Title: COMBINED GENOMICS, TRANSCRIPTOMICS AND PROTEOMICS APPROACH REVEALS A HIGH NUMBER OF TRANSCRIBED AND UNTRANSLATED PSEUDOGENES IN *Francisella noatunensis* subsp. *orientalis*

Authors: Tavares, G.C.[1], Carvalho, A.F.[2], Costa, F.A.A.[1], Pereira, F.L.[2], Dorella, F.A.[2], Soares, S.C.[2], Leal, C.A.G.[2], Figueiredo, H.C.P.[2]

Institutions: [1] AQUAVET, Laboratory of Aquatic Animal Diseases (Veterinary School, Federal University of Minas Gerais, Belo Horizonte, MG 30123-970, Brazil), [2] AQUACEN – National Reference Laboratory of Aquatic Animal Diseases, Ministry of Fisheries and Aquaculture (Veterinary School, Federal University of Minas Gerais, Belo Horizonte, MG 30123-970, Brazil).

Abstract:

Francisella noatunensis subsp. *orientalis* (FNO) is an emerging bacterial pathogen that affect Nile tilapia farms around the world. Outbreaks are characterized by a systemic granulomatous infection causing high mortality rates among diseased fish. FNO is a facultative intracellular pathogen and the genome seems to be undergoing a reductive evolution, with a high number of pseudogenes (over three hundred per genome). However, there are no studies that validate the occurrence of this large number of pseudogenes through transcriptomics and proteomics. The aim of this study was to validate the number of the pseudogenes annotated in FNO strain genome through combination of transcriptomic and proteomic analysis. The FNO-12 strain, isolated from diseased fish in Brazil, was selected for this study. This bacterium was cultured in CHAH agar at 28^{o}C for 96 h and thereafter used to inoculate triplicate of MMH broth equilibrated at 28^{o}C for 24 h. Thereafter, two aliquots of each replicate were collected for 2 trials. In the first trial we evaluated whole-genome transcriptomic through microarray-based gene expression analyses with a custom-made Agilent slide formulated based on the genome sequence of FNO-12 strain. The second trial evaluated the protein expression by liquid chromatography separation in nanoAcquity UPLC system. Transcriptome and proteome results were compared to verify if the pseudogenes were transcribed and further translated. The genome of FNO-12 strain contains 363 annotated pseudogenes, and were identified the transcription of 325 of these in microarray trial, in the same time as only 8 related proteins were further translated in proteomic study. Thus, we verified that almost all pseudogenes annotated are really nonfunctional genes. The result support the *in silico* genome analysis, where these 8 pseudogenes maybe should have been wrongly annotated when observed the sequences already deposited. This result also corroborates the fact that the reductive evolution occurs in FNO and that matches with intracellular lifestyle of this pathogen in the fish host.

Keywords: Francisellosis, genome, transcriptome, proteome, pseudogene

Research supported: Ministry of Fisheries and Aquaculture, CNPq, CAPES and FAPEMIG.