combine the computational modeling in biology network

Dates: ***August 15–19, 2012***
Location: ***Toronto, Canada***

WiFi

Network name – UofT

Or Network name - UTORwin
Security type - WEP
Security key - UToronto1home

Once connected to the network to get online use the following information via a web browser:
UTORid: COMBINE
Password: 2012med1

Sponsors

Donnelly Centre for Cellular + Biomolecular Research
UNIVERSITY OF TORONTO

OntarioGenomicsInstitute

OICR
Ontario Institute for Cancer Research

Contents

1 Keynotes

1.1 How to be an educated ontology consumer

Presenting author: Michael Grüninger

1.2 Comprehensive measure and modeling of *C. elegans* behavior

Presenting author: Will Ryu

1.3 Not your mama's network biology – dynamic cell-cell interaction networks in development and disease

Presenting author: Peter Zandstra

1.4 Model-based engineering of metabolism

Presenting author: Radhakrishnan Mahadevan, University of Toronto

Bioprocess development for biofuels and biochemicals typically requires several rounds of metabolic engineering to meet process targets including product yield, titer and productivity, all of which impact the process economics. Recent advances in experimental and computational technologies have enabled the detailed characterization of biological systems. In particular, the molecular components of these systems including the list of genes, proteins they encode, and compounds that interact with these proteins can be determined. Similar advances in computational modeling techniques have allowed the development of genome-scale models of metabolism in several organisms. In this talk, the use of such models for metabolic engineering will be presented. Model refinement through the incorporation of a fundamental physical constraint that accounts for membrane area will be described. In the first part, a rational approach based on bi-level optimization to enhance bioprocess productivity by forcing co-utilization of substrates will be shown. Experimental results from the application of this approach to enforce substrate co-utilization in Escherichia coli will be discussed. In the next part of the talk, a novel nested nonlinear optimization method for metabolic engineering resulting in over hundred different strain design strategies for biochemicals production will be presented. We will also examine the role of the redundant pathways from a design perspective and present computational results on how these pathways is valuable for robust design. Finally, we will present a synthetic biology approach for dynamic control of metabolism to improve productivity.

2 Biological Models and Simulations

2.1 Update on SBML level 3 and the SBML test suite

Presenting author: Mike Hucka

2.2 Draft for discussion / SBML specifications for revised multi, simple spatial and multi-spatial extensions

Presenting author: Fengkai Zhang, Laboratory of Systems Biology, NIAID/NIH

The working draft of "multi" (Multistate and Multicomponent Species, proposal tracking number: 2430531) is a proposal for the SBML extension handling "multistate and multicomponent" species. The "multi" extension, if complete and stable, provides a community standard for the exchange of sbml multi models. However, the current "multi" extension specification is still under development and tests are being performed to assess the practical applicability for defining models with multi-molecular/multi-state complexes. We propose some additions/revisions to the "multi" extension which we wish to put forward as an SBML "Revised Multi" proposal. We think that the modifications we suggest help to achieve the goals described in the SBML "multi" extension in a simpler and more efficient manner. In order to simplify the definition of spatial models, we also suggest a new SBML "Simple Spatial" extension (we use "Simple Spatial" to distinguish the existing "Spatial" proposal). The "Simple Spatial" extension is independent from the "Revised Multi" extension, but they can be used together. Finally, we suggest another SBML extension called "Multi-spatial". This extension requires both "Revised Multi" extension and "Simple Spatial extension". The detailed specifications about the three extensions as well as three adapted examples from the original multi, one bionetgen example and two Simmune model examples are provided as illustration.

2.3 The Open Source Brain initiative

Presenting author: Padraig Gleeson, University College London

The Open Source Brain repository (http://www.opensourcebrain.org/) aims to be a public repository to allow collaborative development of detailed neuronal models in standardised formats, with curated, stable releases which will evolve in line with new experimental findings, the latest modelling paradigms and simulator technology development. While the models can be collaboratively developed in any simulator format, the ultimate aim is to get as much of the model as possible into NeuroML format to ensure modularity, accessibility and cross simulator portability.

2.4 Experience with SBML packages within iBioSim

Presenting author: Chris Myers, University of Utah

This talk will describe our experiences with various SBML Level 3 packages within the latest version of our iBioSim tool. In particular, our tool now uses the hierarchical composition and layout packages as well as prototypes for array and dynamic packages. These packages coupled with some custom annotations to support the synthetic biology open language (SBOL) have allowed us to eliminate all use of our custom genetic circuit modeling language. The resulting models are, therefore, more readily exchangeable. We hope that our experiences described in this talk will help others who are considering the use of SBML packages in the near future.

