

## TOOLS AND RESOURCES

### **Getting databases talking: Enabling high-throughput data integration through web services**

**H. Schoof**

*Max Planck Institute for Plant Breeding Research, Plant Computational Biology, Cologne, Germany*  
schoof@mpiz-koeln.mpg.de

Key Words:- BioMoby, webservices, workflow and data integration

Is your research data driven? Or is finding the data driving you mad? Despite the fact that we all make diligent efforts to publish our results and to distribute our data via the web, it is still difficult for the average user to collect all the data on a particular gene, pathway or trait without visiting several web sites, and even then one cannot be sure that important pieces of information have not been overlooked. And centralized databases can hardly keep up with the pace of new data being generated.

Distributed but interoperable data resources provide a flexible and scalable architecture to facilitate high-throughput data integration. To enable automated discovery of services and on-the-fly creation of analysis workflows, the BioMOBY project extends the web service paradigm with a biological data object ontology. This creates new possibilities for automatic querying of multiple databases and on-the-fly integration of data. Using point-and-click tools like Taverna, analysis workflows can easily be built and executed remotely, bringing comparative analyses to any biologist's desktop without the need for programming or warehousing. A project initiated by the Multinational Arabidopsis Steering Committee and funded by the NSF and the DFG builds on experiences from the European PlaNet project and demonstrates interoperability in a network of European and US Arabidopsis data providers.

#### References

1. PlaNet: <http://www.eu-plant-genome.net>, funded by the 5th Framework Programme of the European Commission, QLRI-CT-2001-00006
2. *Arabidopsis* data integration through webservices: <http://bioinfo.mpiz-koeln.mpg.de/araws>, funded by Deutsche Forschungsgemeinschaft and National Science Foundation.
3. Wilkinson M, Schoof H, Ernst R, Haase D: BioMOBY successfully integrates distributed heterogeneous bioinformatics Web Services. The PlaNet exemplar case. *Plant Physiol* 2005, 138:5-17.
4. Oinn T et al.: Taverna: A tool for the composition and enactment of bioinformatics workflows. *Bioinformatics* 2004, 20:3045-3054