

Arabidopsis Community Projects and Resources

The Arabidopsis community has always been very open, so today researchers and funding bodies can look back on more than 20 years of strong international collaboration and data sharing. The efforts of the Arabidopsis community have always been guided by long term plans, which have led to the establishment of many Arabidopsis community projects and resources:

The Arabidopsis genome research project (1990-2001) led to the completion of the Arabidopsis genome. During this decade two out of three stock and resource centers ABRC (Arabidopsis Biological Resource Center, US) and NASC (Nottingham Arabidopsis Stock Center, UK) were founded.

The Multinational Coordinated Arabidopsis Functional Genomics Project (2002-2011) led to the functional annotation of most of the Arabidopsis genes, in that at least providing expression data for them. Alongside, The Arabidopsis Information Resource (TAIR) was founded in 2001 to meet the needs of the growing Arabidopsis research community.

From Bench to Bountiful Harvests (2012-2021) aims to obtain in-depth knowledge of how the genome is translated into a continuum of processes, from the single molecule to cells and tissues, the whole plant, plant populations, and fields of plants, to be able to build a predictive model of an Arabidopsis plant.

The directors of Arabidopsis community projects and resources have been contributing to the MASC reports for several years, by presenting their respective goals, progress and news. Since 2014, general plant projects and resources have also been included, reflecting the growing connections between researchers focussing on different plant species.

Resource and Stock Centers

Arabidopsis Biological Resource Center (ABRC)
(page 24)

Nottingham Arabidopsis Stock Centre (uNASC)
(page 25)

RIKEN BioResource Center (BRC)
(page 26)

The Arabidopsis Information Resource (TAIR)
(page 26)

International Arabidopsis Informatics Consortium (IAIC)
(page 28)

Plant Projects and Resources with Strong Participation of the Arabidopsis Community

Bio-Analytic Resource for Plant Biology (BAR)
(page 29)

International Plant Phenotyping Resources
(page 31)

Gramene
(page 32)

Besides the above listed projects and resources, there are many other international and multinational initiatives with major contributions from Arabidopsis researchers

-The 1001 Arabidopsis Genomes Project
www.1001genomes.org

-The Epigenomics of Plants International Consortium EPIC;
www.plant-epigenome.org

- The Plant and Microbial Metabolomics Resource
<http://metnetdb.org/PMR/>

- The SUBA4: The SubCellular Proteomic Database
<http://suba.plantenergy.uwa.edu.au/>

- 1001 Proteomes
<http://1001proteomes.masc-proteomics.org/>

- AraGWAS
<https://aragwas.1001genomes.org/>

- GENIUS machine learning algorithm to infer functional gene networks
<http://networks.bio.puc.cl/genius>

, Arabidopsis hormone database 2.0
<http://ahd.cbi.pku.edu.cn/>

- PlantGSEA: a gene set enrichment analysis
<http://structuralbiology.cau.edu.cn/PlantGSEA/>

- Arabidopsis Protein Phosphorylation Site Database
<http://phosphat.uni-hohenheim.de/index.html>

- 3DCellAtlas
<http://www.georgebassellab.com/3d-cell-atlas/>



Resource and Stock Centers

The Arabidopsis Biological Resource Center (ABRC)

www.abrc.osu.edu

David Somers, somers.24@osu.edu, ABRC Director (since April 2018)
 Jelena Brkljacic, brkljacic.1@osu.edu, ABRC Associate Director
 Keith Slotkin, slotkin.2@osu.edu, ABRC Director (until April 2018)



Open Tools and Resources for Arabidopsis Researchers

ABRC provides seed, clones, vectors, host strains, antibodies, cell lines and education resources for species ranging from *Arabidopsis thaliana* and other Brassica species to maize and other crops.

Recent activities of your project or resource.

We are pleased to announce that we secured additional support for our operation in 2017 through an NSF RAPID funding program, adding to the current CSBR NSF support (April 2016 – March 2019).

Last year we expanded and diversified the ABRC's collection. Within Arabidopsis, we acquired a number of cutting-edge resources such as a set of HALO-tagged transcription factors (TFs) for identification of DNA binding regions, CrY2H vectors and host strains for high-throughput yeast two hybrid assays by deep sequencing, and a comprehensive toolkit comprising 162 vectors for advanced genome engineering via CRISPR/Cas9 mutagenesis.

In addition to Arabidopsis resources, the ABRC has continued to diversify our collection through the acquisition of a soybean TF clone set, seeds from diverse Brassica species, and a comprehensive toolkit for CRISPR/Cas9 across monocots and dicots.

The ABRC has continued to expand and enhance our education and outreach program, reaching out to larger and more diverse audiences. Our education kits using Arabidopsis continue to be highly ordered and most of the kits are available free of charge. Two education papers were published (see Additional Information) and a revamped education website is about to be released (see Future Activities), making this a very good year for our education and outreach program.

