Preface

Information and Simulation Systems for the Analysis of Gene Regulation and Metabolic Pathways

Ralf Hofestädt[1], Nikolay Kolchanov[2] and John Reinitz[3]

[1]*AG Bioinformatics / Medical Informatics, Technical Faculty, University Bielefeld, P.O. Box 100131, 33501 Bielefeld, Germany*
E-mail: ralf.hofestaedt@uni-bielefeld.de
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, RussiaAcad. Lavrentiev ave., 10, Novosibirsk, 630090, Russia*
E-mail: kol@bionet.nsc.ru
[3]*Department of Applied Mathematics and Statistics, The University at Stony Brook, Stony Brook NY 11794-3600, USA*
E-mail: reinitz@ams.sunysb.edu

This special issue comprising the ISB articles 0004 to 0013 [1,2,3,4,5,6,7,8,9] represents selected papers of the third Dagstuhl Seminar for Information and Simulation Systems for the Analysis of Gene Regulation and Metabolic Pathways, which was held from June 24 to 29, 2001. It was a multidisciplinary seminar with participants from 11 different countries. Schloss Dagstuhl workshops in general emphasize computer science, and we were delighted to focus on the rapidly developing links between biosciences and computer sciences. The 2001 meeting is a sequel to the 1995 and 1998 meeting on similar topics (see http://www.techfak.uni-bielefeld.de/ags/bi/events/index.html).

Molecular biology and biotechnology have begun to focus sharply on the problem of gene regulation. This problem is inescapable, because no open reading frame (ORF) will be expressed without the appropriate regulatory sequences. Moreover, some genes code for proteins whose function is to turn other genes on and off. Groups of these genes form networks with complex behaviors. These networks control other genes whose protein products catalyze specific biochemical reactions, and the small molecules which are substrates or products of these reactions can in turn activate or deactivate proteins which control transcription or translation. For that reason, gene regulation can be said to indirectly control biochemical reactions in cellular metabolism, and cellular metabolism itself exerts control on gene expression. For these reasons, the interdependent biochemical processes of metabolism and gene expression can and should be interpreted and analyzed in terms of complex dynamical networks. Hence modeling and

simulation are necessary. Two earlier Dagstuhl seminars (1995 and 1998) have already dealt with modeling and simulation of biochemical networks. Both sought to bridge two divides by both bringing together scientists in the disciplines of gene regulation and metabolic pathways, and within and across both of these areas bringing together experimentalists and theoreticians. Often there had been little previous contact among these groups, but clearly the integration of metabolic and gene expression models as well as the cooperation of theorists and experimentalists is essential in order to solve these complex problems.

Apart from theoreticians and experimentalists, a third group has emerged since 1995 which is centered around databases and the internet. Many molecular biologists turned towards informatics and systematically collected results relating to specific problems. These data have been and will be stored systematically in specific databases, which nowadays are accessible via the Internet. Recently many firms have been founded which provide data essential for the solution of scientific and industrial problems, and even more importantly the corresponding infrastructure. As a result, there are databases available via the Internet for all known sequenced genes (e.g. EMBL), proteins (e.g. SWISS-PROT, PIR, BRENDA), transcription factors (TRANSFAC), biochemical reactions (KEGG) and signal induction reactions (TRANSPATH, GeneNet). Beyond databases, simulators for metabolic networks which employ most of the currently popular modeling methods are also available via the Internet. In addition to the classical methods of differential equations, discrete methods have become quite important. Examples are the object-oriented approach, rule-based systems, Petri Nets, graphs, and Boolean nets.

These recently implemented tools on the Internet are the basic components of the informatic and analytical infrastructure of biotechnology. Clearly the next evolutionary stage of development will be the implementation of integrated molecular information systems (e.g. SRS). The first step to reach that goal is the integration of databases under a specific biological perspective. The next step will be user-defined molecular information fusion. Up to now, there are no standard tools available in order to successfully separate both methods and databases. Exactly for that reason it is imperative to develop uniform intersections at this stage. To discuss properties of these intersections was one major issue of this seminar.

REFERENCES

[1]   Stephanik, A., Freier, A., Hofestädt, R., Lange, M. and Scholz, U. (2002). BioDataServer: A SQL-based service for the online integration of life science data. In Silico Biol. **2**, 0005.
[2]   Schreiber, F. (2002). High Quality Visualization of Biochemical Pathways in BioPath. In Silico Biol. **2**, 0006.
[3]   Rojas, I., Bernardi, L., Ratsch, E., Kania, R., Wittig, U. and Saric, J. (2002). A database system for the analysis of biochemical pathways. In Silico Biol. **2**, 0007.
[4]   Moreno-Hagelsieb, G. and Collado-Vides, J. (2002). Operon conservation from the point of view of Escherichia coli, and inference of functional inter-dependence of gene products from genome context. In Silico Biol. **2**, 0008.
[5]   Kolchanov, N. A., Nedosekina, E. A., Ananko, E. A., Likhoshvai, V. A., Podkolodny, N. L., Ratushny, A. V., Stepanenko, I. L., Podkolodnaya, O. A., Ignatieva, E. V. and Matushkin, Y. G. (2002). GeneNet database: Description and Modeling of Gene Networks. In Silico Biol. **2**, 0009.
[6]   Jaeger, J. and Goodwin, B. C. (2002). Cellular Oscillators in Animal Segmentation. In Silico Biol. **2**, 0010.
[7]   Kozlov, K., Myasnikova, E., Pisarev, A., Samsonova, M. and Reinitz, J. (2002). A method for two-dimensional registration and construction of the two-dimensional atlas of gene expression patterns in situ. In Silico Biol. **2**, 0011.
[8]   Arrigo, P., Ivaldi, G. and Cardo, P. P. (2002). In silico determination of potential antisense targets for human beta-globin variants. In Silico Biol. **2**, 0012.
[9]   Cooper, M., Chapman, S. C., Podlich, D. W. and Hammer, G. L. (2002). The GP problem: Quantifying gene-to-phenotype relationships. In Silico Biol. **2**, 0013.