Newcastle University

I(CO)$_2$S

EPSRC

NSF

combine the computational modeling in biology network

2016

Dates: September 19-24, 2016
Location: Newcastle upon Tyne, UK
Hosted by: Newcastle University

The "Computational Modeling in Biology" Network (COMBINE) is an initiative to coordinate the development of the various community standards and formats in systems biology and related fields. COMBINE is a workshop-style event with oral presentation, posters, and breakout sessions. The five meeting days will include talks about the COMBINE standards and associated or related standardization efforts, as well as presentations of tools using these standards. Oral presentations will be selected from the submitted abstracts. In addition to oral presentations, poster sessions will allow people to inform each other about their software and other projects in a setting that fosters interaction and in-depth discussion.
The 2016 COMBINE meeting will be held in Newcastle upon Tyne, UK from September 19 to 24, 2016: the main Workshop will run 19-23 September and an SBGN Workshop will be held on September 24.

Organisers

Local Chairs

◊ Paolo Zuliani, Newcastle University, UK

◊ Anil Wipat, Newcastle University, UK

COMBINE Coordinators

◊ Gary Bader, University of Toronto, Canada.

◊ Martin Golebiewski, Heidelberg Institute for Theoretical Studies (HITS), Germany.

◊ Michael Hucka, California Institute of Technology, USA.

◊ Nicolas Le Novère, Babraham Institute, UK.

◊ Chris J. Myers, University of Utah, USA

◊ David Nickerson, University of Auckland, New Zealand

◊ Falk Schreiber, Monash University, Australia

◊ Dagmar Waltemath, University of Rostock, Germany

Local Organisers

◊ Keith Flanagan, Newcastle University, UK

◊ Claire Smith, Newcastle University, UK

◊ Mandy Williams, Newcastle University, UK

Funding acknowledgements

Support for COMBINE 2016 is being provided by the following organisations:

◊ EPSRC (AUDACIOUS project EP/J004111/2 and NUFEB project EP/K039083/1)

◊ Interdisciplinary Computing and Complex BioSystems Group

◊ School of Computing Science, Newcastle University

◊ National Science Foundation

Invited Speakers

Alfonso Bueno-Orovio (Oxford)
Carole Goble (Manchester)
Martin Golebiewski (HITS gGmbH)
Dagmar Iber (ETH Zurich)
Sarah Keating (EMBL-EBI)
Andrew Millar (Edinburgh)
Chris J. Myers (Utah)
David Nickerson (Auckland)
Daryl P. Shanley (Newcastle)
Yujiang Wang (Newcastle)

Agenda

All meetings will be held at the Crowne Plaza, Newcastle.

Day 1 (Monday, September 19)

	Speaker	Title
8:15am	Paolo Zuliani & Anil Wipat, Newcastle	Welcome
8:30am - 10:00am	**Session 1** Chair: *Chris J. Myers*	**Modeling**
8:30am	Sarah Keating, EMBL-EBI	SBML in 2016
9:15am	David Nickerson, Auckland	CellML, SED-ML, and the Physiome Model Repository
10:00am	*Break*	
10:30am - noon	**Session 2** Chair: *Herbert Sauro*	**Modeling**
10:30am	Dagmar Iber, ETH Zurich	From Networks to Function - Computational Models of Organogenesis
11:15am	Andrew Millar, Edinburgh	What infrastructure to cross the genotype-to-phenotype gap: foundations or flying buttresses?
noon	*Lunch*	
1:30pm - 3:00pm	**Session 3** Chair: *Paolo Zuliani*	**Medical Applications**
1:30pm	Yujiang Wang, Newcastle	Control of epileptic seizures - a computational perspective
2:15pm	Alfonso Bueno-Orovio, Oxford	Delivering cardiac computational biology to industry and the clinic: Advances towards reshaping future health and care
3:00pm	*Break*	
3:30pm - 5:00pm	**Session 4** Chair: *David Nickerson*	**Models and Data Exchange**
3:30pm	Carole Goble, Manchester	FAIRDOM - FAIR asset management and sharing experiences in Systems and Synthetic Biology
4:15pm	Chris J. Myers, Utah	A Standard Enabled Workflow for Synthetic Biology
5:00pm	Wrapup	
5:30pm	*Social Event*	
7:00pm	Banquet (Crowne Plaza)	

Day 2 (Tuesday, September 20)

	Speaker	Title
8:30am - 10:00am	**Session 5** Chair: *Goksel Misirli*	**Software**
8:30am	David Nickerson, Auckland	OpenCMISS-Osmium: helping PMR support the VPH requirements for identifiable and discoverable computational models
8:45am	Herbert Sauro, Washington	pathwayDesigner: A SBML compliant network simulation and design tool
9:00am	Matthias König, Humboldt	cy3sbml: a Cytoscape app for SBML
9:15am	Curtis Madsen & Nicholas Roehner, Boston	Nona Research Foundation - Creating a SynBio Software Commons
9:30am	Discussion	
10:00am	*Break/Posters*	
10:30am - noon	**Session 6** Chair: *Curtis Madsen*	**Synthetic Biology**
10:30am	Goksel Misirli, Newcastle	Data Integration and Mining for Synthetic Biology Design
10:45am	Hiroyuki Kuwahara, KAUST	A new thermodynamics-based approach to aid the design of natural-product biosynthetic pathways
11:00am	Nicholas Roehner, Boston	How to Remember and Revisit Many Genetic Design Variants Automatically
11:15am	Bryan Bartley, Washington	Version and Variant Control for Synthetic Biology
11:30am	Discussion	
noon	*Lunch*	
1:30pm - 3:00pm	**Session 7**	**Breakouts**
	Leader: *Lucian Smith*, Track: SBML	SBML L3v2 (and libsbml/test suite support)
	Leader: *Anil Wipat*, Track: SBOL	SBOL: Host Context
	Leader: *David Nickerson*	SED-ML: L1V3 Status/Review
3:00pm	*Break/Posters*	
3:30pm - 5:00pm	**Session 8**	**Breakouts**
	Leader: *Leandro Watanabe*, Track: SBML	SBML Distributions
	Leader: *Herbert Sauro*, Track: SBOL	SBOL: Website, FAQ, and Tutorials
	Leader: *David Nickerson*	SED-ML: L1V3 Status/Review
5:00pm	*Wrapup*	

Day 3 (Wednesday, September 21)

	Speaker	Title
8:30am - 10:00am	**Session 9** Chair: *Anil Wipat*	**Invited Talks**
8:30am	Martin Golebiewski, HITS	COMBINE and ISO/TC 276 Biotechnology: From grassroots community standards to official ISO standards
9:15am	Daryl Shanley, Newcastle	Modelling ageing to enhance a healthy lifespan
10:00am	*Break/Posters*	
10:30am - noon	**Session 10** Chair: *Frank Bergmann*	**Web**
10:30am	Vincent Noel, Instituto Butantan	SigNetSim : A web-based framework for designing kinetic models of molecular signaling networks
10:45am	Andreas Draeger, Tuebingen	The ZBIT Systems Biology Software and Web Service Collection
11:00am	Martin Peters, Rostock	SED-ML support in JWS Online
11:15am	Martin Golebiewski, HITS	The NormSys registry for modeling standards in systems and synthetic biology
11:30am	Discussion	
noon	*Lunch*	
1:30pm - 3:00pm	**Session 11**	**Breakouts**
	Leader: *Sarah Keating*, Track: SBML	SBML Math packages
	Leader: *Robert Sydney Cox*, Track: SBOL	SBOL Visual / SBGN
3:00pm	*Break/Posters*	
3:30pm - 5:00pm	**Session 12**	**Breakouts**
	Leader: *James McLaughlin*, Track: SBOL	SBOL: Ontologies and URI Best Practices
	Leader: *Kieran Alden*	Standards for storing and exchanging agent-based models
5:00pm	*Wrapup*	

Day 4 (Thursday, September 22)

