Guijun Zhang

College of Information Engineering

Zhejiang University of Technology

Hangzhou, 310023, China

Email: xxx@zjut.edu.cn

October 14, 2016

Dear Editor,

We are submitting a manuscript entitled “NN-BAYES: Protein contact map prediction using neural network training coupled with naïve Bayes classifiers” for consideration as a Research Article in *Bioinformatics*.

Exciting progress has been recently witnessed on *ab initio* folding of large-size proteins with the assistance of co-evolution based contact predictions. But the success is still anecdotal, which is mainly because the co-evolution based methods request for a high volume of sequence homologs to generate accurate contact map predictions, which are not available to most of the non-humongous protein targets. Therefore, development of efficient methods that can generate balanced and reliable contact maps for different type of protein targets is essential to improve the success rate of the *ab initio* protein structure prediction.

In this study, we developed a new composite pipeline, NN-BAYES, which extends the naïve Bayes classifier (NBC) theorem to combine eight state of the art contact methods that are built from co-evolution and machine learning approaches. The posterior probabilities of the NBC model are then trained with intrinsic structural features through neural network learning for the final contact map prediction. NN-BAYES was tested on both benchmark and CASP targets, which demonstrated significant improvement over the currently best contact map predictors, in particular for the targets that have few sequence homologs. The program has been made publicly available through our website, which should result in a broad impact to the field of *ab initio* protein structure prediction. Based on the importance of the problem addressed and the high efficiency of the method developed, we believe the manuscript is worthy of consideration by *Bioinformatics*.

Sincerely,

XXX, XXX, XXXX, XXX