Data-driven drug discovery and healthcare by machine learning

Yoshihiro Yamanishi, Nagoya University

In recent years, drug discovery has become increasingly difficult. Computational approaches are expected to promote the efficiency of drug development processes. Recent developments in biotechnology have contributed to the increase in the amounts of high-throughput data in the genome, transcriptome, proteome, interactome, phenome and diseasome. These biomedical big data can be useful resources for drug development processes. Machine learning methods are expected to play key roles in the big data analysis. In this study, we developed novel machine learning methods to predict therapeutic targets of diseases, to search for drug candidate molecules, and to design new chemical structures of drug candidate molecules, by integrating various biomedical data on compounds (e.g., chemical structures, clinical phenotypes, gene expression patterns, target molecules) and diseases (e.g., disease-causing genes, environmental factors, and clinical information). A unique feature of our data-driven approach is that it clarifies all target proteins of each drug including off-targets, estimates the mechanisms of action at the pathway level, and generates molecular structures of drug candidates by deep learning. In my talk at the conference, we will show some of the applications to therapeutic target identification, large-scale compound screening, combination therapy, and drug molecular structure design for a variety of diseases.