Planned future activities of your project or resource.

For the past 17 years, The Arabidopsis Information Resource (TAIR) has provided services for searching and ordering ABRC stocks. For the past two years, the ABRC has been collaborating with the College of Arts and Sciences Technology Services (ASC Tech) at Ohio State to design an independent ABRC database and ordering system. The ABRC database provides superior searching capabilities and access to complete stock information, including quality control (QC) and additional features. In addition, the new system is much more flexible and provides full access to a growing number of the ABRC's non-Arabidopsis resources. The ABRC database will be released for testing in the spring of 2018. The ABRC stock ordering will still be available through TAIR during the testing stage. After successful testing of the new database, ABRC stocks will only be available for ordering through our system, and no longer through TAIR.

We have recently received a donation of Brassica species as part of a collection of "Top 20 Brassicales" sequenced through the Brassicales Map Alignment Project (BMAP) sponsored by the U.S. Department of Energy Joint Genome Institute (DOE JGI). The propagation and distribution of this resource is planned for 2018.

More bioenergy-related and other Brassica species, including 100 *Brassica rapa* stocks and a GWAS population of ~500 sequenced accessions of *Boechea stricta*, are expected in the near future. The ABRC plans to continue serving a wide Brassica community by providing stock propagation, maintenance and QC. In addition, our new database with a user-friendly online search tool will allow users to search exclusively for Brassica stocks, making these resources much easier to find and order.

Conferences, Workshops and Training events

The ABRC shared a booth with The Nottingham Arabidopsis Stock Centre (NASC) and participated in the bioinformatics resources workshop at the International Conference on Arabidopsis Research (ICAR) in St. Louis in 2017. The ABRC also had a booth at the annual Plant Biology meeting of the American Society of Plant Biologists (ASPB) in Hawaii. In 2018, ABRC participated in the Plant and Animal Genome conference in San Diego, by presenting a talk focused on the ABRC education program and by promoting donation and distribution of other species' resources. Current conference plans include ICAR, ASPB and the International Plant Molecular Biology (IPMB) meeting in the summer of 2018.

Additional Information

The ABRC published a methods paper (Lindsey B, Rivero L, Calhoun C, Grotewold E, and Brkljacic J (2017) Standardized Method for High-throughput Sterilization of Arabidopsis Seeds. *J Vis Exp* E56587, doi:10.3791/56587) and an education paper (Mann J, Larson J, Pomeranz M, Knee E, Shin D, Miller J, Price C, Crist D, Grotewold E, and Brkljacic J (2017) Life in Bloom: Using Arabidopsis Mutants to Study Plant Development. *Course Source*, <https://www.coursesource.org>) last year and one this year (Price C, Knee E, Miller J, Shin D, Mann J, Crist D, Grotewold E, and Brkljacic J (2018) Following Phenotypes: An Exploration of Mendelian Genetics Using Arabidopsis Plants. *American Biology Teacher*, doi: 10.1525/abt.2018.80.4.291).

We are very grateful to Dr. Keith Slotkin for his service as an ABRC Director and welcome Dr. David Somers, an esteemed OSU Professor, as our new Director.



The Nottingham Arabidopsis Stock Centre (uNASC)

<http://www.Arabidopsis.info>

Sean May
 Director
Sean.May@nottingham.ac.uk



Marcos Castellanos-Uribe
 Operations Manager.
marco.castellanos@nottingham.ac.uk



Another productive year sending over 150,000 tubes of seed to the worldwide plant community sees a slight change in the international profile of our users. Out of a total of 43 countries ordering from NASC during 2017, the top five countries (in order by individual stock numbers) are China (No.1 for the third year in a row), the UK, Australia (first time in the top 5), Germany, and France.

For entertainment purposes, we shared (@NascArabidopsis - Jan 2018) a map of locations in China and the UK (by institute) that had ordered seed from us in 2017.

Of course, our broader catalogue access maps for all countries can always be found at <http://clustrmaps.com/map/Arabidopsis.info>.

For up-to-date details on stock donations or anything else that you wish to know, please do visit the NASC site, or contact curators@arabidopsis.info at any time.

If we (NASC and ABRC) can save you time, effort and promote your research impact by distributing seed on your behalf to the wider plant community then please do holler (or email) - don't wait for us to come to you.

See you in Turku!



RIKEN BioResource Center (RIKEN BRC)

<http://epd.brc.riken.jp/en/>

Masatomo Kobayashi
(RIKEN coordinator)



Open Tools and Resources for Arabidopsis Researchers

Seeds of Arabidopsis (natural accessions, mutants, and transgenic lines)
Plant DNA materials including Arabidopsis
Plant cultured cell lines including Arabidopsis T87 cell
<http://epd.brc.riken.jp/en/>

Recent activities

The Japanese government launched the fourth term of the five-year National BioResource Project (NBRP) in April 2017 (<http://www.nbrp.jp/index.jsp>). RIKEN BRC was selected as the core facility of Arabidopsis, cultured plant cells and genes. We are reconstructing the web catalog and accumulating useful information for research communities.

