Optimization of Aptamer Screening Techniques with Deep Learning Methods

Yan-Xin Lin[1], Hsiang-Yu Wang[1, 2], Min-Kung Hsu[2, 3, *]

[1] Graduate Institute of Animal Vaccine Technology, National Pingtung University of Science and Technology, Pingtung 91201, Taiwan
[2] International Degree Program in Animal Vaccine Technology, National Pingtung University of Science and Technology, Pingtung 91201, Taiwan
[3] General Research Service Center, National Pingtung University of Science and Technology, Pingtung 91201, Taiwan

* Correspondence: mkh@mail.npust.edu.tw

Abstract

Aptamers are single-stranded short chains of DNA, RNA, or XNA. Their structures allow them to bind to target protein molecules with high affinity and specificity, functioning similarly to antibodies. With advancements in DNA and RNA synthesis technologies in recent years, the applications of aptamers have become increasingly widespread. If applied to developing detection assays for emerging infectious diseases such as SARS-CoV-2, rapid diagnostic assays can be developed in the shortest time possible after a disease outbreak. The traditional method uses the exponential enrichment technique known as Systematic Evolution of Ligands by EXponential Enrichment (SELEX) to screen for aptamer candidates with high affinity for target proteins. However, SELEX is time-consuming and costly. Additionally, SELEX is only suitable for the screening of aptamers with short-length sequences, limiting its applicability. To minimize the costs and duration associated with traditional SELEX experiments, this study developed an *in-silico* method for predicting and screening aptamers. This method has been developed into a candidate aptamer sequence prediction online tool, TAP, available at https://hywlab.npust.edu.tw/tap/. The core of this online tool utilizes a Transformer encoder architecture integrated with natural language processing (NLP) technology, enabling it to generate high-affinity candidate aptamer sequences more efficiently. Finally, it's supplemented by experiments to verify its effectiveness. In the future, this approach is expected to be applied to developing rapid diagnostic assays for emerging diseases, achieving the goal of rapidly developing detection assays in the shortest time.