Cell Illustrator 3.0: A Platform for Biopathway Modeling and Simulation

Masao Nagasaki  Atsushi Doi  Ayumu Saito
masao@ims.u-tokyo.ac.jp  doi@ims.u-tokyo.ac.jp  ayumu@ims.u-tokyo.ac.jp

Emi Ikeda  Kazuko Ueno
e-ikeda@ims.u-tokyo.ac.jp  uepi@ims.u-tokyo.ac.jp

Euna Jeong  Christopher J. Savoie  Satoru Miyano
euna@ims.u-tokyo.ac.jp  savoie@gene-networks.com  miyano@ims.u-tokyo.ac.jp

1 Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo, 108-8639, Japan
2 Toranomon 4-chome MT Bldg., 4-2-12 Toranomon, Minato-ku, Tokyo, 105-0001 Japan

Keywords: biopathways, simulation, database, Petri net, Cell Illustrator, pathway modeling

1 Introduction

In the post-genome era, the system level understanding of life is one of the most important issues in bioinformatics and systems biology. As one of the solutions, we are developing a software named Cell Illustrator (CI) [1]. This software aims at modeling and simulating structurally complex dynamic causal interactions and processes, e.g. metabolic pathways, signal transduction cascades, gene regulatory networks, and cell-cell interactions [1, 2]. The architecture of this software is based on Hybrid Functional Petri net with extension (HFPNe) [2], which inherits features of original Petri net - a visual modeling language and suitable for modeling parallel, asynchronous, distributed, non-deterministic, and stochastic processes - and acquire a new feature - two special components generic entity and generic process are introduced. These new elements enable HFPNe to faithfully represent objective notions as other programming languages, e.g. Java and C++. This feature enables the following advances in modeling processes of pathways: (i) An entity can contain more than one value, e.g. list and pair. Proteins often have many modified states, e.g. p53 has known sixteen phosphorylation positions and two acetylation positions and modified states of p53 can be \(2^{18}\). These states are modeled with one generic entity. (ii) An entity can handle more advanced primitive type, e.g. boolean and string. It is especially useful for sequence based modeling of translation and transcription mechanism, since these models need elements that represent sequence information similar to string. (iii) An entity can take a more advanced type, e.g. object that consists of fields and methods and a process can operate more complex reactions. Many parts in a cell, e.g. DNA, mRNA and protein have known specific operations, e.g. translation, transcription, degradation and modification. Thus, if an entity takes the type mRNA, each process that connects to the entity simply needs to specify a method from the methods of the mRNA.

2 Cell Illustrator 3.0

With the HFPNe architecture, we have released a biological pathway modeling and simulating workbench named Cell Illustrator 3.0 [1] with the following features: (i) More than 350 biological figures are newly created for easy pathway modelings. All figures are scalable vector graphics and can keep
Figure 1: (a) Cell fate model of gustatory neurons with microRNA regulations in CI3.0 (pathway modeling mode) and biological figures with ontology associations. (b) An inferred regulatory network on CI3.0 (gene network analysis mode).

high quality for graphical operations (see Figure 1(a)). (ii) Every 350 biological figure has an ontology term in Cell System Ontology [2]. With the functionality, user can create an ontology based pathway knowledgebase without much effort. The knowledgebase enhances the reusability and allows high knowledge-processing technologies. (iii) Major chemical equations are included for better compatibility with other differential equation based softwares, (iv) Automatic pathway layout functionalities are implemented for large-scale pathway modeling and simulation that was proposed in Kato et al. [2]. (v) The application supports the one of open XML formats for pathway modeling and simulation named Cell System Markup Language (CSML) [2]. The converters of other major XML formats, e.g. SBML and CellML, are developed for better reusability of pathways in other formats. (vi) Probe IDs in major microarray data can be mapped to entities in pathway models. In addition, useful functionalities that are necessary for gene network analysis are equipped, e.g. comparisons of gene networks, reachability analyses, and plots of microarray expression data (see Figure 1(b).)

On the new version, we have modeled new pathways, e.g. a cell fate determination model with microRNA in C. elegans (see Figure 1(a)) and a p53 stabilization pathway in mouse [2] with ontology information. The free licensed softwares, Cell Illustrator Player 3.0 and Cell Illustrator Draw 3.0 are available from [1].

References