Planned future activities

RIKEN starts the fourth 7-year term plan from April 2018. During this period, RIKEN BRC enriches biological materials, technologies and information which are crucial requirements for both basic and advanced researches. Revision of the Bioresource Frontier Programs contributes the global issues such as human health, food supply and environmental protection. In the revised Program, Plant-Microbe Symbiosis Research and Development Team engages in the isolation, characterization and utilization of rhizosphere microbes to innovate agricultural technologies.

Conferences, Workshops and Training events

The 29th International Conference on Arabidopsis Research (ICAR2018), June 25-29, 2018, Turku, Finland
The 9th International Meeting of Asian Network of Research Resource Centers (ANRRC), Sept. 20-22, 2017, Beijing, China
The 10th International Meeting of Asian Network of Research Resource Centers (ANRRC), Sept. 5-7, 2018, Seoul, Korea

Additional Information

RIKEN BioResource Center is renamed to RIKEN BioResource Research Center to increase the R & D activities on biological resources.



Arabidopsis Informatics and Data Sharing Resources

The Arabidopsis Information Resource (TAIR)

www.Arabidopsis.org

Leonore Reiser (TAIR curator)
Eva Huala (Director)



Activities

The Arabidopsis Information Resource (TAIR) is a continuously updated, manually curated, online database of genetic and molecular biology data for the model plant *Arabidopsis thaliana*. TAIR's biocurators extract, organize, and interconnect experimental data with computational predictions, community submissions, and high throughput datasets to provide a 'gold standard' annotation for this important reference genome. TAIR provides tools to visualize and analyze data and enables online ordering of seed and DNA stocks, protein chips and other experimental resources from the Arabidopsis Biological Resource Consortium (ABRC).

These activities are made possible by support from the international Arabidopsis and plant research community whose contributions ensure that TAIR remains available and an up to date resource for plant biologists around the world.

New Tool: Gene Ontology (GO) Term Enrichment

One of the most common applications of the Gene Ontology is for term enrichment; finding groups of terms that are over or under-represented in a gene set. The accuracy of enrichment results is, to a large extent, dependent upon the underlying data sets; specifically, the annotations and ontologies. Changes to the Arabidopsis annotations are frequent; new annotations are added on a weekly basis and occasionally annotations are deleted. Changes to ontologies include addition of new terms or pruning of nodes. It is therefore critical that term enrichment tools use current annotation and ontology datasets to produce the most reliable results. To assist the Arabidopsis community in performing term enrichment, we have added the GO Term Enrichment tool developed by PANTHER (http://www.arabidopsis.org/tools/go_term_enrichment.jsp). The tool takes a list of gene IDs (e.g. AGI locus codes, UniProt or GenBank IDs) and sends the data to the PANTHER Classification System (www.pantherdb.org). PANTHER GO term enrichment data sources are updated monthly.

In addition to performing term enrichment for Arabidopsis, the tool can also be used to perform enrichment on 9 other species included in the PANTHER database: Brachypodium, Chlamydomonas, soybean, rice, Physcomitrella, poplar, tomato, sorghum and grape.

Genome and Gene Function Curation

A major goal of TAIR curation is to present the most up to date and accurate picture of the Arabidopsis genome. All gene models, assignments, computational descriptions and sequences in TAIR have been updated to Araport11. In addition, we have updated and added thousands of T-DNA polymorphisms using genome coordinates and locus assignments provided by the Ecker Lab. Many TAIR tools (e.g. gene descriptions, sequences, locus history) now use Araport11 as the underlying data source and the remaining tools will follow. The addition of new gene models has increased the pool of 'unknown' or 'unannotated' genes slightly, as reflected in the updated genome snapshot (http://www.arabidopsis.org/portals/genAnnotation/genome_snapshot.jsp). Identifying and extracting knowledge about these two classes of gene products continues to be a major activity at TAIR.

TAIR continuously updates information about Arabidopsis gene function, expression and mutant phenotypes using data extracted from recent literature. When choosing articles TAIR curators prioritize those that allow us to add functional information about genes that had not previously been described in the literature ('unknown' genes) thus increasing the breadth of functional annotation. We capture this experimentally derived knowledge in the form of GO and Plant Ontology (PO) annotations, individually composed gene summaries and phenotype descriptions, and new links between articles and genes, which are added to TAIR on a weekly basis.

Between March 2017 and February 2018, we added 662 new gene symbols to TAIR and added or updated 701 gene summaries. We added 7376 new GO and PO annotations, generated by TAIR curators, TAIR community, UniProt and the GO consortium, to 1907 loci, incorporating experiment-based data from 943 research articles. We linked 2571 articles to 4261 loci, added or updated descriptions for 147 alleles, and added or updated 146 phenotypes from the literature. In 2017 we processed 607 GO/PO annotations contributed by 86 community members covering 99 publications.

