

Identification of genes under positive selection in *Corynebacterium pseudotuberculosis*

Marcus Vinicius Canário Viana¹, Henrique Figueiredo², Felipe Luiz Pereira²,
Fernanda Alves Dorella², Anne Cybelle Pinto Gomide¹, Alice Rebecca Wattam³,
Vasco a de C Azevedo¹,

1 Federal University of Minas Gerais

*2 National Reference Laboratory for Aquatic Animal Diseases of Ministry of Fisheries and
Aquaculture, Federal University of Minas Gerais*

*3 Biocomplexity Institute of Virginia Tech, Virginia Tech, Blacksburg, Virginia, United
States of America*

Abstract

Corynebacterium pseudotuberculosis is a Gram-positive, intracellular pathogen, close related to the diphtheria etiological agents *C. diphtheria* and *C. ulcerans*. The biovar *Ovis* infects mainly small ruminants causing Caseous Lymphadenitis, while biovar *Equi* infects larger animals, causing different diseases. The species virulence mechanisms and biovar differentiation are not fully understood, and drugs and vaccines are not effective for all hosts. The goal of this work is to identify genes under positive selection in *C. pseudotuberculosis* to better understand its evolution and collaborate with the development of control measures. Due to computational costs, 29 strains (16 *Ovis* and 13 *Equi*) from different hosts and countries were selected for a genome scale analysis using the POTION pipeline. We used FastOrtho for ortholog assignment, a cutoff of 50% for sequence identity within ortholog groups, PRANK for sequence alignment, and dnaml for phylogeny reconstruction. For the positive selection analysis, the pairs of null/positive selection site models M1a/M2a and M7/M8 were compared. Eight genes were identified: uncharacterized sigma 70, adhesin, manganese ABC transporter, fatty acid synthase, lambda repressor-like, tyrosine-tRNA ligase, and two uncharacterized transmembrane and secreted proteins. In addition, 14 other genes had evidence of recombination, including sialidase, ferroxidase, adhesin and sortase. These proteins are related to metabolism and host colonization processes as adhesion, metal uptake, protein synthesis and gene regulation. The adaptations provided by the identified mutations will be investigated and a preliminary analysis of protein sequences shows correlation between biovar and specific amino acids.

Funding: CAPES, CNPq, FAPEMIG, UFMG