	Speaker	Title
8:30am - 10:00am	**Session 13** Chair: *Martin Golebiewski*	**Ermerging Standards**
8:30am	Fengkai Zhang, NIH	SBML Multi - From Specification to Application
8:45am	Samuel Friedman, USC	MultiCellDS: a community-developed standard for curating microenvironment-dependent multicellular phenotype data as digital cell lines, snapshots, and collections
9:00am	David Nickerson, Auckland	Exploiting Electronic Health Record Standard openEHR to Manage Experimental Data in Computational Physiology
9:15am	Kieran Alden, York	Approaches for Developing Computational Models of Immune System Formation and Function
9:30am	Discussion	
10:00am	*Break/Posters*	
10:30am - noon	**Session 14** Chair: *Andreas Draeger*	**Simulation**
10:30am	Kiri Choi, Washington	Tellurium: A Python Based Modeling and Reproducibility Platform for Systems Biology
10:45am	Leandro Watanabe, Utah	Modeling and Simulating Hybrid SBML Models
11:00am	Fabian Fröhlich, Helmholtz Zentrum München	AMICI: An ODE simulation framework for sensitivity analysis of large-scale models
11:15am	J Kyle Medley, Washington	High-performance Model Simulation with libRoadRunner
11:30am	Discussion	
noon	*Lunch*	
1:30pm - 3:00pm	**Session 15**	**Breakouts**
	Leader: *Bryan Bartley*, Track: SBOL	SBOL Sequence Layer
	Leader: *Dagmar Waltemath*, Track: SED-ML/SBML	SED-ML/SBML & qual
3:00pm	*Break/Posters*	
3:30pm - 5:00pm	**Session 16**	**Breakouts**
	Leader: *Tramy Nguyen*, Track: SBOL	SBOL Interaction Definitions
	Leader: *Fengkai Zhang*, Track: SBML	SBML Multi Package
5:00pm	*Wrapup*	

Day 5 (Friday, September 23)

	Speaker	Title
8:30am - 10:00am	**Session 17** Chair: *Dagmar Waltemath*	**Model Exchange**
8:30am	Maciej J Swat, EMBL-EBI	PharmML - exchange format for models used in quantitative system pharmacology and pharmacometrics
8:45am	Augustin Luna, BioPAX Consortium	Biological Pathway Exchange (BioPAX) Format and Pathway Commons Update
9:00am	Miguel de Alba, Federal Institute for Risk Assessment (BfR)	Introducing FSK-ML - an SBML derivate for description and exchange of executable script based models
9:15am	Tramy Nguyen, Utah	Interconversion Between BioPAX and SBML
9:30am	Discussion	
10:00am	*Break/Posters*	
10:30am - noon	**Session 18** Chair: *Nicholas Roehner*	**Visual Standards**
10:30am	Alexander Mazein, European Institute for Systems Biology and Medicine	Comprehensive representation of disease mechanisms on multiple layers of granularity in SBGN PD and AF
10:45am	Robert Sidney Cox, Caltech	The Visual Protein Design Language
11:00am	Jacob Beal, BBN	Progress Report on Development of SBOL Visual 2.0
11:15am	Vasundra Toure, Rostock	Visualising differences in SBML models using SBGN and BiVeS
11:30am	Discussion	
noon	*Lunch*	
1:30pm - 3:00pm	**Session 19**	**Breakouts**
	Leader: *Dagmar Waltemath*, Track: SBML/SBOL/SBGN	COMBINE Annotations (usage of SBO, Miriam, URIs, etc.)
	Leader: *David Nickerson*, Track: SED-ML	SED-ML: New Output Types, Plots, Roadmap
3:00pm	*Break/Posters*	
3:30pm - 5:00pm	**Session 20**	**Breakouts**
	Leader: *Dagmar Waltemath*	COMBINEd Annotations
5:00pm	*Wrapup*	

Day 6 (Saturday, September 24)

The SBGN Workshop/Hackathon will be held in the same venue (Crowne Plaza).

Invited Speakers

Sarah Keating, EMBL-EBI

SBML in 2016

SBML (the Systems Biology Markup Language) was created in 2000 and has since then become a lingua franca for encoding and exchanging computational models in biology. Since its introduction, SBML has gained support from a range of software packages for modeling, and numerous databases today make publicly available published models encoded in the format. These developments continue to promote and drive the evolution of SBML. This talk will briefly outline the format as it stands in 2016 (SBML Level 3) and give an overview of the capabilities that have been added to the format and those that are in the developmental stages. It will also provide an update on the software packages provided by the SBML Team to facilitate the infrastructure necessary to support adoption and use of SBML.

David Nickerson, University of Auckland

CellML, SED-ML, and the Physiome Model Repository

CellML is an XML-based protocol for storing and exchanging computer-based mathematical models in an unambiguous, modular, and reusable manner. In addition to introducing CellML, in this presentation I will provide some of physiological examples that have help drive the development and adoption of CellML. I will also present the proposed changes being introduced in the next version of the specification, CellML 2.0. The Simulation Experiment Description Markup Language (SED-ML) is an XML-based format for encoding simulation setups to ensure exchangeability and reproducibility of simulation experiments. Here I will present an overview of the current version of SED-ML (Level 1 Version 2) and discuss the additional features introduced in the draft Level 1 Version 3 specification, as well as some of other upcoming features often requested. I will also demonstrate some of the functionality and content being added to the Physiome Model Repository which makes extensive use of both CellML and SED-ML to provide a rich repository of reproducible model descriptions.

Dagmar Iber, ETH Zurich

From Networks to Function - Computational Models of Organogenesis

One of the major challenges in biology concerns the integration of data across length and time scales into a consistent framework: how do macroscopic properties and functionalities arise from the molecular regulatory networks and how do they evolve? Morphogenesis provides an excellent model system to study how simple molecular networks robustly control complex pattern forming processes on the macroscopic scale in spite of molecular noise, and how important functional variants can evolve from small genetic changes. Recent advances in 3D imaging technologies, computer algorithms, and computer power now allow us to develop and analyse increasingly realistic models of biological control. In my talk, I will show how data-based modelling can be used to define mechanisms for fundamental developmental processes such as the control of branching processes and the scaling of developmental pattern on differently sized embryonic domains.

Yujiang Wang, Newcastle University

Control of epileptic seizures - a computational perspective

Epilepsy is a crippling disease that affects over 50 million people world-wide. Currently available clinical treatments are only effective in about 70% of people, and far less are treated without detriment to their quality of life. My work aims to understand the mechanisms of the disease better, and to suggest novel treatment to control epileptic seizures. To achieve this I use neuroimaging data and signal processing methods to extract relevant information, which inform dynamic computational models of epileptic seizures. Such models can be used to investigate mechanisms of the disease, to find new treatment strategies, or to tailor existing treatment to patient-specific needs. In my talk I will give a few examples of this approach, and also an outlook of how my proposed approaches can beneficial in the treatment of a wider range of neurological disorders.

Andrew Millar, Edinburgh University

What infrastructure to cross the genotype-to-phenotype gap: foundations or flying buttresses?

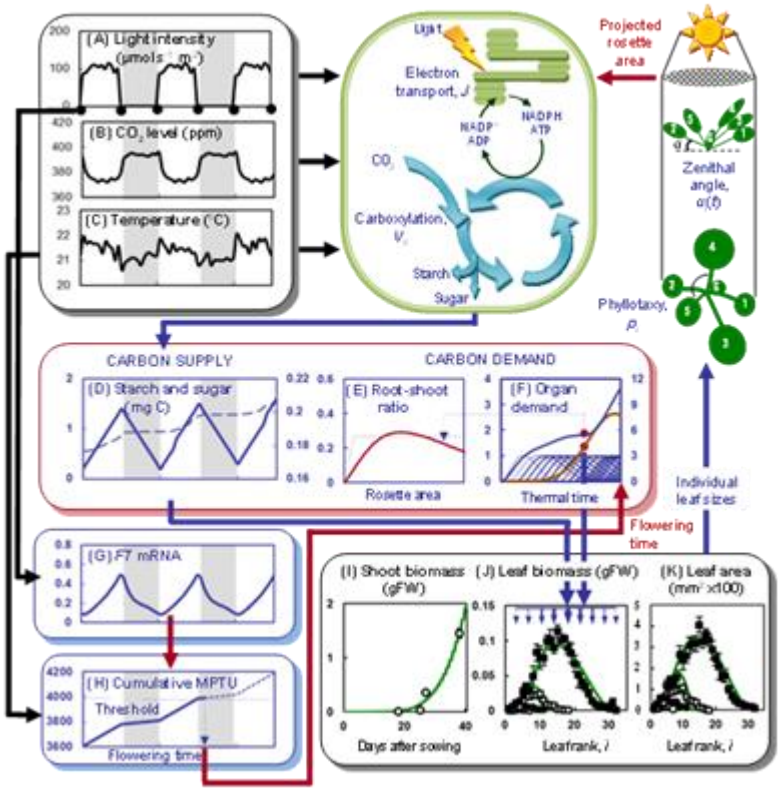
Andrew J. Millar (1), Ricardo Honorato-Zimmer (2), Ally Hume (1,3), Robert Muetzelfeldt (1,4), Gordon Plotkin (2), Uriel Urquiza (1), Argyris Zardilis (1), Tomasz Zielinski (1)

1. SynthSys and School of Biological Sciences, University of Edinburgh, Edinburgh, Scotland
2. SynthSys and School of Informatics, University of Edinburgh, Scotland
3. EPCC, University of Edinburgh
4. Simulistics Ltd., Edinburgh, Scotland.