2.5 Multiscale multicellular simulations using Cell Based Chaste

Presenting author: James Osborne, University of Oxford

Cell Based Chaste (www.cs.ox.ac.uk/chaste) is a multiscale framework for the mathematical modelling of multicellular biological systems. Utilising the natural structural unit of the cell, the framework consists of three main scales: the tissue level (macro-scale); the cell level (meso-scale); and the sub-cellular level (micro-scale), with interactions occurring between all scales. The cell level is central to the framework and cells are modelled as discrete interacting entities using one of a number of possible modelling paradigms, including lattice based models (cellular automata and cellular Potts)

and off-lattice models (cell centre and vertex based representations). The sub-cellular level concerns numerous metabolic and biochemical processes represented by interaction networks rendered stochastically or into ODEs. The outputs from such systems influence the behaviour of the cell level affecting properties such as adhesion and also influencing cell mitosis and apoptosis. Tissue level behaviour is represented by field equations for nutrient or messenger concentration, with cells functioning as sinks and sources.

This modular approach enables competing models to be rigorously compared and for emergent models to be implemented. Moreover, by considering alternative representations of the same processes, simulations can be used to draw connections between competing modelling frameworks promoting model exchange (for example between on and off lattice representations of cells).

2.6 Towards CellML 1.2

Presenting author: Jonathan Cooper, University of Oxford

Since COMBINE 2011, the CellML community has elected a five-member editorial board (http://www.cellml.org/community/editoria
The initial goal of the board has been to gather contributions from the community in the work toward the release of CellML 1.2 and to develop the processes by which future development of the specification can be continued in a sustainable manner. This has involved separating the normative parts of the specification from the informative part to ensure that there is no ambiguity in the actual definition of the specification, thus providing accurate points of reference for proposals to change the specification. The editorial board has also been discussing restricting the set of proposed changes for the next version of CellML to a minimal set that directly addresses current shortcomings of the specification which are inhibiting users from encoding their models in CellML. Here, we will present the work of the editorial board in addressing these and our plans for moving forward with the release of CellML 1.2.

2.7 The new JWS Online simulation interface: SBGN schema generation, MIRIAM annotation and SBML model specification

Presenting author: Franco B du Preez, University of Manchester

The simulation interface to the JWS Online model repository has been upgraded to include automatic SBGN schema generation. This new interface is part of the SysMO-SEEK platform used in the Pan-European SysMO consortium, and supports the import of SBML models, the display of MIRIAM compliant annotation as well as the export of SED-ML scripts for reproducing user-defined model simulations. Support for events has been added to the SBML import functionality, making the vast majority of biomodels simulable in JWS Online. Linking experimental data to models is a priority and a prototypical support system for automatically combining model and experimental results known as DataFuse will be presented. JWS Online currently includes a number of manually created and curated SBGN schemas and preliminary work on capturing and converting these schemes to SBGN-ML, to further enable the sharing and exchange of modeling research, will be demonstrated.

2.8 Tutorial: BioModels database

Presenting author: Camille Laibe, EMBL-EBI

BioModels database (http://www.ebi.ac.uk/biomodels/) is a free and open resource for dynamical models of biological processes, phenomena and systems.

A general introduction of the resource will be provided, explaining the curation and annotation processes. Then, the various ways to browse the content of the database will be demonstrated. This involves first finding models of interest by either using the model-tree, built based on Gene Ontology terms, or performing searches using complex queries. And second, exploring the content of the retrieved model(s) using the different facets of the resource.

More advanced features will be covered as well, like submission of new models, creation of a sub-model from a larger model according to the user's selections, launch of on-line simulations, ...

3 Simulation Algorithm, Experiments and Results

3.1 Introduction to SED-ML

Presenting author: Dagmar Waltemath

3.2 Nested tasks proposal for SED-ML

Presenting author: Frank T. Bergmann, Caltech / University of Heidelberg

In this presentation I will present the Nested Tasks proposal for extending SED-ML beyond the basic simulation classes it uses today.

3.3 Possibilities for SED-ML

Presenting author: Jonathan Cooper, University of Oxford

We have been working on defining and using virtual experiments within a variety of contexts, including cardiac electrophysiology, multi-cellular tissue dynamics, immunology, and synthetic biology. In doing so we have found that extensions are required if SED-ML is to address our needs. This talk will introduce some of our use cases, highlight our functional curation framework (https://chaste.cs.ox.ac.uk/cgi-bin/trac.cgi/wiki/FunctionalCuration) for running experiments on a range of models, and outline our proposals for SED-ML.

3.4 DDMoRe - developing new representation standards in PK/PD

Presenting author: Maciej J Swat, EMBL-EBI

PK and PD are branches of science dealing with drug related simulation and modelling. The former explains the time course of concentration in the human body, the latter the drug action. This can be done using both phenomenological or mechanism based models.

