De Novo Nanobody Design With Neural Networks, AlphaFold, and Docking Algorithms

Venkatraman, Anirudh (School: Homestead High School)

Purpose: COVID-19 has killed more than 5.6 million. Nanobodies that retain the binding affinity of antibodies despite being one-tenth the size, are emerging as viable treatment options to counter the mushrooming viral variants. However, discovering effective nanobodies remains a time consuming process, mostly involving extensive wet-lab procedures. My goal was to develop bio-computational techniques to accelerate the determination of nanobodies with higher specificity and affinity to a pathogenic target, reducing costs and time-to-market and thereby saving lives. Procedure: My unique computational approach used ensemble-stacking to develop nanobodies by sequencing the CDR-H3 region of nanobodies - sequence of amino acids in the variable binding region critical to determining specificity. In a search space of about 2020 different CDR-H3 sequences, the model predicted 50 effective sequences against an antigen target, by training on previous phage display data to accurately predict and rank the affinity of CDR-H3 sequences. Results: Ensemble-stacking achieved highly accurate predictions of binding affinity, with R^2 of 0.75 and a Pearson correlation coefficient of 0.87 in addition to a classification AUROC of 0.965. I tested the top 50 algorithm-generated nanobodies by using 3D computational methods to measure their binding affinity. My approach yielded nanobody candidates with a lower Gibbs Free energy value (a measure of stronger specificity and binding) of -10.24 kcal/mol compared to the best in vitro derived candidates (-9.65 kcal/mol). Conclusion: Results indicate that bio-computational methods using ensemble-stacking can be successfully leveraged to quickly and accurately generate nanobody candidates thereby accelerating development of successful vaccines.

Awards Won:

Fourth Award of $500

Fourth Award of $500