Determination of BarCodes for the specific identification of different arboviroses

Araújo GLT [1]; Moraes AM [1] ; Queiroz PRM [1]

[1] UniCEUB, Brasília/DF

Arboviruses are viral diseases transmitted by arthropods. There are a variety of this type of infection spread around the world; among them it is possible to mention: Chikungunya, Dengue Fever, Yellow Fever, Mayaro Fever, Zika and West Nile Fever. These diseases have very similar symptoms, a factor that complicates the symptomatic diagnosis of their infections. In addition, the flaviviruses antibodies commonly cross-react in serological tests. After all, these tests are not particularly sensitive or specific. Therefore, molecular tests are a way to eliminate these problems. The implementation of BarCodes, short and highly specific genetic sequences, makes it possible to optimize the identification process in a fast and sensitive way. The objective of this study was to determine BarCodes for the arboviruses mentioned in order to improve their identification methods. Complete genome sequences and pertinent data of the viruses were collected from the NCBI. In "The Sequence Manipulation Suite" an individual enzyme cut profiles were generated and those profiles, through the PAST program, generated a phylogenetic clustering tree. Statistical analyses were performed in the R Program. Bibliographic data was used to select suitable genes for barcode mining and they were aligned using the Geneious Prime program. Checking for any undue modifications in those processes was done in the NCBI BLAST tool and obtaining the potential BarCODES, in the Geneious Prime program. A new BLAST was performed to evaluate the effectiveness of the barcodes that, once determined, were localized through protein modeling performed in the SWISS-MODEL site. Through the phylogenetic tree, clustering profiles, indicative of high stability and conservation of Dengue type 4, Chikungunya and Zika viruses, mainly, were observed. These clustering patterns indicate less tendency to mutations, therefore, potentially greater ease in identifying barcodes in these viral groups. The stability and conservation required for selection of the prospecting genes was vital for determination of the BarCODES. Subsequently, in the genes selected for prospecting, after alignment and checking for inappropriate results, at least two efficient barcodes were determined for identification of the viruses studied. The least promising results were from Dengue 4, Yellow Fever and Zika. Yet, it was possible to perform protein modeling and identify the region where the markers are found in all viruses. Even though they are shorter in length than the standard, the size of the determined identification regions proved to be more appropriate for viral communities. Under these circumstances, the high identity value and coverage indicates efficiency of the defined markers and may represent higher precision, accuracy and specificity than conventional identification methods. The barcodes defined here aim both to provide a means of more efficient identification of the viruses in question, and to generate positive impacts on epidemiology and taxonomy by promoting development of the viral barcoding system that remains stagnant in the scientific community.

Keywords: Arboviruses; BarCode; Viral Identification.

Funding: CNPq