Genome Sequences of the High-Acetic Acid-Resistant Bacteria *Gluconacetobacter europaeus* LMG 18890<sup>T</sup> and *G. europaeus* LMG 18494 (Reference Strains), *G. europaeus* 5P3, and *Gluconacetobacter oboediens* 174Bp2
(Isolated from Vinegar)<sup>V</sup>

Cristina Andrés-Barrao,1 Laurent Falquet,2 Sandra P. Calderon-Copete,2 Patrick Descombes,3 Ruben Ortega Pérez,¹ and François Barja¹*  

Department of Botany and Plant Biology, Microbiology Unit, University of Geneva, Chemin des Embrouichis 10, CH-1254 Jussy-Geneva, Switzerland, Vital-IT, Swiss Institute of Bioinformatics, Quartier Sororge—Bâtiment Genopode, CH-1015 Lausanne, Switzerland,¹ and Genomics Platform, CMU, University of Geneva, Rue Michel-Serret 1, CH-1211 Geneva 4, Switzerland³

Received 17 February 2011/Accepted 11 March 2011

Acetic acid bacteria (AAB) are widespread microorganisms that play an important role in multiple natural processes leading to the production of chemicals of industrial interest and high-value food and beverage products (9). *Gluconacetobacter europaeus* (formerly *Acetobacter europaeus*) is one of the most prominent AAB species isolated from industrial submerged vinegar fermenters, with high resistance to acetic acid (more than 18%) (5, 10). The adaptation to these extreme conditions must be the consequence of genome mutations and rearrangements, but little is known about this genetic instability. The determination of the complete genome sequence for this microorganism is absolutely necessary for the identification of genes and proteins involved in such a high levels of acetic acid adaptation. The genome sequences of other industrially important AAB, such as *Gluconobacter oxydans, Acetobacter pasteurianus,* and *Gluconacetobacter hansenii,* have recently been published (1, 6, 8). The genome sequences of two reference strains of *Gluconacetobacter europaeus,* as well as those of *G. europaeus* and *G. oboediens* strains, isolated from submerged red wine and spirit vinegar, respectively, are herein presented.

Seventy-six paired-end indexed libraries were prepared from purified, ~400-bp DNA fragments using the paired-end DNA sample prep kit (Illumina). Enrichment of GC-rich 400-bp DNA fragments by two-step PCR was then performed using AccuPrime Pfx DNA polymerase (Invitrogen), the multiplexing sequencing primer, and the PhiX control kit v2 (Illumina). Libraries were purified on AMP XP beads (Agencourt). Sequencing was performed on an Illumina Genome Analyzer Ix. The reads were quality controlled with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc) and trimmed on both ends, leaving 60 high-quality nucleotides, by using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit). The contigs were assembled from the trimmed reads using Velvet and ABysS varying k-mers (11, 12). The best assemblies were annotated using an internally developed pipeline (2) and the Priam prediction tool (3). The genomic sequences of the four *Gluconacetobacter* strains were 4 to 4.2 Mb in size with a G+C content of 58 to 59%. Seventy-six to 77% of each genome contained 3,434 to 3,771 coding sequences (CDSs) (348 to 388 hypothetical unique proteins). *G. europaeus* LMG 18890<sup>T</sup> and 5P3 strains had 53 and 60 tRNAs and 4 and 5 rRNA operons, respectively, whereas the third *Gluconacetobacter* strain had 67 tRNAs and 9 rRNAs. *G. oboediens* 174Bp2 had 68 tRNAs and 8 rRNAs. *G. europaeus* LMG 18890<sup>T</sup> had 62 transposases and insertion sequence elements. The other two *G. europaeus* strains had 29 and 33 of these elements, but *G. oboediens* had only 9. Due to the repetitive nature of rRNA and transposases genes and the fact that draft genomes contain a large number of gaps, these numbers might be underestimated but are an indication of the high genetic instability of AAB (1). The four draft genomes shared 90.2 to 95% sequence identity. Several sets of the enzymes for the synthesis of dTDP-rhamnose (*polABCD*), involved in the production of capsular polysaccharides (4), were found in the four strains. They also showed type I and type II cellulose-synthesizing operons (absent in *Acetobacter* sp.), to produce cellulose as the main exopolysaccharide.

**Nucleotide sequence accession numbers.** The draft genome sequences and the original short reads of *G. europaeus* LMG 18890<sup>T</sup>, *G. europaeus* LMG 18494, *G. europaeus* 5P3, and *G. oboediens* 174Bp2 were deposited in the European Nucleotide Archive (EMBL/GenBank/DDBJ) (7) database under study accession numbers ERP000517 (CADP01000001-000321), ERP000516 (CADR01000001-000216), ERP000519 (CDS01000001-000256), and ERP000515 (CADP01000001-000321).
(CADT01000001-000200) and project ID numbers 61321, 61325, 61329, and 61333. Isolates have been publicly deposited in the Belgian Coordinated Collections of Microorganisms (BCCM/LMG) under the following accession numbers: *G. europaeus* 5P3, LMG 26311; *G. oboediens* 174Bp2, LMG 26312.

We thank Marie-Louise Chappuis for her skillful technical assistance.

This research was supported in part by the CRAFT contract 017269 from the European Commission, Academic Society of Geneva, and Department of Botany and Plant Biology of Geneva University.

REFERENCES