[BIOINFO 2023]

AI-based Design of Augmented 5'UTR with a High Translational Efficiency

Solbeen Kim[1], and Jin-Wu Nam[1,2,*]

[1] *Department of Life Science, College of Natural Sciences, Hanyang University*
[2] *Bio-BigData Center, Hanyang Institute of Bioscience and Biotechnology, Hanyang University*
* *Corresponding author: jwnam@hanyang.ac.kr*

The mRNA vaccines were clinically successful during the COVID-19 pandemic thanks to the rapid processes of their design and synthesis. The design of mRNA vaccines was mainly aimed at identifying sequences with a high translational efficiency (TE) and RNA stability, as well as a high immunogenicity and safety. Despite the descent efficiency and high safety of commercially available mRNA vaccines, their construct has not been fully investigated and optimized, due to the urgent need. Particularly, the optimal design of mRNA structure have remained in the sequence level. Here, we implemented a deep learning (DL) model to predict TE by considering both sequence and RNA structure of mRNAs. For this purpose, we implemented a circular convolutional neural network (CCNN) model that takes into account the secondary structure of the 5' UTR and the 30nt downstream sequence from the start codon. This model showed higher performance compared to previous TE prediction models. Furthermore, based on the CCNN model, we developed a generative DL model based on feedback generative adversarial network (GAN), which provides more diverse 5'UTR sequences with a high TE than a previous genetic algorithm-based model. Three selected 5'UTR sequences were tested in a RFP reporter system, one of which had approximately 50% higher TE than those used in Moderna and Pfizer-BioNTech. Our new DL models provide a computational platform for the design of augmented 5'UTR sequences that can be applied to the development of cancer mRNA vaccines as well as the design of augmented 3'UTR sequences.