Curation and annotation for BioModels Database, a resource of published quantitative models in biology

Vijayalakshmi Chelliah, Lukas Endler, Camille Laibe, Chen Li, Nicolas Le Novère
European Bioinformatics Institute, Hinxton, UK

INTRODUCTION

BioModels Database [1] (<http://www.ebi.ac.uk/biomodels>) is a free resource for storing, viewing and retrieving quantitative, kinetic models described in peer reviewed publications. It uses the XML based *Systems Biology Markup Language* (SBML) [2] for model storage and internal representation, but allows submission and export of models in various other commonly used formats. While models can be submitted publicly, submitted models are not directly included in the database, but checked and annotated both automatically and manually by the BioModels Database curation team. The database consists of two branches - one for **manually curated and annotated** models and one for **non-curated** ones.

Standards for curation and annotation of quantitative kinetic models have been developed quite recently and the **BioModels.net** (<http://biomodels.net>) project is heavily involved in these efforts. Apart from the model database it also encompasses amongst others the **MIRIAM resources** [3] and the **Systems Biology Ontology**(SBO)[4].

ANNOTATION

The different elements of the models are extensively annotated with references to controlled vocabularies and links to other databases to allow for identification and search. Annotation is mostly carried out manually, although we plan on using some semiautomatic support in the future.

Annotations in both branches:

- model submitter and creator contacts
- reference publication data
- creation and last modification dates

Curated branch only:

- **MIRIAM annotations** of different elements of the model
- **Systems Biology Ontology (SBO)** annotations (SBML l2 v2 and above)

Annotations are saved separately as entries in the database and in the model files. **MIRIAM annotations** are included in SBML files using **RDF** in *annotation* elements, **SBO** terms are directly inserted in the corresponding *sboTerm* attribute of the annotated element.

MIRIAM annotations

represent information as triplets:

- annotated element: eg. reaction Notch activation
- resource and identifier in form of a **URN** (*Uniform Resource Name*)
 - resource: eg. Gene Ontology
 - identifier: eg. Notch receptor processing (**GO:0007220**)
- BioModels.net qualifier: eg. *isVersionOf*

one or more **URNs** identify the data-type and the data-set, the **BioModels.net** qualifier indicates the relation between the annotated element and the **URN**. **URNs** following the MIRIAM scheme can be resolved to **URLs** using the **MIRIAM Resources Web Services** (<http://www.ebi.ac.uk/miriam>)

annotated reaction in BioModels Database

CURATION

To offer scientists reliable information, models are checked to comply to a variant of the **MIRIAM** (*Minimal Information Requested In the Annotation of biochemical Models*) [5] requirements:

- the model must be encoded in **valid SBML**
- it must have a **single, peer-reviewed reference publication**
- it must contain the **creator's** and **submitter's** contact details
- the model's **structure** must reflect the **processes** described in the reference publication
- it must be able to **reproduce all/some results** of the reference publication

Any model has to fulfill the first four requirements to be accepted into BioModels Database. **Non-kinetic** models (eg. FBA) and models in **certain modeling frameworks** (eg. boolean or petri nets) are only considered for the **non-curated** branch up to now.

If a model does not reproduce the results in the article, the **curators** contact the **submitter**, **creators** and **authors** of the model to find and correct potential errors. All **deviations** from the published version are recorded in the model's **notes**. A **figure** or **table** of the model's results has to be attached to the model by the curators to move the model to the curated branch.

Curated and fully annotated model entry in BioModels Database.
The curation result and some annotation links are shown.

References

- 1) N. Le Novère et al. : *BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems*. Nucleic Acids Res. **34**:689-691 (2006)
- 2) M. Hucka et al. : *The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models*. Bioinformatics **19**:524-531 (2003)
- 3) C. Laibe and N. Le Novère: *MIRIAM Resources: tools to generate and resolve robust cross-references in Systems Biology*. BMC Syst. Biol. **1**:58 (2007)
- 4) N. Le Novère, M. Courtot and C. Laibe: *Adding semantics in kinetics models of biochemical pathways*. Proc. of the 2nd Int. ESSEC Symp. 137-153 (2006)
- 5) N. Le Novère et al. : *Minimum information requested in the annotation of biochemical models (MIRIAM)*. Nature Biotechnol. **23**:1509-1515 (2005)