Although the complexity of human body is enormous, compartment based methods have been successfully applied in drug discovery and design for decades and remain indispensable today. Their application is straight forward for frequently measured individual data and results in subject specific set of physiologically meaningful parameters.

However, clinical data come often with few measurement point per subject, inter-subject and inter-occasion variability and one has to resort to population approaches which provide both population and individual parameters estimates. The statistical modeling framework called 'Nonlinear Mixed Effects' is the most popular one used in this context.

Surprisingly all this activities in this vital field of research has not seen much of standardisation efforts so far. There is number of tools, such as Monolix, NONMEM or WinBUGS whose code has to manually transcribed in order to use them in parallel. The recently launched DDMoRe (www.ddmore.eu) project is a pioneer IMI founded effort to feel this gap.

EMBL-EBI has the lead in the work package dedicated to the design of a set of system-to-system Modelling Markup Languages, MMLs. The goal is to use where possible existing standards, such as SBML, DED-ML, UncerML etc. The aspects one has to account for in the new standards are the structural model, parameter distribution, covariate model, correlation between parameters, clinical trial design and others. The MMLs will enable researchers in industry and academia to share their models, verify the results obtained with different tools and allow effective control vocabulary/ontology based search.

At EMBL-EBI we also drive the development of an infrastructure for storing and sharing of widely used existing disease models being currently implemented or the new ones being developed by the DDMoRe industrial and academic partners. Our BioModels Database, which is the leading world-wide resource for Systems Biology models, provides an indispensable experience to achieve this task.

The project has already lead to research activities from which both the Pharmacometrics and System Biology community will benefit. For example the work on notation and encoding of probability distributions is crucial for the former and missing in the latter field.

DDMoRe project has already an important impact on the unification efforts in M&S in Life Sciences. We present an overview of our ongoing efforts, highlight the challenges and issues.

3.5 Bridging experiments and modelling: the SABIO-RK reaction kinetics database

Presenting author: Martin Golebiewski, Heidelberg Institute for Theoretical Studies (HITS), Heidelberg, Germany

SABIO-RK (http://sabio.h-its.org) is a curated database for biochemical reaction kinetics data. The system offers standardized and cross-related data manually extracted from the literature or directly submitted from lab experiments. The database content (currently more than 40000 datasets) includes kinetic parameters and corresponding rate equations for biochemical reactions, described in their biological and experimental context. The captured data is standardized by the use of controlled vocabularies and annotations pointing to other resources and biological ontologies.

A new SABIO-RK web interface allows simple full-text search (e.g. for all kinetic data referring to 'homo sapiens liver'). Autocompletion of the search terms by lists of values and optional filtering of the search results facilitate the finding of data. For a more specific advanced search the building of complex queries is also supported by a query builder. Hierarchical search based on organism taxonomy or tissue and cell type ontology facilitates querying for related data (e.g. for all Mammalia or all liver tissues and cell types). Confidential data, such as directly submitted pre-publication data, are hidden from the public and can be accessed by authorized users after login.

We also have implemented easy-to-use RESTful web services allowing the programmatic data access by other tools (e.g. CellDesigner, VCell and others). Web interface and Web Services support the export of the data together with its annotations in SBML and in BioPax/SBPAX3 (Systems Biology Pathway Exchange) format.

SABIO-RK supports the exchange of annotated data in COMBINE standard formats and thereby facilitates the exchange of kinetic data between experimentalists and modellers, supporting the setup of quantitative computer models.

3.6 SEEK and JERM: standards-compliant integration of systems biology data

Presenting author: Olga Krebs, HITS gGmbH

The SEEK (http://www.seek4science.org/) is a web-based platform for the integrated storage and exchange of standardized experimental data and models, designed to facilitate sharing and collaboration in large, distributed projects. SEEK is currently used by different Systems Biology consortia including Systems Biology of Micro-Organisms[1], Virtual Liver Network[2] , RosAge[3], JenAge[4] and EviMalar[5], that all have to manage a large number of various types of experimental and theoretical data. Underlying the SEEK is the JERM (Just Enough Results Model), which is a minimum information model describing the content and structure of datasets as well as relationships between them. The JERM links SEEK data to MIBBI (Minimum Information for Biological and Biomedical Investigations) standards and community ontologies. JERM data compliance is managed by the distribution of spreadsheet templates, which can be semantically enabled with RightField[6], a tool that embeds ontology terms into specific spreadsheet cells. Within the SEEK, gateways have been built to other community resources, such as the JWS Online [7], Biomodels database [8], SYCAMORE [9], and PubMed amongst others.

