Specifications of Standards in Systems and Synthetic Biology: Status and Developments in 2019

Falk Schreiber[1,2*], Gary D. Bader[3], Padraig Gleeson[4], Martin Golebiewski[5], Michael Hucka[6], Sarah M. Keating[7], Matthias König[8], Chris Myers[9], David Nickerson[10], Björn Sommer[1] and Dagmar Waltemath[11]

[1]Dep. of Computer and Information Science, University of Konstanz, Germany

[2]Faculty of IT, Monash University, Clayton, Australia

[3]The Donnelly Centre, University of Toronto, Canada

[4]Dep. of Neuroscience, Physiology and Pharmacology, University College London, UK

[5]Heidelberg Institute for Theoretical Studies (HITS), Germany

[6]California Institute of Technology, USA

[7]Heidelberg University, Germany

[8]Humboldt-University Berlin, Germany

[9]University of Utah, USA

[10]Auckland Bioengineering Institute, University of Auckland, New Zealand

[11]University Medicine Greifswald, Germany

---

[*] To whom correspondence should be addressed. Email:falk.schreiber@uni-konstanz.de

Abstract

This special issue of the *Journal of Integrative Bioinformatics* presents an overview of COMBINE standards and their latest specifications. The standards cover representation formats for computational modeling in synthetic and systems biology and include BioPAX, CellML, NeuroML, SBML, SBGN, SBOL and SED-ML. The articles in this issue contain updated specifications of SBGN Process Description Level 1 Version 2, SBML Level 3 Core Version 2 Release 2, SBOL Version 2.3.0, and SBOL Visual Version 2.1.

1 Introduction

Standards play an important role in Systems and Synthetic Biology. COMBINE ('COmputational Modeling in BIology' NEtwork) [1, 2] is a formal entity that coordinates standards development in these fields of research, fosters and moderates discussions, designs and implements dissemination strategies, and organises two annual community meetings each year. HARMONY (Hackathons on Resources for Modeling in Biology) is a workshop and hackathon for the development of libraries, specifications and tool support. The COMBINE forum brings together experts from associated fields of research, discusses applications and further developments of COMBINE standards and hence offers a platform for communication between standards developers and users.

COMBINE describes itself as *"... a network formed by the communities developing standards and formats to share computational models. Working together, it is expected that the federated projects will develop a set of interoperable standards covering all the aspects of computational modelling. Building on the experience of mature projects, which already have stable specifications, software support, user-base and community* *governance, COMBINE helps foster or support fledging efforts aimed at filling gaps or new needs”* [3].

COMBINE standards and associated initiatives cover a range of topics, see Fig. 1. COMBINE is an open initiative and everybody is invited to join. The COMBINE web site `https://co.mbine.org/` and COMBINE-related publications [1, 2, 4] provide more information.

