

# United States

[http://www.arabidopsis.org/portals/masc/countries/United\\_States.jsp](http://www.arabidopsis.org/portals/masc/countries/United_States.jsp)

## Contacts:

Xing Wang Deng

Yale University

Email: [xingwang.deng@yale.edu](mailto:xingwang.deng@yale.edu)

Joe Kieber

University of North Carolina, Chapel Hill

Email: [jkieber@bio.unc.edu](mailto:jkieber@bio.unc.edu)

Joanna Friesner, MASC Coordinator

University of California, Davis

Email: [jdfriesner@ucdavis.edu](mailto:jdfriesner@ucdavis.edu)

## 2020 Vision for Biology Workshop: The role of plants in addressing grand challenges

The NSF's 2010 Project, designed to elucidate gene function in Arabidopsis, will have its last competition for proposals in 2010. Researchers in the US began discussions in early 2007 to develop a new vision to propel forward the next decade of Arabidopsis research in the US. In January of 2008, an NSF-sponsored workshop was held which included participants invited from different areas within the Arabidopsis research community as well as from other plant and animal model system communities. The goal of the workshop was to be forward-looking about the direction of biology research in the next decade and to discuss where plant biology, Arabidopsis, and model organisms fit into this larger vision.

Participants felt that focusing on Arabidopsis has proven to be highly effective in facilitating numerous breakthroughs and advances in plant biology in a relatively short time. Evaluation of the AT2010 project thus far described several major accomplishments including the development of a large pool of genetic and genomic resources, the demonstration that rapid and groundbreaking achievements can be made using a reference plant system, and the growth of a vibrant research community. In particular, participants noted that over the last two decades, Arabidopsis has nucleated an energetic group of researchers drawing from a wide variety of disciplines, including significant international collaboration. The ability to develop theories and rapidly test them in Arabidopsis was seen as a major strength and it was noted that Arabidopsis may be unique for its capacity to enable investigation from the molecular to the ecosystem level. Recommendations and findings from the workshop will be presented at the MASC annual meeting during the 19<sup>th</sup> Arabidopsis Conference in July 2008. A summary report can be found on the MASC website: <http://www.arabidopsis.org/portals/masc/workshop2020.pdf>

## New Plant Science Cyberinfrastructure Collaborative (PSCIC)

The PSCIC, also known as the iPlant Collaborative (iPC) is a major new NSF-funded project that began in spring 2008. The iPC is centered at the University of Arizona with several additional partner institutions and is funded at the level of 50 million USD for 5 years, with the possibility of a second 5 year funding period. The project goal is to develop a cyberinfrastructure collaborative for the plant sciences that would enable new conceptual advances through integrative, computational thinking. Importantly, the iPC intends to be fluid and dynamic and community-driven. The project encompasses all of the plant sciences. Arabidopsis, with its advanced resources, datasets and extensive research community, is expected to play an integral part. The project's opening conference, held in April, included over 400 participants, including approximately 100 via direct, live participation webcast. The conference included presentations by iPC members, computational and plant scientists, and education experts, as well as a number of smaller, focused discussion sessions. All presentations from the conference are available for public viewing at the iPC website: <http://iplantcollaborative.org/home>.

The iPC will bring together plant biologists, computer and information scientists and engineers, as well as other experts, to address 'Grand Challenges' in the plant sciences. The definition of a 'Grand Challenge' (GC) is an evolving concept but several ideas from the discussion groups include that GCs should have a compelling solution, have a well defined measure for success, give some predictive power, be both visionary and realistic, be compelling to the general public, and motivate and include many different types of researchers. GCs will be proposed by the community and selected by the iPC Board of Directors. The iPC will not directly fund data collection or researchers but is funded to develop the cyberinfrastructure that successfully approved GC proposals require. International collaboration is encouraged but compensation is limited to expenses. In the first year, in addition to GC proposals, the iPC will consider proposals to develop 'foundational tools' which have obvious benefit to the plant biology community and potentially to one or more future GCs. The iPC expects to begin reviewing workshop proposals as early as June 1, 2008 (for up to 6 workshops a year) and anticipates two rounds of reviews annually. Broader impacts of the iPC project are expected to extend beyond simply addressing grand challenge questions on a scientific level as the iPC will involve extensive community building and educational outreach.

