

Tracking SARS-CoV-2 mutations in Russia

Skoltech researchers have traced **THE SPREAD OF VARIANTS IN RUSSIA** by constructing an evolutionary tree of SARS-CoV-2.

The repeated rise of mutations in SARS-CoV-2 is making the COVID-19 pandemic challenging to control. By analysing changes to the viral genome, Skoltech researchers have been tracking the evolution of mutations in order to better understand how the disease spreads in Russia,

with a view to identifying effective measures for containing the virus.

The SARS-CoV-2 genome consists of nearly 30,000 nucleotides. "Mutations occur somewhere in the sequence once every two or three transmissions," explains Georgii Bazykin, an evolutionary

biologist at Skoltech Bio Center. "The differences in the sequence among samples allow us to deduce the order in which mutations accumulated," he says. "Using this information, we can construct an evolutionary tree of the virus."

Russia's first COVID-19 cases appeared much later than neighbouring countries. Bazykin found that the earliest cases originated from multiple introductions from Europe, notably during the spring holidays of 2020. "The earliest sequences from Russia were not at all similar to ones in China at the time, which suggests that the border closure with China delayed the initial spread," he says. "We also saw that the virus was introduced through tens or hundreds, perhaps even thousands, of individual instances. Restricting air travel appears to have had a very limited effect."

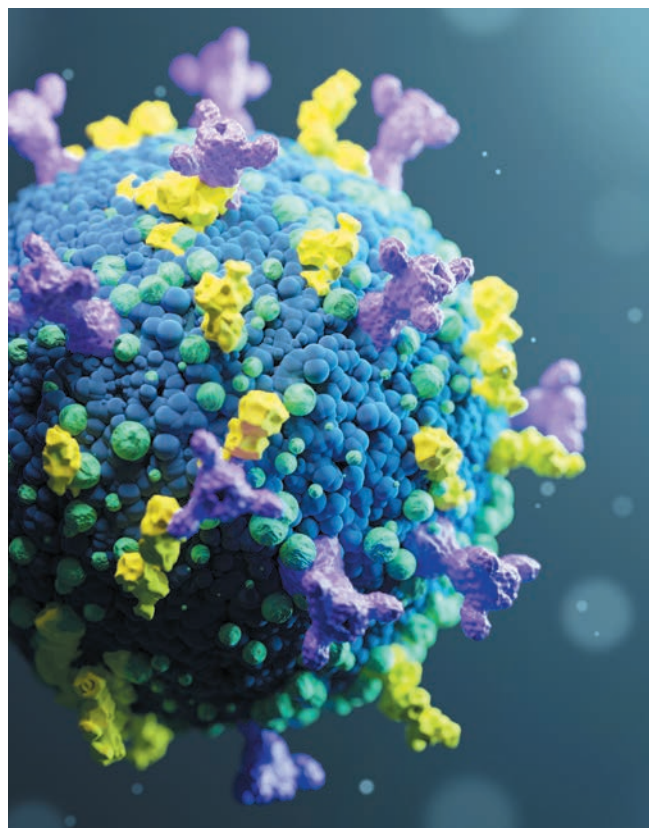
"WE NEED TO IDENTIFY THREATENING NEW VARIANTS ...WHILE THEY ARE STILL RARE."

A boon, but also a challenge, for analysing SARS-CoV-2 was the unprecedented amount of data. "With other viruses like Influenza A, we're lucky if we can get hundreds or thousands of sequences in a year. With SARS-CoV-2, we have millions," says Bazykin, who also traces mutations in viruses such as Influenza A and HIV1. "The algorithms that worked for smaller datasets became unusable, so we had to develop a novel approach to process data of this scale." More importantly, however, the

massive dataset gave the team a high-resolution view of how mutations accumulate.

The wealth of data proved especially useful when examining specific local outbreaks. In mid-2020, Bazykin's team investigated a COVID-19 outbreak at a hospital in Saint Petersburg. When patients began testing positive for the infection, the hospital went into complete lockdown for 35 days, with units isolated from each other and the hospital-wide ventilation system shut down. The team found that the outbreak originated from two or three distinct introductions, since the SARS-CoV-2 genomes formed three distinct groups with different common ancestors. They also concluded that quarantine curbed further transmissions, as the average number of transmissions dropped substantially after imposing the quarantine.

Another study by Bazykin showed that a mix of variants — including those with mutations endemic to Russia — were present in the country until the spring of 2021. By September, however, these variants had been almost entirely replaced with the delta variant. "This variant is accumulating novel mutations, and tracking them will be top priority in the coming months," says Bazykin. "We need to identify threatening new variants and their properties while they are still rare." His group's partnerships with Coronavirus Russian Genetic Initiative consortium, as well as with international organizations, have been critical for success, Bazykin notes. ■



It is critical to detect and track the spread of new variants of the SARS-CoV-2 virus.

Skoltech