References:

[1] http://sysmo.net
[2] http://www.virtual-liver.de/
[3] http://www.sbi.uni-rostock.de/research/research-projects/single/28/
[4] http://www.jenage.de/
[5] http://www.evimalar.org/
[6] http://www.sysmo-db.org/rightfield
[7] http://jjj.biochem.sun.ac.za/
[8] http://www.ebi.ac.uk/biomodels-main/
[9] http://sycamore.eml.org

4 Software Tools and Databases

4.1 BiNoM, a Cytoscape plugin for accessing and analyzing pathways using standard systems biology formats

Presenting author: Eric Bonnet, Curie Institute

An increasing number of detailed large-scale molecular maps have been published in the last years, covering crucial biological processes often disrupted in human diseases. Two examples recently published are the Rb/E2F and mTor maps, which are covering the signaling networks and reactions involved in the regulation of cell growth and proliferation, processes that are almost always dysregulated in oncogenic cells.

However, interpreting such large molecular maps, having hundreds of different molecular species and reactions, is not an easy task. BiNoM (Biological Network Manager) is a Cytoscape plugin software that aims to provide a set of functions to extract condensed and useful information from such maps. More specifically, BiNoM can import data from established standard Systems Biology file formats such as BioPAX, SBML, CellDesigner and CSML. Then a set of functions based on graph-based algorithms allow the user to analyze, decompose and construct modular representation of large scale networks. Such sub-networks can be used to interpret experimental data such as expression sets or can be used for advanced mathematical modeling.

We have implemented a set of novel functions in the latest version of BiNoM. The most recent version of the BioPAX standard (level 3) is now fully supported in BiNoM. BioPAX is very widely used within the Systems Biology community, particularly for large scale on-line repositories such as Reactome. We are also introducing a novel algorithm for the quantitative analysis of pathways in influence networks, taking into account the sign, length and activity of sets of pathways for a selected number of targets of interest. At last, we also present a novel approach for the determination of hit sets within a network of interest. A hit set can be defined as a finite set of nodes to "hit" in order to cut a given number of path from a set of source nodes to a set of target nodes of interest, an approach with evident practical applications in the domain of drug design. In order to demonstrate their utility, we provide concrete examples of the applications of all those novel functions on real biological networks, and more particularly on examples linked to cancer biology. BiNoM is open source, available from the Cytoscape plugin repository and from our website (http://binom.curie.fr).

4.2 Budhat - a tool for SBML model version control

Presenting author: Dagmar Waltemath, University of Rostock

With the application of computational methods in Biology, Systems Biologists deal with increasing numbers of models in their daily schedule. Computational models of biological systems evolve over time; the processes of reusing, expanding or, in general, modifying models is therefore an important aspect of the Systems Biologist's modeling workow.

In this talk, we introduce several building blocks for improved model version control for SBML models: a difference detector, a library for semi-automatic change detections, and a visual tool to display the history of a model to users. We have designed the software in such a modular manner that parts of it can easily be integrated into existing software, and we will demonstrate the working of our tools in a prototype web-based difference detector called Budhat.

4.3 JUMMP just a model management platform

Presenting author: Eils Juergen, DKFZ

Data management and standardization in the area of systems biology is a wide field that has been covered very successfully by international efforts. Biopax, SBGN, SBML as well as SBO, MIRIAM and JUMMP in the context of biomodels.net are major contributions. The DKFZ has initiated in collaboration with the EBI the JUMMP project, which intends the refactoring of the biomodels database. A short summary about recent progress will be presented. Also chances of inclusion of this Combine community into further progress of the JUMMP project (by a provided survey and maybe usability test) will be highlighted.

4.4 PhosphoSitePlus updates

Presenting author: Jon Korhhauser

4.5 Recent developments in libSBML and JSBML

Presenting author: Frank Bergmann

5 Model and Pathway Visualization

5.1 NAViGaTOR

Presenting author: Igor Jurisica, University of Toronto

5.2 SBGN update

Presenting author: Nicolas Le Novère

5.3 Tutorial: SBGN-ED - a tool for the Systems Biology Graphical Notation

Presenting author: Tobias Czauderna, IPK Gatersleben & MLU Halle-Wittenberg

SBGN (Systems Biology Graphical Notation) is an international collaboration to provide standard graphical / visual representations of biochemical and cellular processes in an unambiguous way. SBGN has three different notations: process description (PD), entity relationship (ER), and activity flow (AF). The standard defines a comprehensive set of symbols, with precise semantics, together with detailed syntactic rules defining their usage.

Here we present the SBGN-ED tool including novel developments. SBGN-ED allows to create all three types of SBGN maps from scratch or edit existing maps, to validate these maps for syntactical and semantical correctness, to translate networks from the KEGG database into SBGN, to layout and explore such maps, and to export SBGN maps into several file and image formats. SBGN-ED is freely available from http://www.sbgn-ed.org.

