



Draft Genome Sequence of *Weissella confusa* MBF8-1, a Glucansucrase- and Bacteriocin-Producing Strain Isolated from a Homemade Soy Product

Nicholas C. K. Heng,^a Chia-Wen Yeh,^a Amarila Malik^b

Sir John Walsh Research Institute, Faculty of Dentistry, University of Otago, Dunedin, New Zealand^a; Division of Pharmaceutical Microbiology and Biotechnology, Faculty of Pharmacy, Universitas Indonesia, Depok, Indonesia^b

ABSTRACT We report here the draft genome sequence of *Weissella confusa* MBF8-1, an isolate from a homemade fermented soybean product that produces sucrases and exhibits antibacterial (bacteriocin) activity. The draft genome of *W. confusa* MBF8-1 comprises a 2.2-Mbp chromosome and a 17.8-kbp bacteriocin-encoding plasmid. Two putative glucansucrase genes were also identified.

Weissella confusa (formerly *Lactobacillus confusus*) belongs to the *Lactobacillus-Leuconostoc* branch of the diverse Gram-positive lactic acid bacteria (LAB) group (1). Many LAB are important starter cultures in the fermentation of milk and food substrates and have long been a focus of attention for their production capabilities of either exopolysaccharides or antimicrobial peptides (1). *W. confusa* can be found in a variety of fermented food products and produces various types of novel exopolysaccharides that could be exploited for probiotic applications (1, 2). On the other hand, *W. confusa* is also regarded as an opportunistic pathogen and has been implicated in causing sepsis and bacteremias in humans (2).

W. confusa MBF8-1, a strain isolated from a homemade fermented soybean product in Tangerang, Indonesia, produces multiple exopolysaccharide (EPS) types (3). Homopolymeric EPS, such as glucans and fructans, are synthesized by glucosyltransferases (glucansucrases) and fructosyltransferases (fructansucrases), respectively (4). A previous investigation revealed that *W. confusa* MBF8-1 possesses at least two glucansucrase-encoding (*gtf*) genes in its genome (3). In addition to the production of EPS, *W. confusa* MBF8-1 produces weissellicin MBF, a narrow-spectrum bacteriocin (5, 6). Weissellicin MBF is plasmid encoded, and its biosynthetic locus resides on pWcMBF8-1 (17,643 bp), the first completely sequenced plasmid from *W. confusa* (6). These traits make MBF8-1 an attractive candidate for development as a probiotic strain and hence the rationale for sequencing its genome.

The *W. confusa* MBF8-1 genome was sequenced by a whole-genome shotgun strategy using a Roche GS-FLX+ pyrosequencer (1st BASE, Jakarta, Indonesia). A total of 60,537,716 bp (~28-fold coverage), generated from 149,271 quality-filtered sequence reads (average read length, 406 bp), was assembled using *gsAssembler* version 2.3 (Roche) and *MIRA* version 4.0 (7). The resulting contigs were initially annotated using the Rapid Annotations using Subsystems Technology (RAST) server (8), ordered with reference to the *W. confusa* LBAE C39-2 and *Weissella ceti* WS105 genome sequences (9, 10), and finally annotated by the Prokaryote Genome Annotation Pipeline (PGAP [11]) for deposition in NCBI/GenBank.

The *W. confusa* MBF8-1 draft genome sequence totals 2,199,279 bp comprising 44 chromosomal contigs (2,181,636 bp, with an average G+C content of 44.7% and N_{50}

Received 7 November 2016 Accepted 22 November 2016 Published 26 January 2017

Citation Heng NCK, Yeh C-W, Malik A. 2017. Draft genome sequence of *Weissella confusa* MBF8-1, a glucansucrase- and bacteriocin-producing strain isolated from a homemade soy product. *Genome Announc* 5:e01497-16. <https://doi.org/10.1128/genomeA.01497-16>.

Copyright © 2017 Heng et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Nicholas C. K. Heng, nicholas.heng@otago.ac.nz, or Amarila Malik, amarila.malik@ui.ac.id.

value of 95,287 bp) and plasmid pWcMBF8-1 (17,643 bp). PGAP detected 2,088 coding sequences, five rRNA operons, and 81 tRNA genes in the MBF8-1 chromosome. Using the partial amino acid sequences of glucansucrases Gtf8-1A and Gtf8-1B (3), two complete open reading frames, *gtf8-1A* and *gtf8-1B*, respectively, were identified. A 1,769-amino-acid protein is encoded by *gtf8-1A*, and *gtf8-1B* encodes a 1,652-amino-acid protein. Both Gtf8-1A and Gtf8-1B are predicted to be exported by the secretory (Sec)-dependent pathway (12) and are distinct from the *W. confusa* LBAE C39-2 dextransucrase (13). Gtf8-1A is homologous to GtfKg3 of *Lactobacillus fermentum* Kg3 (14), and Gtf8-1B is more similar to GtfG of *Streptococcus gordonii* (15).

