Scientists identify hundreds of drug candidates to treat COVID-19

August 12 2020, by Iqbal Pittalwala

Credit: Unsplash/CC0 Public Domain

Scientists at the University of California, Riverside, have used machine learning to identify hundreds of new potential drugs that could help treat COVID-19, the disease caused by the novel coronavirus, or SARS-

CoV-2.

"There is an urgent need to identify effective drugs that treat or prevent COVID-19," said Anandasankar Ray, a professor of molecular, cell, and systems biology who led the research. "We have developed a drug discovery pipeline that identified several candidates."

The drug discovery pipeline is a type of computational strategy linked to artificial intelligence—a computer algorithm that learns to predict activity through trial and error, improving over time.

With no clear end in sight, the COVID-19 pandemic has disrupted lives, strained health care systems, and weakened economies. Efforts to repurpose drugs, such as Remdesivir, have achieved some success. A vaccine for the SARS-CoV-2 virus could be months away, though it is not guaranteed.

"As a result, drug candidate pipelines, such as the one we developed, are extremely important to pursue as a first step toward systematic discovery of new drugs for treating COVID-19," Ray said. "Existing FDA-approved drugs that target one or more human proteins important for viral entry and replication are currently high priority for repurposing as new COVID-19 drugs. The demand is high for additional drugs or small molecules that can interfere with both entry and replication of SARS-CoV-2 in the body. Our drug discovery pipeline can help."

Joel Kowalewski, a [graduate student](#) in Ray's lab, used small numbers of previously known ligands for 65 human proteins that are known to interact with SARS-CoV-2 proteins. He generated machine learning models for each of the human proteins.

"These models are trained to identify new small molecule inhibitors and activators—the ligands—simply from their 3-D structures," Kowalewski

said.

Kowalewski and Ray were thus able to create a database of chemicals whose structures were predicted as interactors of the 65 protein targets. They also evaluated the chemicals for safety.

"The 65 protein targets are quite diverse and are implicated in many additional diseases as well, including cancers," Kowalewski said. "Apart from drug-repurposing efforts ongoing against these targets, we were also interested in identifying novel chemicals that are currently not well studied."

Ray and Kowalewski used their machine learning models to screen more than 10 million commercially available small molecules from a database comprised of 200 million chemicals, and identified the best-in-class hits for the 65 human proteins that interact with SARS-CoV-2 proteins.

Taking it a step further, they identified compounds among the hits that are already FDA approved, such as drugs and compounds used in food. They also used the machine learning models to compute toxicity, which helped them reject potentially toxic candidates. This helped them prioritize the chemicals that were predicted to interact with SARS-CoV-2 targets. Their method allowed them to not only identify the highest scoring candidates with significant activity against a single human protein target, but also find a few chemicals that were predicted to inhibit two or more human protein targets.

"Compounds I am most excited to pursue are those predicted to be volatile, setting up the unusual possibility of inhaled therapeutics," Ray said.

"Historically, disease treatments become increasingly more complex as we develop a better understanding of the disease and how individual

genetic variability contributes to the progression and severity of symptoms," Kowalewski said. "Machine learning approaches like ours can play a role in anticipating the evolving treatment landscape by providing researchers with additional possibilities for further study. While the approach crucially depends on experimental data, virtual screening may help researchers ask new questions or find new insight."

Ray and Kowalewski argue that their computational strategy for the initial screening of vast numbers of chemicals has an advantage over traditional cell-culture-dependent assays that are expensive and can take years to test.

"Our database can serve as a resource for rapidly identifying and testing novel, safe treatment strategies for COVID-19 and other diseases where the same 65 target proteins are relevant," he said. "While the COVID-19 pandemic was what motivated us, we expect our predictions from more than 10 million chemicals will accelerate drug discovery in the fight against not only COVID-19 but also a number of other diseases."

Ray is looking for funding and collaborators to move toward testing cell lines, animal models, and eventually clinical trials.

The research paper, "Predicting Novel Drugs for SARS-CoV-2 using Machine Learning from a >10 Million Chemical Space," appears in the journal *Heliyon*, an interdisciplinary journal from Cell Press.

Provided by University of California - Riverside