Subscriptions and Complementary Access

TAIR is supported by a global community of researchers from over 37 countries. As of March 2018, subscribers include: 2 countries (China and Switzerland), 5 academic consortia, 208 academic institutions (list at <http://bit.ly/1RPlaau>), 10 companies and 324 individuals. We continue to provide complementary access for teaching

to instructors who request access using our simple form (<http://bit.ly/2fA5WR1>) and to low GDP countries upon request.

Data Releases

In accordance with our data sharing policy, TAIR continues to publish year-old data in the form of quarterly data releases (http://www.arabidopsis.org/download/index-auto.jsp?dir=/download_files/Public_Data_Releases). These data are freely available and distributed for reuse under the CC-BY4.0 license. They include all GO and PO annotations, gene symbols and names, links between genes and publications, and updated gene descriptions added to TAIR through 12/31/2016. The most current data and tools are available through the website, which is updated on a weekly basis. Anyone can access a limited number of pages per month; unlimited access requires a subscription. Subscribers can also access current data in the form of quarterly releases (http://www.arabidopsis.org/download/index-auto.jsp?dir=/download_files/Subscriber_Data_Releases) and can request custom data sets via our helpdesk (curator@arabidopsis.org).

Outreach: Workshops, Conferences, Presentations and Social Media

In the past year, TAIR has pioneered efforts aimed at increasing community engagement with curation, and training in data management. We organized the first annual 'Annotation Jamboree' at the 2018 Plant and Animal Genome meeting, that focused on gene function annotation using ontologies. Slides from the presentation are available (<http://bit.ly/phoenix-pag-2018>).

To promote Findable, Accessible, Interoperable, and Reusable (FAIR) data publication in plant sciences, we co-presented a seminar on Data Management and Publishing at UC Berkeley with curators from MaizeGDB. Slides from that presentation are available at <http://bit.ly/2pkrlj5>. TAIR is also a member of the newly formed AgBioData consortium of Agricultural Databases (www.agbiodata.org), which aims to develop and promote standards and best practices for the acquisition, display and retrieval of genetic, genomic and breeding data. We plan to continue to offer similar workshops and seminars in the coming year. TAIR staff presented posters, gave talks, and were available for one-on-one interactions at exhibit booths at the following meetings: ICAR 2017 (St. Louis, MO, USA), ASPB 2017 (Honolulu, HI, USA), and PAG 2018 (San Diego, CA, USA). TAIR curators will attend ICAR 2018 in Turku, Finland and ASPB 2018 in Montreal, Canada.

TAIR curators are also available to assist the community through our HelpDesk (curator@arabidopsis.org), and through social media on Facebook (<https://www.facebook.com/tairnews>), Twitter (https://twitter.com/tair_news), and YouTube (<https://www.youtube.com/user/TAIRinfo>).

New User Guide

We published the first major revision of our chapter in Current Protocols in Bioinformatics since 2010. Many sections received significant updates and new protocols were added including: GO Term Enrichment, Gene List Analysis with Bulk Tools, How to submit GO/PO annotations using TAIR's Online Annotation Submission Tool, and Using the Recently Added Literature Browser. The guide is available on line from Current Protocols (doi: 10.1002/cpbi.36; subscription required) or in draft form from TAIR (<http://bit.ly/2Go2wuP>)

Publications

Reiser, L., Subramaniam, S., Li, D., & Huala, E. (2017). Using the Arabidopsis information resource (TAIR) to find information about Arabidopsis genes. Current Protocols in Bioinformatics, 60, 1.11.1–1.11.45. doi: 10.1002/cpbi.36



International Arabidopsis Informatics Consortium (IAIC)

Blake C. Meyers
IAIC Director, BMeyers@danforthcenter.org,
The Donald Danforth Plant Science Center
Joanna D. Friesner
IAIC Assistant, jdfriesner@ucdavis.edu,
North American Arabidopsis Steering
Committee



Open Tools and Resources for Arabidopsis Researchers

The IAIC was formed in 2010 to facilitate a coordinated global Arabidopsis informatics effort to maintain the continuity of key Arabidopsis resources while simultaneously expanding their breadth and depth. Key aims were to include in the IAIC new technologies, resources, and participants on a global scale and advance plant biology while creating novel opportunities for research and education and strengthening international collaborations. The primary accomplishment of IAIC with respect to globally-available resources was to facilitate the collaborative effort to establish and fund a new web-based portal, 'Araport', for Arabidopsis information and digital resources for the global plant biology community via NSF

Award #1262414: ABI Development: The Arabidopsis Information Portal. (www.nsf.gov/awardsearch/showAward?AWD_ID=1262414)

Araport enables biologists to navigate from the *Arabidopsis thaliana* Col-0 reference genome sequence to its associated annotation including gene structure, gene expression, protein function, and interaction networks. Araport offers a single interface through which to access a wide range of Arabidopsis information. Araport will grow through contributions of other labs in the form of modules: data, computation, and visualization tools. (<https://www.araport.org/about>)

Recent activities of your project or resource.

