

# United States

[http://www.Arabidopsis.org/info/2010\\_projects/United\\_States.jsp](http://www.Arabidopsis.org/info/2010_projects/United_States.jsp)

Contact: Philip Benfey

Duke University

Email: [philip.benfey@duke.edu](mailto:philip.benfey@duke.edu)

Mary Lou Guerinot

Dartmouth College

Email: [Mary.Lou.Guerinot@dartmouth.edu](mailto:Mary.Lou.Guerinot@dartmouth.edu)

## **AT2010 midterm progress**

A major event in the U.S. *Arabidopsis* functional genomics community was the marking of the midpoint in the National Science Foundation *Arabidopsis* 2010 program ([www.nsf.gov/bio/pubs/awards/2010awards.htm](http://www.nsf.gov/bio/pubs/awards/2010awards.htm).) The North American *Arabidopsis* Steering Committee (NAASC) organized a two-day workshop in August 2005, to evaluate the progress made and provide guidance for the second half of the program. Workshop participants came from the U.S., Pacific Rim and Europe and represented a wide range of subspecialties within the *Arabidopsis* genomics community. The final report is available in its entirety at [www.nsf.gov/pubs/2006/bio0601/bio0601.pdf](http://www.nsf.gov/pubs/2006/bio0601/bio0601.pdf). Prior to the workshop, a web-based survey was conducted to solicit input from the broader community. In addition, data were collected on the impact of the 86 funded projects to date, including publications, stocks deposited and database submissions. The overall assessment was that most of the initial goals set in 2000 for the first five years had been met or surpassed. Of particular note was the utility of genome-wide resources such as insertion lines. In assessing the funded projects, the panel came to the conclusion that those that used high throughput and/or computational approaches to understanding biological processes had the highest impact. For the upcoming five years of the program, the panel recommended focusing on these 11 areas:

1. Benchmarking gene function
2. Developing genome-wide tools and reagents for analyzing gene function and regulation
3. Improving genome annotation and tools for visualization, annotation and curation
4. Improving database integration and developing new modeling and computational tools
5. Exploring exemplary networks and systems
6. Analyzing non-protein coding genes
7. Leveraging natural variation to understand gene function in *Arabidopsis thaliana*
8. Localizing gene products at the cellular and subcellular level
9. Facilitating metabolomics and ionomics
10. Engaging the broader community
11. Enhancing international collaboration
- 12.

Looking beyond 2010, the panel foresaw a program that "enables the international community of plant biologists to analyze, understand, and manipulate the full spectrum of biological processes required to make a plant that functions effectively and predictably under both standardized laboratory conditions and complex natural environments."

## **Plant Cyberinfrastructure Center Workshop**

Recent progress in plant science has resulted in the development of a wide range of new tools and resources for research and education. Optimal use of these resources requires combining the information represented among them in innovative ways to achieve a better understanding of fundamental principles in plant biology; and it also requires that individuals from multiple fields and disciplines be able to find, understand and effectively employ these resources in novel ways.

To discuss the means by which to achieve such a synthesis of resources, and to design the appropriate cyberinfrastructure for their best utilization, a workshop for a Plant Cyberinfrastructure Center was held at the

National Science Foundation on October 17 and 18, 2005 (complete report at [www.arabidopsis.org/info/2010\\_projects/index.jsp](http://www.arabidopsis.org/info/2010_projects/index.jsp).) Participants represented the widest possible range of plant and computational biologists, with experts in plant genomics, development, metabolism, ecology, and evolutionary biology, ecoinformatics and experts in computational modeling, databases, computer infrastructure, software, and mathematics. The participants achieved consensus on the need for a center to provide the cyberinfrastructure for facilitating interdisciplinary research in plant science and for generating a new kind of plant biologist, and made a number of recommendations and suggestions outlining their vision for such a center.

Workshop recommendations:

1. There is a strong need to create a plant cyberinfrastructure center to promote the integration of diverse and large-scale genomics and other data to address a few fundamental problems in plant biology using multi-disciplinary approaches.
2. A core mission should be to train a new generation of scientists who can combine multi-disciplinary approaches and utilize data in public repositories maximally.
3. The center should be composed of an entity that is connected to a number of adjunct institutions with different types of scientific expertise and facilities.
4. The core infrastructure and staffing should follow the successful models of NCEAS (National Center for Ecological Analysis and Synthesis) and NESCent (National Evolutionary Synthesis Center) and designate enough core staff to facilitate the activities of the center efficiently.
5. The center should provide an umbrella infrastructure to coordinate disparate outreach activities and train students and faculty in achieving a comfort level in interacting with the public about plant science and its importance.
6. The center should maximally leverage existing databases and standards to promote international integration of data.

