genomeA
Journals.ASM.org

Draft Genome Sequence of *Catellicoccus marimammalium*, a Novel Species Commonly Found in Gull Feces

Michael R. Weigand,[a] Hodon Ryu,[c] Laura Bozcek,[c] Konstantinos T. Konstantinidis,[a,b] Jorge W. Santo Domingo[c]

School of Civil and Environmental Engineering[a] and School of Biology,[b] Center for Bioinformatics and Computational Genomics, Georgia Institute of Technology, Atlanta, Georgia, USA; National Risk Management Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio, USA[c]

Catellicoccus marimammalium **is a relatively uncharacterized Gram-positive facultative anaerobe with potential utility as an indicator of waterfowl fecal contamination. Here, we report an annotated draft genome sequence that suggests that this organism may be a symbiotic gut microbe.**

Received 27 October 2012 **Accepted** 21 November 2012 **Published** 7 February 2013

Citation Weigand MR, Ryu H, Bozcek L, Konstantinidis KT, Santo Domingo JW. 2013. Draft genome sequence of *Catellicoccus marimammalium*, a novel species commonly found in gull feces. Genome Announc. 1(1):e00019-12. doi:10.1128/genomeA.00019-12.

Address correspondence to Jorge W. Santo Domingo, santodomingo.jorge@epa.gov.

Catellicoccus marimammalium was first isolated from the carcass of a porpoise that had succumbed to severe enteritis and peritonitis along the coast of Scotland (1). A low-G+C Gram-positive catalase-negative facultative anaerobe, *C. marimammalium* is closely related to, but biochemically distinct from, other genera within the order *Lactobacillales* (1). Recently, 16S rRNA gene-based surveys and species-specific PCR assays have shown high host distribution and an abundance of species closely related to *C. marimammalium* in different waterfowl (2–4). Migratory waterfowl facilitate the spread of microbial pathogens and are a notable source of fecal contamination in recreational waters. Thus, *C. marimammalium*-based methods are relevant to public health and to the environmental monitoring of waterfowl fecal contamination (3–5). However, the study of *C. marimammalium* remains limited, as isolating this species from environmental waters has not been possible using media that support the growth of other facultative anaerobic lactobacilli.

To shed light on the metabolic potential of *C. marimammalium* (strain CCUG 49459 T; Culture Collection, University of Göteborg, Sweden), genome sequence data were generated by a combination of Illumina HiSeq 2000 and Roche 454 platforms. All sequence reads were filtered and trimmed at both the 5′ and 3′ ends based on a threshold of $Q = 20$. Passed Illumina reads were assembled first in parallel runs of Velvet v.1.0.13 (6) with a range of k-mer coverage values (7). The resulting contigs were pooled and then assembled as pseudoreads in conjunction with passed 454 reads using the Genome Sequencer (GS) *de novo* Assembler software v.2.0.01.14 (Roche). This hybrid approach yielded 31 large contigs (≥500 bp each) with an N_{50} size of 153,794 bp, an average length of 41,619 bp, and a maximum length of 308,478 bp. The resulting genome sequence is 1,290,194 bases with 235-fold coverage and has an average G+C content of 42.0%. Annotation by the Rapid Annotations using Subsystems Technology (RAST) server (8), tRNAscan-SE (9), and RNAmmer (10) revealed 1,200 protein-coding regions, 48 tRNA genes, and a single rRNA operon.

Core-genome phylogeny confirmed the placement of *C. marimammalium* within the order *Lactobacillales*, where it forms a unique clade, which is consistent with previous 16S gene phylogenetic analyses (1). Preliminary reconstruction using Pathway Tools v.16 (11) revealed an overall reduced metabolic network with a particularly limited capacity for *de novo* amino acid biosynthesis compared to other closely related organisms. Rather, the genome of *C. marimammalium* encodes functions commonly found in symbiotic gut bacteria, such as bile acid hydrolysis and specialized nutrient transport. Taken together, the reduced genome and metabolic network suggest a symbiotic lifestyle that likely underlies the difficulty of growing *C. marimammalium* in synthetic media. The frequent detection of *C. marimammalium*-like sequences in different waterfowl feces (2–5) suggests that the physiology of this bacterial group is better adapted to waterfowl gut environment than to the gut of other potential hosts. The genome sequence reported here will aid in future studies of *C. marimammalium* to elucidate the nature of its gut symbiosis and to develop additional molecular tools for tracking fecal contamination in recreational waters.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AMYT00000000. The version described in this article is the first version, AMYT01000000.

ACKNOWLEDGMENTS

We thank Eugene Rice for discussions and Jill Hoelle for providing technical assistance.

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded and managed, or partially funded and collaborated in, the research described herein; this work has been subjected to the Agency's administrative review and has been approved for external publication.

Any opinions expressed in this article are those of the authors and do not necessarily reflect the views of the Agency; therefore, no official endorsement should be inferred. Any mention of trade names or commercial products does not constitute endorsement or recommendation for use.

REFERENCES

1. **Lawson PA, Collins MD, Falsen E, Foster G.** 2006. *Catellicoccus marimammalium* gen. nov., sp. nov., a novel Gram-positive, catalase-negative, coccus-shaped bacterium from porpoise and grey seal. Int. J. Syst. Evol. Microbiol. **56**:-429–432.
2. **Lu J, Santo Domingo JW, Lamendella R, Edge T, Hill S.** 2008. Phylogenetic diversity and molecular detection of bacteria in gull feces. Appl. Environ. Microbiol. **74**:3969–3976.
3. **Ryu H, Griffith JF, Khan IU, Hill S, Edge TA, Toledo-Hernandez C, Gonzalez-Nieves J, Santo Domingo J.** 2012. Comparison of gull feces-specific assays targeting the 16S rRNA genes of *Catellicoccus marimammalium* and streptococcus spp. Appl. Environ. Microbiol. **78**: 1909–1916.
4. **Ryu H, Lu J, Vogel J, Elk M, Chávez-Ramírez F, Ashbolt N, Santo Domingo J.** 2012. Development and evaluation of a quantitative PCR assay targeting sandhill crane (*Grus canadensis*) fecal pollution. Appl. Environ. Microbiol. **78**:4338–4345.
5. **Green HC, Dick LK, Gilpin B, Samadpour M, Field KG.** 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. Appl. Environ. Microbiol. **78**: 503–510.
6. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
7. **Luo C, Tsementzi D, Kyrpides NC, Konstantinidis KT.** 2012. Individual genome assembly from complex community short-read metagenomic datasets. ISME J. **6**:898–901.
8. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
9. **Schattner P, Brooks AN, Lowe TM.** 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. **33**:W686–W689.
10. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35**:3100–3108.
11. **Karp PD, Paley SM, Krummenacker M, Latendresse M, Dale JM, Lee TJ, Kaipa P, Gilham F, Spaulding A, Popescu L, Altman T, Paulsen I, Keseler IM, Caspi R.** 2010. Pathway tools version 13.0: integrated software for pathway/genome informatics and systems biology. Brief. Bioinform. **11**:40–79.