Recent IAIC efforts have focused convening relevant attendees in a workshop that will assess the current status of Arabidopsis informatics and the needs and wishes of the Arabidopsis community, looking forward to the next 5-10 years. The 'Future of Arabidopsis Bioinformatics' workshop will build on the project's previous success in enabling the establishment of 'Araport', the Arabidopsis community information resource funded by NSF, which has developed into a well-established portal with a broad range of data sets and functionalities. During the 4+ years since Araport went online the research community continued to generate increasing amounts of data that are challenging to store, manipulate and disseminate.

The broad objectives of the May 2018 IAIC workshop include to: discuss and make plans/recommendations on the directions and sustainability of Arabidopsis Informatics in the coming decade; present new resources inspired by design principles developed at the first project workshop (Dec. 2011); identify informatics challenges that impede discovery and possible community-based solutions for these challenges; brainstorm a list of new tools the community would like to have or existing tools we would like to see expanded; identify stable central data and tool platforms within the community, evaluate how to promote their long-term stability with funding entities, and identify new platforms to facilitate the ever-broadening scope of genome-wide approach data.

Workshop organizers plan to collaboratively develop a white paper publication with outcomes, analysis, and recommendations to disseminate broadly.

Conferences, Workshops and Training events

The project's PI and Araport/AIP PI gave a presentation on aspects of Araport and IAIC efforts entitled "Arabidopsis Informatics" at the 2018 Plant and Animal Genome (PAG) conference.

Plant Projects and Resources with Strong Participation of Arabidopsis Community

Bio-Analytic Resource for Plant Biology (BAR)



Nicholas Provart (Director)
<http://bar.utoronto.ca>

Open Tools and Resources for Arabidopsis Researchers

The Bio-Analytic Resource is a collection of user-friendly web-based tools for working with functional genomics and other data for hypothesis generation and confirmation. Most are designed with the plant (mainly Arabidopsis) researcher in mind. Data sets include:

- * 150 million gene expression measurements (75 million from A.th.), plus "expressologs" (homologs showing similar patterns of expression in equivalent tissues) for many genes across 10 species. View expression patterns with our popular eFP Browser or new ePlant tool.
- * 70,944 predicted protein-protein interactions plus 36,352 documented PPIs (rice interologs also available!) and ~2.8 million protein-DNA interactions.
- * 29,180 predicted protein tertiary structures and experimentally-determined structures for 402 Arabidopsis proteins.
- * Millions of non-synonymous SNPs from the 1001 Arabidopsis Genomes project, delivered through the MASC Proteomics Subcommittee's site at 1001proteomes.masc-proteomics.org.
- * Documented subcellular localizations for 9.3k proteins, predicted localization for most of Arabidopsis proteome, from the SUBA database at the University of Western Australia.

Recent activities of your project or resource.

Work was completed on a "mega-app" from the BAR for Araport.org, ePlant, and a publication came out in the Plant Cell describing it (Waese *et al.*, 2017). ePlant is a zoomable user interface for exploring data across 12 orders of magnitude, from the kilometre scale down to the nanometre scale, encompassing at least 20 different kinds of genome-wide data covering gene expression, subcellular localization, 3D molecular structure, chromatin states, and many others, all in one easy-to-use, extensively tested, open source tool. ePlant is accessible from the Araport.org homepage, and from the Bio-Analytic Resource at <http://bar.utoronto.ca/eplant>.

The first major update to ePlant was released in January 2018, and included a new Plant eFP view for Klepikova *et al.*'s (2016) RNA-seq-based developmental atlas, updated subcellular localization data using SUBA4 data, an ePlant Navigator to easily view expression patterns of homologs in other species, and the highlighting of gene parts in the Gene view.

We also released an "eFP-Seq Browser" for exploring the 113 RNA-seq data sets used to reannotate the Arabidopsis genome for the Araport11 build, or the RNA-seq data sets generated for Klepikova *et al.*'s (2016) developmental atlas, both as read maps and pictographs summarizing the expression levels in various tissues, easily sortable by any desired column: http://bar.utoronto.ca/eFP-Seq_Browser/. In addition to simply sorting the table by expression level, an "eFP overview" provides a table of pictographs to get a feeling for where expression is the strongest or weakest.

We released hundreds of new protein-protein interactions for the extra-cellular domains of leucine-rich repeat receptor kinases into our ePlant and Arabidopsis Interaction Viewer tools (Smakowska-Luzan *et al.*, 2018).

