KP219

Making a Relational Dictionary for NTG Removal Graph and Biological Action of Drug Molecules

Takahiro Katsuda, Yoshimasa Takahashi

Department of Knowledge-based Information Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580 JAPAN

katsuda@mis.tutkie.tut.ac.jp, taka@mis.tutkie.tut.ac.jp

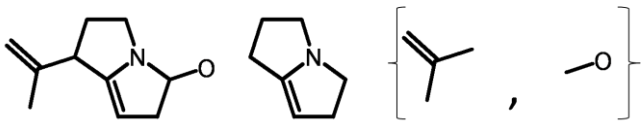
1. Introduction

In drug design and development, one of the most important things is how to find the candidate chemical structure of a lead to new drugs that are expected to have a desirable biological activity. There are various trials for chemical structure data mining aimed for the discovery of association rules for structure-activity knowledge. Ohno and Takahashi proposed a graph mining approach to structure-activity data mining based on non-terminal vertex graph (NTG)[1,2]. The NTG is defined as a vertex graph, which does not have any terminal vertex and any isolated vertex. In our preceding work, we applied the NTG to data mining for a structure database of massive drug molecules and we made a knowledge-base on the NTG-activity relationships[3,4]. However, we cannot always explain their structure-activity relationships using only NTG of the drugs. We also need to pay attention the side chain moieties of the drug molecules when we investigate the structure-activity relationships.

In this study, we defined the side chains as the NTG removal graphs (NTG-RGs) of a molecular graph, and investigated the NTG-RGs in association with the activity classes of drugs. Those NTG-RGs were used for making a dictionary that involves those NTG-RGs and the related knowledge on the drug activities.

2. Methods

The NTG of a molecular graph is defined as a graph which doesn't have any terminal vertex. On the other hand, an NTG-RG is defined with each of the connected subgraphs that are obtained when the NTG is removed from the molecular graph. Generally, most of drug molecular graphs involve an NTG and the NTG-RGs. An example of the NTG and the NTG-RGs of a molecular graph is shown in Fig.1.



Molecular graph NTG NTG-RGs

Fig.1 Example of a molecular graph and its NTG and NTG removal graphs

We used MDDR database for making the dictionary which describes the NTG-RGs and the related drug actions. The drugs that have no NTG in their chemical structures were removed. In this work, 110,890 drugs were employed for the NTG-RG analysis.

3. Results and Discussion

During the work, in the total, 12,151 unique NTG-RGs were identified for 638 drug activity classes. We have much interest in how often individual NTG-RGs appear in each of the activity classes. The statistical analyses were carried out to examine it. The NTG-RGs were further investigated to prepare the dictionary that describes the associations with the drug activities. The dictionary consists of two different dictionary files. One is a file that described about actions of drugs with a particular NTG-RG, and the other is a file that described about NTG-RGs to appear in the compounds that have a particular drug activity.

We also developed a tool to make us easy to use the dictionary. The dictionary tool allows us to view the graphical chart and to know the details of the distribution of each NTG-RG in individual activity classes. Fig.2 illustrates an example of the chart that shows the relative frequency of $–CH_2COOH$ in individual activity classes.

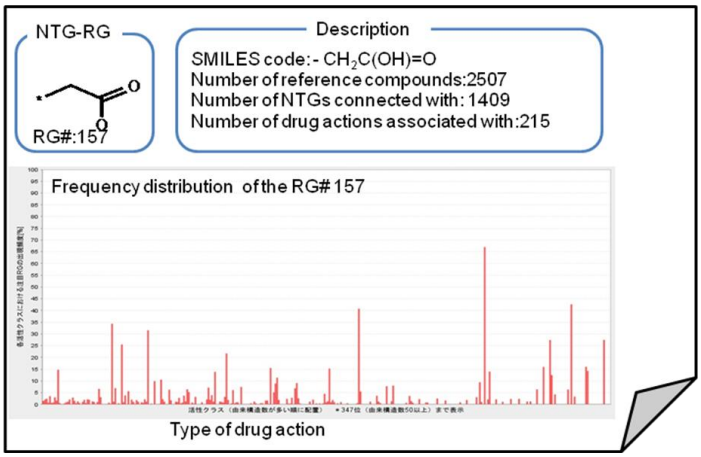


Fig.2 An example of viewing the relative frequency of $–CH_2COOH$ in various drug activity classes

The horizontal axis shows the type of drug action and the vertical axis shows the relative frequency of the substituent (NTG-RG).

It is apparent that the NTG-RG dictionary enables us to choose somehow statistically significant substituent(s) when we consider some side chains on a NTG that is associated with a drug activity of interest.

References

[1] T. Ohno, Y. Takahashi, Structure Feature Analysis of Drugs Using non-terminal Vertex Graph (NTG), Proc. of The 30th Symposium on Structure-Activity Relationships, 2002, pp. 41-42.

[2] Y. Takahashi, Chemical data mining based on non-terminal vertex graph. In *Systems, Man and Cybernetics;* 2004 IEEE International Conference, 2004; Vol. 5, pp. 4583 – 4587.

[3] R. Kuribayashi, Y. Takahashi, Construction of Knowledge-base for Graph Mining of Drug Molecules Using NTG and Application to Activity Prediction, *The 20th Annual Conference of the Japanese Society for Artificial Intelligence*, 2006, 1A2-1.

[4] K. Noto, Y. Takahashi, NTG Knowledge-based System for Candidate Structure Design in Drug Development, *The 22nd Annual Conference of the Japanese Society for Artificial Intelligence*, 2008, 2P2-13.