Investigating carbon cycling microbial communities through probabilistic flux balance analysis of metabolic models: PrFBA

Andrew Freiburger[1,2] (afreiburger@anl.gov), Filipe Liu[1], Keith Tyo[2], and **Christopher Henry (chenry@anl.gov)**[1,]

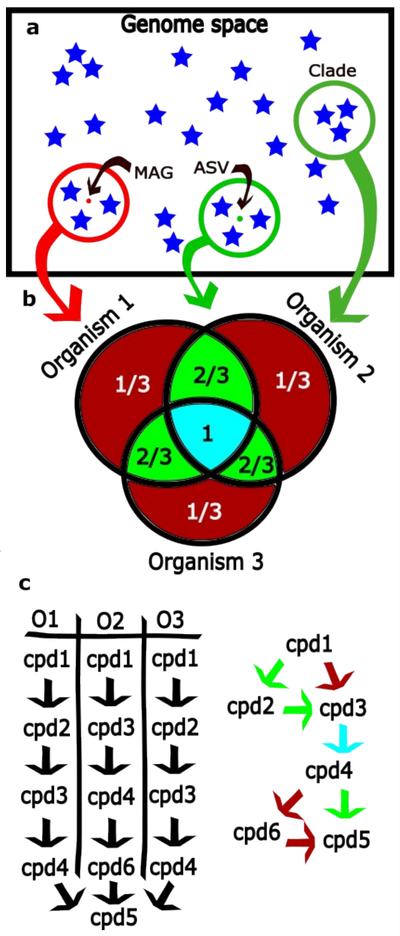
[1]Data Science and Learning Division, Argonne National Laboratory, Lemont, IL; [1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL

INTRODUCTION

Microbial communities are ubiquitous and serve fundamental roles in eukaryotic health, industrial processes, and are under-appreciated ecological agents of biogeochemical cycling and climate regulation. The Anaerobic Methanotrophic Archaea (ANME) and Sulfate-Reducing Bacteria (SRB) marine community, for example, abates gigatons of methane emission by consuming the effluent of deep-sea vents through unique methanotrophic pathways. The delegation of metabolic functions and the ecological principles that determine this delegation among ANME and SRB, however, remain unknown because ANME cannot be cultured and genomic sequenced in laboratory conditions. Shotgun metagenomic sequencing, followed by assembly and binning, is therefore used to genomically capture member genome, but the resultant metagenome assembled genomes (MAGs) are incomplete, often contaminated, have unavoidable genomic uncertainty within a genomic space, as determined by average nucleotide identity (ANI) approaches like GTDBtk.

Genome-scale metabolic models (GEMs) encapsulate metabolic functional knowledge of a genomic sequence, and can be simulated through our flux balance analysis (FBA) to determine pathway activity of the genome in an environment through linear optimization. GEMs can be expanded with 'omics data to further integrate experimental knowledge. We propose a method of integrating genomic uncertainty as functional probability into GEMs (probabilistic GEMs, prGEMS) and can be simulated through probabilistic FBA (prFBA) to prioritize the usage of the most conserved functions among the reference genomes within genomic space to which a MAG is mapped by ANI methods. We recently demonstrate this method with Amplicon Sequence Variants (ASVs, genome fragments like MAGs but are even more incomplete and uncertain) from a mammalian gut microbiome that were sampled over time after anti-biotic treatment and on either low-fat + high-fiber or high-fat + low-fiber diets. Our prGEM and prFBA method resolved the recovery mechanism of the microbiome after antibiotic application, which is under peer-review.