For translational researchers, publications came out for eFP Browser views announced last year: a Developmental Map for Brachypodium distachyon from Marek Mutwil and colleagues (Sibout *et al.*, 2017; http://bar.utoronto.ca/efp_brachypodium/), a Maize Root view, part of the Maize eFP Browser, from Stefan Hey and colleagues (Hey *et al.*, 2017; http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi), and a Classification SuperViewer for Medicago truncatula genomics (Herrbach *et al.*, 2017; http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer_medicago.cgi).

BAR Publications

Herrbach V, Chirinos X, Rengel D, Agbevenou K, Vincent R, Pateyron S, Huguet S, Balzergue S, Pasha A, Provart NJ, Gough C, Bensmihen S (2017). Nod factors potentiate auxin signaling for transcriptional regulation and lateral root formation in Medicago truncatula. *Journal of Experimental Botany* 68: 569-583.

Hey S, Baldauf J, Opitz N, Lithio A, Pasha A, Provart NJ, Nettleton D, Hochholdinger F (2017). Complexity and specificity of the maize (*Zea mays* L.) root hair transcriptome. *Journal of Experimental Botany* 68: 2175-2185.

Sibout R, Proost S, Oest Hansen B, Vaid N, Giorgi FM, Ho-Yue-Kuang S, Legée F, Cézart L, Bouchabké-Coussa O, Soulhat C, Provart N, Pasha A, Le Bris P, Roujol D, Hofte H, Jamet E, Lapierre C, Persson S, Mutwil M (2017). Expression atlas and comparative coexpression network analyses reveal important genes involved in the formation of lignified cell wall in Brachypodium distachyon. *New Phytologist* 215: 1009-1025.

Smakowska-Luzan E, Mott GA, Parys K, Stegmann M, Howton TC, Layeghifard M, Neuhold J, Lehner A, Kong J, Grünwald K, Weinberger N, Satbhai SB, Mayer D, Busch W, Madalinski M, Stolt-Bergner P, Provart NJ, Mukhtar MS, Zipfel C, Desveaux D, Guttman DS, Belkadir Y (2018). An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature* 553: 342-346. doi:10.1038/nature25184.

Waese J, Fan J, Pasha A, Yu H, Fucile G, Shi R, Cumming M, Kelley LA, Sternberg MJ, Krishnakumar V, Ferlanti E, Miller J, Town C, Stuerzlinger W, Provart NJ (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* 29: 1806-1821. doi: 10.1105/tpc.17.00073.

Waese J, Provart NJ (2017). The Bio-Analytic Resource for Plant Biology. Book chapter in *Plant Genomics Databases*, Humana Press, pp. 119-148.

Planned future activities of your project or resource.

A custom eFP view in ePlant for researcher's own RNA-seq data is planned, as part of a new application to Genome Canada. Stay tuned!

Conferences, Workshops and Training events

The BAR participated in the 2017 American Society of Plant Biology (ASPB) Plant Biology conference in Hawaii, as part of the Plant AgData Outreach booth; and Plant and Animal Genomes (PAG) XXVI at the start of 2018 in San Diego, California. It also gave a talk in the Arabidopsis Informatics workshop at ICAR2017 in St. Louis, Missouri.

International Plant Phenotyping Resources

Roland Pieruschka
Julich Plant Phenotyping Centre
r.pieruschka@fz-juelich.de



International Plant Phenotyping Network

IPPN <http://www.plant-phenotyping.org/>

IPPN is a non-profit association that represents the major plant phenotyping centers worldwide. The association aims to provide all relevant information about plant phenotyping. The goal is to increase the visibility and impact of plant phenotyping and enable cooperation by fostering communication between stakeholders in academia, industry, government, and the public. Within the last year, IPPN has established its basic operation by setting up a number of working groups on relevant topics, organizing workshops, meetings, and international symposia.

The next IPPN Symposium will be held in October 2-5th in Adelaide Australia (<http://www.ipps2018.com.au/>).

With the increasing development of a global plant phenotyping community it is also essential to extend the IPPN membership base by implementing and fostering close interaction between academia and industry partners to drive and develop innovation that delivers to academic and industry users. Thus, the IPPN General Assembly has decided to include members from industry and by mid-2018, members from industry can join the association.

EMPHASIS

<https://emphasis.plant-phenotyping.eu/>

The European Strategy Forum for Research Infrastructure (ESFRI) has identified "Plant Phenotyping" as a priority for the European research area and the project EMPHASIS has been listed on the ESFRI Roadmap as an infrastructure project to develop and implement a pan-European plant phenotyping infrastructure.