My group and SynthSys' Systems Biology research includes infrastructure support for data (e.g. www.biodare.ed.ac.uk, refs. 2, 3), models (e.g. www.plasmo.ed.ac.uk) and other intellectual assets (*via* FAIRdom). I will contrast 'foundational' infrastructure, underlying projects at a low level, to support by 'flying buttresses' where infrastructure is added to mature projects. A key use case involves the 24-hour circadian clock, which controls biological processes from the sleep-wake cycle to the cell cycle. We seek to understand how the dynamics of the intracellular clock gene circuit control growth, biomass and life history in whole plants, using the lab model plant species, *Arabidopsis thaliana*. In future, this should help to understand the ecological and breeding pressures on the clock genes.

The rich data of the Arabidopsis community has allowed us and our partners in the BBSRC ROBuST and EU FP7 TiMet projects to model rhythmic processes between germination and flowering, including the crucial, nightly utilisation of starch carbon stores (4). We combined four existing models into the **Arabidopsis Framework Model** (FMv1), which predicts biomass quantitatively (1). We are now using FMv2 to understand quantitatively the pleiotropic phenotypes of a 'slow' clock mutant, bridging genotype to phenotype.

The FM's challenges for simulation tools include hourly data input during simulation, and dynamically creating compartments (leaves). Simulistics' Simile environment supports declarative modelling from a 'systems dynamics' background, with object-oriented features, a graphical interface and model composition. Simile provides full control of the FMv1. I will discuss a different, online approach to support the developing FMv2, and the early development of Chromar, a formal modelling language to express models of this type.



Links to Data/Model/Software

www.simulistics.com www.plasmo.ed.ac.uk/plasmo/models/model.shtml?accession=PLM_76

Selected references, with links

1. Chew YH, Wenden B, Flis A, Mengin V, Taylor J, et al. 2014. Multiscale digital Arabidopsis predicts individual organ and whole-organism growth. Proc Natl Acad Sci U S A 111:E4127-36. http://dx.doi.org/10.1073/pnas.1410238111

2. Flis A, Fernandez AP, Zielinski T, Mengin V, Sulpice R, et al. 2015. Defining the robust behaviour of the plant clock gene circuit with absolute RNA timeseries and open infrastructure. Open Biol 5. http://dx.doi.org/10.1098/rsob.150042

3. Practical evaluation of SEEK and OpenBIS for biological data management in SynthSys. Troup, E., Clark, I., Swain, P., Millar, A.J. & Zielinski, T. SynthSys Technical Report. http://dx.doi.org/1842/12236

4. Seaton DD, Smith RW, Song YH, MacGregor DR, Stewart K, et al. 2015. Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. Mol Syst Biol 11:776. http://dx.doi.org/10.15252/msb.20145766

Alfonso Bueno-Orovio, Oxford University

Delivering cardiac computational biology to industry and the clinic: Advances towards reshaping future health and care.

Both biomedical research and clinical practice rely on complex datasets for the characterization of human health and tailoring and design of therapy. Given the complexity and variety of approaches and recordings, there is an increasing recognition of the need to embed computational methods in medicine and science for analysis, integration and prediction.

This talk will explore the role and maturity of Computational Biology in facilitating a shift towards predictive human-based methodologies in cardiovascular medicine, spurred by advances in computational and experimental techniques, as well as by the growing acknowledgement of the limitations of animal models. Representative examples will illustrate how computational approaches complement, expand, bridge, and integrate experimental and clinical data and methods, and as such they are called to become an integral part of human-based methodologies in pharmacology and medicine.

Carole Goble, Manchester University

FAIRDOM - FAIR asset management and sharing experiences in Systems and Synthetic Biology

Over the past 5 years we have seen a change in expectations for the management of all the outcomes of research – that is the "assets" of data, models, codes, SOPs and so forth. Don't stop reading. Data management isn't likely to win anyone a Nobel Prize. But publications should be supported and accompanied by data, methods, procedures, etc. to assure reproducibility of results. Funding agencies expect data (and increasingly software) management retention and access plans as part of the proposal process for projects to be funded. Journals are raising their expectations of the availability of data and codes for pre- and post- publication. The multi-component, multi-disciplinary nature of Systems Biology demands the interlinking and exchange of assets and the systematic recording of metadata for their interpretation.

The FAIR Guiding Principles for scientific data management and stewardship (http://www.nature.com/articles/sdata201618) has been an effective rallying-cry for EU and USA Research Infrastructures. FAIRDOM (Findable, Accessible, Interoperable, Reusable Data, Operations and Models) Initiative has 8 years of experience of asset sharing and data infrastructure ranging across European programmes (SysMO and EraSysAPP ERANets), national initiatives (de.NBI, German Virtual Liver Network, UK SynBio centres) and PI's labs. It aims to support Systems and Synthetic Biology researchers with data and model management, with an emphasis on standards smuggled in by stealth and sensitivity to asset sharing and credit anxiety.

This talk will use the FAIRDOM Initiative to discuss the FAIR management of data, SOPs, and models for Sys Bio, highlighting the challenges of and approaches to sharing, credit, citation and asset infrastructures in practice. I'll also highlight recent experiments in affecting sharing using behavioural interventions.

http://www.fair-dom.org http://www.fairdomhub.org http://www.seek4science.org

Chris J. Myers, University of Utah

A Standard Enabled Workflow for Synthetic Biology

A synthetic biology workflow is composed of data repositories that provide information about genetic parts, sequence-level design tools to compose these parts into circuits, and system-level design tools to construct and analyze models of complete designs. Data standards enable the ready exchange of information within such a workflow. They allow repositories and tools to be connected from a diversity of sources. This talk will describe one such workflow that utilizes the Synthetic Biology Open Language (SBOL) to describe genetic designs, the Systems Biology Markup Language (SBML) to model these designs, and SBOL Visual to visualize these designs. This workflow includes multiple repositories, such as the Newcastle SBOL Stack and the JBEI ICE Repository, and multiple software tools such as SBOLDesigner and iBioSim, among others. Recently, the ACS Synthetic Biology journal has recommended the use of SBOL in their publications. This talk will demonstrate how this workflow can be used to produce this type of design information.

Martin Golebiewski, HITS gGmbH, Heidelberg Germany

COMBINE and ISO/TC 276 Biotechnology: From grassroots community standards to official ISO standards

The whole data lifecycle in systems biology requires a high degree of standardization of formats for data, models and metadata: Acquisition and description of the data, processing and analysis, their efficient and secure exchange, data integration and incorporation into computational models, as well as the setup, handling and simulation of the models that all have to follow dedicated standards. To this end many grassroots community standards for exchange formats and metadata description have been defined by different scientific communities in the field. However, it often is confusing and cumbersome for the potential users (e.g. experimentalists or modelers) to find the appropriate standards for their tasks and apply them in their workflows.

In light of its rapid growth and market relevance, the International Organization for Standardization (ISO) has identified biotechnology as an area of major potential for standardization. In December 2013 the technical committee of ISO for biotechnology standards (ISO/TC 276) had its inauguration meeting. Where is the journey of ISO/TC 276 Biotechnology heading? How could existing scientific community standards be of use for the definition of official ISO standards? What are the benefits for the grassroots communities in doing so? In my talk I will outline options for classifying and integrating existing community standards into official norms and how that might help in defining a framework and guideline for community standards and their application in different use cases.

Daryl Shanley, Newcastle University

Modelling ageing to enhance a healthy lifespan

The ageing process is complex and multifactorial. The proximate cause of ageing is the accumulation of damage to component molecules, organelles, membranes and cells leading to loss of function in tissues and organs. How such damage arises, how multiple sources of damage affect each other, and how physiological function is affected are all highly complex1. Living systems are evolutionarily programmed for survival, even though such programming does not extend to living forever, and there is plasticity in the capacity to survive, depending on environment, lifestyle, etc. It has long been known that interventions such as caloric restriction and mutations in certain genes reveal plasticity in longevity in short-lived mammals (i.e. mice). To benefit human health and longevity we need to know how this plasticity can be targeted, but again we are faced with complexity. I will present an overview of our efforts to address this challenge using an integrated laboratory and computational modelling approach. The focus of our work has been incremental such as our work on disentangling the dynamics of the nutrient signalling network2-4, integrating mechanism of aging5, and in other work we have made ambitious attempts to model ageing in whole cells6. My presentation will focus mostly on deterministic ordinary differential equation modelling, but will include some work using rule-based approach, stochastic simulation and cover issues such as parameter identifiability and with spatial localisation.

References

1. Kirkwood, T. B. L. A systematic look at an old problem. Nature 451, 644-647 (2008).

2. Dalle Pezze, P. et al. A Dynamic Network Model of mTOR Signaling Reveals TSC-Independent mTORC2 Regulation. Science Signaling 5, ra25-ra25 (2012).

