POSTER - PROTEINS AND PROTEOMICS

A PREVIOUSLY UNKNOWN SMALL AMOEBAE VIRUS SHOWS UNPRECEDENTED PROTEOME STRUCTURE

Ana Karoline Da Nóbrega Nunes Alves (anakarolinealves1@hotmail.com)

Jônatas Santos Abrahão (jonatas.abrahao@gmail.com)

Sávio Torres De Farias (stfarias@yahoo.com.br)

Viruses are a type of mobile genetic element that often encode structural proteins encapsidating their own genomes. Usually, they are seen as simple causative agents of disease, but they show exceptional variation and numerous evolutionary strategies. Several mechanisms related to the increase of total coding potential of a viral genome have been proposed, for example, gene overlapping, polyprotein synthesis, frameshift changes, suppression of stop codons, lack of introns, and short or non-existent intergenic regions. Gene overlap represents a great evolutionary challenge, since two functional proteins overlapping can lead to constrained evolution. These mechanisms reflect the plasticity of viruses' genetic information, and ensure the synthesis of all necessary viral proteins. Here we analyze the Yaravirus brasiliense, an amoeba-infecting 80-nm-sized virus with a 45-kbp dsDNA. Almost all of its 74 genes were previously identified as ORFans, which are coding sequences with few or no homologs in public sequence databases, making it harder to know their origin or function. Proteomic analysis of the viral particle revealed 26 proteins. Considering its unprecedented genetic content, we analyzed Yaravirus genome through threading. Using structural molecular modeling, we seek to understand its genetic organization, its predicted proteome, and how it interacts

with its host. Genomic information is published in the NCBI genome database. The gene sequences were inserted in I-TASSER, a threading-based program. For a deeper understanding of possible protein ligands, we used COACH. For protein visualization, we used PyMol and Swiss-Pdb Viewer. When analyzing structural similarities, we observed that 58 of the encoded proteins had the same function as at least one other protein. From these data, we conducted structural alignments between the Yaravirus proteins and similar proteins from the protein data bank. The results showed that the different proteins that presented similarities with the same function were structurally aligned in different parts at the same proteins deposited in the PDB, suggesting that these small proteins could be dimerizing to reestablish the predicted function. We were able to assemble 26 proteins composed of at least two Yaravirus proteins for the same function. These proteins are involved in many biological processes such as central carbon metabolism, DNA metabolism, RNA metabolism, tRNA metabolism, cell wall breakdown, and cell signalization, including a H(+)/K(+)-exchanging ATPase, a histidine kinase, a serine/threonine kinase, an ABC-type xenobiotic transporter, a farnesyl diphosphate synthase, a heptaprenyl diphosphate synthase, a triacylglycerol lipase, a beta-n-acetylhexosaminidase, a acyl-CoA dehydrogenase, a glutaryl-CoA dehydrogenase, a nitrite reductase, an endo-1,4-ß-xylanase, a chitinase, a phosphatidylinositol-4,5-bisphosphate 3-kinase, a porphobilinogen synthase, a ribonucleoside-diphosphate synthase, a serine tRNA ligase, a fumarase, a malate synthase, a PEP-carboxylase, complex IV, an, a RNA-directed RNA-polymerase, a DNA-directed DNA polymerase, a DNA ligase, a DNA topoisomerase, and a DNA helicase. No known splicing or trans-splicing sites were identified in Yaravirus' genome. Therefore, we propose that Yaravirus has a fragmented proteome, that means, most of the encoded proteins are synthesized as modules and joined together at the protein level.

Palavras-chave: proteomics; threading; genetic structure; viral evolution.