Outline for MM model Paper

* Introduction/Background
  + Methane is an important chemical in the global carbon cycle, it’s a greenhouse gas, and a fuel source. Methanogens make about 1 GT of it per year, so they’re hugely important microorganisms
  + Methanococcus maripaludis S2 is a model methanogen; it possesses a fast doubling time, grows readily in a chemostat, and has a well-developed set of genetic tools.
  + Metabolic models are useful because they serve as organism knowledge bases and because they can be simulated to predict growth phenotypes for potential wet lab experiments. They have promise for guiding metabolic engineering efforts such as harnessing the unique energy metabolism of our hydrogenotrophic methanogen.
  + We have already constructed the most current metabolic models for two Methanosarcina and have also developed a likelihood-based gap filling method for building new metabolic models with increased gene homology. We have combined our expertise in modeling methanogens with our new gapfilling tool to produce the first manually-curated genome scale metabolic model constructed using likelihood-based gapfilling
* Methods
  + Standard model-building methodology
  + Code generating method (codes I want to distribute with the model)
    - maxGrowthOn\_\_ codes
    - simulateKOPanel code
    - switchToFormate code
    - switchToSpecificFerredoxins code
  + Chemostat culture growth method
  + Metabolomics data method
  + Dry cell weight measurement method
* Results
  + Basic model data
    - # rxns, mets, genes (currently 662, 687, 494)
    - % genome covered
    - # gene-associated rxns, transports, exchanges
  + Results of likelihood-based gapfill
    - We have # gapfilled reactions in our model, and unlike other gapfilling, our reactions come equipped with “likelihood” scores. These scores give us and other users insight into the reaction’s place in the model and go beyond the standard method of evaluating a reaction. Now we can easily see the least likely reactions and point ourselves at these entry points as a means of improving the model.
  + Comparison with growth yield data
  + Comparison with knockouts
    - Compared with data from Leigh lab across 6 papers, we match up very well with KO data.
  + Comparison with metabolomics data
    - TBD
* Discussion
  + Another group published a model of M. maripaludis in 2014; this work marked the first effort to represent M. maripaludis metabolism *in silico*, though the model deviated from published literature in several notable ways:
    - Missed electron bifurcation pathway, the essential connection that completes the Wolfe cycle
    - Includes sulfate transport, which M. maripaludis is known not to do
    - Uses methanophenazine, a cofactor we know to be missing from M. maripaludis
  + The other group’s model relied primarily on the KEGG and MetaCyc databases, making use of only 16 other literature sources for their reactions. Though we relied heavily on the DOE Kbase, we also used many literature sources. Furthermore, due to our use of reaction likelihoods, we are able to score each reaction accordingly, as opposed to simply attributing each reaction to the database it was taken from. This gives us a much better way of evaluating how we did in our model.
  + This model represents the first manually curated model that was constructed with likelihood-based gapfilling, at least to our knowledge. The likelihood scores lend an element of accountability to our gapfilling, but we’ve also strived for accountability making our decisions explicit throughout the curation process (this is a tie-in with Ben’s paper).

Figure 1. Comparison of predicted and experimental growth yields

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| KO Genes | H2 | Formate | H2 + Formate | Formate + CO | |
| ∆hmd | N | N | N | N |  |
| ∆mtd | N | N | N | N |  |
| ∆frcA | N | N | N | N |  |
| ∆fruA | N | N | N | N |  |
| ∆frcA∆fruA | N | N | N | N |  |
| ∆vhcAU∆vhuA | N | N | N | N |  |
| ∆fdhA1 | N | N | N | N |  |
| ∆fdhA2 | N | N | N | N |  |
| ∆fdhA1∆fdhA2 | N | L | N | L |  |
| ∆fdhA2∆fdhB2 | N | N | N | N |  |
| ∆ehbF | N | N | N | N |  |
| ∆3H2ase | N | N | N | N |  |
| ∆5H2ase | L | L | N | N |  |
| ∆6H2ase | L | L | N | N |  |
| ∆6H2ase∆cdh | L | L | N | L | **TOTAL** |
| **Total Correct:** | **10 of 10** | **13 of 13** | **2 of 2** | **2 of 2** | **27 of 27** |
|  |  |  |  |  |  |
| Figure 2. Knockout lethality predictions from FBA and agreement with experimental results | | | | | |

|  |  |  |
| --- | --- | --- |
| ***Methanococcus maripaludis S2* model comparison** | | |
| Model | iMR494 | iMM518 |
| Protein Coding Genes | 494 | 518 |
| % ORF Coverage | 29 | 30 |
| Intra/Extracellular Metabolites | 635/52 | 556/49 |
| Dead End Metabolites | 266 | 163 |
| Internal Reactions | 610 | 570 |
| Exchange Reactions | 52 | 49 |
| Gene-Associated Reactions | 485 | 464 |
|  |  |  |
| Table 1. A comparison between iMR494 and iMM518 | |  |