SteadyCom: Predicting Microbial Abundances while Ensuring Community Stability

Siu H. J. Chan[1] (sxc554@psu.edu), Margaret Senftle[1,&], Lin Wang[1], Satyakam Dash[1], and **Costas D. Maranas**[1]

[1]Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802
[&]Current affiliation: Department of Chemical and Biomolecular Engineering, Rice University

http://www.maranasgroup.com/

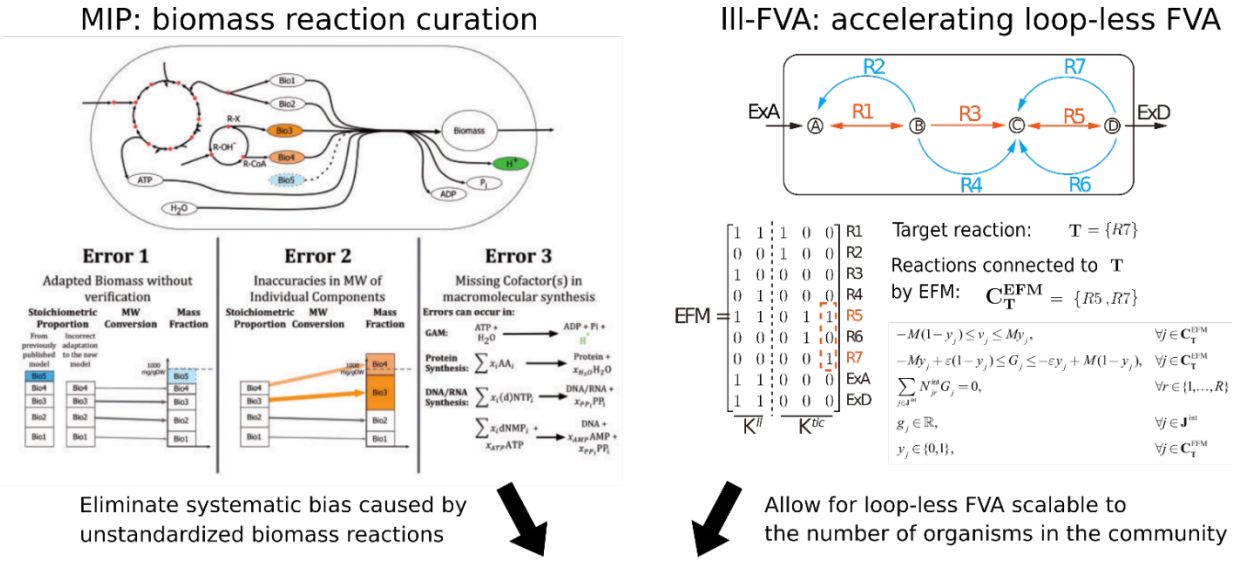
Project Goals: This project aims to apply and extend the genome-scale metabolic modeling approach to microbial communities. To simulate the metabolism of realistic microbial communities which can consist of up to hundreds species, scalable computational methods for simulation and analysis as well as a standardization procedure for unifying single-organism models are required. To this end, an optimization framework termed SteadyCom was developed to simulate the steady-state community metabolism of microbial communities. Together with a procedure introduced to standardize the biomass reactions in single-organism models and a scalable method devised for loop-less flux variability analysis, the results propose a workflow to analyze microbial communities at steady-state that is scalable to a large number of organisms.

Genome-scale metabolic modeling has become widespread for analyzing microbial metabolism. Extending this established paradigm to more complex microbial communities is emerging as a promising way to unravel the interactions and biochemical repertoire of these omnipresent systems. While several modeling techniques have been developed for microbial communities, little emphasis has been placed on the need to impose a time-averaged constant growth rate across all members for a community to ensure co-existence and stability. In the absence of this constraint, the faster growing organism will ultimately displace all other microbes in the community. We introduced the SteadyCom [1] optimization framework for predicting metabolic flux distributions consistent with the steady-state requirement. SteadyCom can be rapidly converged by iteratively solving linear programming (LP) problem and the number of iterations is independent of the number of organisms. A significant advantage of SteadyCom is compatibility with flux variability analysis (FVA). SteadyCom is first demonstrated for a community of four *E. coli* double auxotrophic mutants and is then applied to a gut microbiota model consisting of nine species, with representatives from the phyla Bacteroidetes, Firmicutes, Actinobacteria and Proteobacteria. In contrast to the direct use of flux balance analysis (FBA), SteadyCom is able to predict an abundance profile with a good agreement to experimental gut microbiota. SteadyCom provides an important step towards the cross-cutting task of predicting the composition of a microbial community in a given environment.

