

## **PIERO ontology for analysis of biochemical transformations: Effective implementation of reaction information in the IUBMB Enzyme List**

Masaaki Kotera 1, Yosuke Nishimura 2, Zen-Ichi Nakagawa 1,2, Ai Muto 3, Yuki Moriya 2,4, Shinobu Okamoto 4, Shuichi Kawashima 4, Toshiaki Katayama 4, Toshiaki Tokimatsu 2,4, Minoru Kanehisa 2, and Susumu Goto 2

1 Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Japan

2 Bioinformatics Center, Institute for Chemical Research, Kyoto University, Japan

3 Graduate School of Biological Sciences, Nara Institute of Science and Technology (NAIST), Japan

4 Database Center for Life Science, Research Organization of Information and Systems, Japan

### **Abstract**

Genomics is faced with the issue of many partially annotated putative enzyme-encoding genes for which activities have not yet been verified, while metabolomics is faced with the issue of many putative enzyme reactions for which full equations have not been verified. Knowledge of enzymes has been collected by IUBMB, and has been made public as the Enzyme List. To date, however, the terminology of the Enzyme List has not been assessed comprehensively by bioinformatics studies. Instead, most of the bioinformatics studies simply use the identifiers of the enzymes, *i.e.*, the EC numbers. We investigated the actual usage of terminology throughout the Enzyme List, and demonstrated that the partial characteristics of reactions cannot be retrieved by simply using EC numbers. Thus, we developed a novel ontology, named PIERO, for annotating biochemical transformations as follows. First, the terminology describing enzymatic reactions was retrieved from the Enzyme List, and was grouped into those related to overall reactions and biochemical transformations. Consequently, these terms were mapped onto the actual transformations taken from enzymatic reaction equations. This ontology was linked to Gene Ontology and EC numbers, allowing the extraction of common partial reaction characteristics from given sets of orthologous genes and the elucidation of possible enzymes from the given transformations. Further future development of the PIERO ontology should enhance the Enzyme List to promote the integration of genomics and metabolomics.