Center for Eukaryotic Structural Genomics

Technology Dissemination Report

CESG Tech Report No.	002
Title	Target Trace and PepcDB Generation Software
Research Unit	Bioinformatics
Authors	Pan, X., Wesenberg, G.E., and Bingman, C.A.
Primary Contact	cbingman@biochem.wisc.edu
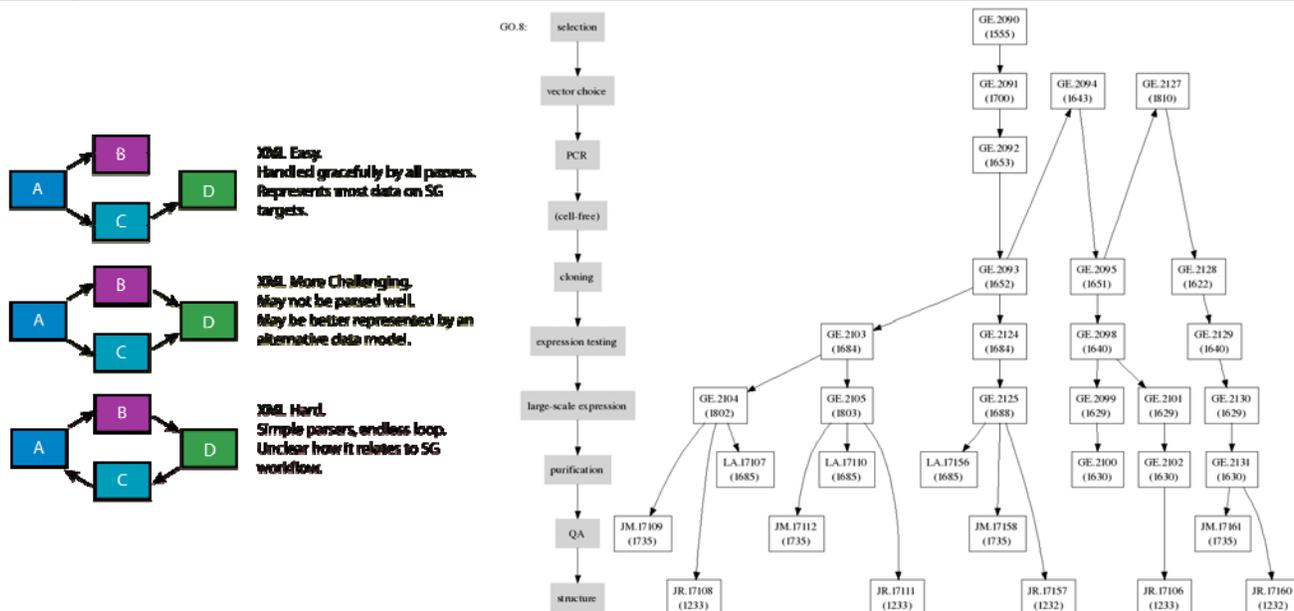
Summary

We formulate both the problem of generating valid and useful PepcDB reports from CESG to the world and tracking the progress of individual targets using graph theory. We suggest that all valid PepcDB XML files can be represented as directed, acyclic graphs. Software has been developed to distill standard Sesame reports containing hierarchical (parent-child) relationships between database objects such as workgroups and sample records into visually appealing and intuitive pictures. These graphs simultaneously display all work done on a given target as one or more graphs rooted in target selection actions, and flowing across our multi-vector, multipass standard platform. There is no need for fixed templates to describe workflow. This allows our platform to be assembled literally "on the fly" based on the response of a given target protein to a battery of experimental trials, while maintaining the ability to report detailed PepcDB trials for each distinct experimental variation. The graph also allows users to easily pinpoint data entry errors, such as mislinked database objects. It has been invaluable in increasing the scientific accuracy and utility of our PepcDB reports.

Although this software was developed in the context of the Protein Structure Initiative, it is trivially extensible to describe the flow of functional proteomics work, or any other high-throughput activity using the Sesame database. With some additional effort, the software could be adapted to produce graphical reports from other project databases. The software was written using freely distributed components, such as Perl, DOT, and GraphViz software, and displays results via an Apache web server to a user's web browser.

Publication:

[1] Pan, X., Wesenberg, G.E., Markley, J.L., Fox, B.G., Phillips, G.N., Jr., and Bingman, C.A.. A graphical approach to tracking and reporting target status in structural genomics. *JSFG* 8(4):209-16.

Acquiring the Technology	Contact Craig Bingman cbingman@biochem.wisc.edu.
Other Acknowledgements	The Sesame programming team, especially Zsolt Zolnai and all CESG Sesame users.