NuFold: Development of RNA Tertiary Structure Prediction Method Using Machine Learning

Yuki Kagaya [1], Zicong Zhang [2], Nabil Ibtehaz [2], Xiao Wang [2], Tsukasa Nakamura [1], and Daisuke Kihara[1,2]

[1] Department of Biological Sciences, Purdue University, West Lafayette, Indiana, 47907, USA
[2] Department of Computer Science, Purdue University, West Lafayette, Indiana, 47907, USA

Abstract:
RNA is one of the essential biomacromolecules not only because it plays a role in the Central Dogma, but also because it has unique and various biological functions. Recent studies have shown that non-coding RNAs, which do not encode proteins, also have essential roles in gene regulation and other biological functions. However, only a limited variety of the tertiary structures of those non-coding RNAs are available, since experimental methods for determining RNA structures are generally challenging and time-consuming. As one of the methods to solve this problem, an efficient and accurate computational prediction method is needed.
In this presentation, we report the development of a method to predict the tertiary structure of RNA from its sequence using state-of-the-art deep learning architecture. Our approach builds on recent advances in deep learning, which has shown great performance in predicting protein structures from sequence information. Specifically, we adapted the AlphaFold deep learning architecture, which has achieved state-of-the-art performance in protein structure prediction, to work with RNA sequences. We will discuss the details of our developed method and demonstrate its preliminary performance on several test cases.