References:

[1] Le Novère et al.: The Systems Biology Graphical Notation. Nature Biotechnology, 27: 735-741, 2009.

[2] Czauderna et al.: Editing, validating, and translating of SBGN maps. Bioinformatics, 26 (18): 2340-2341, 2010.

[3] Junker et al.: Creating interactive, web-based and data-enriched maps using the Systems Biology Graphical Notation. Nature Protocols, 7: 579-593, 2012.

5.4 Biology graphical notation

Presenting author: Tobias Czauderna

5.5 Cytoscape.js - network and pathway visualization on the web

Presenting author: Max Franz, University of Toronto

6 Biological Pathways and Networks

6.1 Pathway Commons and BioPAX, validator update + Factoid/Crowd source curation, Human computation and Games

Presenting author: Gary Bader, University of Toronto

6.2 Paxtools

Presenting author: Emek Demir

6.3 Multifunctional proteins in multilevel networks

Presenting author: Raúl A. Ortiz-Merino, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México

To discern if a protein performs one or many functions within a given biological process may represent a major challenge. For instance, the Saccharomyces Genome Deletion Project have identified ≈20% of the yeast genome as essential or indispensable for this organism to grow in complex medium. However, neither the activity performed by the protein product of those genes or its relation to this phenotype are always clear. Therefore we developed a network-based procedure aimed to classify essential genes involved in S. cerevisiae primary metabolism. It implies the integration and analysis of public biological data using different centrality measures to characterize each node's (gene) role within the resulting network structure. Consequently, once a centrality-network pair shows its utility to categorize metabolism-related indispensable genes (MIGs) as such, both the centrality index and the interactions used to build the network may be used explain their role. This procedure was performed over multi-level networks built from single-level metabolite-mediated interactions (connecting enzyme-coding genes trough the metabolites they use or synthesize; MMIs) and other gene-gene interactions (GGIs). Our method allows the detection of MIGs with high probability and statistical confidence providing a framework to explain their functional impact within metabolism. This detects at least 3 sets of MIGs: 1)not predicted, including YLR355C/ILV5 involved in aminoacid synthesis and mtDNA maintenance whose indispensability cannot be explained by its MMIs nor its GGIs, 2)predicted as indispensable because of its MMIs even when other GGIs are included such as YKL060C/FBA1 that is associated both with glycolysis and V-ATPase assembly and its metabolic activity alone may be related to its indispensability and 3)predicted only when considering multilevel networks, where MMIs are not enough to explain their phenotype. Genes from the third set may codify multifunctional proteins involved in metabolism where both an additional activity (as protein-protein interactions) and/or the catalytic one may explain why they are indispensable for yeast to grow. An important distinction between multifunctional proteins has to be based upon the presence of one or several protein domains. That is because while multi-domain proteins are formed from independently evolved domains, with single-domain proteins the addition of a new activity over a preexisting one carries an "adaptive conflict"; such conflict is based on the notion that most protein mutations aimed to improve the new activity would reduce the original one. In this case, about the half of all the multifunctional proteins we found show to have only one domain. Now we are looking forward to determine how those additional activities could override this adaptive conflict.

6.4 Cytoscape 3.0

Presenting author: Gary Bader, University of Toronto

6.5 Integration of external references for BioPAX pathways and models

Presenting author: Michel Dumontier, Carleton University

BioPAX is a key format for the representation of pathway, reaction and molecule data and provides the ability to associate external references as datatype specified database and identifier pairs. Here, we convert these into a standard set of Bio2RDF URIs in order to enable integration across these resources.

7 Software Tools and Databases

7.1 Reforming Reactome

Presenting author: Robin Haw, OICR

The Reactome Knowledgebase (http://www.reactome.org) of human biological pathways and processes is a curated and peer-reviewed knowledgebase available online as an open access resource that can be freely used and distributed by all members of the biological research community. Recent extensions of our data model accommodate the annotation of disease processes, allowing us to represent the altered biological behavior of mutant variants frequently found in cancer, and to describe the mode of action and specificity of anti-cancer therapeutics. Reactome pathways are available on our web site for browsing, downloading, and manipulation by in-house and third party online analysis tools.