During simulating the community metabolism, an unstandardized biomass reaction of any of the organisms in the community that produces biomass with a molecular weight (MW) different from the defined standard 1 g $mmol^{-1}$ introduces a systematic error to the simulation. We developed the systematic procedure termed Minimum Inconsistency under Parsimony (MIP) [2] for checking the biomass weight and curating the biomass reaction so as to eliminate the systematic error in community simulations. We demonstrated that if the biomass reactions are not standardized, biomass MW discrepancies are accentuated in microbial community simulations as they can cause significant and systematic errors in the community composition. Microbes with underestimated biomass MWs are overpredicted in the community whereas microbes with overestimated biomass weights are underpredicted. The observed departures in community composition are disproportionately larger than the discrepancies in the biomass weight estimate. The procedure represents an important preprocessing step to ensure unbiased simulations of community metabolism.

To effectively perform FVA in the absence of thermodynamically infeasible cycles (TICs) in community models, we devised a method termed localized loop-less flux variability analysis (lll-FVA) [3] with significantly improved computational performance. We identified the fewest needed

constraints sufficient for optimality under the loop-less requirement in terms of elementary flux modes and put forth the concept of localized loop-less constraints to enforce this minimal required set of loop-less constraints. The computational time for loop-less FVA is reduced by a factor of 10-150 compared to the original loop-less constraints and by 4-20 times compared to the currently fastest method Fast-SNP with the percent improvement increasing with model size. Importantly, lll-FVA offers a scalable strategy for loopless flux calculations for multi-compartment/multi-organism models of very large sizes (e.g. $>10^4$ reactions) not feasible before. The Matlab implementations for SteadyCom, MIP and lll-FVA are available at https://github.com/maranasgroup.



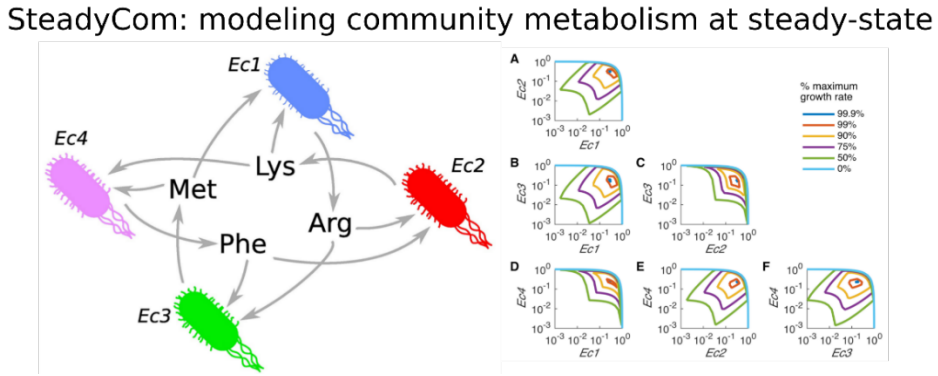


Figure 1. Modeling microbial communities using the SteadyCom framework. By integrating individual models standardized and verified through the MIP procedure [2], SteadyCom [1], which is compatible with flux variability analysis (FVA), can predict metabolic interactions in microbial communities at steady-state. The lll-FVA technique allows for loop-less flux analysis in microbial communities that is scalable to the increasing number of community members [3].

References

1. Chan S.H.J., Simons-Senftle M.N., and Maranas C.D. (2017) SteadyCom: Predicting microbial abundances while ensuring community stability. *PLoS Comput Biol* 13(5): e1005539.
2. Chan S.H.J., Cai J., Wang L., Simons-Senftle M.N., and Maranas C.D. (2017) Standardizing biomass reactions and ensuring complete mass balance in genome-scale metabolic models. *Bioinformatics*, 33:3603-3609.
3. Chan S.H.J., Wang L., Dash S., and Maranas C.D. (2017) Accelerating flux balance calculations in genome-scale metabolic models by localizing the application of loopless constraints. *arXiv preprint arXiv*:1711.04084

Supported by funding from the U.S. Department of Energy by grant DE-FG02-05ER25684 and *DE-SC0008091*.