



Draft Genome Sequence of *Leuconostoc citreum* CW28 Isolated from Pozol, a Pre-Hispanic Fermented Corn Beverage

Clarita Olvera,^a Rosa I. Santamaría,^b Patricia Bustos,^b Cristina Vallejo,^a Juan J. Montor,^a Carmen Wachter,^c Agustín Lopez Munguia^a

Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico^a; Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico^b; Facultad de Química, Universidad Nacional Autónoma de México, CDMX, Mexico^c

ABSTRACT *Leuconostoc citreum* CW28 was isolated from pozol, a Mayan fermented corn beverage. This strain produces a cell-associated inulosucrase, the first described in bacteria. Its draft genome sequence, announced here, has an estimated size of 1.98 Mb and harbors 1,915 coding genes, 12 rRNAs, 68 tRNAs, 17 putative pseudogenes, and 1 putative phage.

Pozol is a nonalcoholic lactic acid beverage of Mayan origin obtained from the natural fermentation of nixtamal (heat- and alkali-treated maize). After lime treatment, the corn kernels are dehulled, washed, and ground, and the resulting dough is shaped into balls that are wrapped in banana leaves, which are fermented at room temperature for 2 days (1). The predominant bacterial groups during all stages of pozol fermentation are lactic acid bacteria (LAB), including *Lactobacillus*, *Weissella*, and *Leuconostoc* spp. (2). It has been demonstrated that some of the LAB involved in pozol fermentation may have probiotic properties and produce oligosaccharides (OS) and polysaccharides with potential biotechnological applications.

Leuconostoc citreum CW28, a strain isolated from pozol, produces high-molecular-weight (HMW) inulin from sucrose through a cell-associated multidomain inulosucrase, IsIA (EC 2.4.1.9). The HMW inulin may be directly used as soluble fiber or hydrolyzed to obtain fructooligosaccharides (FOS) and oligofructose (3, 4). Until now, 9 genomes of *L. citreum* have been reported, including the complete genome of strain KM20, isolated from kimchi (5); draft genomes of strains LBAE E16, LBAE C10, and LBAE C11, isolated from wheat sourdough (6); strains NRRL B-742 and B-1299 (7); strain DmW_111, isolated from wild *Drosophila* (8); and two strains, 1300_LCIT and 1301_LGAS, isolated from intensive care units (9).

L. citreum CW28 was grown in MRS medium at 30°C (4). The genomic DNA was prepared using the phenol-chloroform method and sequenced in an Illumina HiSeq 2000 sequencer at Macrogen (Seoul, South Korea). Paired-end reads of 101 bp were assembled using SPAdes 3.5.0 (10), SSPACE-Basic 2.0 (11), and Consed v.23 (12), with a coverage of 80×. The contigs obtained were aligned to the nearest reference genome (*Leuconostoc citreum* KM20) using the NUCmer program (13). Open reading frames (ORFs) were predicted with Glimmer 3.0 (14); manual annotations were carried out in Artemis 12.0 (15), with comparisons with the GenBank (16), Interpro (17), and ISfinder (<http://www-is.biotoul.fr>) databases. The draft genome of *L. citreum* CW28 contains nine contigs, with a total length of 1,982,921 bp and 38.7% G+C content. A total of 1,915 protein-coding sequences (CDSs), 12 rRNAs, 68 tRNAs, and 17 putative pseudogenes were identified. The PHAST server (18) and manual annotation revealed one putative intact prophage of 34.8 kb with 37 CDSs. An analysis of similarity among the *L. citreum*

Received 16 October 2017 Accepted 23 October 2017 Published 30 November 2017

Citation Olvera C, Santamaría RI, Bustos P, Vallejo C, Montor JJ, Wachter C, Lopez Munguia A. 2017. Draft genome sequence of *Leuconostoc citreum* CW28 isolated from pozol, a pre-Hispanic fermented corn beverage. Genome Announc 5:e01283-17. <https://doi.org/10.1128/genomeA.01283-17>.

Copyright © 2017 Olvera 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 Agustín Lopez Munguia, agustin@ibt.unam.mx.

genomes, measured as average nucleotide identities (ANI), shows that the genome of the *L. citreum* CW28 strain is more similar to B-1299, B-742, and LBAE E16 (99.42, 99.41, and 99.30% similar, respectively) than to KM20, LBAE C11, and LBAE C10 (98.57, 98.47, and 98.33% similar, respectively). Extensive genome analysis of this and other *L. citreum* strains will facilitate the identification of new genes encoding enzymes involved in polymer and OS synthesis with potential biotechnological applications.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [MWJP0000000](https://doi.org/10.1128/MWJP0000000). The version described in this paper is the first version, MWJP01000000.

