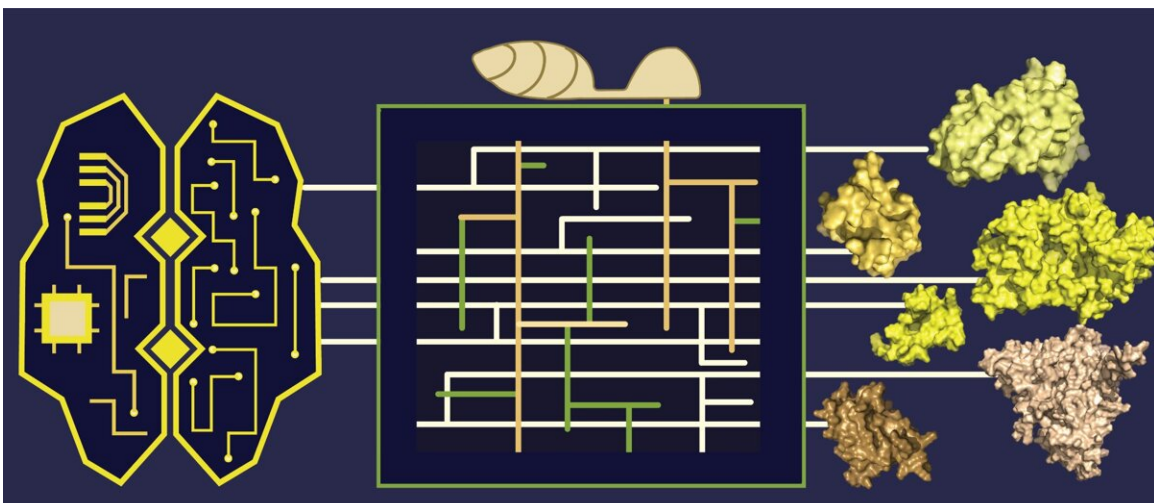# New structure prediction model has mapped 500 previously unsolved proteins

January 19 2022

Elucidating the structures of phytopathogens' secreted proteins with machine learning-based structure prediction tools. Machine learning and plant-pathogen interaction commonly have a black box. During the prediction from the input primary sequences to protein structures, we don't exactly know what happens. Similarly, we do not fully understand the complex interaction at the interface of plants and pathogens. The box in the middle captures the complexity within this black box. Credit: Kyungyong Seong and Ksenia V. Krasileva

Scientists at the University of California, Berkeley, have recently published work that lays the foundation for new ways of thinking about pathogen evolution. "Our research highlights that template-free

modeling that uses machine learning is indeed superior to template-based modeling for the secreted proteins of the destructive fungal pathogen Magnaporthe oryzae," said Kyungyong Seong, first author of the paper published in the *Molecular Plant-Microbe Interactions (MPMI)* journal.

Pathogens use virulence factors known as effectors, which are important for the pathogen's survival. Homology modeling is one of the most widely used methods, but this requires the use of templates of solved effector structures, and solving all the effector structures is too daunting of a task. There are too many effector proteins encoded in pathogens' genomes to simply rely on experimentally solving each one of the structures.

Seong and colleague Ksenia V. Krasileva used a new structure prediction method that was able to model 500 secreted proteins previously not predicted by the template-based method.

"About 70% out of the 1,854 secreted proteins were modeled in our study, and their structures provide an extra layer of information about the effectors based on their similarity to each other or other solved protein structures," said Krasileva. "We demonstrate that new structure prediction methods apply well to the problem of deciphering pathogen virulence factors and other secreted proteins that often have little sequence similarity among themselves or to other proteins."

This new method allows scientists to map thousands of secreted proteins and establish missing evolutionary connection among them. "We believe our research was the first to apply the concept of structural genomics on a plant pathogen in the new era of machine-learning structure prediction," said Seong.

"As the accuracy of structure prediction improves further, it will become more common to see articles that incorporate large-scale protein

structure prediction data," predicted Krasileva. "Our article may spark some ideas of how to use such data, leading some scientists to explore opportunities ahead of other."

They also found that there are many novel sequence-unrelated structurally similar effectors in M. oryzae, and structurally similar effectors are found in other phytopathogens. This suggests that pathogens may be relying on a set of effectors that commonly originated but largely diverged in sequences in the course of evolution to infect plants.

More information: Kyungyong Seong et al, Computational Structural Genomics Unravels Common Folds and Novel Families in the Secretome of Fungal Phytopathogen Magnaporthe oryzae, *Molecular Plant-Microbe Interactions* (2021). DOI: 10.1094/MPMI-03-21-0071-R

Provided by American Phytopathological Society