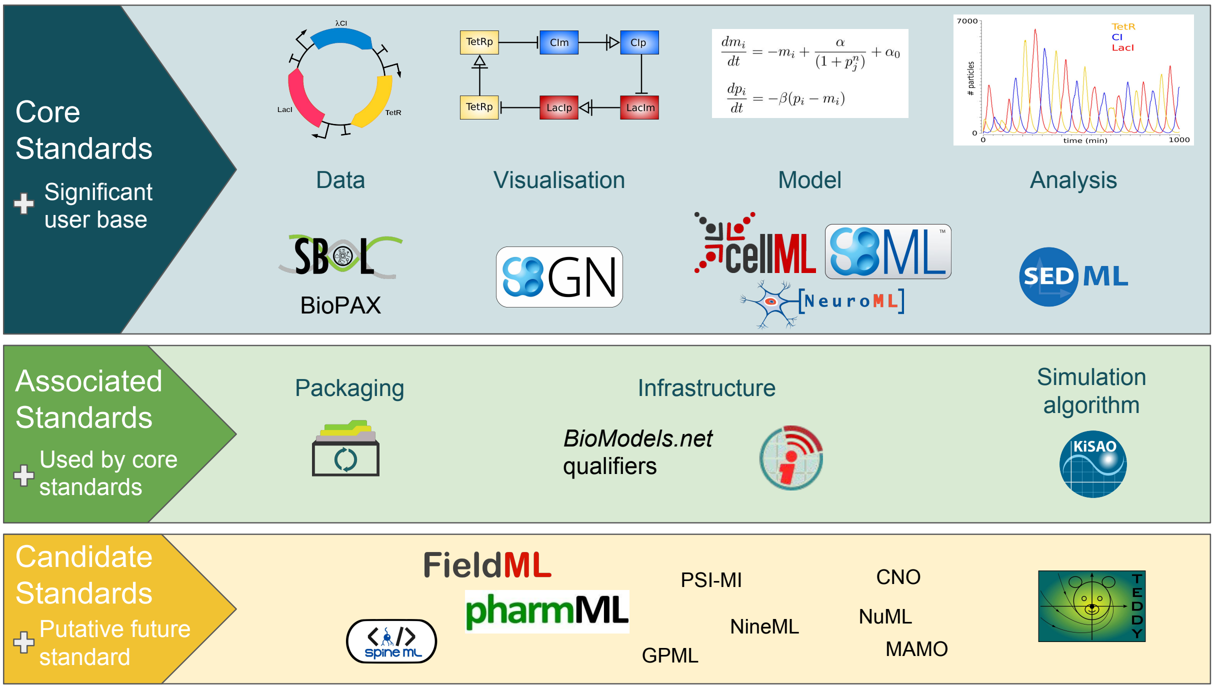

Figure 1: Overview of standards and associated efforts in Systems and Synthetic Biology provided by COMBINE (from [5]).

COMBINE published the first collection of Systems and Synthetic Biology standards as a special issue of the Journal of Integrative Bioinformatics in 2015 [6]. Since then a regular special issue of JIB provides updates to COMBINE standards. The latest update was published in 2018 [7]. This special issue presents developments of standards and related initiatives in 2018/19.

2 Latest versions of COMBINE standards

When using and implementing COMBINE standards, please refer to the following specifications. Note: New specifications are highlighted with **NEW**.

2.1 Core Standards

2.1.1 BioPAX (Biological PAthway eXchange)

BioPAX (Biological PAthway eXchange) [8], specification is BioPAX [9].

2.1.2 CellML

CellML [10], specifications are

- CellML 1.1 [11];
- CellML Metadata Framework 2.0 [12].

2.1.3 NeuroML

NeuroML (Neural Open Markup Language) [13, 14], specification is

- NeuroML version 2.0 [13].

2.1.4 SBGN (Systems Biology Graphical Notation)

SBGN (Systems Biology Graphical Notation) [15], specifications are

- **NEW** SBGN Process Description Level 1 Version 2 [16] - which includes new developments such as the addition of equivalence operator, subunit, and annotation glyphs, modification to the usage of submaps, and updates to clarify the use of various glyphs;
- SBGN Entity Relationship Level 1 Version 2.0 [17];
- SBGN Activity Flow Level 1 Version 1.2 [18].

2.1.5 SBML (Systems Biology Markup Language)

SBML (Systems Biology Markup Language) [19], specifications are

- **NEW** SBML Level 3 Core, Version 2 Release 2 [20] - which corrects some errors and clarifies some ambiguities discovered in Release 1. It defines validation rules that determine the validity of an SBML document and provides many examples of models in SBML form;
- SBML Level 3 Core, Version 2 [21];
- SBML Level 3 Package: Flux Balance Constraints, Version 2 [22];
- SBML Level 3 Package: Hierarchical Model Composition, Version 1 [23];
- SBML Level 3 Package: Multistate, Multicomponent and Multicompartment Species, Version 1 Release 1 [24];
- SBML Level 3 Package: Render, Version 1, Release 1 [25];
- SBML Level 3 Package: Qualitative Models, Version 1 [26];

- SBML Level 3 Package: Layout, Version 1 [27];
- SBML Level 3 Package: Groups, Version 1 [28].

2.1.6 SBOL (Synthetic Biology Open Language)

SBOL (Synthetic Biology Open Language) [29], specifications are

- **NEW** SBOL Version 2.3.0 [30] - which includes representing sequence modifications such as insertion, deletion, and replacement, supporting organisation and attachment of experimental data derived from designs and describing numerical parameters of design elements. It includes specifying types of synthetic biology activities, unambiguous locations for sequences with multiple encodings, refinement of a number of validation rules, improved figures and examples, and clarification on issues related to the use of external ontology terms;
- **NEW** SBOL Visual Version 2.1 [31] - which extends the diagram syntax to include methods for showing modular structure and mappings between elements of a system, interactions arrows that can split or join , and new glyphs for indicating genomic context and for stop codons.

2.1.7 SED-ML (Simulation Experiment Description Markup Language)

SED-ML (Simulation Experiment Description Markup Language) [32], specification is

- SED-ML Level 1 Version 3 [33].