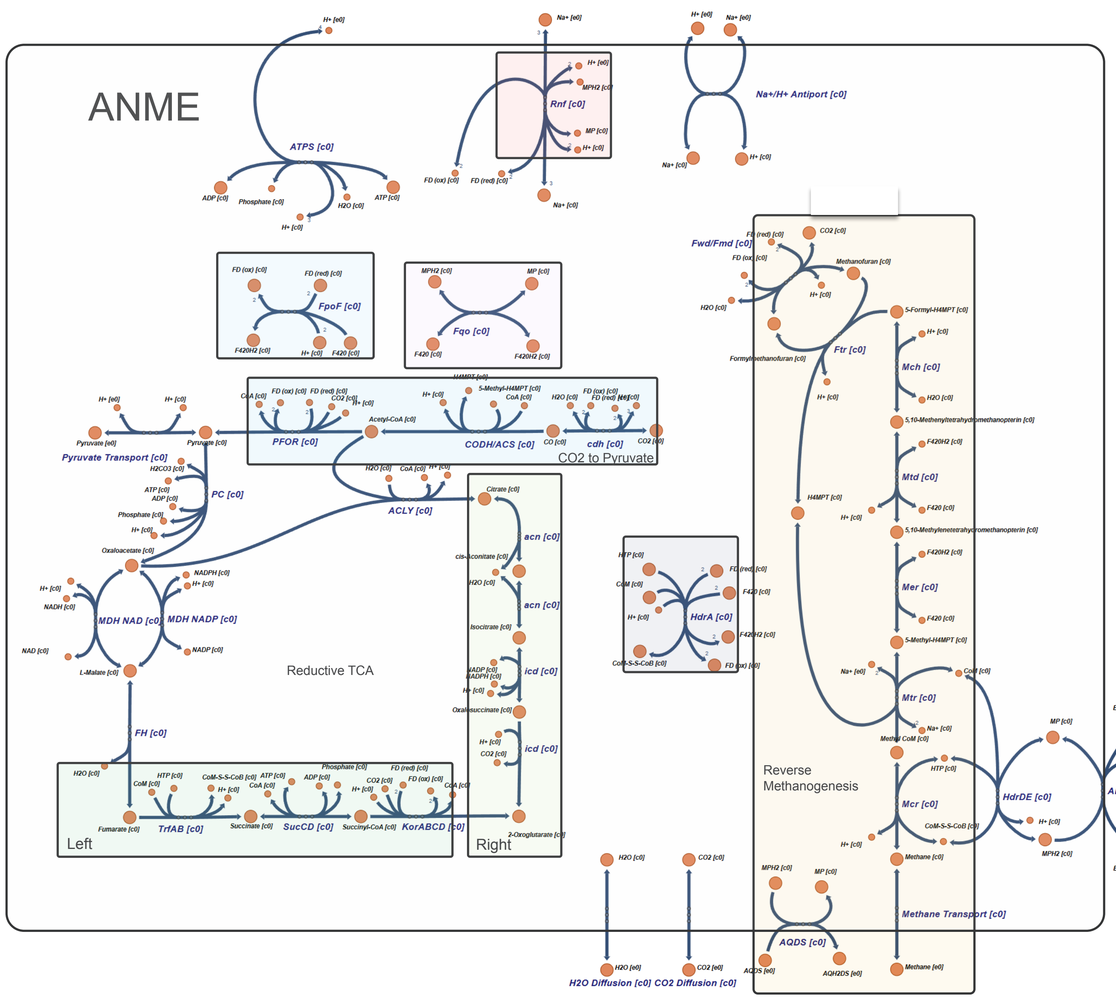
We plan to integrate our following improved biochemical templates for Archaea and methanotrophic pathways into ANME and SRB prGEMs to facilitate the most comprehensive simulations of how these communities metabolically function and importantly respond to environmental perturbations. These results will reveal resiliency of these communities to continue performing their biogeochemical ecological service, and also evince the value of prFBA for community modeling.



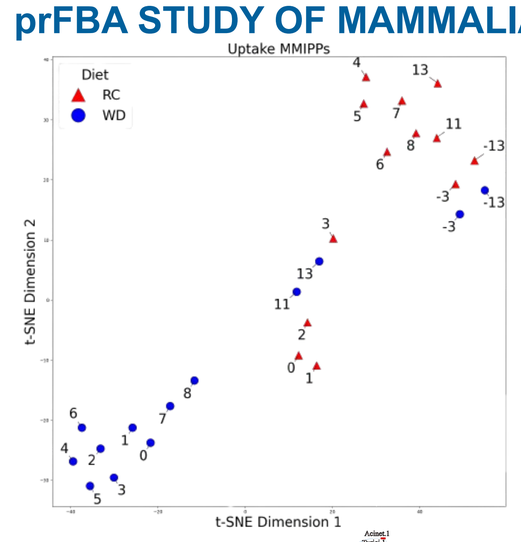
ANME & SRB METABOLIC MODELS

The essential metabolisms of the ANME and SRB community members, who perform crucial biogeochemical services of metabolizing gigatons of methane from deep-sea ocean vents, are represented in the following metabolic maps. Prominent pathways are highlighted and labeled for emphasis. The maps are the result of curated biochemistry from literature, annotations of >40 genomes, assembled pangenomes that minimize inconsistencies, and are coupled with new biochemical templates to capture more archaea-specific metabolism. The Electron Transport Chain (ETC) is believed to the primary syntrophic exchange between the ANME & SRB community members, instead of carbon exchange that was previously proposed in previous studies of *Desulfovibrio vulgaris Hildenborough* [Ferreira et al. **2023.** *Environmental Microbiology*, 25(5), 962–976].

