# Artificial intelligence in drug combination therapy

## Abstract

Currently, the development of medicines for complex diseases requires the development of combination drug therapies. It is necessary because in many cases, one drug cannot target all necessary points of intervention. For example, in cancer therapy, a physician often meets a patient having a genomic profile including more than five molecular aberrations. Drug combination therapy has been an area of interest for a while, for example the classical work of Loewe devoted to the synergism of drugs was published in 1928—and it is still used in calculations for optimal drug combinations. More recently, over the past several years, there has been an explosion in the available information related to the properties of drugs and the biomedical parameters of patients. For the drugs, hundreds of 2D and 3D molecular descriptors for medicines are now available, while for patients, large data sets related to genetic/proteomic and metabolomics profiles of the patients are now available, as well as the more traditional data relating to the histology, history of treatments, pretreatment state of the organism, etc. Moreover, during disease progression, the genetic profile can change. Thus, the ability to optimize drug combinations for each patient is rapidly moving beyond the comprehension and capabilities of an individual physician. This is the reason, that biomedical informatics methods have been developed and one of the more promising directions in this field is the application of artificial intelligence (AI). In this review, we discuss several AI methods that have been successfully implemented in several instances of combination drug therapy from HIV, hypertension, infectious diseases to cancer. The data clearly show that the combination of rule-based expert systems with machine learning algorithms may be promising direction in this field.

# Open-source chemogenomic data-driven algorithms for predicting drug–target interactions

## Abstract

While novel technologies such as high-throughput screening have advanced together with significant investment by pharmaceutical companies during the past decades, the success rate for drug development has not yet been improved prompting researchers looking for new strategies of drug discovery. Drug repositioning is a potential approach to solve this dilemma. However, experimental identification and validation of potential drug targets encoded by the human genome is both costly and time-consuming. Therefore, effective computational approaches have been proposed to facilitate drug repositioning, which have proved to be successful in drug discovery. Doubtlessly, the availability of open-accessible data from basic chemical biology research and the success of human genome sequencing are crucial to develop effective in silico drug repositioning methods allowing the identification of potential targets for existing drugs. In this work, we review several chemogenomic data-driven computational algorithms with source codes publicly accessible for predicting drug–target interactions (DTIs). We organize these algorithms by model properties and model evolutionary relationships. We re-implemented five representative algorithms in R programming language, and compared these algorithms by means of mean percentile ranking, a new recall-based evaluation metric in the DTI prediction research field. We anticipate that this review will be objective and helpful to researchers who would like to further improve existing algorithms or need to choose appropriate algorithms to infer potential DTIs in the projects. The source codes for DTI predictions are available at: <https://github.com/minghao2016/chemogenomicAlg4DTIpred>.

# Microarray is an efficient tool for circRNA profiling

## Abstract

Circular RNAs (circRNAs) are emerging as a new class of endogenous and regulatory noncoding RNAs in latest years. With the widespread application of RNA sequencing (RNA-seq) technology and bioinformatics prediction, large numbers of circRNAs have been identified. However, at present, we lack a comprehensive characterization of all these circRNAs in interested samples. In this study, we integrated 87 935 circRNAs sequences that cover most of circRNAs identified till now represented in circBase to design microarray probes targeting back-splice site of each circRNA to profile expression of those circRNAs. By comparing the circRNA detection efficiency of RNA-seq with this circRNA microarray, we revealed that microarray is more efficient than RNA-seq for circRNA profiling. Then, we found ∼80 000 circRNAs were expressed in cervical tumors and matched normal tissues, and ∼25 000 of them were differently expressed. Notably, many of these circRNAs detected by this microarray can be validated by quantitative reverse transcription polymerase chain reaction (RT-qPCR) or RNA-seq. Strikingly, as many as ∼18 000 circRNAs could be robustly detected in cell-free plasma samples, and the expression of ∼2700 of them differed after surgery for tumor removal. Our findings provided a comprehensive and genome-wide characterization of circRNAs in paired normal tissues and tumors and plasma samples from multiple individuals. In addition, we also provide a rich resource with 41 microarray data sets and 10 RNA-seq data sets and strong evidences for circRNA expression in cervical cancer. In conclusion, circRNAs could be efficiently profiled by circRNA microarray to target their reported back-splice sites in interested samples.

# Inferring and analyzing module-specific lncRNA–mRNA causal regulatory networks in human cancer

## Abstract

It is known that noncoding RNAs (ncRNAs) cover ∼98% of the transcriptome, but do not encode proteins. Among ncRNAs, long noncoding RNAs (lncRNAs) are a large and diverse class of RNA molecules, and are thought to be a gold mine of potential oncogenes, anti-oncogenes and new biomarkers. Although only a minority of lncRNAs is functionally characterized, it is clear that they are important regulators to modulate gene expression and involve in many biological functions. To reveal the functions and regulatory mechanisms of lncRNAs, it is vital to understand how lncRNAs regulate their target genes for implementing specific biological functions. In this article, we review the computational methods for inferring lncRNA–mRNA interactions and the third-party databases of storing lncRNA–mRNA regulatory relationships. We have found that the existing methods are based on statistical correlations between the gene expression levels of lncRNAs and mRNAs, and may not reveal gene regulatory relationships which are causal relationships. Moreover, these methods do not consider the modularity of lncRNA–mRNA regulatory networks, and thus, the networks identified are not module-specific. To address the above two issues, we propose a novel method, MSLCRN, to infer and analyze module-specific lncRNA–mRNA causal regulatory networks. We have applied it into glioblastoma multiforme, lung squamous cell carcinoma, ovarian cancer and prostate cancer, respectively. The experimental results show that MSLCRN, as an expression-based method, could be a useful complementary method to study lncRNA regulations.