3. Sonntag, A. G., Dalle Pezze, P., Shanley, D. P. & Thedieck, K. A modelling-experimental approach reveals insulin receptor substrate (IRS)-dependent regulation of adenosine monosphosphate-dependent kinase (AMPK) by insulin. The FEBS journal 279, 3314-3328 (2012).

4. Dalle Pezze, P. et al. A systems study on amino acid stimulation reveals concurrent activation of AMPK and mTOR converging on ULK1 and autophagy. Nature Communications, in press (2016).

5. Dolan, D. W. P. et al. Integrated Stochastic Model of DNA Damage Repair by Non-homologous End Joining and p53/p21- Mediated Early Senescence Signalling. PLoS Comput Biol 11, e1004246 (2015).

6. Dalle Pezze, P. et al. Dynamic Modelling of Pathways to Cellular Senescence Reveals Strategies for Targeted Interventions. PLoS Comput Biol 10, e1003728 (2014).

Contributed Talks

Approaches for Developing Computational Models of Immune System Formation and Function

Kieran Alden

York University

The application of computational modelling is becoming prevalent in studies that aim to understand how complex interactions between diverse cell types and their environment influence our health. Work in the York Computational Immunology Lab focuses on both the development of computational models of immune system formation, such as the development of lymphoid organs, and function, particularly in Inflammatory Bowel Diseases and Leishmaniasis, in addition to the development of software tools and approaches that assist in the composition and analysis of these models. Although the majority of our models adopt the agent-based modelling paradigm, as the complexity of our models increases, we are beginning to adopt a hybrid approach, integrating other models (such as SBML models) within the agent-based structure. As such, and to better describe our in silico experimentation, we are beginning the process of making our tools (Spartan, ASPASIA, and others) compatible with appropriate COMBINE standards. In this talk an overview of the model design and analysis approaches we adopt will be provided, focusing on a short case study on how we have used ASPASIA, a tool-kit for evaluating the effects of biological interventions on SBML model behaviour, to gain insights into how Th17-cell plasticity is controlled, prior to integrating the SBML model into an individual focused one that permits an exploration of Inflammatory Bowel Diseases.

Version and Variant Control for Synthetic Biology

Bryan Bartley (1), Michal Galdzicki (2), John Gennari (3), Herbert Sauro (1)

1. University of Washington, Dept. of Bioengineering
2. Arzeda Corp., Seattle, WA
3. University of Washington, Biomedical & Health Informatics

Version control systems have become invaluable for managing code complexity in software development projects. Version control for synthetic biology could prove similarly valuable, as synthetic biology projects grow in complexity. We discuss how versioning data for a genetic circuit may be captured using the Synthetic Biology Open Language and present a use case based on construction of a 4-module circuit. Versioning requires description of provenance information as well as details about the biological host of a circuit. Additionally we introduce variant control, an entirely new type of management system that addresses the needs of synthetic biologists to manage variant designs. Variant control leverages semantic annotations captured in SBOL to cluster similar designs together in design space. Variant design could help synthetic biologists explore design space and discover design rules.

Progress Report on Development of SBOL Visual 2.0

Bryan Bartley *(University of Washington)*, Jacob Beal *(Raytheon BBN Technologies)*, Mike Bissell *(Amyris)*, Kevin Clancy *(Thermo Fisher Scientific)*, Robert Sidney Cox *(Kobe University)*, Raik Grunberg *(University of Montreal)*, Chris Myers *(University of Utah)*, Nicolas Le Novere *(Babraham Institute)*, Matthew Pocock *(Turing Ate My Hamster, Ltd.)*, Chris Voigt *(MIT)*, Zach Zundel *(University of Utah)*

People engineering biological organisms often need to communicate in diagrams, both about the structure of nucleic acid sequences and about functional relationships between components of such sequences with one another or with other biological substances. This talk provides an update on progress in developing SBOL Visual 2.0, which extends the prior structure-focused SBOL Visual 1.0 toward these goals, as well as drawing on SBGN in its approach to interactions.

Some typical practices and conventions have begun to emerge for such diagrams. SBOL Visual 2.0 aims to organize and systematize such conventions in order to produce a coherent language for expressing both structure and function of genetic designs. At the same time, the language's design aims to maximize the applicability of this language by allowing as much flexibility and freedom as possible in how such diagrams are organized, presented, and styled: in particular, it should be readily possible to create diagrams both by hand and with a variety of software programs. Finally, means will be provided for extending the language with new and custom diagram elements, and for adoption of useful new elements into the language.

Every element of a SBOL Visual 2.0 should correspond to an element of the SBOL 2 data model. As such, an SBOL Visual diagram will potentially be able to be parsed to create an SBOL data construct (though this may be impractical for many formats in practice). Not all SBOL data constructs will be able to be mapped into SBOL visual, however, and in fact attempting a complete mapping is generally undesirable, as effective communication will often involve selecting or emphasizing only certain aspects of a complete design, particularly when dealing with complex or multi-layered systems.

Tellurium: A Python Based Modeling and Reproducibility Platform for Systems Biology

Kiri Choi (1), J. Kyle Medley (1), Caroline Cannistra (1), Matthias König (2), Lucian Smith (1), Kaylene Stocking (1), and Herbert M. Sauro (1)

1. Department of Bioengineering, University of Washington
2. Institute for Theoretical Biology, Humboldt University of Berlin

Tellurium is a powerful Python-based integrated environment designed for model building, analysis, simulation and reproducibility in systems and synthetic biology. Tellurium is a modular, cross-platform, and open-source integrated development environment (IDE) composed of multiple libraries, plugins, and specialized modules and methods. Tellurium ensures exchangeability and reproducibility of computational models by supporting SBML (Systems Biology Markup Language), SED-ML (Simulation Experiment Description Markup Language), the COMBINE archive, and SBOL (Synthetic Biology Open Language). Tellurium is a self-contained modeling platform which comes with a fully configured Python distribution independent of other local Python installations on the target machine. Tellurium comes with various libraries, including but not limited to: libRoadRunner, Antimony, phraSED-ML, COBRApy, NumPy, SciPy, and etc. Additionally, we include several user-friendly plugins and advanced modules for a wide-variety of applications, ranging from visualization tools to complex algorithms for bifurcation analysis and multi-dimensional parameter scanning. By combining multiple libraries, plugins, and modules into a single platform, Tellurium provides a unified but extensible solution for biological modeling and simulation.

Introducing FSK-ML - an SBML derivate for description and exchange of executable script based models

Miguel de Alba, Guido Correia Carreira, Alexander Falenski, Annemarie Käsbohrer, and Matthias Filter.

Bundesinstitut für Risikobewertung

SBML is the leading standard for models in Systems Biology. Unlike there, other scientific areas face the challenge that the majority of models have been developed in specific scripting languages like R or Matlab. In order to promote the standardized description and exchange of models in the domain of microbial risk assessment (MRA) we developed the Food Safety Knowledge Markup Language (FSK-ML) by extension and adaption of SBML. FSK-ML now allows to export and exchange executable script-based MRA models with well-structured meta-data. Built on top of the Combine Archive concept, the FSKX archive serves as a container to provide meta-data on MRA models, default values, simulation scenarios, experimental data or data generated by running a model simulation. Models programmed in scripting languages, such as R, Python or Matlab, are stored together with linked resources (e.g. R libraries) allowing to execute the original model scripts in a proper simulation environment. Finally a reference implementation using JSBML has been developed and will be showcased. A number of example applications have been deployed in the open-source data analytics platform Konstanz Information Miner (KNIME) making use of it.

The ZBIT Systems Biology Software and Web Service Collection

Andreas Dräger

Center for Bioinformatics Tübingen, Applied Bioinformatics Group, University of Tübingen, Germany

Background: Networks and models in systems biology today have been scaled up to the sizes of full genomes. They comprise thousands of reactions, metabolites, regulatory events, and many further biochemical components. In order to build, analyze, and explore networks and models on this scale, highly specific software solutions are required.

Results: We introduce the systems biology software collection that has been developed and continuously maintained at the Center for Bioinformatics Tübingen (ZBIT) over more than one decade now. All tools have been created to solve research questions in ongoing large-scale systems biology projects with numerous national and international collaborators. These tools cover all aspects of the model life cycle. KEGGtranslator and BioPAX2SBML can be used to gather data for building draft networks based on the KEGG PATHWAY database and files in BioPAX format. SBMLsqueezer generates kinetic equations and derives the units for the parameters therein. SBMLsimulator unifies the Simulation Core Library and the heuristic optimization toolbox EvA2 for model simulation and calibration. ModelPolisher accesses the BiGG Models knowledge-base for model annotation. SBML2LaTeX documents models by creating human-readable reports.

