

## 2009 Report: Phenomics

Prepared by Eva Huala (Co-chair, [huala@acoma.stanford.edu](mailto:huala@acoma.stanford.edu))

### Phenomics seed and DNA resources

The Salk Institute Genome Analysis Laboratory (SIGnAL) continues its progress toward making a set of genetically purified, confirmed T-DNA insertion lines consisting of two alleles for each Arabidopsis gene available for ordering from the Arabidopsis Biological Resource Center (ABRC). When completed, this phenome-ready population will consist of 50,000 lines, the goal being to confirm and purify two insertion alleles each for approximately 25,000 loci. To date, 28,558 confirmed lines have been received by ABRC. This set of completed lines provides two homozygous insertion alleles for 8,517 genes, and a total of 16,856 genes are currently covered by at least one homozygous insertion. A complete set of the lines as well as a one-allele per locus ('unigene') set are being made available from ABRC at an economical price. The confirmed population is also being organized into pools of different sizes to allow efficient forward phenotypic screening for traits that can be identified within larger populations.

ABRC has also recently begun to distribute a collection of approximately 8,000 artificial microRNA (amiRNA) clones developed by Greg Hannon, Dick McCombie, Rob Martienssen, Detlef Weigel and Ravi Sachidanandam. Production of this resource was funded by an Arabidopsis 2010 grant, A Comprehensive Resource for Analysis of Arabidopsis Gene Function (<http://2010.cshl.edu/scripts/main2.pl>), the focus of which was to develop a genome wide resource for RNAi interference in Arabidopsis using amiRNAs with altered targeting capacity that can efficiently and specifically silence any chosen gene. The clones received by ABRC target approximately 8,000 genes, and were previously distributed exclusively by Open Biosystems. The ultimate goal of the project is to develop several clones for each gene and these will be made available through Open Biosystems and ABRC. The clones received to date can be found and ordered through the TAIR web site.

### Tools for Association Mapping

Detlef Weigel, Max Planck Institute:

The 1001 Genomes Project has completed Illumina sequencing of 80 accessions at 6-12x coverage and 4 additional accessions to much higher coverage (<http://1001genomes.org>). This material will soon be available through the stock centers for association mapping, to complement the collection genotyped with the 250k SNP chip by Borevitz and Nordborg. All 250k SNPs are being called in the 84 accessions mentioned above, and they can be seamlessly integrated into any association mapping project that uses the Borevitz/Nordborg collection, or they can be used on their own. With this collection, nearly complete sequence information will be available for any haplotype that comes out of any 250k SNP association mapping project.

### Phenotype annotation tools and ontologies

The Plant Ontology Consortium (<http://www.plantontology.org/>) led by Pankaj Jaiswal (Oregon State University) has received renewed funding from NSF to continue developing the plant structure and growth/developmental stage ontologies for use in describing plant phenotypes. Other useful ontologies for describing plant phenotypes and related data can be found at the Open Biomedical Ontologies website (<http://www.obofoundry.org/>). These include the GO biological process and cellular component ontologies, the PATO ontology for phenotypic qualities, and ontologies for describing units of measurement (UO), experimental conditions (OBI), and chemical entities of biological interest (CHEBI).

Phenote, a curation tool to facilitate annotating phenotypes using ontologies, was developed by National Center for Biomedical Ontology (NCBO) in collaboration with Berkeley Bioinformatics and Ontologies Project (BBOP) ([www.phenote.org](http://www.phenote.org)). It's currently in use by several model organism databases including FlyBase, ZFIN, WormBase and others. Data annotated with Phenote is based on the EQ model for representing phenotypes, combining entities from any ontology (for example GO or PO) with qualities (such as those found in PATO).

### Phenomics community events

The 1st International Plant Phenomics Symposium was held in Canberra on April 22-24, 2009. This symposium, the first of its type, was focused on the use of plant phenomics and functional genomics to boost crop productivity and its proceedings will be published in a special issue of the journal Functional Plant Biology. Symposium participants, led by Jülich Plant Phenomics Centre (JPPC) and the Australian Plant Phenomics Facility (APPF), have begun to organize an International Plant Phenomics Initiative to provide stronger vehicle for international collaboration. Members of this Initiative are currently working to develop an agenda and will organize a meeting later in 2009 to decide on priorities and actions. The agenda will likely include exchanging protocols, validating systems, exchanging staff for technical education and developing collaborative funding bids. Please contact Bob Furbank ([Robert.Furbank@csiro.au](mailto:Robert.Furbank@csiro.au)) or Frank Gilmer ([jppc@fz-juelich.de](mailto:jppc@fz-juelich.de)) for further information on the Initiative.

### High throughput phenotyping projects and data

Bob Furbank, Australian Plant Phenomics project:

The Australian Plant Phenomics Facility (APPF, [www.plantphenomics.org.au](http://www.plantphenomics.org.au)) consists of two facilities, an Arabidopsis screening module in Canberra (medium throughput growth and chlorophyll fluorescence screening with mathematical morphological

analysis and phenomic database capability) and the Plant Accelerator automated glasshouse facility in Adelaide. The NCRIS funded National Facility will be available to researchers at the marginal cost of running the facility and several international collaborations are being established and encouraged. The resulting phenotype data will be freely released to the international community following a quarantine period (6-12 months) to allow data to be prepared for publication. In addition, a new Australian project, Phenomics Ontology Driven Data Management (PODD), has received funding to provide data management services to the APFF and other Australian phenomics projects as well as providing access to phenotype data for the broader community. This project plans to collaborate with TAIR, Gramene, and the Plant Ontology Consortium in the areas of data exchange, data standards and ontology development.

