Updated TAXI, a taxonomic innovations database depicting operons structure and evolution

Lucas Ferreira[1], José Miguel Ortega[2]

1 SGC - STRUCTURAL GENOMICS CONSORTIUM, UNICAMP
2 UFMG. LABORATÓRIO DE BIODADOS

Abstract

The structure of operons, navigation from one operon to another that shares orthologous genes and the analysis of the clade of origin of the operon and their component genes comprise the information present in TAXI database. TAXI stands for Taxonomic Innovations, being the main goal to understand along evolution how the operons have formed, which one is the most recent and most ancient gene in it, and understanding the funcions that are restricted to anym microbial clade, e.g. a family, a species or even a strain. Update of TAXI database presents a new set of fully indexed tables. The present number of organisms, transcription units, clusters of orthologous genes and genes add up to, respectively: 1753, 3343458, 551692, and 6732117. The most important updates refer to integration of identifiers with external databases UniProt and Kegg. TAXI information can be accessed through an organism of interest, gene symbol, UniProt accession or Kegg Orthologues group (KO). Queries to the database return operons where the most recent or the most ancient gene is in the first position, aiming to facilitate the study of the evolution of regulation in operons. By using the navigation though orthologue groups it is possible to verify the distinct operon compositions of the gene of interest, helping the study of co-functionalities. TAXI database is now available in an instance located at a CPD at bioinfo.icb.ufmg.br/taxi.