Title	RNA
Author(s)	Bui, Ngoc Thang
Citation	
Issue Date	2015-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/12762
Rights	
Description	Supervisor: Ho Bao Tu, ,

Abstract

In 2006, Fire and Mello received their Nobel Prize for their contributions to research on RNA interference (RNAi). Their work and that of others on discovery of RNAi have had an immense impact on biomedical research and will most likely lead to novel medical applications to design novel drugs for treating many kinds of diseases such as influenza A virus, HIV, Hepatitis B virus, cancer and so on. RNAi is the biological process in which short interfering strand RNA (siRNA) target and silence the target gene (mRNA). In RNAi, siRNAs can be synthesized and injected in to the cell to silence mRNAs, i.e, to control the diseases. However siRNAs can target and silence the same mRNA different efficacy and siRNAs can also silence unrelated mRNAs. Therefore synthesizing highly effective siRNAs to design novel drugs is one of the most crucial issues on RNAi research.

Research on siRNAs can be seen by consecutive generations each characterized by its typical problems. The first generation focuses on the problem of finding effective siRNA design rules where each effective siRNA design rule is composed by important characteristics of siRNAs influencing to their knockdown efficacy. In this generation, many effective siRNA design rules were found out by biological empirical processes and applying machine learning techniques. The second generation focuses on the problem of building predictive models to predict knockdown efficacy of siRNAs. Machine learning techniques have been alternatively and mostly employed to solve this problem. However, following limitations remain: most of siRNA design tools have low performance and many siRNAs generated by these effective siRNA design rules are inactive or ineffective. Performance of the proposed models is also still low and decreases when tested on independent datasets. As a result, finding solutions for the two above problems in order to generate highly effective siRNAs is still a great challenge. Due to those limitations, the next generation of methods for generating highly effective siRNAs has mostly not appeared.

Our research focuses on contribution to overcome the above–mentioned limitations in the first two generations. On the first problem, we proposed two effective siRNA design rules by developing a new descriptive method. This method not only detected characteristics of previous design rules but also discovered new positional characteristics to design effective siRNAs. On the second problem, we proposed computational methods to build better predictive models for predicting the siRNA knockdown efficacy. The key idea is not only focusing on learning algorithms but also exploiting results of the empirical processes to enrich the siRNA representation by incorporating siRNA design rule, and using labeled as well as scored datasets. Based on experimental evaluation, our proposed predictive models achieved better performance than all models recently reported in the literature.

Keywords: RNAi, siRNA, siRNA design rule, semi–supervised learning, bilinear tensor regression.