

Systems Biology in Supercomputing Environment

Satoru Miyano

Human Genome Center, Institute of Medical Science, The University of Tokyo
4-6-1 Shirokanedai, Minatoku, Tokyo 108-8639, Japan

We developed a software platform XiP (eXtensible integrative Pipeline) that is a flexible, editable and modular environment with a user-friendly interface for systems biology. With XiP, we can build various analysis workflows including simulation, visualization, pathway modeling, etc. Construction of a workflow under XiP follows the intuitive notion of dragging and dropping. XiP already equips with several "ready to use" pipeline flows for the most common analysis and more than 300 statistical/computational analysis components, and XiP also recognizes components written in R (<http://www.r-project.org>). It runs under multiprocessor environments. The source code is available as open source under the Lesser General Public License (LGPL) (download: <http://xip.hgc.jp/wiki/en/>).

On the other hand, we have been developing a modeling and simulation tool Cell Illustrator Online 5.0 (CIO, <https://cionline.hgc.jp/>) that enables us to draw, model, elucidate and simulate complex biological processes and systems such as metabolic pathways, signal transduction cascades, gene regulatory pathways and dynamic interactions of various biological entities. Cell System Ontology (CSO) (<http://www.csml.org/>) and Cell System Markup Language (CSML) constitute the basis of CIO. CIO uses Hybrid Functional Petri Net with extension (HFPNe) which is defined by enhancing some functions to hybrid Petri net so that various aspects in biological pathways can be intuitively modeled. CIO has been used to build, for example, a simulatable macrophage pathway knowledge base (MACPAK) (<http://macpak.csml.org/>). A key technology for driving systems biology is a method for automatic parameter estimation for models. For this purpose, we developed a method for HFPNe, namely CIO, by using a technology called data assimilation which "blends" simulation models and observational data "rationally". This data assimilation method requires super-computer systems.

We are currently creating the next generation supercomputing environment for systems biology on XiP by enrolling CIO, data assimilation, model checking, and various systems biology tools such as SiGN (large-scale gene network estimation software package (<http://sign.hgc.jp/>)) on XiP.