In silico characterization of non-synonymous substitutions and immune recognition of SARS-CoV-2: potential vaccine target for development in the Americas

Andres F. Cuspoca-Orduz[1], Laura L. Díaz-Lache[1], Álvaro F. Acosta-Costilla[1], Yardany R. Mendez-Fandiño[1]

Escuela de Medicina, Universidad Pedagógica y Tecnológica de Colombia (UPTC), Tunja – Colombia

Background

Worldwide cooperative effort among SARS-CoV-2 during the global spread had allowed researchers to advance rapidly in the knowledge of the virus. Repositories of whole genome sequences of the virus are available online, in which frequent non-synonymous mutations (MNS) have been identified, allowing the research community to infer changes in the pathogenicity and immune recognition of the virus. Prevention and control strategies are needed in order to control the virus' spread, in which the development of a vaccine is necessary worldwide although it is a challenge for the research community. This study was aimed at proposing a vaccine candidate for SARS-CoV-2 considering LATAM circulating viruses, and non-synonymus mutation regions among the proteins of the virus.

Materials and methods.

From the identification of MSNs, collected worldwide by China National Center for Bioinformation (CNCB) and presented by the Global Initiative on Sharing All Influenza Data (GISAID), regions without presence of non-synonymous substitutions (SNS) of the SARS-CoV-2 proteome were identified. Agglutination predictions were made using bioinformatic tools among the most frequent human leukocyte antigen (HLA) class I and II haplotypes among Latin American countries. Also, an enrichment was carried out on the related protein domains, considering the domains prone to post-translational modifications. The smallest number of peptides capable of binding to the most frequent HLAs of LATAM studied were part of a multiepitope construct, to which we added adjuvants towards the N and C terminal, to achieve a complete immunogenic response performing molecular dynamics with Toll-like receptor (TLR-4) to test a safe and stable interaction.

Results.

The construct has 25 sequences with vaccine potential obtained from the regions without SNS, capable of being the most frequent HLA-I binders in the Americas. These are part of

domains related to viral replication and evasion of the immune response by the host. The proteins NSP3 ,SP, NSP4, NSP2 and others were found to be binders in HLA-II and HLA-I molecules. Also, areas of recognition of linear epitopes and conformations were identified, conserved in the nucleoprotein and Spike. These sequences, were introduced as part of a vaccine construct, showing by molecular dynamics an adequate and safe interaction with the Toll-4 receptor.

Conclusions.

Based on SNS so far , and an own bioinformatic pipeline, we propose a vaccine that satisfies the requirements for in vivo and in vitro experimental assays.