



New York University

A private university in the public service

Prof. Dennis Shasha
Courant Institute of Mathematical Sciences
Department of Computer Science
New York University
251 Mercer Street
New York, N.Y. 10012, USA
Telephone: (212) 998-3086
Electronic mail: shasha@cs.nyu.edu
Web: <http://cs.nyu.edu/cs/faculty/shasha/index.html>

September 14, 2005

Dr. Rodrigo A. Gutierrez
Department of Biology
New York University

Dear Rodrigo,

I am writing to support your HHMI application. Our collaboration in the coming years will focus on our ongoing work in machine learning approaches to map information flow and discover functionality in gene networks. As an underpinning to that work, we will continue to collaborate on the NSF Virtual Plant Bioinformatics project on which we are co-Pis.

It has been a pleasure to work with you for the last two years. That we have accomplished so much in such a short time is largely due to your intelligence, hard work, and clear-headedness. You are able to take a difficult problem involving experimentation, statistics, and computation and synthesize the essential patterns. You are strong in the lab, you have a wide scientific culture, and you are even a credible programmer.

During our collaboration, we have worked on a systems biology approach to understand regulation by Carbon and Nitrogen in Arabidopsis. This has involved:

- Integrating gene expression with metabolic networks in a collection of cooperating databases and analytical tools.
- Introducing new visualization methods to understand expression data.
- Modeling gene network states to allow predictions about untested experimental conditions.

Our visualization work known as Sungear has been aimed at integrating the results of different experiments. Suppose one has N experiments and each results in a set of genes that have some property of interest (e.g. they are all induced at a certain level or above). One would like to answer certain questions about the relationships among the experiments such as: How much of an intersection is there between one experiment's genes of interest and

one, two, three or more other experiments? Does any subset of the experiments show an overrepresentation in some functional category, e.g. nitrogen metabolism? In general, how can one relate experiments to function?

Finally, you have pioneered a database/tool system called Virtual Plant which includes visualization tools like Sungear and call-outs to Cytoscape as well as analytical tools like your widely used Biomaps software.

In all of our work, we have tried to build frameworks that apply across plants and even to mammals and bacteria. Sungear for example can apply equally well to baseball and proteomics. I believe this work will be of great general interest and anticipate that we will continue to collaborate on many projects including Virtual Plant as you continue your promising career.

On a personal level, you are a real pleasure to work with. You are thorough, open, knowledgeable, and creative. When you present material, you do so in a lucid and graphically appealing way. I foresee outstanding work from your future research and expect you to be a leader in your field.

Regards,

Dennis Shasha