What is SBGN?

The Systems Biology Graphical Notation (SBGN, see also Le Novère et al., 2009) is a standard notation for the visual representation of biochemical and cellular processes. SBGN maps can be used to exchange complex biological knowledge in a clear, concise and unambiguous manner. See figure 1 for an example of the SBGN Process Description (PD) language.

Implementation

To meet this need, we defined a new markup language named SBGN-ML. This file format is based on an XML Schema definition (XSD). See figure 2 for an example.

In addition we developed a supporting software library called LibSBGN. Besides reading and writing SBGN-ML files, this library will also be used to validate SBGN maps against the specifications, and convert to- and from related systems biology standards, such as SBML (with layout extension) and BioPAX. LibSBGN is currently implemented in Java and C++. LibSBGN is still under development, but it is already being adopted by several tools (see the project wiki for an up-to-date list).

A test suite has been created; it consists of dozens of reference maps covering every aspect of SBGN. Corresponding reference documents have been encoded in XML. To test the adherence of supporting tools to the SBGN-ML specification, a rendering comparison pipeline was set up (see figure 3); it automatically renders SBGN maps through different programs, thus enabling comparison against the reference map.

What is LibSBGN?

As SBGN is becoming adopted more widely, there is a need for exchanging maps electronically between the various systems biology tools. Exchange using graphics-only file formats (such as SVG) is often insufficient, because biological meaning of elements is lost. There is a need for a toolset to exchange maps while preserving biological meaning and relations.

Roadmap

Milestone 1: (completed Feb 2011)
- Implement semantics of Process Description language Level 1 Version 1.1.
- Only high-level specifications.
- Low-level validation with XML Schema
- Beta support for C++

Milestone 2:
- Implement all three languages: PD, ER and AE.
- Support for third party extensions
- High-level validation with Schematron

Milestone 3:
- Pixel-perfect graphics specification
- And beyond...
  - Conversion to and from SBML (with layout extension)
  - Conversion to and from BioPAX

Acknowledgements

The authors acknowledge the support of a BBSRC United States Partnering Award (BB/G530441/1) for travel and other collaboration costs.

References

N. Le Novère et al., The Systems Biology Graphical Notation, Nature Biotechnology 2009, 27(8):735-41