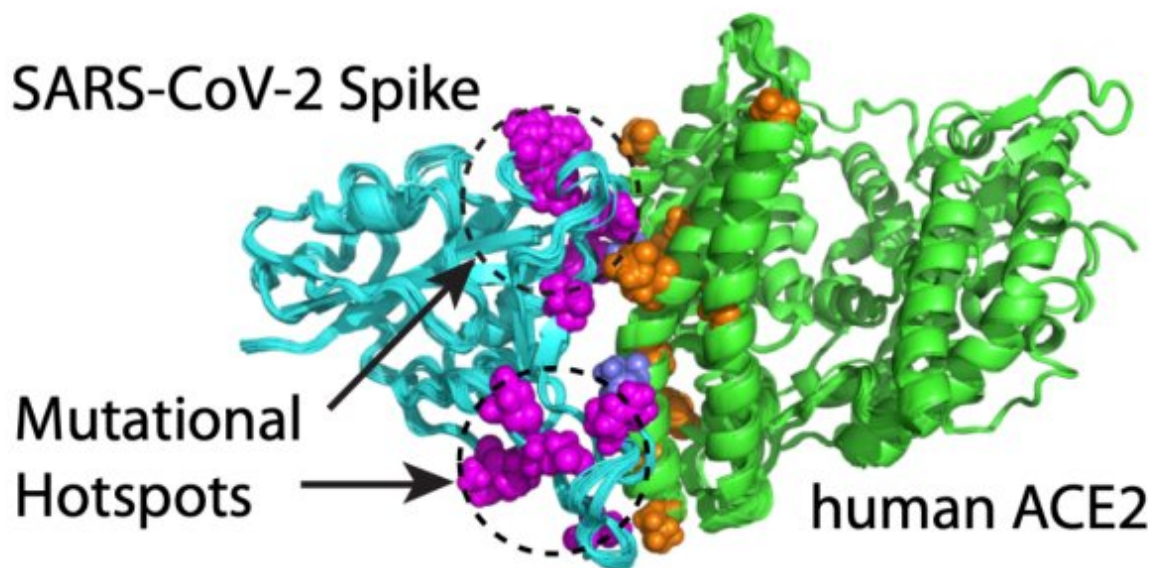# Computational modeling can predict mutation hotspots in SARS-CoV-2 spike protein

May 19 2021, by Rachel Harrison

Computational modeling shows that mutations on SARS-CoV-2's spike protein that enhance the virus' ability to bind to the ACE2 receptor occur in two clusters or mutation "hotspots." Credit: Hin Hark Gan and Kristin Gunsalus, NYU's Department of Biology

SARS-CoV-2 has evolved to acquire mutations on the spike protein—the part of the virus that protrudes from its surface and latches onto cells to infect them—that enhance the coronavirus' ability to bind to human cells or evade antibodies. A new study from the Centers for Genomics and Systems Biology at New York University and NYU Abu Dhabi uses computational modeling to assess the biological significance of spike protein mutations, uncovering versions of the virus that bind more tightly or resist antibodies and offering a promising public health surveillance tool.

The study, which appears in the *Journal of Molecular Biology*, also suggests that these mutations on the spike protein are a key reason for the virus' rapid spread in parts of the world.

New and more transmissible COVID-19 variants have emerged in recent months, fueling surges of cases in countries like India and Brazil. As a public health measure, rapid surveillance methods are needed to determine the biological effects of variants and to help anticipate emerging infectious viral strains. But monitoring new variants is no small task; genome sequencing shows that the SARS-CoV-2 spike protein alone, for example, has about 5,000 possible variants.

"Screening such a large set of variants poses a tremendous challenge for conventional experimental methods," said Hin Hark Gan, a senior research scientist at NYU's Center for Genomics and Systems Biology and the study's lead author. "An advantage of computer-based modeling is that a hundred mutations can be readily assessed in a few days."

Gan and his colleagues turned to a computational method that models how the SARS-CoV-2 spike protein recognizes the ACE2 receptor—a protein on the surface of many types of cells—to gain entry into host

cells. Studies of coronaviruses indicate that spike-ACE2 recognition is the basis for infection.

The researchers focused on screening the mutations located where the spike protein and ACE2 receptor meet. They assessed 1,003 mutation combinations in the spike and ACE2 proteins, including those resulting in the fast-spreading spike variants that have originated in Brazil, South Africa, the U.K., and India.

Their systematic assessment of variants uncovered that spike mutations that bind tightly to the ACE2 receptor occur in two clusters or mutation "hotspots" on the binding interface. These hotspots are located in structurally flexible regions, indicating that mutations that increase binding effectively reprogrammed the spike conformation to enhance its recognition of the ACE2 receptor.

The researchers also looked at single, double, and triple mutations in the critical spike interface region, which make up some of the recently emerged infectious variants. Their modeling analysis suggests the spike variants S477N, N501Y, and S477N + E484K and E484K + N501Y—fast-spreading double mutants found in Brazil, South Africa, the U.S., and U.K.—have increased binding to the ACE2 receptor relative to the original coronavirus that emerged in Wuhan.

Gan and colleagues observed that the E484K and E484Q mutations found in some recent fast-spreading variants are not only predicted to bind more strongly to ACE2 but have also been shown to confer antibody resistance. Neutralizing antibodies are produced in response to viral infection and target different sites on the spike protein to prevent the virus from invading host cells. This prompted the researchers to look at another factor contributing to viral transmission: antibody resistance of individual spike mutations.

In particular, the variant circulating in India has two mutations in the spike interface region: L452R and E484Q. This variant is not predicted to bind to the ACE2 receptor more tightly than the virus that originated in Wuhan, likely because the individual mutations have opposing effects (the L452R mutation binds less easily while the E484Q mutation binds more easily). Strikingly, however, both of these mutations are strong antibody evaders, a scenario not found in other recent variants.

"As more of antibody target sites become resistant to antibodies due to viral mutations, the efficacy of existing antibodies and vaccines may diminish," added Gan. "This scenario is a likely cause for the rapid spread of the variant in India."

The study not only provides explanations for the coronavirus' rapid spread—both mutations that enhance binding to human cells and help evade antibodies—but also points to a promising predictive tool in the ongoing public health fight against SARS-CoV-2.

"Our computational modeling method can be used as a real-time surveillance tool to screen for emerging infectious COVID-19 variants. It allows for a more timely response to emerging outbreaks and could be used to guide the development of new vaccines," said Kristin C. Gunsalus, professor of biology at NYU, faculty director of bioinformatics at NYU Abu Dhabi, and the study's senior author.

Provided by New York University

Citation: Computational modeling can predict mutation hotspots in SARS-CoV-2 spike protein (2021, May 19) retrieved 28 April 2024 from https://phys.org/news/2021-05-mutation-hotspots-sars-cov-spike-protein.html