Conclusion: All tools presented are freely available and can either be used as online programs at http://webservices.cs.uni-tuebingen.de using any common web browser, or for download as desktop programs from http://www.cogsys.cs.uni-tuebingen.de/software/. These tools have been used in several research projects, including community efforts, such as the path2models project. Their development has often been a driving force for further community efforts, such as the JSBML project.

MultiCellDS: a community-developed standard for curating microenvironment-dependent multicellular phenotype data as digital cell lines, snapshots, and collections

Samuel Friedman, David Agus, Paul Macklin

University of Southern California

Medicine and biology have become increasingly multi-institutional and collaborative. This accelerates scientific progress when research communities pool their resources where consortia can recruit more patients to clinical studies, create large datasets and assemble measurements from instruments that are prohibitive for individual institutions to build, and integrate more diverse sets of expertise.

Research communities gain further when adopting centralized data repositories and standardized data elements. Data and insight sharing are less efficient when communicated by direct member-to-member interactions with non-standard data, and scientific advancement often suffers. Shared repositories spread data archiving and maintenance costs, while data standards allow shared data analysis and visualization tools; together, they allow the broader community to bring fresh expertise and insights to the field.

The work has not yet reached its fullest potential. Lacking a standardized data representation and centralized repositories to systematically report these measurements, data becomes undiscoverable, difficult to adapt for computer simulations, and precludes further community-driven extensions of the original data. This reflects an unmet need in biology at any non-molecular scale: phenotypic measurements from multicellular experiments remain trapped in individual papers.

We present MultiCellDS (MultiCellular Data Standard): a community-developed standard that functionally describes cell phenotypes with contextual information from the microenvironment. The refined standard supports a library of over 200 digital cell lines: a hierarchical, extensible representation of a biological cell line's phenotype in one or more microenvironmental contexts; of digital snapshots: a hierarchal representation of multicellular data of various forms coming from clinical, experimental, and simulation domains; and of digital collections: a way to organize the wide variety of data from a multitude of sources. With customized levels of detail, MultiCellDS enables the pooling and curation of measurements and insights across many sources into a standardized, searchable repository.

AMICI: An ODE simulation framework for sensitivity analysis of large-scale models

Fabian Fröhlich (1,2) and Jan Hasenauer (1,2)

1. Institute of Computational Biology, Helmholtz Zentrum München, Ingolstädter Landstraße 1 85764 Neuherberg
2. Chair of Mathematical Modeling of Biological Systems, Center for Mathematics, Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany

Ordinary differential equations (ODE) are commonly used to model the dynamic behaviour of biochemical reaction networks. Forward sensitivities of ODE models can give additional insight into the model and are often used in parameter estimation and experimental design. For large scale networks with hundreds or thousands of parameters, forward sensitivities are typically computationally not tractable as the computation time scales linearly with respect to the number of parameters. For a set of benchmark problems we show that, in contrast to forward sensitivities, adjoint sensitivities scale almost constant with respect to the number of parameters and are tractable to compute for systems with thousands of parameters and state variables. The respective methods are implemented in the Advanced MATLAB Interface for CVODES and IDAS (AMICI). AMICI is a MATLAB toolbox, which implements both forward and adjoint sensitivities as well as second order sensitivities. AMICI can import SBML models and supports sensitivity computation for models with events and assignment rules. Optimized Native C code is generated for every model and efficient simulation routines are compiled as standalone .mex files and could be used by other toolboxes. These routines and enable model analysis such as parameter estimation or experimental design for pathway- to genome-scale models.

The NormSys registry for modeling standards in systems and synthetic biology

Martin Golebiewski

HITS gGmbH, Heidelberg, Germany

To survey standard formats for computational modeling in biology such as the COMBINE core standards and others we have developed a registry which not only lists the standards, but also compares their major features, their possible fields of biological application and use cases (including model examples), as well as their relationships, commonalities and differences. This NormSys registry for modeling standards (http://normsys.h-its.org) provides a common entry point for modelers and software developers who plan to apply the standards for their respective case of application, and serves them with detailed information and links to the standards, their specifications and APIs.

cy3sbml: a Cytoscape app for SBML

Matthias König

Humboldt-Universität zu Berlin

cy3sbml is an app for the work with SBML in Cytoscape 3 providing the following functionality: SBML import; support of SBML Core and the Layout, Qualitative Model, Groups, Hierarchical Model Composition, and Flux Balance Constraints packages; reaction-species and extended SBML graphs; access to MIRIAM and SBO-based annotations; use of SBML attributes for visual mappings, network filtering and visualization; SBML validation. cy3sbml includes importer for BioModels and BiGG models to load SBML from standard repositories.
In addition to porting all functionality to Cytoscape 3 many new features have been implemented since the initial Cytoscape 2 release:
- support of the additional SBML packages comp and groups as well as fbc version 2
- implementation of kinetic graphs in addition to species-reaction graphs, providing insights into the relationship between parameters, kinetic laws, reactions, etc.
- resolving additional annotation information from primary annotations using the ontology lookup service (OLS)
- Integration with additional apps and systems biology workflows via OSGI APIs and commands accessible via cyREST
- redesign of GUI using web technology (JavaFx WebView)
- integration with BiGG database
- simple installation via app store

Availability and implementation: Freely available for non-commercial purposes from the Cytoscape App Store at: http://apps.cytoscape.org/apps/cy3sbml
Supplementary information: Source code and additional information at https://github.com/matthiaskoenig/cy3sbml/

A new thermodynamics-based approach to aid the design of natural-product biosynthetic pathways

Hiroyuki Kuwahara, Xuefeng Cui, Meshari Alazmi, and Xin Gao

King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, 23955, Saudi Arabia

Recent advances in genome editing and metabolic engineering enabled the engineering of de novo biosynthetic pathways for high-value natural products such as those for biofuels, cosmetics, perfumes, and drugs. Because of a large search space for potential metabolic routes to produce a given natural product, the development of a computational method to rank biosynthetic pathways for a given chassis is essential to the engineering of productive heterogeneous biosynthesis systems. One important design criterion for biosynthetic pathway designs is often concerned with the conversion rate of a given starting material to produce a target product in a host organism. Since quantitative information required to construct kinetic models of such metabolic systems is seldom available, commonly applied approaches to assess this productivity criterion have been based on the net thermodynamic favorability of pathways using the standard Gibbs energy of reactions. While ranking schemes based on the net favorability may be seen as good heuristics based on this criterion at first glance, they may not correspond well with the true picture of the titer of the target product, especially when a given pathway has strong competing reactions in the chassis organism. This is because these approaches do not take competing endogenous metabolic reactions into account and can only quantify the ratio of the target concentration to the source concentration at equilibrium. In this talk, we introduce a novel thermodynamics-based approach that searches for potential biosynthetic pathways and ranks them by considering the endogenous metabolic infrastructure. The core of our new method is a new host-dependent reaction weighting scheme, which captures the competition for a metabolic precursor with endogenous reactions using a statistical mechanical model. Thus, our reaction weighting scheme quantifies, for each metabolic reaction, a fraction of a given precursor that is converted into next intermediate metabolites by considering all competing endogenous metabolic reactions. Our pathway score is based on the sum of all reaction weights in a given pathway. Thus, it can better characterize the effects of the endogenous metabolic reactions and capture the productivity of each pathway in the given host more appropriately. Our algorithm exhaustively searches for biosynthetic routes from the given starting metabolite to the given target metabolite and generates top-K biosynthetic paths, each of which has at most n reaction steps. Unlike Gibbs energy-based reaction weighting schemes, our scheme guarantees that each reaction weight has the same sign. This allows for an efficient way to search for top-K routes by detecting and avoiding many unproductive routes quickly.

Biological Pathway Exchange (BioPAX) Format and Pathway Commons Update

Augustin Luna

The BioPAX Consortium

BioPAX is a standard language that aims to enable integration, exchange, visualization and analysis of biological pathway data. Specifically, BioPAX supports data exchange between pathway data groups and thus reduces the complexity of interchange between data formats by providing an accepted standard format for pathway data. It is an open and collaborative effort by the community of researchers, software developers, and institutions. We will present recent work and future prospects related to BioPAX and its usage in the Pathway Commons effort, which aggregates BioPAX-formatted data, and advances in usages with the Systems Biology Graphical Notation (SBGN).

