

Combining Pairwise Structural Similarity and Deep Learning Interface Contact Prediction to Estimate Protein Complex Model Accuracy in CASP15

Jianlin Cheng¹, Raj Roy¹, Jian Liu¹, Nabin Giri¹, and Zhiye Guo¹

¹University of Missouri

March 13, 2023

Abstract

Estimating the accuracy of quaternary structural models of protein complexes and assemblies (EMA) is important for predicting quaternary structures and applying them to studying protein function and interaction. The pairwise similarity between structural models is proven useful for estimating the quality of protein *tertiary* structural models, but it has been rarely applied to predicting the quality of *quaternary* structural models. Moreover, the pairwise similarity approach often fails when many structural models are of low quality and similar to each other. To address the gap, we developed a hybrid method (MULTICOM_qa) combining a pairwise similarity score (PSS) and an interface contact probability score (ICPS) based on the deep learning inter-chain contact prediction for estimating protein complex model accuracy. It blindly participated in the 15th Critical Assessment of Techniques for Protein Structure Prediction (CASP15) in 2022 and ranked first out of 24 predictors in estimating the global accuracy of assembly models. The average per-target correlation coefficient between the model quality scores predicted by MULTICOM_qa and the true quality scores of the models of CASP15 assembly targets is 0.66. The average per-target ranking loss in using the predicted quality scores to rank the models is 0.14. It was able to select good models for most targets. Moreover, several key factors (i.e., target difficulty, model sampling difficulty, skewness of model quality, and similarity between good/bad models) for EMA are identified and analyzed. The results demonstrate that combining the multi-model method (PSS) with the complementary single-model method (ICPS) is a promising approach to EMA.

Hosted file

MULTICOM_EMA.docx available at <https://authorea.com/users/302597/articles/629409-combining-pairwise-structural-similarity-and-deep-learning-interface-contact-prediction-to-estimate-protein-complex-model-accuracy-in-casp15>