Vzyme: A Template-Based Method to Predict Reactions between Two Chemical Compounds

Yasushi Okuno okuno@kuicr.kyoto-u.ac.jp
Masahiro Hattori hattori@kuicr.kyoto-u.ac.jp
Masaaki Kotera kot@kuicr.kyoto-u.ac.jp
Yoshinobu Igarashi igarashi@kuicr.kyoto-u.ac.jp
Susumu Goto goto@kuicr.kyoto-u.ac.jp
Minoru Kanehisa kanehisa@kuicr.kyoto-u.ac.jp

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Keywords: metabolomics, network inference, enzyme, Vzyme

1 Introduction

Metabolomics plays an important role for linking genomes and cellular functions, since metabolites are components of cellular regulatory processes and their levels can be related to responses of biological systems to genetic or enviromental changes. Compared to genome, proteome, and transcriptome, however, metabolome is just beginning to be unraveled because of the difficulties of experimental measurements. Especially, secondary metabolites are chemically instable and have complicated structures and their pathways are less well understood. We have developed a virtual enzyme system (Vzyme), which constructs metabolite networks by predicting whether each chemical compound has a reactive connection with each other. Vzyme is a template-based search program utilizing chemical knowledge and consisting of two steps. In the first step, empirical knowledge of reactant-product relationships is acquired from the REACTION section of the LIGAND database and organized in the form of a template library. In the second step, a query compound pair is checked with the template library and if any template is found an edge is made between the pair. We applied this method to infer reaction pathways of phytosterols which are plant secondary metabolites.

2 Method

From an atomic-level representation of chemical structure transformation before and after the reaction process, we can extract features to identify possible enzymatic reactions. That is, we can use the replacement pattern of each atom obtained by the comparison of two molecular structures and consider it as a candidate of the template characterizing each reaction. Here a replaced atom (or atom group) through a reaction process is defined as its reaction center position which the enzyme directly attacks. In general, the number of reaction centers is very small (one or two) in most enzymatic reactions.

In order to substantialize above rules, Vzyme computes the best molecular alignment of two molecules which has the smallest number of reaction centers and the maximum number of aligned atoms, using the association graph method [1] and depth first search algorithm. The reaction template is represented as an atomic matching pattern of the reaction center and its vicinal atoms between two compounds.

3 Results

3.1 Number of Reaction Templates

We used KEGG/REACTION database [2, 3, 4] and manually compiled 3955 binary (reactant-product) relations. We then extracted 1465 reaction templates from the binary relational pairs, which forms the current version of Vzyme.

3.2 Template Search and Prediction of Enzymatic Reaction

Given two query chemical compounds, Vzyme picks up reaction template candidates between the two and checks them with the template library. If all of the candidates are present in the library, it is highly possible that the two compounds have reactant-product relationships of enzymatic reactions.

3.3 Application: Prediction of Phytosterols Metabolic Pathway

Phytosterol family (32 entries in KEGG/LIGAND) was selected as an example of plant secondary metabolites. Inferred metabolic networks consisted of 16 phytosterol compounds and 23 possible enzymatic reactions links. As shown in Figure 1, Vzyme returns not only enzymatic relations of two query compounds but also the most suited EC numbers to the relations. In order to construct global metabolic network, i.e. metabolome, all the compound entries in KEGG/LIGAND are also under calculation.

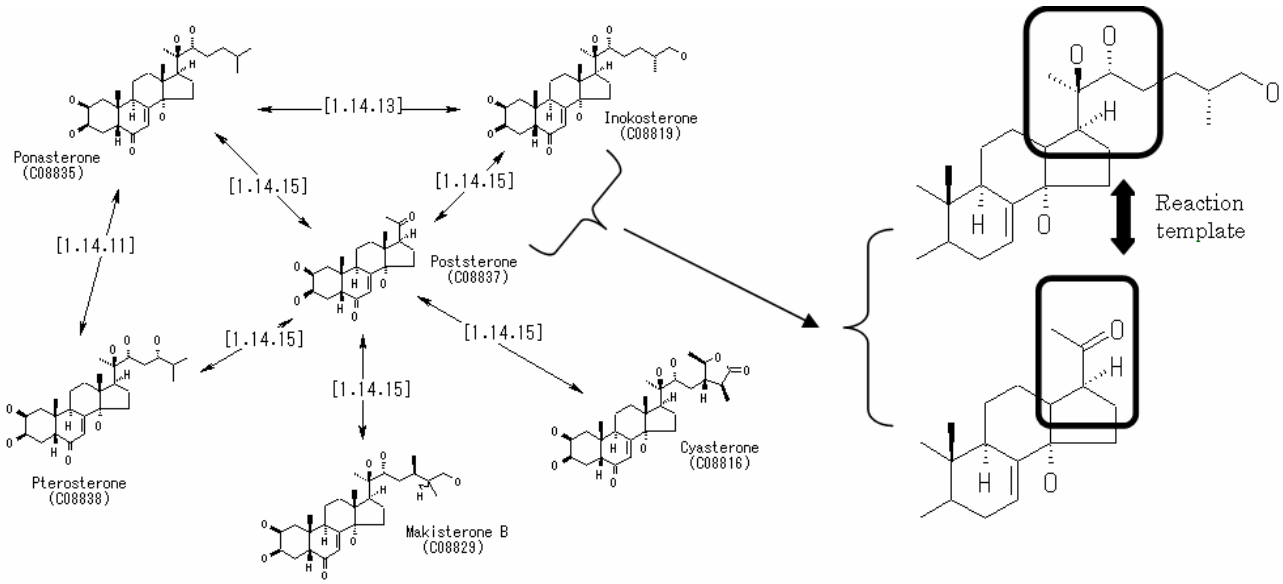

Figure 1: A snapshot of predicted phytosterols metabolic network (left) and a example of a reaction template (right).

Acknowledgments

We thank Rumiko Yamamoto, Tomoko Komeno and Koichiro Tonomura for improvement and maintenance of LIGAND database. This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan and from the Japan Society for the Promotion of Science.

References

[1] Hattori, M., Okuno, Y., Goto, S., and Kanehisa, M., manuscript in preparation.

[2] Kotera, M., Okuno, Y., Hattori, M., Goto, S., and Kanehisa, M., Classification and analysis of KEGG/REACTION database, *Genome Informatics*, 13:490–491, 2002.

[3] Yamamoto, R., Goto, S., Okuno, Y., Hattori, M., Komeno, T., and Kanehisa, M., Improvement of the LIGAND chemical database, *Genome Informatics*, 13:492–493, 2002.

[4] `http://www.genome.ad.jp/ligand/`