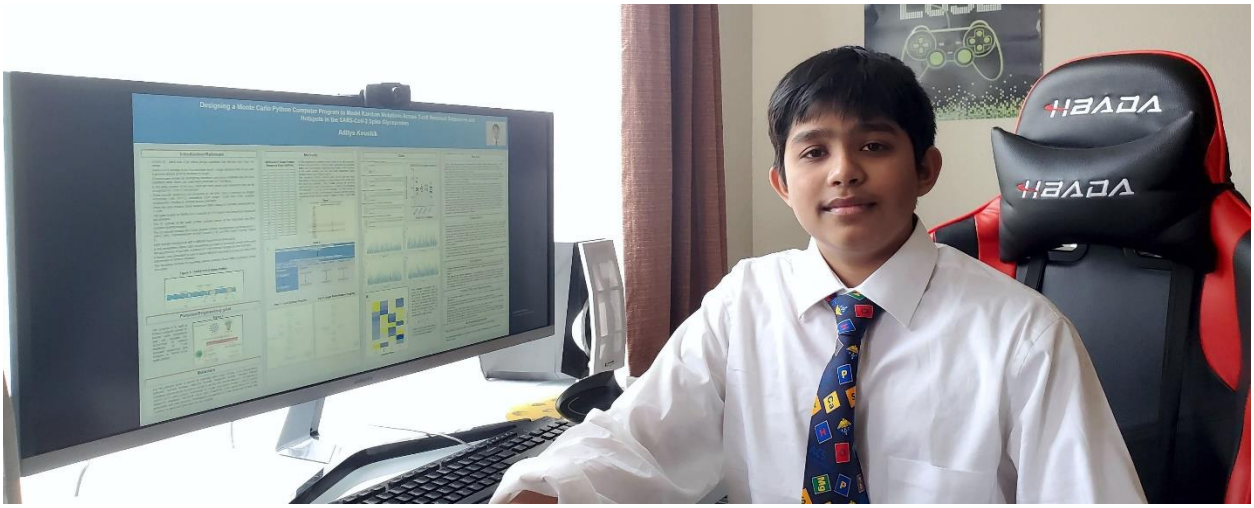

Designing a Monte Carlo Python Computer Program to Model Random Mutations Across T-cell resistant Sequences and Hotspots in the SARS-CoV-2 Spike Glycoprotein

Aditya Kiran Koushik
La Cueva High School, Albuquerque

ABSTRACT

The purpose of this experiment is to use Monte Carlo simulations in the Python programming software to predict sporadic mutations in T-cell resistant sequences and hotspots of the SARS-CoV-2 spike glycoprotein. First, the Immune Epitope Database (IEDB) library was used to acquire 48 well-validated T-cell epitopes in the SARS-CoV-2 spike protein. Using Monte Carlo simulations in Python (MC-P), the number of random mutations needed to change a 9-mer T-cell resistant sequence into a T-cell epitope was recorded. Next, in a separate MC-P model, the number of cycles required to randomly replace a single amino acid in the 9-mer stretch of spike protein to a new highly transmissible variant of SARS-CoV-2 (D614G, V483A, G4765, and L54F) was also recorded. My results show that; 1) randomly altering amino acids in the 9-mer T-cell resistant sequence ("VLYQDVNCT") for 10,000,000 cycles didn't match any of the 48 T-cell epitopes. 2) single point mutations (D614G, V483A, G4765, and L54F) took an average of 6-7 cycles when ran through the program. Therefore, the conversion of a T-cell resistant sequence into a T-cell epitope is rare, and single point mutations in spike protein are more frequent, which could result in highly transmissible variants of the virus.