

The DREAM Project: Defining Comparative Analysis Metrics for Systems Biology

The National Center for Biomedical Computing at Columbia University (MAGNet), in collaboration with the IBM Computational Biology Center (CBC), is starting a new initiative to create a yearly workshop and database for the comparative analysis of systems biology methods and algorithms, specifically in the area of reverse engineering.

A small planning workshop has been organized, which will be hosted in March by the New York Academy of Sciences in the context of their Systems Biology SIG. The meeting will convene a small number of colleagues to further explore and define the details and formats of this initiative. The goal of this planning workshop is to explore the scientific community consensus and potential interest in creating a community-driven database called DREAM (Database for Reverse Engineering Analysis and Methods) and an associated recurring workshop that will use this database for the quantitative and systematic comparison of systems biology methods and algorithms.

The DREAM project is intended to fill up a gap in the world of reverse engineering of biological circuits, also known as pathway/cellular-network inference. It will consist of two interdependent features: the DREAM workshops, and the DREAM database.

The DREAM workshop. The proposed workshop is modeled after CASP (Critical Assessment of Structure Prediction, Moult J., et al., *Proteins*, 2005, Sep 26), a bi-yearly meeting to assess the state of the art in algorithms for structure prediction. The DREAM workshops will provide a useful platform to the systems biology community to understand the specific strengths and limitations of the multitude of pathway inference algorithms available in the literature. The organizers will curate a community-driven dataset containing biological data (microarray, proteomic, sequence, etc.) created from known, but not publicly available pathway models. Data may be either derived from biological experiments, or from simulations of known, validated biological models, or from completely synthetic models. The challenge to the systems biology community would be to infer the underlying pathway model from the data. The ultimate goal of this exercise is to gain insight on the merits and areas of strength (or lack thereof) of the different algorithms and methods and, more importantly, on the specific context(s) in which they work best. Thus, the workshop may have different sub-categories of analysis (i.e. reverse engineering from microarray expression profile, from protein concentration profiles, from promoter sequence data, etc.). The specific scope and nature of the categories will be an important subject of debate at the planning meeting.

The DREAM database. A second aim of this project, which will also be discussed at the planning workshop, is to create the DREAM database, a repository of computational analysis tools, biological and synthetic models, data, and literature on reverse engineering. The database will be made available to the community by the MAGNet Center, via a web site where researchers can either deposit their algorithms, models, and data or can download data to test their algorithms. Similarly, the experimental biologists can use the site either to find the most appropriate algorithms to dissect their experimental data or to upload their data and models to provide new scientifically

relevant challenges to computational biologists.

We believe that these ideas could find a substantial acceptance and resonance in the computational and systems biology community. But to serve its intended audience, it is critical that the format of the workshop is both enticing and scientifically sound. With the first workshop scheduled for the fall of 2006 in the context of the DIMACS meetings, we would like to invite you to participate in this key planning workshop where we hope to discuss the specifics of both the workshop and the database.

Funding for both the workshop and the database has been made available in the context of the NIH roadmap initiative for the National Centers for Biomedical Computing. The planning workshop will take place on March 9-10, 2006 at the New York Academy of Sciences. For the first DREAM workshop, which will occur on Sept. 7-8 '06, we have already secured logistic and financial support from DIMACS (Ctr for Discrete Math and Theoretical Computer Science: <http://dimacs.rutgers.edu/Workshops/ReverseEng/>)

So far, the following researchers have accepted to be part of this initiative:

Organizers:

Andrea Califano, Dr. (Columbia University) Co-Chair
Gustavo Stolovitzky, Ph.D. (IBM) Co-Chair

Steering Committee:

Gary Bader, Ph.D. (Memorial Sloan Kettering)
Joel Bader, Ph.D. (Johns Hopkins)
Hamid Bulouri, Ph.D. (Institute for Systems Biology)
Harmen Bussemaker, Ph.D. (Columbia)
Jim Collins, Ph.D. (Boston University)
Diego Di Bernardo, Ph.D. (TIGEM)
Tim Gardner, Ph.D. (Boston University)
Andre' Levchenko, Ph.D. (Johns Hopkins)
Eran Segal, Ph.D. (Weitzmann Inst.)
Andrey Rzhetsky, Ph.D. (Columbia)
Marc Vidal, Ph.D. (Harvard)