7.2 Tutorial: Access Reactome database via application programming interfaces

Presenting author: Guanming Wu, OICR

Reactome is an open-source, free access, manually curated and peer-reviewed biological pathway knowledgebase. As one of most popular free access pathway databases, it has been used widely in many pathway based data analysis projects. In this tutorial, we are going to introduce the data model used behind the Reactome database, the mapping from our object-oriented data model to a relational database, data exchange via BioPAX and SBML, new methods we have developed recently to annotate, and visualize disease pathways. We will spend most of time to introduce several application programming interfaces (APIs) that can be used to access Reactome contents programmatically: Perl API has been used mainly by the Reactome web application; Java API has been used by the Reactome curator tool and several Java based software utilities; SOAP API was developed several years ago and is being used by third parties to integrate Reactome contents into their own databases; recently developed RESTful API will be used as backend services for future Reactome web based applications. We will present examples to show how to use these APIs.

8 Interoperability Between Standards

8.1 Semantic interoperability framework

Presenting author: Sarala Wimalaratne, EBI

I will present an infrastructure for supporting the semantic interoperability of biomedical resources based on the management (storing and inference-based querying) of their ontology-based annotations. The infrastructure consists of: (i) a repository to store and query ontology-based annotations; (ii) a knowledge base server with an inference engine to support the storage of and reasoning over ontologies used in the annotation of resources; (iii) a set of applications and services allowing interaction with the integrated repository and knowledge base.

8.2 Towards ML-agnostic modelling

Presenting author: Mike Cooling, University of Auckland

8.3 MIRIAM Registry and Identifiers.org

Presenting author: Nick Juty, EMBL-EBI

The MIRIAM Registry is a catalogue of data collections (corresponding to controlled vocabularies or databases), their URIs and corresponding physical URLs or resources. Identifiers.org is a resolving layer built upon the information stored in the Registry. Together, they provide an annotation and cross-referencing framework which provides perennial, unambiguous and resolvable identifiers. We will describe both components, and how they are used.

8.4 Reusing SB standards to exchange PK/PD models

Presenting author: Stuart Moodie, EBI

In this talk I will describe how we are developing an exchange language for PK/PD modelling software as part of the DDMoRe project. Rather than create a new standard we in tend to use existing standards as much as possible. I will present our analysis of how SBML and SED-ML could be extended to support the exchange of PK/PD models in the future.

8.5 The COMBINE archive

Presenting author: Nicolas Le Novère

8.6 Bio2RDF's namespace SPARQL endpoint

Presenting author: François Belleau, Centre de recherche du CHUQ

We know thousand of bioinformatics databases, many web site like PathGuide and Bioinformatics Links Directory publish a collection of them. Bio2RDF project need to build URI to those database, since the Banff Manifesto an official resource is still needed.

In this presentation we explain the process by which an RDF mashup of namespace database was produce using ETL software. We explain our way to publish it on the web using the Virtuoso triplestore. Finally we demonstrate how its data can be consumed via SPARQL, SOAP or with new tools like RelFinder. We conclude by answering two questions about namespace. Which namespace are the most popular to identify database ? How far is the BioPAX community is to adopt MIRIAM new namespace standard?

9 ICSB/COMBINE Tutorials

9.1 Modelling and simulation of quantitative biological models

Presenting authors: Frank Bergmann, Michael Blinov, Akira Funahashi, Martin Golebiewski, Noriko Hiroi, Stefan Hoops, Ion Moraru, Franco du Preez, Sven Sahle

In this tutorial participants will learn setting-up quantitative computer models of biological networks using experimental kinetic data and simulating them in different systems biology platforms. Hands-on sessions, lectures and software demonstrations will be included providing attendees with the necessary skills to enable them to access experimental kinetics data from available resources, assembling computer models with these data and finally simulating the models within different tools.

Agenda 1) Accessing biochemical reaction kinetics data via the SABIO-RK database 2) Setting up quantitative models in CellDesigner 3) Setting up quantitative models in Virtual Cell (VCell) 4) Setting up models using JWS online 5) Simulating models in COPASI, CellDesigner, VCell and JWS online 6) Parameter estimation, optimization and model fitting 7) Model visualization and analysis 8) Integrated management of data, models and processes via the SEEK platform

9.2 Introduction to the statistical inference of regulatory networks

Presenting author: Frank Emmert-Streib

10 Posters

10.1 The VANTED(v2) framework for systems biology applications and its SBGN-ED Add-on for editing, validating and translating SBGN maps

Presenting author: Tobias Czauderna, IPK Gatersleben & MLU Halle-Wittenberg

Systems biology approaches follow an iterative cycling between experimental (wet-lab) and computational (dry-lab) methods with the aim of generating a holistic understanding of biological systems. The reconstruction of different kinds of networks based on experimental datasets allows for the representation of the diverse nature of biological systems on a global scale.

