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

Laurent Heirendt[1] & Sylvain Arreckx[1], Thomas Pfau[2], Sebastián N. Mendoza[3,18], Anne Richelle[4], Almut Heinken[1], Hulda S. Haraldsdóttir[1], Jacek Wachowiak[1], Sarah M. Keating[5], Vanja Vlasov[1], Stefania Magnúsdóttir[1], Chiam Yu Ng[6], German Preciat[1], Alise Žagare[1], Siu H.J. Chan[6], Maike K. Aurich[1], Catherine M. Clancy[1], Jennifer Modamio[1], John T. Sauls[7], Alberto Noronha[1], Aarash Bordbar[8], Benjamin Cousins[9], Diana C. El Assal[1], Luis V. Valcarcel[10], Iñigo Apaolaza[10], Susan Ghaderi[1], Masoud Ahookhosh[1], Marouen Ben Guebila[1], Andrejs Kostromins[11], Nicolas Sompairac[22], Hoai M. Le[1], Ding Ma[12], Yuekai Sun[12], Lin Wang[6], James T. Yurkovich[13], Miguel A.P. Oliveira[1], Phan T. Vuong[1], Lemmer P. El Assal[1], Inna Kuperstein[22], Andrei Zinovyev[22], H. Scott Hinton[14], William A. Bryant[15], Francisco J. Aragón Artacho[16], Francisco J. Planes[10], Egils Stalidzans[11], Alejandro Maass[3,18], Santosh Vempala[9], Michael Hucka[17], Michael A. Saunders[12], Costas D. Maranas[6], Nathan E. Lewis[4,19], Thomas Sauter[2], Bernhard Ø. Palsson[13,21], Ines Thiele[1], **Ronan M.T. Fleming**[1]* (ronan.mt.fleming@gmail.com)

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