Genome Sequence Announcement of *Lactobacillus vaginalis* LMG S-26419, Isolated from a Healthy Woman

Edoardo Puglisi,a Serena Galletti,b Valeria Sagheddu,b Marina Ellib

aDepartment for Sustainable Food Processes, Università Cattolica del Sacro Cuore, Piacenza, Italy
bAAT-Advanced Analytical Technologies, Fiorenzuola d'Arda, Piacenza, Italy

ABSTRACT   The *Lactobacillus vaginalis* LMG S-26419 strain, also named CBA-L88 (BV2), was isolated at the AAT-Advanced Analytical Technologies laboratories from a vaginal swab obtained from a healthy woman. The total genome size is 1,806,242 bp with a G+C content of 40.6%.

*L*actobacillus vaginalis was first isolated and described by Embley and colleagues in 1989 (1), and it has been reported to be one of the predominant species colonizing the vaginal microbiota of healthy women (2). However, the onset of bacterial vaginosis significantly reduces the level of *L. vaginalis* (3). In this paper, we report the complete draft genome sequence of *L. vaginalis* strain LMG S-26419, isolated by plating a vaginal swab taken from a healthy woman donor on Rogosa agar medium (BD, Difco, NJ, USA) and incubated under anaerobic conditions at 37°C. Sampling was carried out in compliance with the Helsinki Declaration, and an informed written consent was signed. A single isolated colony on Rogosa agar medium (BD, Difco) was submitted for DNA extraction performed with a Whatman CloneSaver card system (96-well format) (VWR, PA, USA) and processed following the manufacturer's instructions. The P0 (5′-GAGAG TTTGATCCTGGCTCAG-3′) (4) and P2 (5′-GGTGAGTAACGCGTGGGGAA-3′) (5) primer sets were used for the partial amplification of the 16S rRNA gene that was then sequenced (BMR Genomics, Padua, Italy). The nucleotide sequence was used to search the GenBank database (https://www.ncbi.nlm.nih.gov/) using BLAST (6), the blastn algorithm, and the Ribosomal Database Project (7).

For total DNA extraction, a single colony of *L. vaginalis* was inoculated in MRS-Cys (De Man Rogosa Sharpe BD, Difco) supplemented with 0.5 g/liter of cysteine (Sigma-Aldrich, MO, USA) under anaerobic conditions for 18 h at 37°C, and 1 ml of the liquid culture was aliquoted for the DNA extraction using the MasterPure Gram-positive DNA purification kit (Epicentre, Cambio, UK) following the manufacturer's instructions. The genomic library was constructed using the Nano-DNA Illumina kit, and the Illumina MiSeq v3 platform for 2 × 300-bp paired-end reads was used to sequence the complete genome (Fasteris, Switzerland). A total of 1,695,030 high-quality reads (i.e., with a Phred score above 30) were obtained after sequencing and quality filtering. KmerGenie v1.6213 software (8) was applied to determine the most appropriate number of k-mers for the assembly, which was then performed with Velvet v1.2.10 (9). Assembly metrics were evaluated with QUAST v4.0 (10), highlighting an N_{50} value of 47,415 bp and a largest contig of 1,879,874 bp. Assembly resulted in 123 contigs; the genome has a total length of 1,879,874 bp with a coverage of 80× and a G+C content of 40.6%. Rapid Annotations using Subsystems Technology (RAST) (11) functional annotation of the genome resulted in 1,801 coding sequences and 73 combined rRNAs and tRNAs. The genome was interrogated for the presence of antibiotic resistance genes through the ResFinder database (12), and no positive significant hits were found. Bacteriocin analysis was conducted with BAGEL4 (13), which identified a potential region of interest

Address correspondence to Marina Elli, marina.elli@aat-taa.eu.

