

Genome Sequence of *Clostridium paraputrificum* 373-A1 Isolated in Chile from a Patient Infected with *Clostridium difficile*

Enzo Guerrero-Araya,^{a,b} Angela Plaza-Garrido,^{a,b} Fernando Díaz-Yañez,^a Marjorie Pizaro-Guajardo,^{a,b} Sandro L. Valenzuela,^b Claudio Meneses,^{c,d} Fernando Gil,^a Eduardo Castro-Nallar,^b Daniel Paredes-Sabja^{a,b}

Microbiota-Host Interactions and Clostridia Research Group, Departamento de Ciencias Biológicas, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Santiago, Chile^a; Center for Bioinformatics and Integrative Biology, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Santiago, Chile^b; Centro de Biotecnología Vegetal, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Santiago, Chile^c; FONDAF Center for Genome Regulation, Santiago, Chile^d

***Clostridium paraputrificum* is a gut microbiota member reported in several cases of bacteremia and coinfections. So far, only one genome sequence of a *C. paraputrificum* (AGR2156) isolate is available. Here, we present the draft genome of *C. paraputrificum* strain 373-A1, isolated from stools from a patient with *C. difficile* infection.**

Received 6 September 2016 Accepted 12 September 2016 Published 3 November 2016

Citation Guerrero-Araya E, Plaza-Garrido A, Díaz-Yañez F, Pizaro-Guajardo M, Valenzuela SL, Meneses C, Gil F, Castro-Nallar E, Paredes-Sabja D. 2016. Genome sequence of *Clostridium paraputrificum* 373-A1 isolated in Chile from a patient infected with *Clostridium difficile*. *Genome Announc* 4(6):e01178-16. doi:10.1128/genomeA.01178-16.

Copyright © 2016 Guerrero-Araya 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 Eduardo Castro-Nallar, eduardo.castro@unab.cl, or Daniel Paredes-Sabja, daniel.paredes.sabja@gmail.com.

Clostridium paraputrificum is a Gram-positive, endospore-forming, strictly anaerobic, and chitinolytic bacterium (1). Few studies have addressed features of *C. paraputrificum*. Early work associated *C. paraputrificum* with an increased risk of colon cancer (2). *C. paraputrificum* has also been isolated from healthy individuals (3), as well as from cases of bacteremia (4) and patients with AIDS (5). *C. paraputrificum* contributes to colonic epithelium maturation and development (6); however, its role in infectious diseases remains unclear.

Genomic data for strain 373-A1 was generated using Illumina technology (7) on an Illumina MiSeq platform (600 incorporated cycles; 2 × 300 bp). Library preparation was carried out according to the TruSeq DNA kit, generating 3,165,742 paired-end reads and 1.47 GB of information. Next, raw data were filtered (>Q20) and merged (minimum overlapping of 20 bp) in PEAR version 0.9.8 (8). The resulting set of reads was re-cleaned using Pathoscope version 2.0.6 (9) in order to obtain only the reads mapping against bacterial genomes. Data were assembled using SPAdes version 3.5.0 (10) with an auto cutoff coverage. The final assembly contained 41 contigs, and an N_{50} equal to 259,466 bp with a mean of coverage of 472×.

The genome annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline (11) and this annotation was saved in GenBank (MAPZ01). The assembly of the draft genome sequence consists of 41 contigs amounting to 3,488,595 bp with a G+C content of 29.8%. Of the 3,351 predicted genes, 3,220 were protein-coding genes, and 106 RNAs and 24 rRNA operons were identified. The majority of the protein-coding genes (60.19%) were assigned a putative function, while the remaining ones were annotated as hypothetical proteins.

For a comparison, the assembly report ASM42402v1 (*C. paraputrificum* AGR2156) was used. The genome of *C. paraputrificum* 373-A1 is smaller than *C. paraputrificum* AGR2156 (i.e., 3,488,595 bp versus 3,561,289 bp, respectively). *C. paraputrificum* 373-A1 had fewer predicted genes and genes encoding

putative proteins (i.e., 3,351 genes and 3,220 proteins) than *C. paraputrificum* AGR2156 (i.e., 3,457 genes and 3,345 proteins). On the other hand, the G+C content was similar in both strains, with 29.8% and 29.6% for *C. paraputrificum* 373-A1 and *C. paraputrificum* AGR2156, respectively. When we focus on the number of contigs and N_{50} values, *C. paraputrificum* 373-A1 has 41 contigs with an N_{50} value of 259,466, in contrast with *C. paraputrificum* AGR2156, which has 32 contigs with an N_{50} value of 320,228, which indicates the high quality of our genome sequence.