**Young Researcher Exchange Program**

In collaboration with the German *Arabidopsis* Functional Genomics program, AFGN, an NSF-funded Young Researcher Exchange Program was started. This program provides stipends and travel expenses for short-term research visits (up to three months) for graduate students (and postdoctoral fellows in the U.S.) to work in German laboratories and for German students to work in U.S. labs. For more information please see ([www.uni-tuebingen.de/plantphys/AFGN/CallforUScandidates.pdf](http://www.uni-tuebingen.de/plantphys/AFGN/CallforUScandidates.pdf)) and ([www.uni-tuebingen.de/plantphys/AFGN/yrep.htm](http://www.uni-tuebingen.de/plantphys/AFGN/yrep.htm)).

**Major funding sources for Arabidopsis functional genomics:**

- NSF: National Science Foundation ([www.nsf.gov/](http://www.nsf.gov/))
- USDA: U.S. Department of Agriculture ([www.usda.gov/wps/portal/usdahome](http://www.usda.gov/wps/portal/usdahome))
- DOE: U.S. Department of Energy ([www.energy.gov/](http://www.energy.gov/))
- NIH: National Institutes of Health ([www.nih.gov/](http://www.nih.gov/))

**Arabidopsis genomics tools and resources:**

- TAIR (The *Arabidopsis* Information Resource, <http://arabidopsis.org>) collects information and maintains a database of genetic and molecular biology data for *Arabidopsis thaliana*, a widely used model plant. TAIR is produced by the Carnegie Institution of Washington Department of Plant Biology, Stanford, California, and the National Center for Genome Resources (NCGR), Santa Fe, New Mexico. Funding is provided by the National Science Foundation, (Grant No. DBI-9978564) and DBI-0417062. TAIR collaborates with the *Arabidopsis* Biological Resource Center (ABRC) to provide researchers with the ability to search and order stocks. The data in TAIR can be searched, analyzed, downloaded, and viewed using the interactive SeqViewer tool. In addition, pages on news, information on the *Arabidopsis* Genome Initiative (AGI), *Arabidopsis* lab protocols, and useful links are provided.

TAIR's first genome release, version 6.0, has been integrated into the TAIR database, SeqViewer, MapViewer and sequence analysis datasets, and is also available from NCBI. The TAIR6 release contains 26,751 protein coding genes, 3818 pseudogenes and 838 non-coding RNA genes (31,407 genes in all). A total of 437 new genes were added and nine genes were made obsolete. Updates were made to 973 genes including 831 updates to coding sequences, 14 gene splits and 7 gene merges and the addition of 1200 new splice variants. A total of 3159 *Arabidopsis* genes (10%) now have annotated splice variants. No changes were made to the chromosome sequences for this release. Access to the fully annotated chromosome sequences in TIGR xml format as well as FASTA files of cDNA, coding, genomic and protein sequences and lists of added, deleted and changed genes are available at the TAIR website. FASTA formatted files for all sequence analysis

datasets including sets of intron, intergenic, UTR, upstream and downstream sequences are also available. Access the new release using TAIR's genome browser (<http://arabidopsis.org/servlets/sv>).

- ABRC (The *Arabidopsis* Biological Resource Center, [www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm](http://www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm)) distributes and is actively collecting stocks that are associated with genomics and phenomics. The NSF grant which supports the Center was recently renewed for a five-year period, through March 2011. Relevant seed stocks include insertion lines (200,000+ from many major contributors as well as smaller sets of purified lines from individual researchers), the lines of the *Arabidopsis* TILLING service, 800+ distinct natural accessions, 9 recombinant inbred populations, related species and RNAi lines including the AGRİKOLA lines. In addition, 1,900 of the ca. 50,000 "Confirmed" (genetically purified) T-DNA lines being produced by the J. Ecker laboratory have been received, and the remainder will be arriving during the next two years. These will be made available as they arrive, and distribution of these as large sets and sub-sets for phenotypic experimentation, etc, is planned. On the DNA side, ABRC presently houses full-length ORF and cDNA clones for 15,000+ genes, BACs covering the entire genome, BACS of four related species, the AGRİKOLA RNAi clones and various sets of Expression and Destination clones. The extensive Expression ORF collection from S. P. Dinesh Kumar is being received, with 1,100 of these currently in-house. It should be emphasized that donation of published mutants and clones, including purified insertion mutants and expression clones are very welcome. The annual distribution of seed and DNA stocks continues to exceed 90,000 in total volume.