The EMPHASIS project (European Infrastructure for Multi-Scale Plant Phenotyping And Simulation for Food Security in a Changing Climate) coordinated by the Institute of Bio- and Geosciences, IBG-2: Plant Sciences at the Forschungszentrum Jülich. In 2017 EMPHASIS has started the Preparatory Phase which is a 4 year H2020 funded project. Within the Preparatory Phase, EMPHASIS will develop a business plan embedded in a sound legal framework for a long term operation of EMPHASIS. One of the major activities in preparation of the business plan consisted in a mapping of the plant phenotyping landscape in Europe to foster the information exchange about infrastructures, major



stakeholders, main projects and networks. Additionally, close interaction between national plant phenotyping communities was established and about 23 national communities have already mandated a representative, to form a link to EMPHASIS by contributing to the so-called "Support Group" and shape the development of EMPHASIS.

EMPHASIS has also initiated a close interaction with the ESFRI project AnaEE (<https://www.anaee.com/>) which offers access to experimental and modelling platforms as well as data on terrestrial and aquatic ecosystems. The activities of both projects and areas of cooperation have been outlined in a recent publication by: Roy *et al* (2017) European infrastructures for sustainable agriculture, *Nature Plants* 3, 756-758.

European Plant Phenotyping Network 2020

EPPN2020, <https://eppn2020.plant-phenotyping.eu/>

The EPPN2020 is a H2020 funded research infrastructure project that will provide European public and private scientific sectors with access to a wide range of state-of-the-art plant phenotyping facilities, techniques and methods, and help boost the exploitation of genetic and genomic resources available for crop improvement that represents a major scientific challenge for this coming decade. EPPN2020 specifically aims to facilitate the community progressing across the whole phenotyping pipeline, involving sensors and imaging techniques, data analysis in relation to environmental conditions, data organization and storage, data interpretation in a biological context and meta-analyses of experiments carried out on different organs at different scales of plant organization.

Specifically EPPN2020 includes:

- i) access to 31 key plant phenotyping installations;
- ii) Joint research activities to develop: novel technologies and methods for environmental and plant measurements, tools for statistical analysis and a European plant phenotyping information systems;
- iii) networking activities to increase integration between phenotyping facilities and users within and outside of EPPN2020.

Within the first year, EPPN2020 initiated a first call for access, which resulted in 33 application. The second call for access was issued in early February 2018 with an application deadline in June 2018. New calls will be released every 6 months.

Gramene: A comparative genomics and pathways resource for plants

Marcela Karey Tello-Ruiz (Project Manager)
mkmonaco.cshl@gmail.com
 Doreen Ware (PI)
<http://www.gramene.org>



Open Tools and Resources for Arabidopsis Researchers



Gramene provides open access to comparative genomics and pathways data, and analysis tools. We host 53 reference genomes including *Arabidopsis thaliana*, *A. lyrata*, Brassicas, sugar beet, Solanaceae, cereals, fruits, and basal plants, each displaying functional gene annotations, genetic and structural variation, gene expression, gene trees with orthologous and paralogous gene classification, whole-genome alignments, and synteny maps. In addition, we host 264 pathways curated in rice and inferred in 74 additional plant species (including the above) by orthology projection. Its integrated search capabilities and interactive views facilitate visualizing gene features, gene neighborhoods, phylogenetic trees, gene expression profiles, pathways, and cross-references. Our tools support comparative analyses of our data as well as data brought in by our users, and include a BLAST/BLAT sequence aligner, a sequence assembly converter, a genetic variant effect predictor, an advanced BioMart-based query interface, data analysis and visualization of OMICS data, and multi-species pathway comparisons. Together these comparative data and resources enable powerful cross-species comparisons.

Recent activities of your project or resource.

BACKGROUND - The Gramene database (<http://www.gramene.org>) is an integrated resource for comparative genome and functional analysis in plants. The database provides the Arabidopsis community, agricultural researchers, and plant breeders access to reference genomes, comparative resources including whole-genome DNA alignments and protein-based gene trees for 53 crops and model species including *A. thaliana*, *A. lyrata*, Brassicas, sugar beet, Solanaceae, cereals, fruits, and lower plants. In addition to 264 pathways curated in rice and inferred in 74 additional plant species (including the above), thus enabling powerful cross-species comparisons.]

The Gramene project has had five data releases since January 2017. Of significance for the Arabidopsis community are 7 new dicot genomes: jute (*Corchorus capsularis*), cucumber (*Cucumis sativus*), cotton (*Gossypium raimondii*), common sunflower (*Helianthus annuus*), blue lupin (*Lupinus angustifolius*), cassava

(*Manihot esculenta*), coyote tobacco (*Nicotiana attenuata*), and common bean (*Phaseolus vulgaris*), as well as one monocot genome: white Guinea yam (*Dioscorea rotundata*). We also released significant updates to the genome assemblies of three major crops: sorghum (V3), soybean (V3), and peach (V2), and updated gene models for maize, barley, and Japonica rice.

