UnFolding Genetic Disorders: Evaluating AlphaFold Metrics on Predicting Pathogenic Missense Mutations

Weinstein, Rayne (School: Rowland Hall- Lincoln Street Campus)

Alphafold, a protein-folding prediction tool created by Google DeepMind, has shown unprecedented ability to accurately simulate the 3D structure of proteins given an amino acid sequence. Based on this technology, tools have been developed to predict the pathogenicity of single-nucleotide mutations. Currently, the majority of tools are based on a per-residue individual confidence metric (pLDDT) that AlphaFold produces in its simulation or a predicted change in Gibbs Free Energy. However, many AlphaFold metrics have not yet been evaluated for their ability to predict pathogenicity, and could significantly improve current models. One such metric is the Predicted Aligned Error (PAE), which represents the predicted distance between any two residues on separate monomers. This metric is often used to predict if and where two proteins will bind by identifying regions on each with close contact. This study hypothesized that the change in PAE produced by introducing a missense mutation can predict the mutation's pathogenicity, since the protein's primary pathway will be significantly disrupted upon a pathogenic mutation. Additionally, since the metric is evaluating the effect on a specific interface, a PAE-based model may be able to identify which disrupted interactions tend to drive pathogenic effects. For this study, I used the protein p53 as a model due to its large dataset of mutation pathogenicity classifications. Specifically, I evaluated the disruption of PAE on the interfaces of p53 binding to TAZ2 and DNA. Lastly, I compared the performance of a PAE-based model to the FoldX and AlphaMissense models to assess how PAE-based models compare to current models.

Awards Won: