## Preface

# The Second Workshop on Ontology and Genome

Takako Takai-Igarashi[1], Toshihisa Takagi[2] and Edgar Wingender[3,4]

[1] *Graduate School of Information Science and Technology, University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
E-mail: taka@bi.is.s.u-tokyo.ac.jp*
[2] *Department of Computational Biology, Graduate School of Frontier Sciences, University of
Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba, 277-8562, Japan
E-mail: tt@k.u-tokyo.ac.jp*
[3] *Dept. of Bioinformatics, Medical School, University of Göttingen, Goldschmidtstr. 1, D-37077
Göttingen, Germany
E-mail: e.wingender@med.uni-goettingen.de*
[4] *BIOBASE GmbH, Halchtersche Str. 33, D-38304 Wolfenbüttel, Germany
E-mail: ewi@biobase.de*

In the frequently proclaimed post-genome sequencing era, integration of all the biological knowledge on genes and gene products is required for computational re-engineering of biological mechanisms. To approach this goal, ontologies may play a key role since they provide semantics of biological knowledge in human- and computer-readable form. Just recently, and largely pushed by the development of the widely accepted Gene Ontology (GO; http://www.geneontology.org/), the potential value of properly built ontologies for these purposes is going to be recognized by an increasing number of scientists. It seems, however, that the full potential of ontological approaches in handling biological information is still to be explored.

The "Workshop on Ontology and Genome" (http://www.bi.s.u-tokyo.ac.jp/ontology-workshop/) which was was held in May 2003 in Tokyo aimed at providing an opportunity for discussing these issues in general and, in particular, topics on ontologies and their use for integration of biological data (for the program, see http://www.bi.s.u-tokyo.ac.jp/ontology-workshop/#program). It was the second workshop of this recently launched series. The preceding one was about "Natural Language Processing and Ontology Building for Biology" and was held in Tokyo in February 2002 (http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA/WS.html).

. It considered the necessity that functional genomics have to make efficient use of the overwhelming flow of published data about specific biological systems every day, but also has to make use of the knowledge that has been gathered and published over decades of "pre-genomic" research. For this purpose, efficient text-mining and natural language processing (NLP) tools will become of high value. Thus,

the 2002 workshop aimed at mediating cooperations in building and promoting ontologies for NLP in biology, at planning tutorials, workshops etc. Such promotion has been actually progressed after the workshop.

The second workshop now attempted to find out how ontologies of molecular networks and their cellular contexts (histological and anatomical environments) can be constructed, for which additional areas specific ontologies are urgently needed, what the most appropriate methods are to achieve that, and how they could optimally applied onto the reconstruction of the tremendously complex biological knowledge in this field.

After two days of internal discussions amongst experts in the field, a public Symposium was held to inform a broader audience about the state of the art and some recent achievements in the field. Some of the presentations of both these events were summarized by the authors and are published now as this Special Issue of ISB. After introducing the use of Gene Ontology for protein database annotation by Evelyn Camon *et al.* [1], the development and application of ontologies for annotation of the mouse genome is described by Hideo Matsuda and co-workers [2]. More recently, "immunoinformatics" is an upcoming discipline [3] which makes use of proper ontological formalizations as well as is described in the contribution by Marie-Paule Lefranc [4]. The group of Kousaku Okubo presents its approaches and application of an ontology about anatomical concepts [5]. Other attempts to ontologically model anatomical relations were described at the workshop by John Gennari with the aim of symbolic modeling of the structure of the human body, and by Edgar Wingender in the context of modeling genetic and signaling networks [6]. An ontological systematization of biological processes is detailed by Toshihisa Takagi and colleagues [7], and an ontology for cell signaling networks is presented by Takako Takai-Igarashi and Riichiro Mizoguchi [8]. At the end of this Special, Isabel Rojas and colleagues give a short review on the use of ontologies specifically for metabolic networks with their own works embedded [9].

A couple of summarizing statements and questions were put together at the end of and after the workshop, intended to stimulate the ongoing discussions about proper features and applicability of ontologies (http://www.bi.s.u-tokyo.ac.jp/ontology-workshop/#summary). For the GO team, E. Camon and M. Ashburner have responded to these topics (http://www.bi.s.u-tokyo.ac.jp/ontology-workshop/GO_responses.html)." It is envisaged to continue this discussion during a follow-up workshop in 2004. This third one of the series is planned for July 1-3 in Göttingen focusing on "**Development and Application of Ontologies on OMICS Data** (http://www.bioinf.med.uni-goettingen.de/ontology-workshop/)".

It is intended to provide a platform for experts in genomics, transcriptomics, proteomics, and metabolomics to exchange their experiences in using existing, or their needs in new ontologies for their work, with colleagues involved in ontological research.


## REFERENCES

[1]     Camon, E., Barrell, D, Lee, V., Dimmer, E. and Apweiler, R. (2003). The Gene Ontology Annotation (GOA) Database - An integrated resource of GO annotations to Swiss-Prot and TrEMBL. In Silico Biol. **4**, 0002.
[2]     Kasukawa, T., Bono, H., Hayashizaki, Y., Okazaki, Y. and Matsuda, H. (2003). MaXML: mouse annotation XML. In Silico Biol. **4**, 0003.
[3]     Petrovsky, N., Schönbach, C. and Brusic, V. (2003). Bioinformatic strategies for better understanding of immune function. In Silico Biol. **3**, 0034.
[4]     Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Bosc, N., Folch, G., Guiraudou, D., Jabado-Michaloud, J., Magris, S., Scaviner, D., Thouvenin, V., Combres, K., Girod, D., Jeanjean, S., Protat, C., Yousfi-Monod, M., Duprat, E., Kaas, Q., Pommié, C., Chaume, D. and Lefranc, G. (2003). IMGT-ONTOLOGY for immunogenetics and immunoinformatics, http://imgt.cines.fr. In Silico Biol. **4**, 0004.
[5]     Hishiki, T., Ogasawara, O., Tsuruoka, Y. and Okubo, K. (2003). Indexing anatomical concepts to OMIM Clinical Synopsis using UMLS Metathesaurus. In Silico Biol. **4**, 0005.

[6]     Wingender, E. (2003). TRANSPATH, TRANSFAC and CYTOMER as starting points for an ontology of regulatory networks. In Silico Biol. **4**, 0006.

[7]     Fukuda, K. I., Yamagata, Y. and Takagi, T. (2003). FREX: A database of biological processes with a hierarchical and recursive representation structure. In Silico Biol. **4**, 0007.

[8]     Takai-Igarashi, T. and Mizoguchi, R. (2003). Cell Signaling Networks Ontology. In Silico Biol. **4**, 0008.

[9]     Rojas, I., Ratsch, E., Saric, J. and Wittig, U. (2003).Notes on the use of ontologies in the biochemical domain. In Silico Biol. **4**, 0009.