

## Statistical modeling of SARS-Cov-2 mutation in the U.S.

Yuru Jing <sup>1</sup>, Angela Antonou, Prof. <sup>2</sup>

<sup>1,2</sup>Dept. of Mathematics and Computer Science, University of St. Francis, 500 Wilcox Street, Joliet, IL 60435

YuruJing@stfrancis.edu <sup>1</sup>, Aantonou@stfrancis.edu <sup>2</sup>

**Abstract:** As a result of the severity of the SARS-Cov-2 outbreak worldwide, researchers have invented vaccines to prevent its spread. However, the higher mutation characteristics of an RNA virus such as SARS-Cov-2 may make it more likely that a group of individuals with a higher average and standard deviation mutation frequency will have a shorter duration of immunity against SARS-Cov-2 after vaccination. To investigate this, we analyzed a random sample of thousands of SARS-Cov-2 RNA sequences from infected individuals within the US and aligned them against the reference sequence WIV04, which is the RNA sequence from the first outbreak in Wuhan, China. We determined the number of mutants within the sequences using the fast optimal global sequence alignment algorithm (FOGSAA) for each case in order to determine the mutation frequency. Using this mutation frequency, we identify which characteristics within the population are associated with higher mutation frequency and use this to predict those groups most likely to have a reduced duration of immunity.

**Keywords:** fast optimal global sequence alignment algorithm, branch-and-bound technique, SARS-Cov-2, statistical modeling