

Data availability. This genome sequence has been deposited in GenBank under the accession number [CP053187](https://doi.org/10.1093/nar/gkq1019). The Nanopore reads are available in the NCBI SRA under accession number [SRR11185522](https://doi.org/10.1093/nar/gkq1019).

ACKNOWLEDGMENTS

MinION sequencing was performed at the NADC in Ames, Iowa. We thank Markus Göker (DSMZ) for permitting the use of his Illumina sequencing reads and Stefan Spring (DSMZ) for curating *T. sanguinis* MOL361^T. We thank David Alt (USDA) for sequencing support.

This research was supported by USDA Agricultural Research Service (USDA-ARS) funds. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

REFERENCES

- Bosshard PP, Zbinden R, Altwegg M. 2002. *Turicibacter sanguinis* gen. nov., sp. nov., a novel anaerobic, Gram-positive bacterium. *Int J Syst Evol Microbiol* 52:1263–1266. <https://doi.org/10.1099/00207713-52-4-1263>.
- Verbarq S, Göker M, Scheuner C, Schumann P, Stackebrandt E. 2014. The families Erysipelotrichaceae emend., Coprobacillaceae fam. nov., and Turicibacteraceae fam. nov. p 79–105. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), *The prokaryotes: Firmicutes and Tenericutes*. Springer, Berlin, Germany. https://doi.org/10.1007/978-3-642-30120-9_205.
- Cuiv PÓ, Klaassens ES, Durkin AS, Harkins DM, Foster L, McCarrison J, Torralba M, Nelson KE, Morrison M. 2011. Draft genome sequence of *Turicibacter sanguinis* PC909, isolated from human feces. *J Bacteriol* 193:1288–1289. <https://doi.org/10.1128/JB.01328-10>.
- Köhler T, Dietrich C, Scheffrahn RH, Brune A. 2012. High-resolution analysis of gut environment and bacterial microbiota reveals functional compartmentation of the gut in wood-feeding higher termites (*Nasutitermes* spp.). *Appl Environ Microbiol* 78:4691–4701. <https://doi.org/10.1128/AEM.00683-12>.
- Looft T, Allen HK, Cantarel BL, Levine UY, Bayles DO, Alt DP, Henrissat B, Stanton TB. 2014. Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J* 8:1566–1576. <https://doi.org/10.1038/ismej.2014.12>.
- Auchtung TA, Holder ME, Gesell JR, Ajami NJ, Duarte RTD, Itoh K, Caspi RR, Petrosino JF, Horai R, Zárate-Bladés CR. 2016. Complete genome sequence of *Turicibacter* sp. strain H121, isolated from the feces of a contaminated germ-free mouse. *Genome Announc* 4:e00114-16. <https://doi.org/10.1128/genome.A.00114-16>.
- Kumar S, Chen C, Indugu N, Werlang GO, Singh M, Kim WK, Thippareddi H. 2018. Effect of antibiotic withdrawal in feed on chicken gut microbial dynamics, immunity, growth performance and prevalence of foodborne pathogens. *PLoS One* 13:e0192450. <https://doi.org/10.1371/journal.pone.0192450>.
- Johnson TA, Sylte MJ, Looft T. 2019. In-feed bacitracin methylene disalicylate modulates the turkey microbiota and metabolome in a dose-dependent manner. *Sci Rep* 9:8212. <https://doi.org/10.1038/s41598-019-44338-5>.
- Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, Spector TD, Bell JT, Clark AG, Ley RE. 2016. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* 19:731–743. <https://doi.org/10.1016/j.chom.2016.04.017>.
- BenSassi N, Vas J, Vasdal G, Averos X, Estevez I, Newberry RC. 2019. On-farm broiler chicken welfare assessment using transect sampling reflects environmental inputs and production outcomes. *PLoS One* 14:e0214070. <https://doi.org/10.1371/journal.pone.0214070>.
- Fung TC, Vuong HE, Luna CDG, Pronovost GN, Aleksandrova AA, Riley NG, Vavilina A, McGinn J, Rendon T, Forrest LR, Hsiao EY. 2019. Intestinal serotonin and fluoxetine exposure modulate bacterial colonization in the gut. *Nat Microbiol* 4:2064–2073. <https://doi.org/10.1038/s41564-019-0540-4>.
- Kemis JH, Linke V, Barrett KL, Boehm FJ, Traeger LL, Keller MP, Rabaglia ME, Schueler KL, Stapleton DS, Gatti DM, Churchill GA, Amador-Noguez D, Russell JD, Yandell BS, Broman KW, Coon JJ, Attie AD, Rey FE. 2019. Genetic determinants of gut microbiota composition and bile acid profiles in mice. *PLoS Genet* 15:e1008073. <https://doi.org/10.1371/journal.pgen.1008073>.
- Gao H, Jiang Q, Ji H, Ning J, Li C, Zheng H. 2019. Type 1 diabetes induces cognitive dysfunction in rats associated with alterations of the gut microbiome and metabolomes in serum and hippocampus. *Biochim Biophys Acta Mol Basis Dis* 1865:165541. <https://doi.org/10.1016/j.bbdis.2019.165541>.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network based-calling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
- Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database Collaboration. 2011. The Sequence Read Archive. *Nucleic Acids Res* 39:D19–D21. <https://doi.org/10.1093/nar/gkq1019>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- 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>.