This draft sequence provides a new repertoire of genes and genome information for this strain of the opportunistic *C. paraputrificum*.

Accession number(s). The *C. paraputrificum* 373-A1 genome has been deposited in GenBank under the accession number [MAPZ00000000](https://www.ncbi.nlm.nih.gov/nuclink/117816).

ACKNOWLEDGMENTS

This work was supported by grants from Comisión Nacional de Ciencia y Tecnología de Chile (FONDECYT grant 1151025), from the Research Office of Universidad Andres Bello (DI-641-15/R 2015) to D.P.-S. E.C.-N. was funded by “CONICYT + PAI/Concurso nacional apoyo al retorno de investigadores(as) desde el extranjero, conovatoria 2014 + Folio 82140008.” We acknowledge FONDEQUIP EQM 150104 for funding the Illumina sequencer.

FUNDING INFORMATION

This work, including the efforts of Daniel Paredes-Sabja, was funded by Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) (1151025 and Fondecip 2015 EQM 150104).

REFERENCES

- Šimůnek J, Koppová I, Tiščenko G, Dohnálek J, Dušková J. 2012. Excretome of the chitinolytic bacterium *Clostridium paraputrificum* 14. *Folia Microbiol (Praha)* 57:335–339. <http://dx.doi.org/10.1007/s12223-012-0137-2>.

2. Horie H, Kanazawa K, Okada M, Narushima S, Itoh K, Terada A. 1999. Effects of intestinal bacteria on the development of colonic neoplasm: an experimental study. *Eur J Cancer Prev* 8:237–245. <http://dx.doi.org/10.1097/00008469-199906000-00012>.
3. Šimůnek J, Kopecný J, Hodrová B, Bartonová H. 2002. Identification and characterization of *Clostridium paraputrificum*, a chitinolytic bacterium of human digestive tract. *Folia Microbiol (Praha)* 47:559–564. <http://dx.doi.org/10.1007/BF02818798>.
4. Woo PCY, Lau SKP, Chan K-M, Fung AMY, Tang BSF, Yuen K-Y. 2005. *Clostridium bacteriaemia* characterised by 16S ribosomal RNA gene sequencing. *J Clin Pathol* 58:301–307. <http://dx.doi.org/10.1136/jcp.2004.022830>.
5. Shinha T, Hadi C. 2015. *Clostridium paraputrificum* bacteremia associated with colonic necrosis in a patient with AIDS. *Case Rep Infect Dis* 2015:312919. <http://dx.doi.org/10.1155/2015/312919>.
6. Cherbuy C, Honvo-Houeto E, Bruneau A, Bridonneau C, Mayeur C, Duée P-H, Langella P, Thomas M. 2010. Microbiota matures colonic epithelium through a coordinated induction of cell cycle-related proteins in gnotobiotic rat. *Am J Physiol Gastrointest Liver Physiol* 299:G348–G357. <http://dx.doi.org/10.1152/ajpgi.00384.2009>.
7. Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438. <http://dx.doi.org/10.1517/14622416.5.4.433>.
8. Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics* 30:614–620. <http://dx.doi.org/10.1093/bioinformatics/btt593>.
9. Francis OE, Bendall M, Manimaran S, Hong C, Clement NL, Castornallar E, Snell Q, Schaalje GB, Clement MJ, Crandall KA, Johnson WE. 2013. Pathoscope: species identification and strain attribution with unassembled sequencing data. *Genome Res* 23:1721–1729. <http://dx.doi.org/10.1101/gr.150151.112>.
10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshtkin 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. <http://dx.doi.org/10.1089/cmb.2012.0021>.
11. 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. <http://dx.doi.org/10.1093/nar/gkw569>.