533f A New Milp Based Approach for *in Silico* Reconstruction of Metabolic Networks and Its Application to Marine Cyanobacterium *Prochlorococcus Marinus*

Xiaoxia (Nina) Lin, Aaron Brandes, Jeremy Zucker, and George M. Church

The increasing availability of annotated genomes has rendered the possibility of applying systems approaches, e.g. flux balance analysis (FBA), to the study of a large variety of organisms. To achieve this high-throughput goal, we have been developing an automatic bioinformatics pipeline to facilitate the process of generating *in silico* whole-cell metabolic models from genome annotations. In this work, we present a new computational framework for the key step in this pipeline which constructs metabolic networks by integrating genome annotation, reaction database, and phylogenetic information.

Our goal is to construct metabolic pathways/networks for a new species based on its genome annotation and a multiple-species pathway/reaction database (e.g. BioCyc databases). Using a mixed-integer linear programming (MILP) optimization framework, the new algorithm selects a set of reactions from a universal super-network which can achieve the functionality of a pathway or network to convert specific metabolites or to enable the cell to live and grow. The solution includes not only reactions already identified in the genome annotation but also additional ones required to achieve the functionality which are most possible phylogenetically. Alternative and/or sub-optimal solutions can also be systematically generated to increase the likelihood of identifying the real biological network. In addition, quantitative data such as nutrient condition can be readily incorporated to improve the predictions.

The above approach has been applied to the study of a marine cyanobacterium *Prochlorococcus marinus*, which dominates the phytoplankton in the tropical and subtropical oceans and contributes to a significant fraction of the global photosynthesis. The algorithm automatically generated novel TCA pathways which do not exist in the pathway databases and are consistent with partial knowledge of cyanobacteria. We have also successfully reconstructed the central carbon metabolic network and are currently exploring the whole-genome metabolic network of *Prochlorococcus* as well as using the results from this approach to identify missing genes/enzymes and to refine the genome annotation.