



Draft Genome Sequence of *Adlercreutzia equolifaciens* IPLA 37004, a Human Intestinal Strain That Does Not Produce Equol from Daidzein

Lucía Vázquez,^{a,b} Ana Belén Flórez,^{a,b} Baltasar Mayo^{a,b}

^aDepartamento de Microbiología y Bioquímica, Instituto de Productos Lácteos de Asturias (IPLA-CSIC), Villaviciosa, Asturias, Spain

^bInstituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Asturias, Spain

ABSTRACT Equol is an intestinal bacterial metabolite derived from the isoflavone daidzein and has beneficial health effects. Most equol producers belong to the family *Coriobacteriaceae*, which includes species such as *Adlercreutzia equolifaciens*. Here, we report the draft genome sequence of *A. equolifaciens* IPLA 37004, a human isolate that does not produce equol.

Among isoflavones and their metabolites, the bacterial daidzein-derived compound (S)-(–)-equol has the strongest estrogenic activity (1) and antioxidant action (2), by which it can positively influence human health (3). However, only 30 to 60% of humans harbor equol-producing microbes in their intestines. Most equol-producing bacteria belong to the family *Coriobacteriaceae* of the phylum *Actinobacteria*, which includes species such as *Adlercreutzia equolifaciens*, *Asaccharobacter celatus*, *Enterorhabdus mucosicola*, *Slackia isoflavoniconvertens*, and *Slackia equolifaciens* (3). However, whether equol production is a family-, species-, or strain-specific characteristic is not yet known.

Strain IPLA 37004 was isolated from fecal samples from a woman with an equol-producing microbiota, using agar plates of Gifu anaerobic medium (GAM) (Nissui) supplemented with 5 g liter⁻¹ arginine (GAM-Arg; Merck) and cultivation under strict anaerobic conditions. Sequencing and sequence comparison of the 16S rRNA gene of IPLA 37004 identified the strain as *A. equolifaciens*. Surprisingly, the strain did not produce equol when cultured in GAM-Arg with isoflavones, daidzein, or dihydrodaidzein. To get insights regarding its non-equol-producing phenotype, the genome of *A. equolifaciens* IPLA 37004 was sequenced, and its sequence was analyzed.

After growth of the strain in GAM-Arg, total genomic DNA from IPLA 37004 was extracted and purified using the DNeasy blood and tissue kit (Qiagen), as suggested by the manufacturer. For genome sequencing, a library was constructed using SPRIworks fragment library system I (Beckman Coulter), according to the manufacturer's instructions. Paired-end sequencing (2 × 150-bp reads) was performed at Eurofins Genomics on a HiSeq sequencer (Illumina) (sequence mode, NovaSeq 6000; S2 flow cell; paired-end 150-bp reads; Xp flowchart). The average genome coverage approached 100×, yielding a total of 5,568,276 bp of adapter-free and quality-filtered reads after the use of Trimmomatic (4) and FastQC (5) software packages. *De novo* assembly of quality-filtered nontrimmed reads in contigs was accomplished using SPAdes v3.6.2 (6), setting the parameters with a k value of 125 or 127 and only-assembler. Then, the NCBI Prokaryotic Genome Annotation Pipeline (7) was used to predict and annotate the open reading frames. Whole-genome comparisons were achieved by using the multiple genome alignment software package Mauve v2.3.1 (8). Finally, BLAST v2.10.0 searches (BLASTp suite, NCBI) were used to determine homology between proteins of different strains.

The genome of *A. equolifaciens* IPLA 37004 consisted of 2,664,741 bp in 169 contigs

Citation Vázquez L, Flórez AB, Mayo B. 2020. Draft genome sequence of *Adlercreutzia equolifaciens* IPLA 37004, a human intestinal strain that does not produce equol from daidzein. *Microbiol Resour Announc* 9:e01537-19. <https://doi.org/10.1128/MRA.01537-19>.

Editor David Rasko, University of Maryland School of Medicine

Copyright © 2020 Vázquez 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 Baltasar Mayo, baltasar.mayo@ipla.csic.es.

Received 16 December 2019

Accepted 1 February 2020

Published 20 February 2020

longer than 200 bp, with a G+C content of 63.7%. The maximum contig length was 325,998 bp, and the N_{50} and N_{90} values were 101,827 and 25,297 bp, respectively. A total of 2,310 genes were identified in the IPLA 37004 genome, including 2,251 coding sequences, of which 2,198 encoded complete proteins and 53 were identified as pseudogenes. In addition, 59 genes encoded RNA molecules, including 8 rRNAs, 48 tRNAs, and 3 noncoding RNAs. Genes encoding reductases homologous to those of *A. equolifaciens* DSM 19450^T involved in equol formation (9) were not detected in the IPLA 37004 genome. Whole-genome comparison of DSM 19450^T and IPLA 37004 identified shared flanking genes upstream and downstream of the equol locus in *A. equolifaciens* DSM 19450^T. The deduced amino acids of the shared genes showed an identity range of 80 to 99%. The gene cluster of DSM 19450^T encompassed a region of about 11 kbp, which was absent in IPLA 37004.

Data availability. The genome sequence of *A. equolifaciens* IPLA 37004 was deposited in GenBank under accession number [VJNE00000000](https://www.ncbi.nlm.nih.gov/GenBank/BJNE00000000). The fastq files of the raw reads were deposited in the NCBI SRA under accession number [SRR9657807](https://www.ncbi.nlm.nih.gov/SRA/SRR9657807).

ACKNOWLEDGMENTS

This study was supported by projects from MINECO (grant AGL-2014-57820-R) and Asturias Principality (grant IDI/2018/000114).

REFERENCES

- Selvaraj V, Zakroczymski MA, Naaz A, Mukai M, Ju YH, Doerge DR, Katzenellenbogen JA, Helferich WG, Cooke PS. 2004. Estrogenicity of the isoflavone metabolite equol on reproductive and non-reproductive organs in mice. *Biol Reprod* 71:966–972. <https://doi.org/10.1095/biolreprod.104.029512>.
- Choi EJ, Kim GH. 2014. The antioxidant activity of daidzein metabolites, O-desmethylangolensin and equol, in HepG2 cells. *Mol Med Rep* 9:328–332. <https://doi.org/10.3892/mmr.2013.1752>.
- Mayo B, Vázquez L, Flórez AB. 2019. Equol: a bacterial metabolite from the daidzein isoflavone and its presumed beneficial health effects. *Nutrients* 11:2231. <https://doi.org/10.3390/nu11092231>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov AS, Lesin V, Nikolenko S, Pham S, Prjibelski A, Pyshkin A, Sirotkin A, 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>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
- Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
- Toh H, Oshima K, Suzuki T, Hattori M, Morita H. 2013. Complete genome sequence of the equol-producing bacterium *Adlercreutzia equolifaciens* DSM 19450^T. *Genome Announc* 1:e00742-13. <https://doi.org/10.1128/genomeA.00742-13>.