## Textpresso Scientific Literature Text Mining System for Arabidopsis

Textpresso is a text-mining system for scientific literature (<http://www.textpresso.org>). Its two major elements are (1) access to full text, so that entire articles can be searched, and (2) introduction of categories of biological concepts and classes that relate two objects (e.g., association, regulation, etc.) or describe one (e.g., biological process, etc.). A search engine enables the user to search for one or a combination of these categories and/or keywords within an entire literature. The system was initially developed as part of Wormbase for *C. elegans* but more recently has been extended to additional organisms including Arabidopsis, which now has its own site ([www.textpresso.org/arabidopsis](http://www.textpresso.org/arabidopsis)). TAIR was heavily involved in developing the Arabidopsis Textpresso system; most of the Arabidopsis scientific literature comes from TAIR which sent around 15,000 in-house articles. In addition, TAIR provided the Arabidopsis vocabularies and categories. Currently, Textpresso for Arabidopsis contains information on these data types (data counts in parentheses): abstract (25,877), body (15,276), title (26,834), totaling 67,987 entries.

The system allows users to perform targeted searches of Arabidopsis literature, for example, using gene names, GO terms, or a combination of search terms. The results page lists literature citations containing the search query specified and the entire set can be downloaded into Endnote, or displayed for printing or in xml format. Although the system does not provide the full text of articles it does provide abstracts, and perhaps most useful, displays several sentences from each article that contain the search term(s). This quickly reveals the context of search terms which can help users determine the usefulness of the reference. Gene names or other entities mentioned only in the supplementary materials will not be retrieved because these documents are not currently included in the set of documents indexed and searched by Textpresso. Reference: *Textpresso: an ontology-based information retrieval and extraction system for biological literature*; Müller HM, Kenny EE, Sternberg PW (2004) Textpresso: An Ontology-Based Information Retrieval and Extraction System for Biological Literature. PLoS Biol 2(11): e309

## AT2010 Project

The National Science Foundation (NSF)-sponsored 2010 project aims to determine a function for all genes in *Arabidopsis thaliana* by the year 2010. Since its inception the project has funded proposals in two main areas: proposals that address gene function directly and proposals that develop enabling tools and resources for functional genomics research. Since the first awards were granted in 2001, approximately \$185 million has been allocated for 135 awards encompassing 111 diverse projects, ranging from studies of gene families, development of research tools, analysis of natural variation, transcriptomics, widely-used reverse genetics resources such as sequence-indexed insertion lines, and many others. In 2007, NSF committed approximately \$18 million for 16 awards spanning 12 projects. Project descriptions and funding levels can be found at the NSF AT2010 award page: [www.nsf.gov/bio/pubs/awards/2010awards.htm](http://www.nsf.gov/bio/pubs/awards/2010awards.htm).

This year, as in 2007, the Program will focus on (1) research on exemplary networks using high throughput methods and integrating modeling with experimental data to understand the gene circuitry underlying basic plant processes; (2) projects that will develop experimental and computational methods, tools, and resources for enabling a broad community of scientists to conduct functional genomics research on Arabidopsis; and (3) projects to perform genome-wide analyses of the gene function. Changed for 2008 is that the project will be an NSF-only activity, proposals will not be jointly reviewed by the German Arabidopsis Functional Genomics Network Program. International collaboration is still encouraged.

