

An update of the Bachmann COMBINE archive

Oliver Hölsken^{1,4} , Felix Patricius Hans^{2,4} , Sami Habib^{3,4} , Abel Hodelin Hernandez^{3,4} 

¹Medical School - Charité – University Medicine Berlin, Berlin, Germany

²Medical Center - University of Freiburg, University Emergency Department, Freiburg, Germany

³University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

⁴Master - Biomedical Informatics and Data Science (BIDS), Hochschule Mannheim – University of Applied Sciences in collaboration with the Medical Informatics for Research and Care in University Medicine (MIRACUM Consortium), Germany

Abstract

Introduction: The increase of size and complexity of simulation studies in systems biology and systems medicine proposes new challenges to sharing reproducible results. The Computational Modeling in Biology Network (COMBINE) improves the coordination of standard formats for several features of simulation studies [11]. On the other side, GitHub has been used as an essential common platform for managing software projects and supporting collaborative development [4].

Methods:

Results:

Conclusions:

Keywords: Systems Biology, Computational Biology, Data Science, Medical Informatics, COMBINE, containers, data

Introduction

1. Update von Dagmars paper
2. Orientierung an equator network (STARE-HI)

Scientific background

The increase of size and complexity of simulation studies in systems biology and systems medicine proposes new challenges to sharing reproducible results. The Computational Modeling in Biology Network (COMBINE) improves the coordination of standard formats for several features of simulation studies, such as Systems Biology Markup Language (SBML), CellML, Systems Biology Graphical Notation (SBGN), Systems Biology Graphical Notation (SBGN), and Systems Biology Result Markup Language (SBRML). These standards aim to encode, simulate and visualize biological models [11].

On the other side, GitHub has been used as an essential common platform for managing software projects and supporting collaborative development. Now a day some educational projects have begun to adopt it for hosting and managing course content because it gets transparency features to create, reuse, and remix materials; and to monitor activity on assignments and projects [4].

Rationale

Objectives

Materials and methods

The purpose of this project was divided into four groups. The tasks of each of them were the management of a documentation platform, review of exiting materials, comparison of provided models, graphical representation, and supply of a model script.

To provide the model from Bachmann *et. al.*[1] as a Fully Featured COMBINE Archive, we have created a public repository using the open-source platform GitHub, with a CC0-1.0 License. We choose GitHub as a data management platform to supervise the course of the project because it provides an intuitive and easy customizable environment, along with some features for documentation, and agile project management [4]. This repository contains the proposal directory's structure from Scharm & Waltemath [11] with the directories documentation, model, experiment, and result.

To achieve this goal, we research the literature about the Bachmann model and the COMBINE Archive, along with modeling file formats, checking the reproducibility of the COMBINE structure and software tools that had been used. In addition, we established communication channels, developed a rough schedule, provided a template for documentation, and periodically reviewed intermediate deliverables.

In our research, we found six SBML-Bachmann models in two different repositories, JWS Online and BioModels (see Results). We choose the last model from 14th November 2019 to perform other tasks in our project, because it provides complementary files for the simulation.

We found and tested five software tools for the simulation of biological systems. JWS Online, Systems Biology tool for the construction, modification, and simulation of kinetic models and the storage of curated models [7]; on this repository was one of the found models. Webviewer Uni Rostock (CombineArchiveWeb), a tool to visualize and manage COMBINE files [11]; this is the goal application of our project. COPASI, a software application for the simulation and analysis of biochemical networks and their dynamics [9]. SED-ML, a suite of tools for creating, editing, simulating and validating SED-ML files [13]. Tellurium, a tool to model, simulate and analyze biochemical systems [5].

One of the objectives of our project was to provide a standardized graphical representation of the Bachmann model based on the SBGN. Consequently, we performed research but we do not found any SBGN of this model. Therefore, we decided to create an SBGN network *de novo* based on Le Novère [6] and Touré *et. al.* [12]. In this step, we select the SBGN language, and lastly, we created the Process Description (PD) map with the web tool Newt Editor (v3.0.3) [2]. To validate the SBGN-ML we imported it into several software and libraries, LibSBGN from Newt Editor, the open-source software Visualisation and Analysis of Networks containing Experimental Data (VANTED) [8], Krayon for SBGN [3], and SBGNViz [10]. Lastly, we cleaned up the map and colored the relevant features in the model to improve the developed map.

Results

In our research, we found six SBML-Bachmann models, one of them as support information of Bachmann *et. al.* [1], this was the first delivered model. The others come from two different repositories, JWS Online and BioModels.

We found three models on the repository of BioModels. The first, BIOMD0000000347_url.xml, was submitted on 22nd July 2011 and modified on 31st January 2012. Together with this model were other files in different formats. Most of them were generated by tools to simulate, visualize, validate and document the model, one of them is another SBML model, BIOMD0000000347_urn.xml. The third and newest, Bachmann2011.xml, was posted on 14th November 2019. This one has other complementary files for the simulation of this model. The models in JWS Online do not have any date of building or update, so we do not know when these were built. The first model in JWS online, bachmann.xml, is from *Mus musculus* and represents the STAT's pathway in a cell simulation *in silico*. The second model, bachmann2.xml, was obtained from the BioModels database (BioModels ID: BIOMD0000000347).

Conclusions

Authors' contribution: The authors contributed equally to this work.

Competing interests: The authors declare no competing interests.

Funding: None.

Acknowledgement:...