2.2 Associated Standards

Associated standards provide an additional layer of semantics to COMBINE representation formats. These are:

- COMBINE Archive [34], specification is COMBINE Archive 1.0 [35];
- Identifiers.org URIs [36], community resources to provide persistent identification (no specification);
- Systems Biology Ontology and Kinetic Simulation Algorithm Ontology [37] for controlled vocabularies and semantics in systems biology (no specification);
- BioModels.net qualifiers [38] for representing relation between a model component and the resource used to annotate it (no specification).

References

[1] M. Hucka, D. P. Nickerson, G. D. Bader et al. Promoting coordinated development of community-based information standards for modeling in biology: the COMBINE initiative. *Frontiers in Bioengineering and Biotechnology*, 3:19, 2015.

[2] C. Myers, G. D. Bader, P. Gleeson, M. Golebiewski, M. Hucka, N. Le Novère, D. Nickerson, F. Schreiber and D. Waltemath. A brief history of COMBINE. In *Proceedings of the 2017 Winter Simulation Conference*, pages 884–895. 2017.

[3] Combine, 2017. http://co.mbine.org/.

[4] D. Waltemath, F. T. Bergmann, C. Chaouiya et al. Meeting report from the fourth meeting of the computational modeling in biology network (COMBINE). *Standards in Genomic Sciences*, 9(3):1285–1301, 2014.

[5] F. Schreiber, G. D. Bader, P. Gleeson, M. Golebiewski, M. Hucka, N. Le Novère, C. Myers, D. Nickerson, B. Sommer and D. Waltemath. Specifications of standards in systems and synthetic biology: Status and developments in 2016. *J. Integrative Bioinformatics*, 13(3):289, 2016.

[6] F. Schreiber, G. D. Bader, M. Golebiewski et al. Specifications of standards in systems and synthetic biology. *J. Integrative Bioinformatics*, 12(2):258, 2015.

[7] F. Schreiber, G. D. Bader, P. Gleeson et al. Specifications of standards in systems and synthetic biology: Status and developments in 2018. *J. Integrative Bioinformatics*, 15(1):0013, 2018.

[8] E. Demir, M. P. Cary, S. Paley et al. The BioPAX community standard for pathway data sharing. *Nature Biotechnology*, 28:935–942, 2010.

[9] BioPax, 2017. http://www.biopax.org/.

[10] A. A. Cuellar, C. M. Lloyd, P. F. Nielsen, D. Bullivant, D. Nickerson and P. Hunter. An overview of CellML 1.1, a biological model description language. *Simulation*, 79(12):740–747, 2003.

[11] A. A. Cuellar, W. Hedley, M. Nelson, C. M. Lloyd, M. D. B. Halstead, D. P. Bullivant, D. P. Nickerson, P. J. Hunter and P. M. F. Nielsen. The CellML 1.1 specification. *J. Integrative Bioinformatics*, 12(2):259, 2015.

[12] M. T. Cooling and P. J. Hunter. The CellML metadata framework 2.0 specification. *J. Integrative Bioinformatics*, 12(2):260, 2015.

[13] R. C. Cannon, P. Gleeson, S. Crook, G. Ganapathy, B. Marin, E. Piasini and R. A. Silver. LEMS: A language for expressing complex biological models in concise and hierarchical form and its use in underpinning NeuroML 2. *Frontiers in Neuroinformatics*, 8(79), 2014.

[14] P. Gleeson, S. Crook, R. C. Cannon et al. NeuroML: A language for describing data driven models of neurons and networks with a high degree of biological detail. *PLoS Computational Biology*, 6(6):e1000815, 2010.

[15] N. Le Novère, M. Hucka, H. Mi et al. The Systems Biology Graphical Notation. *Nature Biotechnology*, 27(8):735–741, 2009.

[16] A. Rougny, V. Toure, S. Moodie et al. Systems biology graphical notation: Process description language level 1 version 2. *J. Integrative Bioinformatics*, 16(3):20190022, 2019.

[17] A. A. Sorokin, N. Le Novère, A. Luna, T. Czauderna, E. Demir, R. Haw, H. Mi, S. L. Moodie, F. Schreiber and A. Villéger. Systems biology graphical notation: Entity relationship language level 1 version 2. *J. Integrative Bioinformatics*, 12(2):264, 2015.