BACKGROUND - Our phylogenetic gene trees classifying orthologous and paralogous relationships, provide evolutionary history and insights into speciation and duplication events. These gene trees provide the basis that informs synteny maps, which enable inter-species browsing across ancestral regions. In addition to the protein-based gene trees, our WGA mappings provide a valuable resource to characterize non-coding conservation. The analyses for multiple species can be viewed simultaneously with links showing homologous genes and WGA mappings.

Our existing whole-genome DNA alignments (WGA) collection was built against *A. thaliana* (dicot model crop; 34 alignments) and *Oryza sativa Japonica* (monocot staple food crop; 44 alignments). Currently, our synteny collection includes synteny maps for *A. thaliana* against 14 species: *A. lyrata*, cucumber, soybean, cotton, sunflower, lupin, cassava, tobacco, Japonica rice, common bean, poplar, peach, sorghum, and grape. SNP and structural diversity data, including individual genotypes, are available for *A. thaliana*, and another 11 species. The SNPs are displayed in the context of gene annotation and protein domain structure, along with predicted functional consequences (e.g. missense variant). Gramene hosts several variation data sets for Arabidopsis.

In the past year, we included the full data set (12.9 M SNPs) from the 1001 Arabidopsis Genomes Project, and linked the variants in 95 inbred lines to 107 existing phenotypes from the GWAS study by Atwell *et al* (2010). Genotypes from 1,179 strains from the 2012 study by Horton and collaborators using the Affymetrix 250k Arabidopsis SNP chip, and an updated data set produced through a BBSRC funded multi-institutional collaboration involving resequencing 18 Arabidopsis lines published by Clark *et al* in 2007 were obsolete and archived.

For supporting Arabidopsis researchers, the annotation data hosted at Gramene was integrated with the automated GWAS workflow provided by SciApps (collaboration with the CyVerse project), a ready-to-use workflow platfor, where a list of nearby genes can be automatically retrieved for the most significant associated loci.

In collaboration with the Expression Atlas project (EMBL-EBI), we provide experimental baseline expression data for 23 plant species including *A. thaliana* (185 experiments in various ecotypes, tissues and developmental stages) and *A. lyrata* (3 experiments in 3 tissues), through our Ensembl genome browser and Plant Reactome pathways interfaces.

In addition, we provide direct links to differential gene expression data on the Atlas website for 519 experiments in *A. thaliana*, and 1 in *A. lyrata*, among a total of 720 experiments in 26 plant species. Through Gramene's Plant Reactome, we now offer 264 curated rice pathways, as well as orthology-based pathway projections to 74 plant species including both, *A. thaliana* and *A. lyrata*.

In the past year, our integrated search database and modern user interface (<http://search.gramene.org>) were greatly enhanced to leverage the diverse annotations above described to facilitate finding genes through selecting auto-suggested filters with interactive views of the results. We now offer detailed views for search results featuring genomic location, gene trees and homologs lists, gene expression, and cross-references, while a pathways view is in development.

Our Blast and BioMart interfaces enable complex queries of sequence, annotation, homology, and variation data. The genomes portal of the Gramene project is developed on the Ensembl infrastructure and in collaboration with the Ensembl Genomes project (EMBL-EBI). The pathways portal of the project is the Plant Reactome (<http://plantreactome.gramene.org>). The project supports application programming interfaces to access the data, in addition to the graphical user interface, which supports visual displays, and download of the data in standardized formats for the genome and pathway data, as well as high-resolution, publication-ready, image files.

As we aim to reach out to a larger audience, we continue our monthly webinar series and invite the arab-gen@net.bio.net mailing list whenever topics relevant to the Arabidopsis community are covered. Recorded webinars are available for public view from Gramene's YouTube channel <https://www.youtube.com/channel/UCMtmq20XMccsNUaACuqQJ-w>

Planned future activities

We will continue with activities similar as those described above, and which aim to: 1) improve and expand our reference data collection of plant genomes and standardized comparative annotations, 2) enrich our Plant Reactome pathways data resource, 3) integrate visualization and analysis tools for exploring emerging genomic and pathway data, and 4) transform the community through communication and training opportunities.

Conferences, Workshops and Training events

In the past year, Gramene participated in 16 scientific conferences to present talks, posters, training workshops, hand out brochures and other educational materials. We organized community outreach booths for members of the AgBioData consortium at the Plant and Animal Genomes and Plant Biology conferences. Continued to broadcasted live webinars, and made the video-

recordings available in Gramene's YouTube channel. In collaboration with CyVerse, we organized a bioinformatics workshop at the CSHL Plant Genomes and Biotechnology meeting and together with the MaizeCode project, organized a maize annotation jamboree. Other plant education activities geared to K-12 students and faculty included a STEM summer camp, a DNA workshop for Science Olympics participants, and hands-on activities to celebrate Fascination of Plants Day. We plan on continuing our outreach, education and training activities in the next year.