Nona Research Foundation - Creating a SynBio Software Commons

Andrea Desrosiers (1), Curtis Madsen (2), Nicholas Roehner (2), Douglas Densmore (2)

1. Nona Research Foundation
2. Department of Electrical & Computer Engineering, Boston University

New methods in DNA fabrication and the focus on a design/fabricate/characterize cycle have dramatically increased the amount of data in existing experimental laboratories. However, the software used to manage this information has not scaled appropriately. Additionally, more advanced software tools in automated design synthesis and validation, literature mining, robotic automation and analysis are beginning to emerge in the academic wetlab. Unfortunately, these software tools currently cannot interoperate. They are created in isolation without future integration mechanisms. Interoperability is required in order to create a seamless automated workflow. Therefore, software platform is needed upon which modular tools can be created, shared, and developed, and the relevant data can be managed and distributed.

The Nona Research Foundation (aka Nona www.nonasoftware.org) focuses on addressing these needs and making sure they are met to elevate the entire field. Nona has begun to set up an online hub, organize related national events and programs, and hold local meetings where software developers, researchers, students, and interested third parties come together to share needs and experiences. Through these activities, Nona will serve as a clearinghouse for knowledge and emerging research helping to accelerate the creation of powerful new open-source tools for synthetic biology and bio-design automation. The network aims to bring together academic and research groups adept at designing software and carrying out research-based activities with others who can ensure broad distribution to make these advances widely available and useful to participants around the world.

Comprehensive representation of disease mechanisms on multiple layers of granularity in SBGN PD and AF

Alexander Mazein (1), Marek Ostaszewski (2), Inna Kuperstein (3), Mansoor Saqi (1), Bertrand De Meulder (1), Feng He (4), Irina Balaur (1), Diane Lefaudeux (1), Johann Pellet (1), Piotr Gawron (2), Stephan Gebel (2), Andrew Parton (5), Steven Watterson (5), Nathanaël Lemonnier (6), Pierre Hainaut (6), Markus Ollert (4), Emmanuel Barillot (3), Andrei Zinovyev (3), Rudi Balling (2), Reinhard Schneider (2) and Charles Auffray (1)

1. European Institute for Systems Biology and Medicine, CNRS-ENS-UCBL, Université de Lyon, 50 Avenue Tony Garnier, 69007 Lyon, France
2. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Campus Belval, 7 Avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette, Luxembourg
3. [1] Institut Curie, Paris, France; [2] INSERM, U900, Paris, France; [3] Mines ParisTech, Fontainebleau, France
4. Department of Infection and Immunity, Luxembourg Institute of Health (LIH), House of BioHealth, 29 Rue Henri Koch L-4354 Esch-Sur-Alzette, Luxembourg
5. Northern Ireland Centre for Stratified Medicine, Ulster University, C-Tric, Altnagelvin Hospital Campus, Derry, Co Londonderry, Northern Ireland, BT47 6SB, UK
6. Institut Albert Bonniot - INSERM U823, University Grenoble-Alpes, Grenoble, France

Motivation

Large amount of high-throughput data become available in the effort to better understand diseases. Despite the availability of various network-based approaches, tools for disease-specific functional analysis are greatly underdeveloped and it becomes increasingly important to advance in systematic data interpretation and hypothesis generation. The direct approach to solve this problem is developing highly accurate comprehensive computerized representations of disease mechanisms on the level of cellular and molecular interactions (Fujita et al., 2013, PMID 23832570; Kuperstein et al., 2015, PMID 26192618; Mizuno et al., 2016, PMID 26849355).

Methods

Recent advances in systems biology made it possible to unambiguously represent biological processes in a consistent way (Le Novère, 2015, PMID 25645874), made this information human- and machine-readable so it can be efficiently explored by computational methods. Disease-specific representations are developed in CellDesigner (www.celldesigner.org) following the Systems Biology Graphical Notation standard (www.sbgn.org). The involvement of domain experts from different groups ensures that different points of view are considered and all the disease hallmarks are covered and adequately represented.

Results

We present a concept of the DISEASE MAPS as a community effort and as a collection of reference resources for making sense of omics data in studies focused on a particular disease. Essentially a disease map provides a consensus review on the known disease mechanisms in the format of interconnected metabolic, signalling and gene regulatory pathways,
and can be used as the basis for hypothesis generation. We describe our experience on developing disease maps for Parkinson's disease, asthma and cancer, and demonstrate how these resources can be used for data visualisation and interpretation. Because these maps are developed using a strict computational format, they can be used for developing dynamic predictive mathematical models. While being complementary to generic pathway enrichment tools (such as freely available g-Profiler and DAVID, and commercial Ingenuity Pathway Analysis and MetaCore), the disease maps focus on the integration of information into a single hierarchically-organised network, thus enabling analysis using the full power of advanced systems biology approaches in the area of systems medicine.

Conclusion

To progress with this approach we propose building on the best practices and lessons learned from previous projects and applying shared standards, tools and protocols for generating high-quality representations and enabling the exchange of reusable pathway modules (e.g. inflammation, central metabolism, etc.). We envision this strategy will facilitate powerful advances in systems medicine for understanding disease mechanisms, cross-disease comparison, finding disease comorbidities, suggesting drug repositioning, generating new hypotheses, and after careful validation, redefining disease ontologies based on their endotypes - confirmed molecular mechanisms.

Acknowledgements: Funded by IMI (U-BIOPRED n°115010, eTRIKS n°115446).

High-performance Model Simulation with libRoadRunner

J. Kyle Medley and Herbert M Sauro

University of Washington

Timecourse simulation of kinetic models is an extremely important tool in biology, and has found applications in pharmacology (PK/PD models), rational design of synthetic systems and whole-cell model simulation. Pushing the envelope of high-speed model simulation is key to exploring more complex and diverse biological modeling approaches. libRoadRunner is a high-performance simulator and analysis library for SBML-encoded models. To facilitate maximum performance, libRoadRunner makes use of the LLVM library to compile models to native machine code which can be executed directly on the target CPU. By using an open and transparent development process, making our source code freely available, and working extensively with our collaborators through week-long workshops, we have enabled widespread access to this powerful modeling tool.

Data Integration and Mining for Synthetic Biology Design

Goksel Misirli (1), Jennifer Hallinan (2), Matthew Pocock (3), Phillip Lord (1), James Alastair (1), McLaughlin (1), Herbert Sauro (4), and Anil Wipat (1)

1. Newcastle University
2. Macquarie University
3. Turing Ate My Hamster Ltd.
4. University of Washington

One aim of synthetic biologists is to create novel and predictable biological systems from simpler modular parts. This approach is currently hampered by a lack of well-defined and characterized parts and devices. However, there is a wealth of existing biological information, which can be used to identify and characterize biological parts, and their design constraints in the literature and numerous biological databases. However, this information is spread among these databases in many different formats. New computational approaches are required to make this information available in an integrated format that is more amenable to data mining. A tried and tested approach to this problem is to map disparate data sources into a single data set, with common syntax and semantics, to produce a data warehouse or knowledge base. Ontologies have been used extensively in the life sciences, providing this common syntax and semantics as a model for a given biological domain, in a fashion that is amenable to computational analysis and reasoning. Here, we present an ontology for applications in synthetic biology design, SyBiOnt, which facilitates the modeling of information about biological parts and their relationships. SyBiOnt was used to create the SyBiOntKB knowledge base, incorporating and building upon existing life sciences ontologies and standards. The reasoning capabilities of ontologies were then applied to automate the mining of biological parts from this knowledge base. We propose that this approach will be useful to speed up synthetic biology design and ultimately help facilitate the automation of the biological engineering life cycle.

Interconversion Between BioPAX and SBML

Tramy Nguyen

University of Utah

Standards are an important feature of the systems and synthetic biology fields because they enable exchange and reproducibility of biological models. Interoperability of tools is important because each tool is tailored for a different application. Since there are different tools that utilize different standards, enabling the interconversion among the standards would be a useful feature since more tools would be able to exchange models. The focus of this project entails the interconversion between BioPAX and SBML. This project aims to have a lossless conversion. A key feature of this project is to have consistent terms that are used between the conversions going from both directions. Additionally, this converter is targeting SBML package extensions to support both SBML qualitative and hierarchal models among other things.

Exploiting Electronic Health Record Standard openEHR to Manage Experimental Data in Computational Physiology

Koray Atalag (1), Geoff Williams (2), Gary R. Mirams (2), David Nickerson (1), Jonathan Cooper (2)

1. Auckland Bioengineering Institute, University of Auckland, New Zealand
2. Department of Computer Science, University of Oxford, United Kingdom

In computational physiology, standards to express mathematical and anatomically based geometric models rely heavily on the XML suite of standards and the Semantic Web. The Physiome Model Repository (PMR) provides an infrastructure to manage models and personal workspaces using distributed version control as well as providing ontology-based semantic annotation and advanced semantic searching mechanisms (including a SPARQL endpoint). We have previously described how to link clinical data to this semantic pipeline using the open access electronic health record (EHR) standard, openEHR, such that both models and related clinical data could be discovered and retrieved. This is an important step forward for enabling translational research and creating personalised decision support tools for the computational physiology community. However there is no agreed formalism to manage the structure and semantics of experimental data (e.g. from a wet lab), nor one which supports the semantic linkage of such data to model resources. Yet a set of experimental data is the basis for model development and validation. Linking models and data is therefore vital, but is currently done manually or in an ad hoc process. We have explored the utilisation of the openEHR standard to manage experimental data. openEHR provides a generic model-based approach to data modelling, and a very flexible means to express, persist and query structured data. The main premise of openEHR is to be able to manage heterogeneous data without the need to build custom data models—reusable and modular models of information can instead be represented using high level tools (known as an Archetype) and can be persisted in this form and queried using an openEHR compliant backend system easily. This can simplify data management tasks for the computational physiology community and also enable semantic interoperability.