[18] H. Mi, F. Schreiber, S. L. Moodie, T. Czauderna, E. Demir, R. Haw, A. Luna, N. Le Novère, A. A. Sorokin and A. Villéger. Systems biology graphical notation: Activity flow language level 1 version 1.2. *J. Integrative Bioinformatics*, 12(2):265, 2015.

[19] M. Hucka, A. Finney, H. M. Sauro et al. The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.

[20] M. Hucka, F. T. Bergmann, C. Chaouiya et al. The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 2 Core Release 2. *Journal of Integrative Bioinformatics*, 16(3):20190021, 2019.

[21] M. Hucka, F. T. Bergmann, A. Dräger et al. The systems biology markup language (SBML): Language specification for level 3 version 2 core. *Journal of Integrative Bioinformatics*, 15(1):20170081, 2018.

[22] B. G. Olivier and F. T. Bergmann. SBML level 3 package: Flux balance constraints, version 2. *J. Integrative Bioinformatics*, 15(1):20170082, 2018.

[23] L. P. Smith, M. Hucka, S. Hoops, A. Finney, M. Ginkel, C. J. Myers, I. I. Moraru and W. Liebermeister. SBML level 3 package: Hierarchical model composition, version 1 release 3. *J. Integrative Bioinformatics*, 12(2):268, 2015.

[24] F. Zhang and M. Meier-Schellersheim. SBML level 3 package: Multistate, multicomponent and multicompartment species, version 1, release 1. *Journal of Integrative Bioinformatics*, 15(1):20170077, 2018.

[25] F. T. Bergmann, S. M. Keating, R. Gauges, S. Sahle and K. Wengler. SBML level 3 package: Render, version 1, release 1. *Journal of Integrative Bioinformatics*, 15(1):20170078, 2018.

[26] C. Chaouiya, S. M. Keating, D. Bérenguier, A. Naldi, D. Thieffry, M. P. van Iersel, N. Le Novère and T. Helikar. The systems biology markup language (SBML) level

3 package: Qualitative models, version 1, release 1. *J. Integrative Bioinformatics*, 12(2):270, 2015.

[27] R. Gauges, U. Rost, S. Sahle, K. Wengler and F. T. Bergmann. The systems biology markup language (SBML) level 3 package: Layout, version 1 core. *J. Integrative Bioinformatics*, 12(2):267, 2015.

[28] M. Hucka and L. P. Smith. The Systems Biology Markup Language (SBML) Level 3 Package: Groups, Version 1 Release 1. *Journal of Integrative Bioinformatics*, 13(3):290, 2016.

[29] M. Galdzicki, K. P. Clancy, E. Oberortner et al. The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nature Biotechnology*, 32(6):545–550, 2014.

[30] C. Madsen, A. Goni-Moreno, U. P et al. Synthetic Biology Open Language (SBOL) Version 2.3. *Journal of Integrative Bioinformatics*, 16(3):20190025, 2019.

[31] C. Madsen, A. Goni-Moreno, Z. Palchick et al. Synthetic Biology Open Language Visual (SBOL Visual) Version 2.1. *Journal of Integrative Bioinformatics*, 16(3):20180101, 2019.

[32] D. Waltemath, R. Adams, F. T. Bergmann et al. Reproducible computational biology experiments with SED-ML – the Simulation Experiment Description Markup Language. *BMC Systems Biology*, 5(1):198, 2011.

[33] F. Bergmann, J. Cooper, M. König, I. Moraru, D. Nickerson, N. L. Novère, B. Olivier, S. Sahle, L. Smith and D. Waltemath. Simulation Experiment Description Markup

Language (SED-ML) Level 1 Version 3 (L1V3). *Journal of Integrative Bioinformatics*, 15(1):20170086, 2018.

[34] F. T. Bergmann, R. Adams, S. Moodie et al. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics*, 15(1):369, 2014.

[35] F. T. Bergmann, N. Rodriguez and N. Le Novère. COMBINE archive specification version 1. *J. Integrative Bioinformatics*, 12(2):261, 2015.

[36] N. Juty, N. Le Novère and C. Laibe. Identifiers. org and miriam registry: community resources to provide persistent identification. *Nucleic Acids Research*, 40(D1):D580–D586, 2012.

[37] M. Courtot, N. Juty, C. Knüpfer et al. Controlled vocabularies and semantics in systems biology. *Molecular Systems Biology*, 7(1):543, 2011.

[38] BioModels.net, 2017. http://co.mbine.org/standards/qualifiers.