References

- [1] Julie Bachmann, Andreas Raue, Marcel Schilling, Martin E Böhm, Clemens Kreutz, Daniel Kaschek, Hauke Busch, Norbert Gretz, Wolf D Lehmann, Jens Timmer, and Ursula Klingmüller. "Division of labor by dual feedback regulators controls JAK2/STAT5 signaling over broad ligand range". In: *Molecular Systems Biology* 7.1 (2011), p. 516. DOI: <https://doi.org/10.1038/msb.2011.50>.
- [2] Hasan Balci, Metin Can Siper, Nasim Saleh, Ilkin Safarli, Ludovic Roy, Merve Kiliarslan, Rumeysa Ozaydin, Alexander Mazein, Charles Auffray, Özgün Babur, Emek Demir, and Ugur Dogrusoz. "Newt: a comprehensive web-based tool for viewing, constructing and analyzing biological maps". In: *Bioinformatics* 37.10 (Nov. 2020), pp. 1475–1477. DOI: [10.1093/bioinformatics/btaa850](https://doi.org/10.1093/bioinformatics/btaa850).

- [3] Frank T. Bergmann, Tobias Czauderna, Ugur Dogrusoz, Adrien Rougny, Andreas Dräger, Vasundra Touré, Alexander Mazein, Michael L. Blinov, and Augustin Luna. “Systems biology graphical notation markup language (SBGNML) version 0.3”. In: *Journal of Integrative Bioinformatics* 17.2-3 (2020). DOI: doi:10.1515/jib-2020-0016.
- [4] Joseph Feliciano, Margaret-Anne Storey, and Alexey Zagalsky. “Student Experiences Using GitHub in Software Engineering Courses: A Case Study”. In: *2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C)*. 2016, pp. 422–431. ISBN: 978-1-5090-2245-8.
- [5] J. Medley, Kiri Choi, Matthias König, Lucian Smith, Stanley Gu, Joseph Hellerstein, Stuart Sealfon, and Herbert Sauro. “Tellurium notebooks—An environment for reproducible dynamical modeling in systems biology”. In: *PLOS Computational Biology* 14 (2018). DOI: 10.1371/journal.pcbi.1006220.
- [6] Nicolas Le Novère, Michael Hucka, Huaiyu Mi, Stuart Moodie, Falk Schreiber, Anatoly Sorokin, Emek Demir, Katja Wegner, Mirit I. Aladjem, Sarala M. Wimalaratne, Frank T. Bergman, Ralph Gauges, Peter Ghazal, Hideya Kawaji, Lu Li, Yukiko Matsuoka, Alice Villéger, Sarah E. Boyd, Laurence Calzone, Melanie Courtot, Ugur Dogrusoz, Tom C. Freeman, Akira Funahashi, Samik Ghosh, Akiya Jouraku, Sohyoung Kim, Fedor Kolpakov, Augustin Luna, Sven Sahle, Esther Schmidt, Steven Watterson, Guanming Wu, Igor Goryanin, Douglas B. Kell, Chris Sander, Herbert Sauro, Jacky L. Snoep, Kurt Kohn, and Hiroaki Kitano. “The Systems Biology Graphical Notation”. In: *Nature Biotechnology* 27.8 (2009), pp. 735–741. DOI: 10.1038/nbt.1558.
- [7] Brett G. Olivier and Jacky L. Snoep. “Web-based kinetic modelling using JWS Online”. In: *Bioinformatics* 20.13 (Apr. 2004), pp. 2143–2144. DOI: 10.1093/bioinformatics/bth200.
- [8] Hendrik Rohn, Astrid Junker, Anja Hartmann, Eva Grafahrend-Belau, Hendrik Treutler, Matthias Klapperstück, Tobias Czauderna, Christian Klukas, and Falk Schreiber. “VANTED v2: a framework for systems biology applications”. In: *BMC Systems Biology* 6.1 (2012), p. 139. ISSN: 1752-0509. DOI: 10.1186/1752-0509-6-139.
- [9] Sven Sahle, Pedro Mendes, Stefan Hoops, and Ursula Kummer. “A new strategy for assessing sensitivities in biochemical models”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366.1880 (2008), pp. 3619–3631. DOI: 10.1098/rsta.2008.0108.
- [10] Mecit Sari, Istemi Bahceci, Ugur Dogrusoz, Selcuk Onur Sumer, Bülent Arman Aksoy, Özgün Babur, and Emek Demir. “SBGNViz: A Tool for Visualization and Complexity Management of SBGN Process Description Maps”. In: *PLOS ONE* 10.6 (2015), pp. 1–14. DOI: 10.1371/journal.pone.0128985.
- [11] Martin Scharm and Dagmar Waltemath. “A fully featured COMBINE archive of a simulation study on syncytial mitotic cycles in Drosophila embryos”. In: *Eurosurveillance* 5.2421 (2016). DOI: <https://doi.org/10.12688/f1000research.9379.1>.
- [12] Vasundra Touré, Nicolas Le Novère, Dagmar Waltemath, and Olaf Wolkenhauer. “Quick tips for creating effective and impactful biological pathways using the Systems Biology Graphical Notation”. In: *PLOS Computational Biology* 14.2 (2018). DOI: 10.1371/journal.pcbi.1005740.

- [13] Dagmar Waltemath, Richard Adams, Frank T. Bergmann, Michael Hucka, Fedor Kolpakov, Andrew K. Miller, Ion I. Moraru, David Nickerson, Sven Sahle, Jacky L. Snoep, and Nicolas Le Novère. “Reproducible computational biology experiments with SED-ML - The Simulation Experiment Description Markup Language”. In: *BMC Systems Biology* 5.198 (2011). DOI: 10.1186/1752-0509-5-198.