Skysong Innovations
Advancing the Arizona State University Knowledge Enterprise

1475 N. Scottsdale Road, Suite 200
Scottsdale, AZ 85287-3538
Phone: 480 884 1996 Fax: 480 884 1984

Case ID:M22-087LC
Published: 8/1/2022

## Inventors

**Petr Sulc**

**Remi Monasson**

**Simona Cocco**

## Contact

Jovan Heusser
jovan.heusser@skysonginnovations.com

# Machine Learning for Aptamer Design and Analysis

-Discovery and design of novel molecules with the ability to bind to targets of interest is a key problem in diagnostics, therapeutics and molecular biology in general. There are many different classes of molecules and each have advantages and disadvantages. While aptamers are limited in composition to four possible types of bases, they have been shown to bind with high affinity to a variety of targets of interest including proteins, nucleic acids, viruses, exosomes, metabolites and so on. And, new research into chemical modification of the bases may increase the chemical space of aptamers and provide greater diversity in sequence libraries. However, novel aptamers are often discovered after multiple rounds of selection and amplification which is time consuming and requires a lot of reagents.

Professor Petr Sulc in collaboration with researchers at the French National Centre for Scientific Research have developed a machine learning model for analyzing sequence datasets obtained from binding selection experiments. With a dataset of sequences obtained from DNA aptamer selection experiments, the model can be trained to predict what is a good or bad binder to a target as well as identify the sequence motif that is contributing most to the binding ability. It can also be used to generate new sequences that have not been encountered in the experiment which can be used in therapeutic and diagnostic applications.

This model can build on aptamer or phage selection experiments and generate novel binders that will bind to a target molecule of interest.

Potential Applications

- Machine learning model used as a classifier and generator of novel aptamers
  - Therapeutics
  - Diagnostics
  - Research – reagents, imaging, etc.

Benefits and Advantages

- Can identify sequence motifs that contribute the most to the binding ability
- Can generate novel binders based on data from the experiment
- Can interpret the dataset as well
- Can act as a generator and classifier at the datasets
- Specifically developed for datasets obtained from evolutionary selection, where successful sequences are amplified and selected in the next round
- Utilizes few parameters making it interpretable and allowing for the identification of what sequence inputs lead to the highest score
- In comparison experiments, this model performs well as both a classifier and generator, while other architectures had problems generalizing on the dataset

For more information about this opportunity, please see

Gioacchino et al - BioRxiv - 2022

For more information about the inventor(s) and their research, please see

Dr. Sulc's departmental webpage

Dr. Sulc's laboratory webpage