TABLE 1 Relevant genes of the *L. vaginalis* LMG S-26419 secretome

Category	Gene ID no.[a]	Gene function
Cell wall and capsule	PROKKA_01729	Putative glycosyltransferase EpsH
	PROKKA_00446	Putative sugar transferase EpsL
	PROKKA_00596	Putative glycosyltransferase EpsJ
Adhesion	PROKKA_00005	Putative mucus binding protein
	PROKKA_00115	MucBP domain protein
	PROKKA_00581	MucBP domain protein
	PROKKA_01171	MucBP domain protein
	PROKKA_01173	MucBP domain protein
Glycerol metabolism	PROKKA_01202	Glycerol-3-phosphate dehydrogenase
Lipoteichoic acid metabolism	PROKKA_00017	Lipoteichoic acid synthase
	PROKKA_00368	Lipoteichoic acid synthase
	PROKKA_00431	Lipoteichoic acid synthase
	PROKKA_00571	D-Alanyl-lipoteichoic acid biosynthesis protein DltD
	PROKKA_00573	D-Alanyl-lipoteichoic acid biosynthesis protein DltB
	PROKKA_00804	Lipoteichoic acid protein A
Peroxidase production	PROKKA_00726	NADH oxidase
	PROKKA_00956	NADH flavine reductase

[a] ID, identification.

with two uncharacterized proteins (YqbO and BasT) whose roles as bacteriocins still need to be confirmed.

A BLAST search of the genome sequence of *L. vaginalis* LMG S-26419 against the LAB-Secretome database (14) resulted in 144 positive hits, which indicates a number of features of potential interest listed in Table 1.

Data availability. The whole-genome shotgun project has been deposited in NCBI under the BioProject accession number PRJNA504605 with the corresponding GenBank accession number RJTI00000000 and the SRA accession number SRR8206196.

REFERENCES

1. Embley TM, Faquir N, Bossart W, Collins MD. 1989. *Lactobacillus vaginalis* sp. nov. from the human vagina. Int J Syst Bacteriol 39:368–370. https://doi.org/10.1099/00207713-39-3-368.
2. Vásquez A, Jakobsson T, Ahrné S, Forsum U, Molin G. 2002. Vaginal *Lactobacillus* flora of healthy Swedish women. J Clin Microbiol 40:2746–2749. https://doi.org/10.1128/JCM.40.8.2746-2749.2002.
3. Jespers V, Kyongo J, Joseph S, Hardy L, Cools P, Crucitti T, Mwaura M, Ndayisaba G, Delany-Moretlwe S, Buyze J, Vanham G, van de Wijgert JHHM. 2017. A longitudinal analysis of the vaginal microbiota and vaginal immune mediators in women from sub-Saharan Africa. Sci Rep 7:11974. https://doi.org/10.1038/s41598-017-12198-6.
4. Di Cello F, Fani R. 1996. A molecular strategy for the study of natural bacterial communities by PCR-based techniques. Minerva Biotecnol 8:126–134.
5. Cocolin L, Manzano M, Cantoni C, Comi G. 2001. Denaturing gradient gel electrophoresis analysis of the 16S rRNA gene V1 region to monitor dynamic changes in the bacterial population during fermentation of Italian sausages. Appl Environ Microbiol 67:5113–5121. https://doi.org/10.1128/AEM.67.11.5113-5121.2001.
6. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
7. Maidak BL, Larsen N, McCaughey MJ, Overbeek R, Olsen GJ, Fogel K, Blandy J, Woese CR. 1994. The Ribosomal Database Project. Nucleic Acids Res 22:3485–3487. https://doi.org/10.1093/nar/22.17.3485.
8. Chikhi R, Medvedev P. 2014. Informed and automated *k*-mer size selection for genome assembly. Bioinformatics 30:31–37. https://doi.org/10.1093/bioinformatics/btt310.
9. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
11. 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.
12. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261.
13. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: a user-friendly Web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res 46:W278–W281. https://doi.org/10.1093/nar/gky383.
14. Zhou M, Theunissen D, Wels M, Siezen RJ. 2010. LAB-Secretome: a genome scale comparative analysis of the predicted extracellular and surface-associated proteins of lactic acid bacteria. BMC Genomics 11:651. https://doi.org/10.1186/1471-2164-11-651.