THE USE OF AI AND ADVANCED COMPUTER SYSTEMS TO DEVELOP DRUGS AGAINST NEW EMERGING THREATS

Machine Learning Guided Prediction Of Receptor Binding And Immune Escape For SARS-CoV-2

Le Thanh Mai Pham Sandia National Laboratories **Thomas Sheffield** Sandia National Laboratories **Ryan Bruneau** Sandia National Laboratories **Stephen Won** Sandia National Laboratories **Kenneth L. Sale** Sandia National Laboratories **Brooke Harmon** Sandia National Laboratories

The impacts of recent pandemics such as 2002-2003 SARS-CoV pandemic, the 2009 H1N1 Influenza pandemic, and the ongoing SARS-CoV-2 Pandemic have highlighted an urgent need to predict future viral pandemics and preemptively work towards producing medical countermeasures (MCM) such as vaccines and antibody therapeutics. In this study, we used a combination of high throughput screening and machine learning (ML) models to identify heavy-chain only antibodies (HCAbs) with potent efficacy against all circulating SARS-CoV-2 viral variants of interest (VOIs) and concern (VOCs). Using the binding HCAb affinity data against SARS-CoV-2 VOI and VOCs (pre-Omicron variants) and model features from other published data, we were able to develop an ML model that successfully identified HCAbs with efficacy against Omicron variants, independent of which were not included in our experimental biopanning workflow. This ML-coupled biopanning approach reduced the experimental screening burden by 78% to 90% for the Omicron BA.5 and Omicron BA.1 variants, respectively. We also successfully trained models to predict receptor binding and immune evasion for future SARS-CoV-2 variants. Our workflow can profile existing variants and predict potential antigenic changes for never-before-seen SARS-CoV-2 variants, which aids in the development of pathogen-agnostic capabilities, capitalizes on AI/ML to optimize MCM development, and more rapidly deliver capabilities to the Total Force for managing future pandemics and preventing future surges.

Take Home Message:
-Different sets of mutations in the viral genome and their interactions with receptors/antibodies can be predicted through protein sequence-based protein-protein interactions.
-A combination of experimental data and a modeling-informed data analytics platform improves the prediction of antigenic evolution and immune evasion for existing and future SARS-CoV-2 variants.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.