Here we present VANTED (v2) (http://www.vanted.org), a framework for systems biology applications. The VANTED framework aims at the integration, analysis and visual exploration of experimental data in the context of biological networks. The main functions of VANTED can be divided into the six areas: 1) Visualization and Exploration, 2) Data Mapping and Integration, 3) Data and Network Analysis, 4) Network Simulation, 5) Data Handling, and 6) Knowledge Representation. These areas provide an advanced analysis pipeline for experimental datasets from the field of systems biology. The VANTED framework comprises the VANTED core as well as a framework for various extensions called add-ons. The add-ons are being developed for the needs of life scientists and extend the functionality of the VANTED core towards various tasks and topics in systems biology. One of these add-ons is SBGN-ED for editing, validating and translating SBGN maps. The Systems Biology Graphical Notation (SBGN) is a standard for the visual representation of biochemical and cellular processes and networks. Three different languages (PD - Process Description, ER - Entity Relationship, and AF - Activity Flow) cover several aspects of biological systems in different levels of detail. SBGN helps to communicate knowledge more efficient and accurate between different research communities. SBGN-ED allows creating all types of SBGN maps from scratch, to validate these maps for syntactical and semantical correctness, to translate maps from the KEGG database into SBGN, to explore SBGN maps, and to export them into several file and image formats.

10.2 MIRIAM Identifiers.org and the MIRIAM Registry: annotation and cross-referencing framework

Presenting author: Nick Juty, EMBL-EBI

We describe our work to provide the community with directly resolvable URIs through Identifiers.org., using information stored in the MIRIAM Registry. Identifiers.org offers an annotation and cross-referencing framework which provides perennial, unambiguous and resolvable identifiers, which fit well with Linked Data initiative and Semantic Web vision.

10.3 The Phage and Prophage database

Presenting author: Devon Radford, University of Toronto

Phage have a substantial influence on bacterial evolution, and contribute a broad range of important proteins to bacterial proteomes when they integrate as prophages. Research continues to identify bacterially encoded phage-like operons effecting prokaryotic physiology both from within prophage and in distinct non-prophage regions. Novel and biologically important functionalities are being elucidated for these regions throughout the prokaryotic super-kingdom. These include tail-like bacteriocins selectively killing other microbes; capsid-like encapsulin nanocompartments with a variety of applications in stress resistance and metabolism; integrase-like recombinases; type III, type IV and type VI secretion systems which incorporate phage tail-like proteins to transmit bacterial factors; and pathogenicity islands converting otherwise harmless bacteria into highly effective pathogens. These proteins represent many of the key differences between pathogenic and benign bacterial strains, novel sources of therapeutics, important regulators of bacterial ecology, and potentially key components of bacterial physiology. Thus a definitive understanding of the content and distribution of phage and phage-like proteins is profitable to addressing a variety of important questions. The Phage and Prophage Database addresses this deficit providing an important tool for advancing bacteriology and virology research. The PPD interface is highly intuitive and easy to use for a broad range of bacterial and viral research problems.

10.4 Development of new representation standards in PK/PD

Presenting author: Maciej J Swat, EMBL-EBI

Pharmacokinetics and pharmacodynamics are branches of science dealing with drug related simulation and modelling. The former explains the time course of concentration, the latter the drug action. This can be done using both phenomenological or mechanism-based models. Despite the complexity of human body compartment based methods have been successfully applied in drug discovery and design for decades and remain indispensable today. Their application is straightforward for frequently measured individual data. However, clinical data come often with few measurements per subject, inter-subject and -occasion variability and one has to resort to population approaches. All this activities in this vital field have not seen much of standardization efforts so far. There is number of tools available but tool specific codes have to be manually transcribed in order to use them in parallel. The recently launched DDMoRe (www.ddmore.eu) project is a pioneer IMI founded effort to feel this gap. EMBL-EBI has the lead in the work package dedicated to the design of system-to-system Modelling Markup Languages, MMLs. The goal is to use where possible existing standards, such as SBML, SED-ML or UncerML and extend them. MMLs will enable researchers in industry and academia to share their models, verify the results obtained with different tools and allow effective ontology based search. We also drive the development of an infrastructure for storing and sharing of widely used existing disease models or new ones being developed by the DDMoRe partners. Our BioModels Database, leading worldwide resource for Systems Biology models, provides an indispensable experience to achieve this task. The project has already led to research activities from which both the Pharmacometrics and System Biology community will benefit. E.g. the work on encoding of probability distributions is crucial for the former and missing in the latter field. We present the ongoing efforts, and highlight the challenges and issues.

10.5 SED-ML level 1 version and tool support

Presenting author: Dagmar Waltemath, University of Rostock

SED-ML is an XML-based format for encoding simulation setups, to ensure exchangeability and reproducibility of simulation experiments. It follows the requirements defined in the MIASE guidelines.