OpenCMISS-Osmium: helping PMR support the VPH requirements for identifiable and discoverable computational models

Tommy Yu, Peter Hunter, David Nickerson

Auckland Bioengineering Institute, University of Auckland, New Zealand

As community demand for reproducible and reusable science has pushed the evolution of model and simulation encoding standards, so has it driven the desire for the Physiome Model Repository (PMR) to support collaborative, identifiable, and discoverable archiving of computational models. The OpenCMISS-Osmium project will move the software stack behind PMR to a more modern and forward looking platform as well as addressing user feedback regarding the currently deployed generation of PMR instances. Osmium is the open-source software platform being developed to expand the capabilities of the PMR and support the upcoming PHYSIOME portal.

OpenCMISS-Osmium originated as a refactoring of PMR2, the software stack behind the current generation of PMR, and the additional new features to meet the growing demands for collaborative, reproducible, and reusable modelling in the Physiome/VPH communities. Supporting the PHYSIOME portal curation and publication workflows will, furthermore, require new features be implemented in Osmium beyond the current capabilities of PMR. Enabling the integration of published models in the PHYSIOME portal, and other collaborating community repositories, will similarly require extensions to existing features in PMR and the addition of new features.

As of May 2016, there were 940 workspaces in PMR, of which 273 are private. In the last year (since 2015-06), 101 workspaces have been created by 20 active users. We currently average 1600 visitors per month of which approximately 33% are new visitors.

SigNetSim : A web-based framework for designing kinetic models of molecular signaling networks

Vincent Noël (1), Marcelo S. Reis (1), Matheus H.S. Dias (1), Lulu Wu (1,2), Amanda S. Guimarães (1,2) Daniel F. Reverbel (2), Junior Barrera (1,2), and Hugo A. Armelin (1,3)

1. LETA-CeTICS, Instituto Butantan, Brazil
2. Instituto de Matemática e Estatística, Universidade de São Paulo, Brazil
3. Instituto de Química, Universidade de São Paulo, Brazil

Molecular biology is experiencing a revolution, thanks to data acquisition becoming increasingly cheaper, and to the development of Systems Biology, an emergent research field that shows new ways to study such high-throughput data. However, biologists need a new generation of tools capable of handling large quantitative data sets, and also performing rigorous mathematical analysis of the kinetic models that may be derived from them.

SigNetSim is a versatile framework coded in Python, whose main component is a library designed for building, adjusting and analyzing quantitative biological models, compatible with the latest standard for biological models (SBML 3.1 core). Once loaded, the model is converted into a formal mathematical representation, using the computer algebra system Sympy, thus allowing usage of formal methods on the model. When numerical methods are necessary, SigNetSim can generate C code, which can be executed in parallel. While designed to be run on computational servers and used remotely, for instance, via Jupyter notebook, SigNetSim also comes with an intuitive Django web interface made for non-specialists in programming languages. Through this interface, a user can input experimental data into a simple database, design models and optimize them, and perform simulation of multiple conditions. SigNetSim was successfully used to model both the Ras GTPase switch and the MAPK cascade in the mouse Y1 adrenal tumor cell line, and the in vitro competition between telomeric sequences and NAD+ for binding to GAPDH protein of Trypanosoma cruzi.

Results showed that the SigNetSim framework is a very useful tool for designing kinetic models of molecular signaling networks. Through its user-friendly web interface, even researchers and students that are inexperienced in programming can build, adjust and simulate models for both scientific and didactic purposes. We are currently adding for hierarchical model composition into this framework, enabling us to easily build larger models.

SED-ML support in JWS Online

Martin Peters (1,2) Johann Eicher (2), David van Niekerk (2), Kathleen Green (2), Stuart Owen (3), Natalie Stanford (3), Carole Goble (3), Dagmar Waltemath (1) and Jacky Snoep (2,3)

1. Rostock University, Germany
2. Stellenbosch University, South Africa
3. Manchester University, UK

JWS Online is a public resource of systems biology models (Olivier and Snoep, 2004). It offers a rich set of functionalities to explore, simulate, and modify simulation models. JWS Online is also an integral part of the SEEK data management system (Wolstencroft et al., 2015) that is the backbone of the FAIRDOMHub (https://fairdomhub.org). We present recent progress in the support for standardised simulation descriptions and export of COMBINE archives (Bergmann et al., 2014) in the latest version of JWS Online. A SED-ML (Waltemath et al., 2011) database was built as part of JWS to store SED-ML scripts for the reproduction of model-based simulations. The scripts can be constructed from scratch, or generated automatically from a simulation in the JWS online simulator. Besides downloading the simulation scripts, JWS Online also offers export of complete simulation studies using the COMBINE archive format.
Such an archive can be stored in a private area or, after curation, made public. It may contain the models used in a study, the simulation script, the result data sets, validation data, figures, etc. We make use of SED-ML and COMBINE archives to reproduce figures in scientific manuscripts. Model references can be made to JWS Online or Biomodels, and data can be linked from SEEK on the FAIRDOM data and model management platform, or other public web-sites. Furthermore, MATH-ML scripting gives a lot of versatility to precisely reproduce model simulation results. The COMBINE archives can be accessed via URLs which makes it possible to reproduce manuscript figures via live simulations simply by linking a figure to a URL. We envision to implement a system, in close collaborations of JWS with scientific journals, that ensures a rapid population of the SED-ML database. Most exciting illustrations of the power of the archives are examples that link experimental data and models, and examples that simulate several different model scenarios and plotting the results simultaneously. A set of reproducible studies are available for exploration at: http://jjj.bio.vu.nl/models/experiments/

References

- Bergmann, F.T., Adams, R., Moodie, S., Cooper, J., Glont, M., Golebiewski, M., Hucka, M., Laibe, C., Miller, A.K., Nickerson, D.P., Olivier, B.G., Rodriguez, N., Sauro, H.M., Scharm, M., Soiland-Reyes, S., Waltemath, D., Yvon, F., Le Novère, N. (2014) COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. BMC Bioinformatics, 15:369.
- Olivier, B.G. and Snoep, J.L. (2004). Web-based kinetic modelling using JWS Online. Bioinformatics 20:2143-2144.
- Waltemath, D., Adams, R., Bergmann, F.T., Hucka, M. Kolpakov, F., Miller, A.K., Moraru, I.I., Nickerson, D., Sahle, S., Snoep, J.L. and Le NovÃ¨re, N. (2011) Reproducible computational biology experiments with SED-ML - The Simulation Experiment Description Markup Language. BMC Systems Biology 5:198.
- Wolstencroft, K., Owen, S., Krebs, O., Nguyen, Q., Stanford, N.J., Golebiewski, M., Weidemann, A., Bittkowski, M., An, L., Shockley, D., Snoep, J.L., Mueller, W., Goble, C. SEEK: a systems biology data and model management platform (2015) BMC systems biology, 9,33.

How to Remember and Revisit Many Genetic Design Variants Automatically

Nicholas Roehner and Douglas Densmore

Boston University

As genetic design scales from gene clusters to genomes, it becomes challenging to specify the space of possible designs and track how this space changes as new design rules are learned. Storing individual designs is not a practical solution to this problem, as even relatively small libraries of under 20 parts can be used to construct hundreds of thousands of design variants. While previous rule-based and grammar-based approaches to genetic design have been successful in specifying the composition constraints and patterns of DNA components making up design variants, they have done so without a firm theoretical basis for comparing these specifications. To meet this need, we have developed an approach to genetic version control and implemented it in a software tool called Knox. Knox efficiently stores genetic design spaces as graphs that implicitly encode rules for composing sets of DNA components. Knox also provides methods for merging, querying, and reverting changes to these graphs, thereby making it possible for machines to track different versions of genetic design spaces throughout the history of a project and identify common design motifs among them. As genetic design scales, we anticipate that genetic version control and tools like Knox will play a critical role in supporting applications ranging from statistical design to machine learning.