The abundance of knowledge, and especially, community resources, generated so far by the 2010 project have helped facilitate rapid advances in elucidating the function of many Arabidopsis genes. It is clear that cooperation and collaboration were key factors in the past success of the Arabidopsis Genome Project, and for the success of the current 2010 Project and other large-scale international Arabidopsis functional genomics projects. To maximize the return from such efforts, future projects must similarly be integrated and coordinated. It is also critical that funding levels are maintained.

## NAASC and the 18th and 19th International Conferences on Arabidopsis Research

The eight member North American Arabidopsis Steering Committee (NAASC, [www.arabidopsis.org/portals/masc/countries/NAASC\\_Info.jsp](http://www.arabidopsis.org/portals/masc/countries/NAASC_Info.jsp)) is composed primarily of US researchers and represents Arabidopsis researchers in the United States, Canada and Mexico. Annual elections by North American researchers registered at TAIR provide new members to replace two that rotate off the committee each year. Mark Estelle (Indiana University) and Jane Glazebrook (University of Minnesota) were recently elected to serve a four year NAASC term starting this July. Judith Bender and Xing Wang Deng conclude their term at the 2008 International Conference on Arabidopsis Research (ICAR). Additional continuing committee members include Xuemei Chen, Joe Kieber, Julian Schroeder, Caren Chang, George Haughn, and Scott Poethig.

NAASC provides North American representation to the MASC and serves as the main organizing and fundraising body for the ICAR when it is held in North America. Since 1995 the meeting had been in the US approximately 2 of every 3 years. In 2005 the NAASC changed the format of North American meetings to include sites to alternate with the usual location in Madison, Wisconsin. At the annual MASC meeting held during the 2007 ICAR in Beijing, the MASC adopted a 3 year conference site rotation: North America, Europe, and Asia/Pacific Rim.

The 18<sup>th</sup> ICAR, held in 2007 in Beijing, marked the first time the meeting took place in an Asian country and nearly 1500 attendees participated. NAASC member/MASC co-Chair Xing Wang Deng was the lead scientific organizer for the conference. Building upon increasingly global research efforts, the 19<sup>th</sup> Conference will mark another milestone: in 2008, the 19<sup>th</sup> ICAR will be held in Montreal, Canada, the first time the ICAR has come to Canada. The conference organizing

committee is chaired by NAASC members Joe Kieber and Xuemei Chen, and includes local researchers at McGill University as well as other NAASC members. Joanna Friesner, the MASC Coordinator from UC Davis, provided overall organization of each conference held between 2006-2008. The 2009 ICAR will be held in Edinburgh, Scotland, with lead organization provided by GARNet in the UK and assistance from the MASC Coordinator. The ICAR is expected to be held in Asia in 2010 and to return to the US in 2011.

**TAIR/Plant Physiology Collaboration:** A unique partnership has been formed between the journal Plant Physiology and TAIR to ensure that Arabidopsis gene function data published in the Journal are reliably captured in TAIR's database. This collaboration provides a mechanism for authors to submit Arabidopsis gene function information to TAIR as part of the publication process. It is anticipated that similar collaborations with additional journals will be established in the future.

## **US Young Researcher Exchange Program**

In 2005 a program was established to allow graduate students and post-doctoral fellows from NSF-supported US labs to engage in short-term research visits to German labs. This NSF-funded program is a collaboration with the German Arabidopsis Functional Genomics program, AFGN, which similarly allows German students to work in US labs. Since its inception, the US program has funded research visits to Germany by 2 post-doctoral fellows and 10 graduate students. The program will continue until the end of May, 2009. More information on the US program: [www.arabidopsis.org/portals/masc/NSF\\_Arabidopsis\\_research\\_program.pdf](http://www.arabidopsis.org/portals/masc/NSF_Arabidopsis_research_program.pdf)

Grant information: [www.nsf.gov/awardsearch/showAward.do?AwardNumber=0529918](http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0529918)

More information on the German program: [www.uni-tuebingen.de/plantphys/AFGN/yrep.htm](http://www.uni-tuebingen.de/plantphys/AFGN/yrep.htm)