ACKNOWLEDGMENTS

This work was funded by CB-2013-CONACyT 219728, PAPIIT-UNAM IN223917, and FONCICYT-BMBF-267620.

We thank Maria Elena Rodriguez and Fernando Gonzalez for technical support for the maintenance of the laboratory strain collection.

REFERENCES

- Ben Omar N, Ampe F. 2000. Microbial community dynamics during production of the Mexican fermented maize dough pozol. *Appl Environ Microbiol* 66:3664–3673. <https://doi.org/10.1128/AEM.66.9.3664-3673.2000>.
- Ampe F, Ben Omar N, Moizan C, Wachter C, Guyot J-P. 1999. Polyphasic study of the spatial distribution of microorganisms in Mexican pozol, a fermented maize dough, demonstrates the need for cultivation-independent methods to investigate traditional fermentations. *Appl Environ Microbiol* 65:5464–5473.
- Olivares-Illana V, Wachter-Odarte C, Le Borgne S, López-Munguía A. 2002. Characterization of a cell-associated inulosucrase from a novel source: a *Leuconostoc citreum* strain isolated from pozol, a fermented corn beverage of Mayan origin. *J Ind Microbiol Biotechnol* 28:112–117.
- Olivares-Illana V, López-Munguía A, Olvera C. 2003. Molecular characterization of inulosucrase from *Leuconostoc citreum*: a fructosyltransferase within a glucosyltransferase. *J Bacteriol* 185:3606–3612. <https://doi.org/10.1128/JB.185.12.3606-3612.2003>.
- Kim JF, Jeong H, Lee JS, Choi S-H, Ha M, Hur C-G, Kim JS, Lee S, Park H-S, Park Y-H, Oh TK. 2008. Complete genome sequence of *Leuconostoc citreum* KM20. *J Bacteriol* 190:3093–3094. <https://doi.org/10.1128/JB.01862-07>.
- Laguette S, Amari M, Vuillemin M, Robert H, Loux V, Klopp C, Morel S, Gabriel B, Remaud-Siméon M, Gabriel V, Moulis C, Fontagné-Faucher C. 2012. Genome sequences of three *Leuconostoc citreum* strains, LBAE C10, LBAE C11, and LBAE E16, isolated from wheat sourdoughs. *J Bacteriol* 194:1610–1611. <https://doi.org/10.1128/JB.06789-11>.
- Passerini D, Vuillemin M, Laguerre S, Amari M, Loux V, Gabriel V, Robert H, Morel S, Monsan P, Gabriel B, Fontagné-Faucher C, Remaud-Siméon M, Moulis C. 2014. Complete genome sequence of *Leuconostoc citreum* strain NRRL B-742. *Genome Announc* 2(6):e01179-14. <https://doi.org/10.1128/genomeA.01179-14>.
- Wright SM, Carroll C, Walters A, Newell PD, Chaston JM. 2017. Genome sequence of *Leuconostoc citreum* DmW_111, isolated from wild *Drosophila*. *Genome Announc* 5(24):e00507-17. <https://doi.org/10.1128/genomeA.00507-17>.
- Roach DJ, Burton JN, Lee C, Stackhouse B, Butler-Wu SM, Cookson BT, Shendure J, Salipante SJ. 2015. A year of infection in the intensive care unit: prospective whole-genome sequencing of bacterial clinical isolates reveals cryptic transmissions and novel microbiota. *PLoS Genet* 11:e1005413. <https://doi.org/10.1371/journal.pgen.1005413>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Boetzer M, Henkel CV, Jansen HJ, Butler B, Pirovano W. 2011. Scaffolding preassembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* 8:195–202. <https://doi.org/10.1101/gr.8.3.195>.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with Glimmer. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. *Nucleic Acids Res* 37:D26–D31. <https://doi.org/10.1093/nar/gkn723>.
- Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MD, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJ, Zdobnov EM. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* 29:37–40. <https://doi.org/10.1093/nar/29.1.37>.
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <https://doi.org/10.1093/nar/gkr485>.