

Complete Genome Sequences of Fish Pathogenic *Weissella ceti* Strains WS74 and WS105

H. C. P. Figueiredo,^a C. A. G. Leal,^a F. A. Dorella,^a A. F. Carvalho,^a S. C. Soares,^a F. L. Pereira,^a V. A. C. Azevedo^b

AQUACEN, National Reference Laboratory for Aquatic Animal Diseases, Ministry of Fisheries and Aquaculture, Federal University of Minas Gerais, Belo Horizonte, Brazil^a; Laboratory of Cellular and Molecular Genetics, Institute for Biological Science, Federal University of Minas Gerais, Belo Horizonte, Brazil^b

We describe here the genome sequencing and annotation of *Weissella ceti* strains WS74 and WS105, isolated from diseased rainbow trout in Brazil. The two genomes were sequenced with an Ion Torrent personal genome machine (PGM) using a fragment library. The genomes of strains WS74 and WS105 consist of circular chromosomes 1,389,513 bp and 1,390,396 bp long, respectively, both presenting a G+C content of 40.75%.

Received 3 September 2014 Accepted 10 September 2014 Published 16 October 2014

Citation Figueiredo HCP, Leal CAG, Dorella FA, Carvalho AF, Soares SC, Pereira FL, Azevedo VAC. 2014. Complete genome sequences of fish pathogenic *Weissella ceti* strains WS74 and WS105. *Genome Announc.* 2(5):e01014-14. doi:10.1128/genomeA.01014-14.

Copyright © 2014 Figueiredo 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 H. C. P. Figueiredo, figueiredoh@yahoo.com.

The *Weissella* species belong to the group of lactic acid bacteria (LAB) and have been isolated from a broad variety of nutrient-rich environments (1) and animals (2–6). Recently, the first species associated with diseased fish and beaked whales was reported, named *Weissella ceti* (7, 8). *W. ceti* has been isolated from outbreaks of hemorrhagic disease in rainbow trout in Brazil (9, 10). Although genome data for *W. ceti* strains WS08 and NC36 exist, comparative genome analyses for discovering new virulence factors are partially impaired by the draft status of *W. ceti* NC36 (11). We sequenced *W. ceti* strains WS74 and WS105 isolated from rainbow trout in Brazil during different outbreaks and years and presenting distinct pulsed-field gel electrophoresis (PFGE) types (9).

W. ceti strains WS74 and WS105 were thawed, streaked onto Man Rogosa and Sharpe (MRS) agar (Sigma-Aldrich, USA), and incubated at 25°C for 48 h. The DNA was extracted with a Maxwell 16 DNA purification kit using a Maxwell 16 instrument (Promega, USA). Extracted DNA molecules were quantified with a Qubit 2.0 fluorometer (Life Technologies, USA), and the DNA integrity was assessed using TapeStation (GE, USA). The DNA was sequenced in a personal genome machine (PGM) Ion Torrent sequencing system (Life Technologies, Carlsbad, CA) using a fragment library with an Ion PGM 200 sequencing kit.

The genome sequencing of *W. ceti* WS74 yielded genome coverage of ~210×. An initial assembly of WS74 with Mira 4.0 (12) and Newbler 2.9 resulted in 120 and 19 contigs with N_{50} s of ~68 kb and ~131 kb, respectively. For the genome of *W. ceti* WS105, genome coverage of ~296× was obtained, and the initial assembly with Mira and Newbler resulted in 473 and 20 contigs with N_{50} s of ~24 kb and ~162 kb, respectively. For both strains, the software programs CONTIGuator 2.0 (13) and FGAP 1.7 (14) were used for recursive and reference-based (WS08 strain) overlapping of contigs. The remaining gaps were solved using recursive and reference-based read mapping with CLC Genomics Workbench version 7.0 (CLC bio, Germantown, MD).

Automatic gene and pseudogene prediction and annotation

were performed using the software PROKKA version 1.7 (15) and RAST server version 4.0 (16). All putative frameshifts were manually curated based on the coverage and quality of the given base using genome mapping. The genome of *W. ceti* WS74 comprises 1,389,513 bp with a G+C content of 40.75% and contains 1,438 genes, which represent 1,338 coding sequences (CDSs), 3 pseudogenes, 1 tmRNA, 77 tRNAs, and 19 rRNAs. The genome of *W. ceti* WS105 has 1,390,396 bp with a G+C content of 40.75% and harbors 1,433 genes, which are composed of 1,336 CDSs, 1 tmRNA, 71 tRNAs, and 18 rRNAs. The genome sequences of the 2 strains of *W. ceti* described here open new doors for genome comparison of this bacterium, which will help in elucidating the mechanisms that drove the emergence of an LAB as a new fish pathogen.