Minami Matsui, RIKEN:

A comprehensive new database called the RIKEN Hub Database (<https://database.riken.jp>) has been developed by Tetsuro Toyoda to provide an integrated access point for RIKEN data. The new database includes a section on activation-tagged lines (<http://activation.psc.database.riken.jp>) which contains phenotype data and insertion site information for 500 activation-tagged lines found to have visible phenotypes (Miki Nakazawa, Youichi Kondou and Eli Kaminuma). A set of Ac/Ds transposon lines including 200 visible phenotypes and flanking sequence information for 18,000 integration sites is described at <http://rapid.psc.database.riken.jp> (Takashi Kuromori and Eli Kaminuma). Another section of the new database (<http://arabifox.psc.database.riken.jp>) contains information on 1,500 visible phenotypes and 9,000 integrated full-length cDNAs for Arabidopsis FOX (Full-length cDNA over-eXpressor gene) lines (Takanari Ichikawa, Youichi Kondou and Eli Kaminuma). RIKEN BRC (Bioresource center; Masatomo Kobayashi) has begun to distribute 1,000 of the Arabidopsis FOX lines, see <http://www.brc.riken.go.jp/lab/epd/Eng/species/arabidopsis.shtml> for ordering information. The "Rice FOX Arabidopsis mutant database" is now freely accessible at <http://amber.gsc.riken.jp/ricefox/>. Information on 11,000 rice full-length cDNAs that have been expressed in Arabidopsis is included, along with the resulting phenotypes which have been categorized into subsets using various criteria.

Phenotype data for 140 Ac/Ds transposon insertion lines (RAPID) generated by Dr. Takashi Kuromori and having visible phenotypes (out of 4000 lines examined) are freely available at <http://range.gsc.riken.jp/phenome/> and their associated phenotype data will be integrated into TAIR. An additional set of 500 activation-tagged lines with phenotypes (generated by Dr. Youichi Kondou) can be found at <http://amber.gsc.riken.jp/act/top.php> (the current requirement to sign a Materials Transfer Agreement before viewing the data is expected to be lifted soon). A new project to generate Arabidopsis Full-length cDNA overexpressing (FOX) lines for 13,000 Arabidopsis full-length cDNAs generated by Dr. Takanari Ichikawa is now underway and these will be made accessible from RIKEN BRC. A similar project for overexpressing rice cDNAs in Arabidopsis can be found at <http://ricefox.psc.riken.jp/login/> but is currently only accessible to Japanese scientists (this restriction will be lifted by the end of 2008). A hub database project has been organized by Dr. Tetsuro Toyoda at RIKEN to connect Arabidopsis genome and phenome information, including Ac/Ds and Activation-tagged line projects described above. See <http://omicspace.riken.jp/> for a full description of the new database and <http://www.psc.riken.go.jp/english/database/index.html> for access to current RIKEN databases.

Pierre Hilson, Christine Granier, AGRON-OMICS project:

The AGRON-OMICS project, which stands for Arabidopsis GROWth Network integrating OMICS technologies (<http://www.agron-omics.eu/>), is conducting an in-depth study of leaf growth in the model plant species *Arabidopsis thaliana*. Started in November 2006, this European Integrated Project aims at the identification of the molecular components controlling growth and a better understanding of their interactions. A major asset of the project is PHENOPSIS, an automated platform for Arabidopsis leaf growth phenotyping developed at INRA (Granier et al., 2006 New Phytologist). Within AGRON-OMICS, the high-throughput phenotyping efforts based on PHENOPSIS focus on the measurements of an exhaustive data sets of leaf growth variables from the rosette scale down to the cellular scale in hundreds of genotypes affected in cell cycle, endoreduplication, cell wall properties, metabolism, hormonal status, circadian rhythm and flowering time. At this time, more than 160 genotypes have been grown and phenotyped in the platform for the AGRON-OMICS project. A database has been developed to organize metadata and phenotypic data associated to the PHENOPSIS platform. It includes data from AGRON-OMICS and from other projects: the ERANET-ARABRAS project, on leaf growth in response to environmental stresses in different accessions (100 lines, 2007-2010) and the GENOPLANTE-DNV project, on the identification of leaf growth QTLs in various recombinant inbred line populations grown in a range of environmental conditions (240 recombinant inbred lines, 2007-2010). Access to the database associated to the PHENOPSIS platform is available online (<http://bioweb.supagro.inra.fr/phenopsis/>) and data stored in the database are made publicly available just after publication. Software applications are also being developed by several groups in the context of AGRON-OMICS to exchange metadata and phenotypic data and to automatically measure the growth parameters defining leaf growth at the macro- and microscopic scale.