Automatic generation of genome-scale metabolism and expression models for bacteria

Laura Kutt[1], Matthew DeJongh[2], Christopher S. Henry[3], Ines Thiele[1], and Ronan MT Fleming[1]
(*ronan.mt.fleming@gmail.com)

[1]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg; [2]Hope College, Holland, Michigan. [3]Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois.

http://stanford.edu/group/SOL/multiscale/

Project Goals: The aim of the project is to increase the availability of next-generation genome-scale models that integrate metabolism and macromolecular expression (ME models). As these models can predict the rate of macromolecular synthesis, e.g., protein synthesis, and have wider coverage of cellular processes compared to commonly used genome-scale metabolic models, they facilitate more comprehensive integration of omics data. The overall goal is to make available a large set of draft ME models and enable the community to simulate with ME models on a variety of platforms, including the KBase platform. This will facilitate the extended application of genome-scale ME models in systems biology research.

With growing amounts of omics data, computational models are crucial for derivation of biological insights. Most modelling paradigms use statistical methods to find data patterns. An alternative approach uses comprehensive models built upon prior biological data contextualised using omics data derived from a particular context. Metabolism and Expression models (ME models) integrate transcription and translation processes with metabolic networks and are the new-generation genome-scale models. These models include the metabolic burden associated with expression the proteins required for catalysis of biochemical processes, such as enzymes and translation factors. All the included proteins are produced by the model taking into account the nucleotide and amino acid sequences of the expressed genes. To date, the generation of ME models has been restricted to a few species, e.g., E. coli[1, 2]. Although there are potentially an extensive number of applications for ME models, the lack of readily available draft ME models and suitable solvers has limited their wider application within systems biology. Recently, suitable solvers have been developed[3, 4]. This work describes a pipeline for automatic generation of draft ME models for bacteria using the KBase platform. We demonstrate the utility of this pipeline for generation of draft ME models for 684 strains of bacteria. In addition to examining metabolic fluxes, these models can be utilised for more comprehensive mapping of omics data to examine proteome allocation, calculate core proteome requirements, and to simulate gene-expression profiles under different conditions. We envisage that dissemination of ME models via the KBase will facilitate widespread application of this new generation of genome-scale models to bioenergy research.

References

[1] Thiele I, Fleming RMT, Que R, Bordbar A, Diep D, et al. (2012) Multiscale Modeling of Metabolism and Macromolecular Synthesis in E. coli and Its Application to the Evolution of Codon Usage. PLoS ONE 7: e45635.

[2] Ebrahim A, Brunk E, Tan J, O'Brien EJ, Kim D, et al. (2016) Multi-omic data integration enables discovery of hidden biological regularities. Nature Communications 7: 13091.

[3] Ma D, Yang L, Fleming RMT, Thiele I, Palsson BO, et al. (2016) Reliable and efficient solution of genome-scale models of Metabolism and macromolecular Expression. arXiv:160600054 [q-bio] .

[4] Yang L, Ma D, Ebrahim A, Lloyd CJ, Saunders MA, et al. (2016) solveME: fast and reliable solution of nonlinear ME models. BMC Bioinformatics 17: 391.

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. Matthew DeJongh supported by a Fulbright Research Scholar Grant for Research in Computer Science at the University of Luxembourg.