PathwayDesigner: A SBML compliant network simulation and design tool

Herbert M Sauro and J. Kyle Medley

University of Washington

PathwayDesigner is an SBML compliant cellular network design and simulation tool. PathwayDesigner supports SBML level 2 and 3 as well as a native text based format. PathwayDesigner allows users to draw cellular network diagrams using a simple interface. Users can add species, reactions (and associated rate laws) and compartments to a diagram. Users can also specify additional algebraic equations to supplement the model as well as discrete events and user defined functions (as described in the SBML spec document). Simulations are carried out using the libRoadRunner backend. Events can be attached to species and reactions with visual cues indicating their presence. PathwayDesigner includes a plugin API that allows extensions to be easily written. Included with the distribution are eight plugins that include extensions for interactive simulation, metabolic control analysis, antimony support, parameter scanning, arrow head editor, random network generator and a force directed network layout tool. SBML annotations are used to store all visual aspects related to a model. PathweyDesigner is also SBW compatible. Availability: PathwayDesigner is currently a Windows only tool but a prototype Mac OS version is available (pathwaydesigner.org).

The Visual Protein Design Language

Robert Sidney Cox III and Raik Gruenberg

Kobe University

We present a set of glyphs and graphical objects for representing the designs of engineered proteins, along with a set of drawing constraints. These objects allow the representation of globular protein domains, linkers and loop regions, as well as membrane spanning and anchor domains. Additional glyphs allow for the annotation of protein recognition sequences or tags including those on the N-terminus or C-terminus, enzymatic sites, points of noncovalent contact, covalent modifications by small molecules, cleavage sites, and protein stability elements. We provide a rich set of symbols which can be used with labels to describe many of the possible types of protein designs, while keeping the total number of glyphs and shapes small and easy to draw.

These glyphs are intended to be used in a prescriptive way, and have precise requirements on size scaling, alignment to the drawn line, and line width to distinguish protein design elements from other graphical features shown on the same diagram. We also prescribe a set of label characters for a limited number of variants of each glyph that the user can annotate. Variations are suggested for hand drawn versions of the glyphs and figures. While the drawing constraints are prescriptive, their use or absence is decided by the user or tool depending on the aims of communication. Thus a domain or region might be left unannotated, or several parts of a catalytic site could be annotated at relative locations along the protein line. Our primary aim is to allow direct communication of the most relevant features of a protein design, while providing a clear and rigorous way to draw protein designs which are not too ugly or cluttered.

PharmML - exchange format for models used in quantitative system pharmacology and pharmacometrics

Maciej J Swat

EMBL-EBI, Cambridge, UK

New format enabling the efficient exchange and integration of pharmacometric (PMX) and quantitative system pharmacology (QSP) models across software tools has been defined and implemented as key element of the DDMoRe interoperability platform [1,2]. PharmML has been designed as the storage and exchange medium for mathematical and statistical models [3]. PharmML is a declarative language which development has been based on requirements provided by the DDMoRe community, popPK/PD and QSP partners, and on specific use cases from the main target tools. DDMoRe developed number of translators allowing to run PharmML coded model in softwares and programming languages, such as NONMEM, Monolix, PFIM, BUGS, R, Matlab, SBML, Python etc. PharmML provides a structure for encoding continuous and discrete data models equipped with complex variability structure, covariate, structural and observation models. Definition of complex clinical trial designs and modeling steps is possible as well. As a comprehensive self-contained format, it allows to encode models in tool agnostic manner. To facilitate the encoding of statistical models a database of probability distributions, ProbOnto, has been developed [4]. Moreover, to store the typical results produced in a QSP/PMX workflow a complementary format, Standard Output (SO), has been created [5].

The new exchange formats facilitate
(i) smooth and error-free transmission of models between tools,
(ii) use of complex workflows via standardised model and output definitions,
(iii) reproducibility of research,
(iv) bug tracking,
(v) improved interaction with regulatory agencies regarding modeling and simulation,
(vi) development of new tools and methods and
(vii) reuse of existing model resources, e.g. BioModels database.

References:

[1] DDMoRe project, URL: www.ddmore.eu
[2] Swat et al. (2015). Pharmacometrics Markup Language (PharmML): Opening New Perspectives for Model Exchange in Drug Development. CPT Pharmacometrics Syst Pharmacol, 4(6):316-9. URL: pharmml.org
[3] Lavielle, M. (2014). Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools. Chapman & Hall/CRC Biostatistics Series.
[4] Swat MJ, Grenon P, Wimalaratne S. ProbOnto - ontology and knowledge base of probability distributions, Bioinformatics 2016; doi: 10.1093/bioinformatics/btw170. URL: www.probonto.org
[5] URL: ddmore.eu/projects/so-standard-output

Visualising differences in SBML models using SBGN and BiVeS

Tom Gebhardt, Martin Scharm, Vasundra Toure, Dagmar Waltemath, Olaf Wolkenhauer

University of Rostock

Simulation models change over time. Reasons may include typo corrections, parameter updates, specification updates or improved semantic annotations. The BiVeS tool is a Java library to detect differences between simulation models encoded in SBML or CellML. It returns a list of changes between two versions of a model. BiVeS provides both human-readable and computer-readable output formats. One output type is JSON, a data-interchange format that is based on JacaScript and easy to parse. SBGN is a visual standard to display biological networks. It defines sets of glyphs to use for the representation of biological entities (macromolecules, simple chemicals, complexed etc.) and relations between them (catalysis, inhibition, association, dissociation etc.).

In our talk, we will introduce the BiVeS tool and its output types. We will focus on the novel JSON format. Contrary to the previous formats, BiVeS now exports all information necessary to generate the difference map using SBGN-compliant glyphs. The differences are furthermore highlighted in the SBGN map, using a specific color encoding. The support of SBGN for the visualisation of difference further improves the readability and interpretation of changes in simulation models, and it aids users in understanding the evolution of a model.

Modeling and Simulating Hybrid SBML Models

Leandro Watanabe and Chris Myers

University of Utah

The leading standard for the representation of biological models for simulation is the Systems Biology Markup Language (SBML). Traditionally, one would simulate such models using a single formalism such as ordinary differential equations or stochastic methods. However, more complex models like the whole-cell model requires the coupling of different formalisms since different biological process have been shown to be better represented using different formalism. For example, metabolism is often modelled using flux-balance analysis and some reactions as stochastic processes. Recently efforts have been put forth to encode the whole-cell model in SBML. In this process, a scheme to represent hybrid models have been proposed. As part of our work, a simulation method has been implemented within the tool iBioSim to support such models.

SBML Multi - From Specification to Application

Fengkai Zhang and Martin Meier-Schellersheim

Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, NIH, USA

Domain-detailed reaction rule (or rule-based) modelling is an approach that describes interactions between pairs of molecular components, specifying how the interactions depend on particular states of the molecules and their location in specific compartments. To promote model sharing among the software tools supporting it the SBML Multi Package (Multistate, Multicomponent and Multicompartment Species Package for SBML Level 3) was developed as a standard specifically to satisfy the requirements and reflect the nature of rule-based models. In our presentation, we will briefly review rule-based modelling, provide an update on the development of the SBML Multi package and introduce the main functions of libsbml-multi library with examples. We will also illustrate how to export Simmune models to the SBML Multi format and, conversely, import SBML Multi models into Simmune for further analysis and simulation. This work is supported by the intramural program of the NIAID, NIH.

Posters

Presenter	Title
Nima Afshar, Auckland	Computational Modelling of Glucose uptake in enterocytes using CellML
Kieran Alden, York	Approaches for Understanding the Relationship Between a Computational Model and Captured Real-World Phenomena
Florian Auer, Goettingen	Bringing Pathway Knowledge to Systems Medicine Approaches
Petronela Buiga, Manchester	Application of systems biology regulation of HER2 signalling in breast cancer therapy
Kiri Choi, University of Washington	Using SBML to Model Cellular Automata
Harold Fellerman, Newcastle	A temperature-dependent computational model of the PCR reaction
Matthias König, Humboldt University	Personalized liver function tests: A Multiscale Computational Model Predicts Individual Human Liver Function From Single-Cell Metabolism
Goksel Misirli, Newcastle	Data Integration and Mining for Synthetic Biology Design
Goksel Misirli, Newcastle	Modular Composition of Synthetic Biology Designs using Rule-Based Models
Jeremy Revell, Newcastle	Cross Entropy Method For Stochastic Parameter Inference
James Scott-Brown, Oxford	Visual Comparison of SBML Models
Vasundra Toure, Rostock	A demonstration of a fully-featured COMBINE Archive
Vasundra Toure, Rostock	Designing informative SBGN maps using SBGN-ED
Brijesh S. Yadav, Tel Aviv	Multidimensional patterns of metabolic response in abiotic stress-induced growth of Arabidopsis thaliana