

Creation and Analysis of Biochemical Constraint-based Models: the COBRA Toolbox v3.0

Laurent Heirendt¹ & Sylvain Arreckx¹, Thomas Pfau², Sebastián N. Mendoza^{3,18}, Anne Richelle⁴, Almut Heinken¹, Hulda S. Haraldsdóttir¹, Jacek Wachowiak¹, Sarah M. Keating⁵, Vanja Vlasov¹, Stefania Magnúsdóttir¹, Chiam Yu Ng⁶, German Preciat¹, Alise Žagare¹, Siu H.J. Chan⁶, Maïke K. Aurich¹, Catherine M. Clancy¹, Jennifer Modamio¹, John T. Sauls⁷, Alberto Noronha¹, Aarash Bordbar⁸, Benjamin Cousins⁹, Diana C. El Assal¹, Luis V. Valcarcel¹⁰, Iñigo Apaolaza¹⁰, Susan Ghaderi¹, Masoud Ahookhosh¹, Marouen Ben Guebila¹, Andrejs Kostromins¹¹, Nicolas Sompairac²², Hoai M. Le¹, Ding Ma¹², Yuekai Sun¹², Lin Wang⁶, James T. Yurkovich¹³, Miguel A.P. Oliveira¹, Phan T. Vuong¹, Lemmer P. El Assal¹, Inna Kuperstein²², Andrei Zinovyev²², H. Scott Hinton¹⁴, William A. Bryant¹⁵, Francisco J. Aragón Artacho¹⁶, Francisco J. Planes¹⁰, Egils Stalidzans¹¹, Alejandro Maass^{3,18}, Santosh Vempala⁹, Michael Hucka¹⁷, Michael A. Saunders¹², Costas D. Maranas⁶, Nathan E. Lewis^{4,19}, Thomas Sauter², Bernhard Ø. Palsson^{13,21}, **Ronan M.T. Fleming**^{1*} (ronan.mt.fleming@gmail.com)

¹Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 6 avenue du Swing, Belvaux, L-4367, Luxembourg; ²Life Sciences Research Unit, University of Luxembourg, 6 avenue du Swing, Belvaux, L-4367, Luxembourg; ³Center for Genome Regulation (Fondap 15090007), University of Chile, Blanco Encalada 2085, Santiago, Chile; ⁴Department of Pediatrics, University of California, San Diego, School of Medicine, La Jolla, CA 92093, USA; ⁵European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, CB10 1SD, United Kingdom; ⁶Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, USA; ⁷Department of Physics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA; Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA, USA; ⁸Sinopia Biosciences, San Diego, CA, USA; ⁹School of Computer Science, Algorithms and Randomness Center, Georgia Institute of Technology, Atlanta, GA, USA; ¹⁰Biomedical Engineering and Sciences Department, TECNUN, University of Navarra, Paseo de Manuel Lardizabal, 13, 20018, San Sebastian, Spain; ¹¹Institute of Microbiology and Biotechnology, University of Latvia, Jelgavas iela 1, Riga LV-1004, Latvia; ¹²Department of Management Science and Engineering, Stanford University, Stanford CA 94305-4026, USA; ¹³Bioengineering Department, University of California, San Diego, La Jolla, CA, USA; ¹⁴Utah State University Research Foundation, 1695 North Research Park Way, North Logan, Utah 84341, USA; ¹⁵Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London, United Kingdom; ¹⁶Department of Mathematics, University of Alicante, Spain; ¹⁷California Institute of Technology, Computing and Mathematical Sciences, MC 305-16, 1200 E. California Blvd., Pasadena, CA 91125, USA; ¹⁸Mathomics, Center for Mathematical Modeling, University of Chile, Beauchef 851, 7th Floor, Santiago, Chile; ¹⁹Novo Nordisk Foundation Center for Biosustainability at the University of California, San Diego, La Jolla, CA 92093, United States; ²⁰Latvian Biomedical Research and Study Centre, Ratsupites iela 1, Riga, LV1067, Latvia; ²¹Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitovet, Building 220, 2800 Kgs. Lyngby, Denmark; ²²Institut Curie, PSL Research University, Mines Paris Tech, Inserm, U900, F-75005, Paris, France.

<https://opencobra.github.io/cobratoolbox>

Project Goals: The aim of this project is to expand the functionality of the COBRA Toolbox to cover new modelling methods, provide a comprehensive and consistent documentation as well as a new suite of interactive tutorials.

The COBRA Toolbox [1] is a comprehensive software suite of interoperable COntstraint-Based Reconstruction and Analysis (COBRA) methods to model, analyse, and predict a variety of metabolic phenotypes using genome-scale biochemical networks. By design, its functions can be flexibly combined to implement tailored COBRA protocols for any biochemical network in biology, biomedicine, and biotechnology. Version 3.0 includes new methods for quality controlled reconstruction, modelling, topological analysis, strain and experimental design, network visualisation as well as network integration of cheminformatic, metabolomic, transcriptomic, proteomic, and thermochemical data. New multi-lingual code integration also enables an expansion in application scope via high-precision, high-performance, and nonlinear numerical optimisation solvers for multi-scale, multi-cellular and reaction kinetic modelling, respectively.

A dedicated effort has been made to ensure that all functions in the COBRA Toolbox 3.0 are consistently documented. A novel and comprehensive suite of more than 35 tutorials (opencobra.github.io/cobratoolbox/latest/tutorials/) has been developed to enable beginner, intermediate, and advanced users to practise a wide variety of COBRA

methods. The COBRA Toolbox 3.0 protocol and its tutorials may be adapted for the generation and analysis of a constraint-based model in a wide variety of molecular systems biology scenarios.

The code and the documentation are freely released as part of the openCOBRA project on github.com/opencobra/cobratoolbox.

References

- [1] Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, et al. (2018) Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3.0. *Nature Protocols* (accepted) .

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429. This project was funded by the National Centre of Excellence in Research and has received funding from the European Union's HORIZON 2020 research and innovation programme under grant agreement No 668738. In addition, this work was supported by the Fonds National de la Recherche grant O16/11402054, and the Luxembourg National Research Fund (FNR) ATTRACT program grant (FNR/A12/01). Nathan E. Lewis has been supported by NIGMS (R35 GM119850) and the Novo Nordisk Foundation (NNF10CC1016517). Anne Richelle has been supported by the Lilly Innovation Fellows Award.