Complete genome sequence of *Corynebacterium pseudotuberculosis* 33

Viana, MVC1; Parise, D[1]; Sousa, TJ[1]; Benevides, LJ[1]; Mariano, D[1]; Rocha, FS[1]; Bagano, P[1]; Guimaraes, LC[2]; Pereira, FL[1]; Dorella, FA1; Rammos, R[2]; Silva, A[2]; Selim, SAK[3]; Salaheldean, M[3]; Figueiredo, H[1] and Azevedo, V[1]

[1]*Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Minas Gerais, MG;*
[2]*Universidade Federal do Pará, UFPA, Belém, PA;* [3]*Faculty of Veterinary Medicine, Cairo University, Cairo, Giza, Egypt*

Corynebacterium pseudotuberculosis biovar Equi is a Gram-positive, pleomorphic, facultative intracellular pathogen that causes outbreaks of Oedematous Skin Disease (OSD) in buffalo from Egypt. Isolates from this host harbor the diphtheria toxin gene. Sequencing of *C. pseudotuberculosis* genomes isolated from this host will help to understand the pathogenicity of the species, including weather the diphtheria toxin is required for infection. *C. pseudotuberculosis* 33 was isolated from a buffalo diagnosed with OSD in Egypt, in 2008. The genome was sequenced by Ion Torrent platform, using Ion PGM Template OT2 400 Kit and Ion PGM Hi-Q Sequencing Kit. The sequencing resulted in a number of 1,439,326 reads and a total of 250,725,445 bases that were checked for quality by FastQC v0.10.1. A *de novo* assembly was done with Newbler v2.9 and generated 14 contigs with a N50 of 536,189. Scaffolding was done by CONTIGuator v2.7, using *C. pseudotuberculosis* 31 (CP003421.2) as the reference genome. An in-house script established the *dnaA* gene as the beginning of the chromosome. These steps have been integrated in the webtool Simba v1.2.1. The gaps were closed by a reference assembly using CLC Genomics Workbench v6.5 and *C. pseudotuberculosis* 31 as the reference. The genome was annotated using RASTtk. Sequencing errors causing frameshifts were curated by checking for indel errors in the reads, using CLC Genomics Workbench v6.5 and Artemis v16.0.0. The assembled genome has 2,403,550 bp, 52.08 % of GC content, and a mean sequencing coverage depth of 104.11x. The annotation has 2,281 CDSs, 52 tRNAs, and 12 rRNAs. An insertion of 40Kb containing the diphtheria toxin was found, as in *C. pseudotuberculosis* 31. This genome and other buffalo isolates will be used in a comparative genomic study to better understand the pathogenic mechanisms of OSD.