**Brief Bio and (PR)<sup>2</sup>: Problems & Pitches – Rants & Raves by *{Richard Bonneau}***

Dr. Bonneau is a faculty member at New York University's Center for Comparative Functional Genomics, where he is a joint member of both the Biology and Computer Science departments. His affiliate position with the Institute for Systems Biology offers many exciting possibilities for exchange of ideas between these two centers, both focused on applying functional genomics to a wide range of systems. His research efforts are focused on making computational tools, algorithms and methods for systems-wide elucidation of biological systems. His research aims to develop computational methods at the intersection of two interrelated fields: protein structure and functional genomics. Rich is also currently the technical lead on two grid computing collaborations with IBM -- the first and second phases of the Human Proteome Folding Project

Dr. Bonneau did his doctoral work at the University of Washington in Seattle working with David Baker on the the protein prediction platform, Rosetta. Rich is currently a member of the Rosetta-commons and continues to develop and work with Rosetta as part of several projects in the lab. Before joining NYU he worked as a Senior Scientist with Leroy Hood at the Institute for Systems Biology. Rich also oversees TACITUS's (www.tacitus.com) approach to data gaming (3D visualization) for all applications that focus on genomics, computational biology and cell biology.

Lab website:
http://homepages.nyu.edu/%7erb133/R.Bonneau-lab.html

Viz tools and lab projects:
http://err.bio.nyu.edu/cytoscape/bionetbuilder/
http://gaggle.systemsbiology.org/docs/
http://www.worldcommunitygrid.org/projects_showcase/viewHpf2Research.do

**General Questions**

1) What is (are) your main interest(s) in attending the workshop?

Well, honestly, you've assembled a great group and I want to see what people are up to. But after that I'd like to figure out if there are some better ways to coordinate how we organize our efforts to develop visualization tools. I've got a few ideas about how to organize visualization without needing to have each of us hire 100 programmers and I'd like to discuss them. I'd like to talk about how to get some of these ideas funded.

2) What information/knowledge management needs do you have?
Explain your 'dream tool' for scientific discovery and innovation.

Between our genome annotation project (using structure prediction to annotate proteins) and our network inference/functional-genomics projects we have viz needs that fall into two main categories:

I'm separating the issue of analysis, reconstructing casualty or biological structures and focusing on data integration and visualization.

2.1. protein centric explorations of annotations and protein associates. Ideally what I want is a web/list/table/query driven site where people look at proteins they are interested in. At the same time networks of protein associations, processes , context are displayed as part of a separate tool. The web-annotation view and the context-network tool intercommunicate so that all queries in one are echoed in the other so that people can explore putative annotations in a network context. We have a system that does this, but we need to work on it to get the communication more symmetric.

2.2. functional genomics data integration to explore the results of biclustering, network inference, etc. We have new methods in the lab that integrate lots of datatypes to predict regulatory networks, co-regulated groups, active pathways, and TF and microRNA binding sites. Thus we need people to be able to explore the results of the analysis (a diverse set of data-structures) in the context of the original data (also a diverse set of data types). So far we've done this with the Gaggle (Paul Shannon) and had ok success. The gaggle is built with the principle that we need independent tools for different views / queries/ datatypes, and that communication between the tools should be a compact controlled vocabulary with a simple API. What would be nice to do here is to work on making it simpler to hook in new tools, and simpler to set up. Translating all the synonymous names genes have in different datasets for the same organism is a real pain! In the end there are some real backend problems that should be solved …we need a naming and data standards enforcer.

3) What is the most insightful visualization of static or dynamic phenomena you know?
 *[Ideally this visualization led to a major discovery/innovation. Examples could come from science, art, or any other field of human endeavor. Note that we plan to use this visualization on your name card.]*

A *good* subway map.

4) What would you like to learn / achieve at the workshop?

I think that for viz it not just the difficult of coding, or the difficulty of laying out large graphs or problems like that, but we really need to be more creative in how we present.  I want to get ideas along those lines. Also, I would like to be given a large grant without having to write it.

*Please send the completed document by <u>February 20th, 2008</u>*
*to Katy Borner <katy@indiana.edu> and Elisha Hardy <efhardy@indiana.edu>*