Our latest ANME pathway contains 43 elemental flux modes (EFMs, an alternative to FBA for simulating GEMs) of which 17 are ATP generating modes. The reverse methanogenesis branch is essential for every calculated EFM including non-ATP modes thus guaranteeing that Methane is essential for energy. The AQDS is replaced by the DIET matrix and forces our models to form an obligate syntrophic relationship to cycle methanophenazine. The improved annotation accuracy of these models empowers community simulations towards evaluating the "reverse methanogenesis" hypothesis (the leading theory of how these communities metabolize methane by reversing the methane generation pathway). These simulations may ideally also explain the selective stability of these communities and clarify their vulnerabilities to anthropogenic influences.

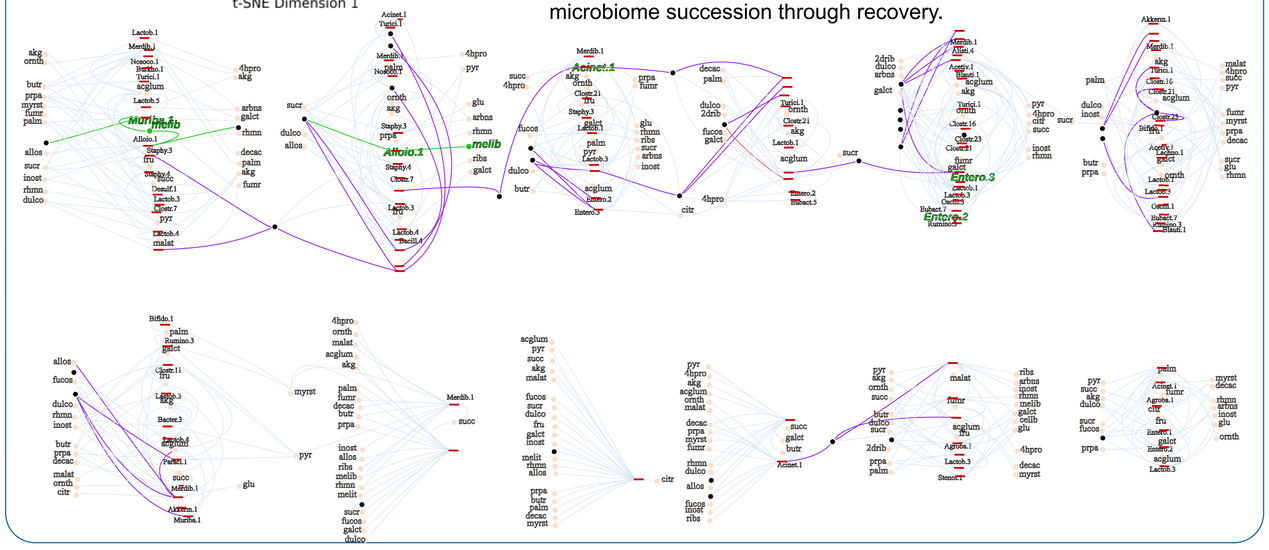


prFBA STUDY OF MAMMALIAN GUT MICROBIOME RECOVERY

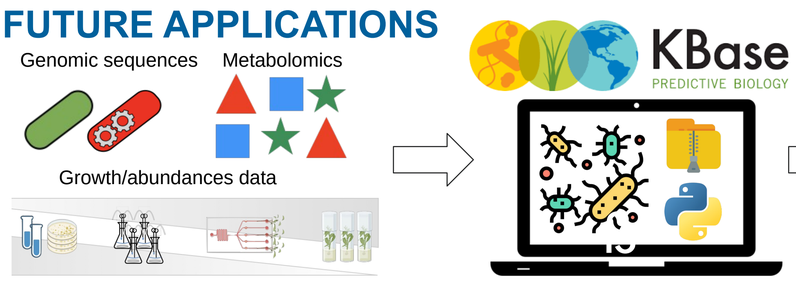


prGEMs were created from ASVs derived of mice gut microbiome after antibiotics was given to 2 sets of mice: one fed a high-fat & low-fiber diet (WD) and one fed a low-fat & high-fiber diet (RC). PrFBA-generated fluxes were are visualized in the left figure that plots the prGEM of each sample according to their metabolic function. Clustering of the pre-antibiotic communities and the "recovered" RC communities in the top-right corner demonstrates a "healthy" metabolic profile. The middle cluster reveals that RC microbiomes recover much faster than and are not perturbed as much as WD microbiomes. The lower-left cluster exhibits the vulnerability and dysbiosis of WD microbiomes that manifests in greater antibiotics perturbation and slower recovery.