Nucleotide sequence accession numbers. The whole-genome projects of *W. ceti* strains WS74 and WS105 have been deposited at Genbank under the accession numbers CP009223 and CP009224, respectively.

ACKNOWLEDGMENTS

This research was supported by the Ministry of Fisheries and Aquaculture-MPA, the National Center for Animal Information-INCT/CNPq/UFGM, and FINEP. We also thank FAPEMIG and CAPES.

REFERENCES

- Björkroth KJ, Schillinger U, Geisen R, Weiss N, Hoste B, Holzapfel WH, Korkeala HJ, Vandamme P. 2002. Taxonomic study of *Weissella confusa* and description of *Weissella cibaria* sp. nov., detected in food and clinical samples. *Int. J. Syst. Evol. Microbiol.* 52:141–148.
- Flaherty JD, Levett PN, Dewhirst FE, Troe TE, Warren JR, Johnson S. 2003. Fatal case of endocarditis due to *Weissella confusa*. *J. Clin. Microbiol.* 41:2237–2239. <http://dx.doi.org/10.1128/JCM.41.5.2237-2239.2003>.
- Lawhon SD, Lopez FR, Joswig A, Black HC, Watts AE, Norman TE, Porter BF. 2014. *Weissella confusa* septicemia in a foal. *J. Vet. Diagn. Invest.* 26:150–153. <http://dx.doi.org/10.1177/1040638713510516>.
- Lee MR, Huang YT, Liao CH, Lai CC, Lee PI, Hsueh PR. 2011. Bacteraemia caused by *Weissella confusa* at a university hospital in Taiwan, 1997–2007. *Clin. Microbiol. Infect.* 17:1226–1231. <http://dx.doi.org/10.1111/j.1469-0691.2010.03388.x>.
- Lee W, Cho SM, Kim M, Ko YG, Yong D, Lee K. 2013. *Weissella confusa*

- bacteremia in an immune-competent patient with underlying intramural hematomas of the aorta. *Ann. Lab Med* 33:459–462. <http://dx.doi.org/10.3343/alm.2013.33.6.459>.
6. Vela AI, Porrero C, Goyache J, Nieto A, Sánchez B, Briones V, Moreno MA, Domínguez L, Fernández-Garayzábal JF. 2003. *Weissella confusa* infection in primate (*Cercopithecus mona*). *Emerg. Infect. Dis.* 9:1307–1309. <http://dx.doi.org/10.3201/eid0910.020667>.
 7. Liu JY, Li AH, Ji C, Yang WM. 2009. First description of a novel *Weissella* species as an opportunistic pathogen for rainbow trout *Oncorhynchus mykiss* (Walbaum) in China. *Vet. Microbiol.* 136:314–320. <http://dx.doi.org/10.1016/j.vetmic.2008.11.027>.
 8. Vela AI, Fernández A, de Quirós YB, Herráez P, Domínguez L, Fernández-Garayzábal JF. 2011. *Weissella ceti* sp. nov., isolated from beaked whales (*Mesoplodon bidens*). *Int. J. Syst. Evol. Microbiol.* 61: 2758–2762. <http://dx.doi.org/10.1099/ijs.0.028522-0>.
 9. Costa FA, Leal CA, Schuenker ND, Leite RC, Figueiredo HC. 24 March 2014. Characterization of *Weissella ceti* infections in Brazilian rainbow trout, *Oncorhynchus mykiss* (Walbaum), farms and development of an oil-adjuvanted vaccine. *J. Fish Dis.* 2014. <http://dx.doi.org/10.1111/jfd.12236>.
 10. Figueiredo HC, Costa FA, Leal CA, Carvalho-Castro GA, Leite RC. 2012. *Weissella* sp. outbreaks in commercial rainbow trout (*Oncorhynchus mykiss*) farms in Brazil. *Vet. Microbiol.* 156:359–366. <http://dx.doi.org/10.1016/j.vetmic.2011.11.008>.
 11. Ladner JT, Welch TJ, Whitehouse CA, Palacios GF. 2013. Genome sequence of *Weissella ceti* NC36, an emerging pathogen of farmed rainbow trout in the United States. *Genome Announc.* 1(1):e00187–12. <http://dx.doi.org/10.1128/genomeA.00187-12>.
 12. Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. Proceedings of the German Conference on Bioinformatics, GCB '99. GCB, Hannover, Germany.
 13. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A. 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. Source code. *Biol. Med.* 6:11. <http://dx.doi.org/10.1186/1751-0473-6-11>.
 14. Piro VC, Faoro H, Weiss VA, Steffens MB, Pedrosa FO, Souza EM, Raittz RT. 2014. FGAP: an automated gap closing tool. *BMC Res. Notes* 7:371. <http://dx.doi.org/10.1186/1756-0500-7-371>.
 15. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.
 16. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42: D206–D214. <http://dx.doi.org/10.1093/nar/gkt1226>.