Creating a Deep Learning Model to Predict off-target DNA sites using prime editing technique

By
Khaled Mohammad Ahmad Jum'ah

Supervision
Dr. Lama Hamadneh

Co-supervisor
Dr. Krzysztof Poterlowicz

Al-Zaytoonah University of Jordan, 2021

Abstract

Prime editing is one of genome modifying systems, it shows a promising expectation to replace CRISPR due to their lower off-target activity. Until now, there is no predictive model designed specially to predict the off-target activity for this technique. In this study, a convolutional neural network model was developed to predict the type of binding between PegRNA and the target DNA, and if it will bind correctly to the target site (on-target) or bind to other sites (off-target). An additional class (no-binding) was added to reduce PegRNA-DNA off-target false positive predictions. The data were extracted from experiments that used prime editing to induce mutation on HEK293T cells to study their off-target activity. Then the data were divided into 3 classes based on the class of the binding between the PegRNA and the target DNA. The data was repossessed and converted into vectors to be used as an input for CNN deep learning model. The results show a high model prediction capability to detect the on-target and the no binding sites with a 100% present accuracy per class, and for the off-target class

the prediction accuracy was 63.6%. the reduction in the prediction accuracy is likely due to the low number of samples in the off-target data.