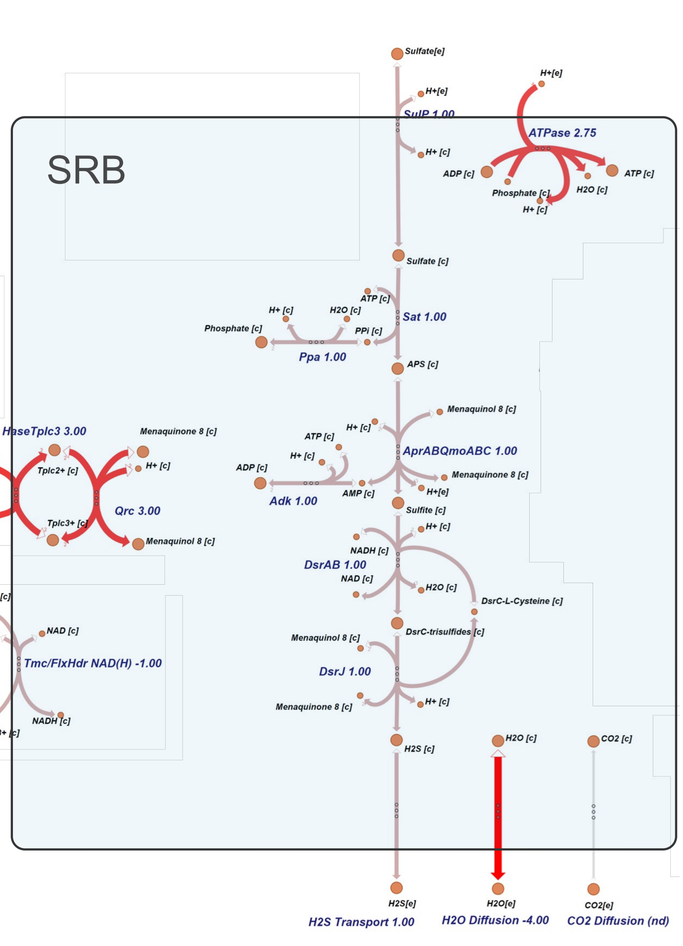
The Escher Map below illustrates the simulated community at each day interval. Dashed edges denote consumption and solid edges denote production, and highlighted edges and emphasized labels communicate the metabolic patterns that we propose indicate ecological succession. These simulations results are consistent with the experimental conclusions that complex carbohydrates metabolism is the crux of microbiome succession through recovery.



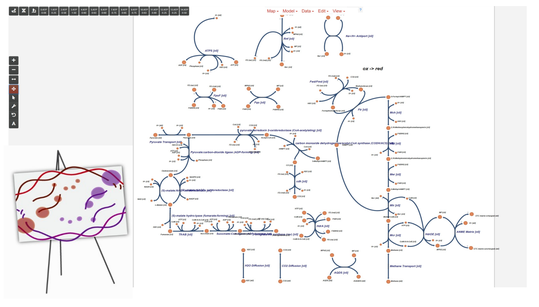
FUTURE APPLICATIONS

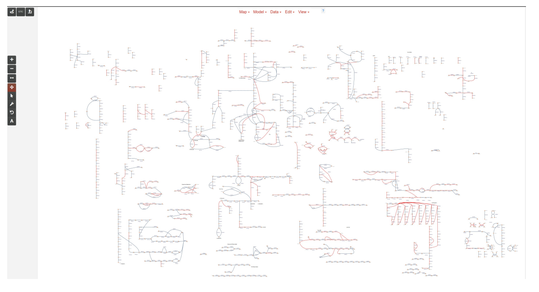


We plan to connect our augmented metabolic templates with the prGEM & prFBA pipeline to generate uniquely accurate simulations of these ANME communities for basic understanding of biogeochemical processes pertaining to climate change. We further intend to expand the pipeline to integrate more data types and thereby improve our ability to recapitulate experimental systems. The subsequent step is to automate the mature pipelines in KBase Apps so investigators can reproducibly use this same analysis pipeline to study community-derived genomic fragments (MAGs and/or ASVs). A final aspiration is to wrap our code in the ModelSEEDpy API so that computational researchers can utilize our methods in their own high-throughput pipelines.



Data Visualization: Network visualization of the metabolic templates are available at https://modelseed.org/annotation/projects/anme/. The ETC map displays all individual EFM of the ANME ETC network, while the Genome-Scale map shows the metabolism that was common across all ANME clades or strains.





U.S. DEPARTMENT OF ENERGY
Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.
KBase
Northwestern University
Argonne NATIONAL LABORATORY