

MASC Bioinformatics- 2006 Subcommittee Report

Prepared by Chris Town (Chair) and Heiko Schoof

Last year's report described the first major step towards improved data integration in the *Arabidopsis* community in the form of an NSF-Funded Workshop on "Data Integration" at TIGR in April, 2005. This was followed by a workshop at the International *Arabidopsis* Meeting in Madison, Wisconsin in June 2005 at which the results of the TIGR workshop were presented to the research community. The Madison workshop was well attended (over 100 people) and there was considerable enthusiasm for the prospects of Web Services to help solve some of the issues of data integration. Based upon this community response, both the Deutsche Forschungsgemeinschaft (DFG) and the National Science Foundation (NSF) indicated their willingness to support a pilot project to begin to develop web services. In the Fall of 2005, proposals were submitted to DFG (by Heiko Schoof) and to NSF (by Chris Town) and both were funded (see <http://bioinfo.mpiz-koeln.mpg.de/araws>). Each proposal aims to implement web services that will provide and facilitate the integration of data that is not part of existing data warehouses like TAIR, NASC or MIPS. In addition to developer's workshops to be held both in Germany and the U.S., each group will provide a "flying geek" service to provide on-site assistance to clients with limited informatics expertise.

The goals of this integrated pilot project are twofold. The first is to educate a limited number of groups in web services technology, to define the IT requirements necessary to deploy web services at any site and to develop a set of Standard Operating Procedures that can be widely used by the *Arabidopsis* community. The second is to actually have some more web services and workflows up and running in time for demonstrations at the International *Arabidopsis* meeting in Madison in the summer of 2006. At this point, participation in the pilot project has, for logistical reasons, been limited to relatively small groups of people in the EU and U.S. However, this is not meant to be an exclusive process and we hope that these initial trials will allow us to more rapidly propagate web services throughout the *Arabidopsis* community in the months and years to come.

The first workshop was held at MPIZ Cologne from March 14-18, 2006 and was attended by representatives of MIPS, MPIEB-Tuebingen, MPIMP-Golm, RZPD-Berlin, University of Cologne, VIB/PSB-Ghent and Centre for Biotechnology in Turku, Finland as well as the TIGR partners. In addition to instruction and demonstration of the core elements of BioMoby web services technology, topics covered included ontologies, data standards, and types of data that might be served. Although many areas requiring a significant amount of work were identified, the group was optimistic about having new services in place by the summer. The parallel U.S. workshop took place at TIGR from April 24-28, 2006 and built upon the experiences of the Cologne meeting.

In August, 2005 the NSF convened a meeting of prominent researchers from both within and outside the *Arabidopsis* community to assess the progress of the 2010 project at the mid-point of its 10-year period. Embedded in the report (www.arabidopsis.org/info/2010_projects/AT2010WorkshopFinal.pdf) are a significant number of comments and recommendations relating to perceived bioinformatics needs. Some of these had already been voiced and discussed at the TIGR workshop in April 2005. Practical steps towards accomplishing some of the goals were again discussed at the Cologne workshop and will continue to be the focus of the BioMoby pilot project. A successful outcome of this project, including the enabling of web services in the broader *Arabidopsis* community, will go a long way towards satisfying the recommendations of the midterm report.