

In silico genomic analysis of the endophytic bacterium *Bacillus* *amyloliquefaciens* 629

Sant'Anna, Brena M. M.; Queiroz, Artur L.; Roque, Milton R. A.

Universidade Federal da Bahia, CPqGM/Fiocruz-Bahia, Universidade Federal da Bahia

The endophytic bacteria *Bacillus amyloliquefaciens* 629 isolated from *Theobroma cacao* L. was used to sequencing and genome annotation. Sequencing was performed using NGS Ion Torrent PGM platform (Life Technologies) 318 chip, the genome was assembled using SPAdes Genome Assembler version 3.5.0 and ordered by CONTIGuator 2.3. To evaluate and close gaps were performed manual curation with sequence alignment and editing tools, and also automatic annotation by RAST version 2.0 server. In order to build the phylogenetic tree, conserved sequences of 16S rRNA gene and rpoD were aligned and whole genome was used to DNA-DNA hybridization by ggdc (genome-to-genome distance calculator) and ANI (average nucleotide identity). Complementary analysis was carried out by Tetra Correlation Search (TCS). Here we describe the genome of *Bacillus amyloliquefaciens* 629, with 16 contigs, containing 3,903,367 bp, composed of 4,013 predicted genes, including 3,912 protein-coding sequences, 82 tRNAs and 19 copies of the genes for 5S, 16S, and 23S rRNA. Our comparative genome analysis reveals a new classification to strain 629, in the plantarum subspecies group, of *Bacillus amyloliquefaciens* subsp. plantarum, that shows synonymia with *Bacillus velezensis*. We used a target approach including a plant-host interaction profile, involved in chemotaxis, adhesion, colonization and motility genes in comparative analysis, to understand the essential metabolic systems for endophytic bacteria. The strain 629 showed 132 singletons compared with other strains of *Bacillus velezensis*, 9 genomic islands were predicted and 440 shared genes with other genomes of endophytic bacteria. Some of these regions of the genome establish advantages in endophytism process, to aid in the adaptation and colonization of different environments and host plants.