The current version is SED-ML Level 1 Version 1 and covers the description of time course simulation experiments.SED-ML is a community project that is under continuous development, currently exploring the extension of SED-ML to cover more simulation experiment types. The SED-ML community comprises simulation software developers and modelers from the different communities applying computational methods on biology, medicine, neuroscience etc.

On this poster we introduce the current SED-ML language, and we provide an overview of existing SED-ML software tools and libraries.

10.6 Virtual cell modeling framework: synergy of multiple modeling tools and data sources

Presenting author: Michael Blinov, University of Connecticut Health Center

The Virtual Cell (VCell) is a modeling and simulation framework designed for building of a wide range of mathematical models (e.g. compartmental and spatial), performing deterministic, stochastic, and rule-based simulations, as well as analysis of simulation results. VCell has more than 3,000 users who created more than 29,000 models. Providing such a large number of users with different modeling capabilities is a daunting task. However, many features that may be of value to VCell users have been developed by other software teams. Thus, we took an approach of providing (fully acknowledged) access to these capabilities from within VCell interface. For the last view years we offered rule-based modeling capabilities of BioNetGen (http://bionetgen.org) software integrated within VCell. Recently we have succeeded in several other integration efforts, offering users parameter estimation capabilities developed for the Copasi (http://copasi.org) simulation and modeling software, as well as spatial stochastic particle-based simulation capability that we developed based on and in collaboration with developers of SmolDyn (http://smoldyn.org), a computer program for cell-scale biochemical simulations. On the other hand, the VCell database has more than 500 public models and many more "semi-private" models shared among groups of collaborators. There is a strong need for tools that enable users to annotate models that will be shared. To this end, we provide users with the access to the Pathway Commons (http://pathwaycommons.org) collection of databases, allowing the user to search pathway databases, creating a new model based on the pathway data, or linking model elements to entities in public databases. This feature helps our users to make annotated models that are easier to understand by others and shared with the community. To complement this capability, we provide the import of fully-annotated models from BioModels.net (http://biomodels.net) database, preserving annotations and thus providing users with web-links to original resources. We are now working on an interface for SABIO-RK database of reaction kinetics (http://sabio.villa-bosch.de/), enabling import of reaction kinetics constants directly into VCell models.

10.7 Multi-omics pathway visualization: integration of interaction and flux data in PathVisio

Presenting author: Martina Kutmon, Maastricht University

PathVisio is an open source pathway visualization and analysis tool. It allows users to not only draw and edit biological pathways, but also to visualize and analyze their data.

Currently, it is possible to visualize a single data set on the data nodes in a pathway, apply different visualization options and perform pathway statistics. Here, we want to present the new visualization extension to also show data on the lines of the pathways, namely the reactions and interactions, to allow, for example, the analysis of fluxomics data. Furthermore, we are working on making multi-omics data visualization more intuitive by providing a framework to load multiple data sets simultaneously.

Our new approaches will allow researchers to get a system wide, graphical overview of biological data in contrast to tables with thousands of numbers.

10.8 The FieldML format for representing, storing, and exchanging fields

Presenting author: Poul Nielsen, Auckland Bioengineering Institute

FieldML is an open format for interchanging field descriptions and data. It is based on a minimal set of concepts, yet is able to support arbitrarily complicated domains and field functions. The format places an emphasis on extensibility, reusability, efficiency. An API and I/O library is available to facilitate integration into software. The API currently supports serialisation is in XML plus external bulk data formats, such as HDF5.

10.9 The Physiome Model Repository 2

Presenting author: Poul Nielsen, The University of Auckland

Physiome Model Repository 2 (PMR2) is a modular software system under active development that provides a solution for the management and presentation of quantitative biological models. PMR2 achieves this by extending the framework provided by the Zope web application framework while working alongside with the Content Management System (CMS) Plone. Storage of the raw working model is governed by the Distributed Version Control System (DVCS) Mercurial, and this system is integrated as an extension module within PMR2. Advantages of using a DVCS allows a group of collaborators to more easily combine their incremental changes to their model. It also allow both centralised and peer-to-peer modes, allowing both the creation of a centralised repository or private exchanges of work-in-progress between a small subset of collaborators. The CMS side of PMR2 contains extension modules that can process the raw models within the DVCS into more readily consumed sets of pages of description and documentation pages based on the metadata stored within the raw models. Results of the simulation of a model, generated code and curation results can be generated based on the model type and the extension modules that were used. The current set of extension modules provide basic visualisation functionality, but can be further extended or created anew to to provide mission specific presentation features.

10.10 Produce, publish and consume semantic data: the Bio2RDF way

Presenting author: Francois Belleau, Centre de recherche du CHUQ

The poster describe in detail the process used to create a new SPARQL endpoint.