

# B-cell epitopes prediction in trypanosomatids genome core

Anderson Coqueiro Santos<sup>1</sup>, Leandro Martins de Freitas<sup>2</sup>

<sup>1</sup>*Institute of Biological Sciences at UFMG and* <sup>2</sup>*Multidisciplinary Institute of Health at UFBA*

The trypanosomatidae family is one part of the phylum euglenozoa including species such as *Leishmania* and *Trypanosoma* spp. There are some potent vector-borne diseases for humans, and others mammals, mainly in subtropical and tropical countries. Leishmaniasis, Chagas diseases, and human African trypanosomiasis (HAT) are some pathologies infecting people worldwide, responsible for approximately 9 million people infected in all world. The observation of orthologous - shared genes present in a common ancestor among species - or paralogous genes - duplicated shared genes - has been one of the main research conducted. All genes shared by different groups of organisms are defined as "genome core". In our study, Trypanosomatids orthologous proteins groups were retrieved from OrthoMCL database. These proteins are conserved in eukaryotic organisms and normally two paralogs sequences were found, as observed in this work for *T. cruzi*, *T. congolense* and *T. brucei*. The orthologous group were submitted to predictions using Bepipred software to find possible B-cell epitopes using a threshold for epitope assignment greater than or equal to 0.35. The statistical analyses were performed using R software (p-value  $\leq 0.05$ ). It was considered as predicted b-cell epitopes each amino acid assigned by the software. The predictions for b-cell resulted in a total 5,865,620 possible epitopes (each single amino acid) with an average of 243 epitopes by proteins. A total of 177,946 b-cell predicted regions (more than 5 amino acids grouped linearly). The regions showed variation in amino acid length, ranging from 5 to 428. Analyzing the proportion of predicted amino acids over the total protein length, we observed that *Leishmania* proteins have a larger number of epitopes when compared with *Trypanosoma* proteins. Some *Leishmania* proteins have more than 80% constituted by b-cell predicted epitopes. When we compared the number and length of b-cell epitopes regions, the *Trypanosoma* species has lower number compared with *Leishmania* species with statistical significance. We also investigated epitopes position and conservation in the aligned sequences in each orthologous group and it was found they are conserved. With this work we observed that *Leishmania* species presented linear epitopes length greater than *Trypanosoma* species. These results point the importance of immune response against these parasites, while *Leishmania* has an obligatory intracellular phase, not being accessible to b-cell, *Trypanosoma* exposed more time during infection could evolve to less b-cell regions. Which may lead to selection of trypanosomes that has a lower number and length of epitopes for b-cell explaining the epitope differences among species.