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession no. [MNBZ00000000](https://doi.org/10.1093/nar/gkz000). The version described in this paper is the first version, MNBZ01000000. The GenBank accession number of plasmid pWcMBF8-1 is [KR350502](https://doi.org/10.1093/nar/gkz000).

ACKNOWLEDGMENTS

This work was supported by TWAS Basic Science Research Grants RGA: 10-088 RG/BIO/AS_G-UNESCO, FR: 3240246003 (2011), and RGA: 14-094 RG/BIO/AS_G-UNESCO. FR: 324028606 (2015) was awarded to A. Malik. C.-W. Yeh was supported by an E. K. Crow Summer Scholarship (Division of Health Sciences, University of Otago).

REFERENCES

- Fusco V, Quero GM, Cho GS, Kabisch J, Meske D, Neve H, Bockelmann W, Franz CMAP. 2015. The genus *Weissella*: taxonomy, ecology and biotechnological potential. *Front Microbiol* 6:155. <https://doi.org/10.3389/fmicb.2015.00155>.
- Fairfax MR, Lephart PR, Salimnia H. 2014. *Weissella confusa*: problems with identification of an opportunistic pathogen that has been found in fermented foods and proposed as a probiotic. *Front Microbiol* 5:254. <https://doi.org/10.3389/fmicb.2014.00254>.
- Malik A, Radji M, Kralj S, Dijkhuizen L. 2009. Screening of lactic acid bacteria from Indonesia reveals glucansucrase and fructansucrase genes in two different *Weissella confusa* strains from soya. *FEMS Microbiol Lett* 300:131–138. <https://doi.org/10.1111/j.1574-6968.2009.01772.x>.
- van Hijum SA, Kralj S, Ozimek LK, Dijkhuizen L, van Geel-Schutten IG. 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol Mol Biol Rev* 70:157–176. <https://doi.org/10.1128/MMBR.70.1.157-176.2006>.
- Sari R, Anita C, Radji M, Malik A. 2011. Bacteriocin screening of lactic acid bacteria strains from local isolates of genus *Streptococcus* and *Weissella*. *Indo J Pharm Sci* 9:116–121.
- Malik A, Sumayyah S, Yeh CW, Heng NCK. 2016. Identification and sequence analysis of pWcMBF8-1, a bacteriocin-encoding plasmid from the lactic acid bacterium *Weissella confusa*. *FEMS Microbiol Lett* 363:fnw059. <https://doi.org/10.1093/femsle/fnw059>.
- Chevreaux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB '99. GCB, Hannover, Germany.
- 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. <https://doi.org/10.1186/1471-2164-9-75>.
- Amari M, Laguerre S, Vuillemin M, Robert H, Loux V, Klopp C, Morel S, Gabriel B, Remaud-Siméon M, Gabriel V, Moulis C, Fontagné-Faucher C. 2012. Genome sequence of *Weissella confusa* LBAE C39-2, isolated from a wheat sourdough. *J Bacteriol* 194:1608–1609. <https://doi.org/10.1128/JB.06788-11>.
- Figueiredo HC, Leal CA, Dorella FA, Carvalho AF, Soares SC, Pereira FL, Azevedo VA. 2014. Complete genome sequences of fish pathogenic *Weissella ceti* strains WS74 and WS105. *Genome Announc* 2(5):e01014–14. <https://doi.org/10.1128/genomeA.01014-14>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufu S, Li W. 2013. Prokaryotic genome annotation pipeline. *In* The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>.
- Amari M, Arango LF, Gabriel V, Robert H, Morel S, Moulis C, Gabriel B, Remaud-Siméon M, Fontagné-Faucher C. 2013. Characterization of a novel dextransucrase from *Weissella confusa* isolated from sourdough. *Appl Microbiol Biotechnol* 97:5413–5422. <https://doi.org/10.1007/s00253-012-4447-8>.
- Kralj S, van Geel-Schutten GH, Dondorff MM, Kirsanovs S, van der Maarel MJ, Dijkhuizen L. 2004. Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology* 150:3681–3690. <https://doi.org/10.1099/mic.0.27321-0>.
- Vickerman MM, Sulavik MC, Nowak JD, Gardner NM, Jones GW, Clewell DB. 1997. Nucleotide sequence analysis of the *Streptococcus gordonii* glucosyltransferase gene, *gtfG*. *DNA Seq* 7:83–95. <https://doi.org/10.3109/10425179709020155>.