

Reports of the MASC Subcommittees

Bioinformatics

<http://Arabidopsisresearch.org/index.php/subcommittees/bioinformatics>



Compiled by Nicholas Provart with input from MASC Bioinformatics Subcommittee members and the wider Arabidopsis community.

Araport.org: From not existing a couple of years ago, Araport has become one of the most popular sources for Arabidopsis researchers according to a survey on Arabidopsis bioinformatics conducted by MASC in April 2018. It offers a multitude of ways to query various genomic data sets from Arabidopsis, both for researchers accessing it via its detailed web pages and programmatically. It is seeking renewed funding this year through the U.S. National Science Foundation.

TAIR: TAIR staff have continued updating Arabidopsis gene data in TAIR from published literature over the past year. Between March 2017 and February 2018, TAIR added 662 new gene symbols to TAIR and added or updated 701 gene summaries. A total of 7376 new GO and PO annotations, generated by TAIR curators, TAIR community, UniProt and the GO consortium, were added to 1907 loci, incorporating experiment-based data from 943 research articles. TAIR linked 2571 articles to 4261 loci, added or updated descriptions for 147 alleles, and added or updated 146 phenotypes from the literature. In 2017 TAIR processed 607 GO/PO annotations contributed by 86 community members covering 99 publications. The recently updated TAIR user guide (<http://bit.ly/2Go2wuP>) has detailed protocols for using TAIR's search, browse and analysis capabilities to access these data. Finally, TAIR continues to release year old data with no restrictions to the public. The 14th public release from TAIR contains recent annotation data up to March 31, 2017. Educators can continue to request access to the "full" version of TAIR for teaching purposes.

DNA and RNA resources: The Bailey-Serres Lab and collaborators used a combination of mRNA-seq and ribo-seq to identify vast and selective translational regulation in Arabidopsis in response to phosphate starvation (Bazin *et al.*, 2017). Dolf Weijer's Lab published an "atlas" of early embryo development based on the INTACT method for isolating cell-type-specific nuclear RNA, hybridized to Arabidopsis Gene 1.1 ST microarrays (Palovaara *et al.*, 2017). The Searle Lab reported a transcriptome-wide study mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and ncRNAs (David *et al.*, 2017). The Yamamoto Lab used pair-end sequencing of transcription start site tags

(TSS tags) to identify 324,461 promoters in Arabidopsis, including many not associated with existing gene models (Tokizawa *et al.*, 2017). The Vandepoele Lab reported differential transcript usage events – related to alternative splicing – for 8,148 Arabidopsis genes across 206 public RNA-seq data sets, with protein sequences being changed in 22% of the cases (Vanechoutte *et al.*, 2017). The Van de Peer Lab used gene pairs resulting from whole genome duplication events to show a "remarkable" 92 homoeologous gene pairs with similar patterns of tissue-specific gene expression where one homoeolog exhibits more expression in aerial tissues and the other homoeolog exhibits more expression in tip-growth tissues (Smet *et al.*, 2017).

Gene regulatory networks/codes: Roger Deal's group used ATAC-seq to interrogate chromatin accessibility in stem cells and mesophyll cells to help delineate cell-type-specific gene regulatory networks acting in each (Sijacic *et al.*, 2018). The Muday Lab identified transcriptional and receptor networks controlling root responses to ethylene (Harkey *et al.*, 2018). Several groups published methods for inferring or exploring gene regulatory networks in Arabidopsis including GRACE (Banf and Rhee, 2017), GENIST (Balaguer *et al.*, 2017), TF2Network (Kulkarni *et al.*, 2018), Expresso (Aghamirzaie *et al.*, 2017), and SeqEnrich (Becker *et al.*, 2017). The Wigge Lab published a paper describing the G-box regulatory code in Arabidopsis to identify the set of bZIP and bHLH transcription factors that bind to G-boxes (CACGTG) that are most predictive of expression of genes with G-boxes in their promoters (Ezer *et al.*, 2017).

Other "big data": The 1001 (Arabidopsis) Genomes Consortium continues to make leveraging the power of natural variation for elucidating gene function easier than ever. The recent paper describing the global pattern of polymorphisms across 1,135 Arabidopsis accessions (1001 Genomes Consortium, 2016) provides the underlying data, while useful tools like AraGWAS – a "major new resource" that contains hundreds of thousands of linkages between phenotypes and polymorphisms (Togninalli *et al.*, 2018) – and easyGWAS, published last year (Grimm *et al.*, 2017), enable easy identification of potentially causal SNPs. Text mining: the Sternberg lab at the California Institute of Technology published Textpresso Central, which contains many new associations for Arabidopsis genes (Müller *et al.*, 2018), while eGenPub selects articles that are about specific plant proteins (from 8 species) in the UniProt database (Ding *et al.*, 2017). The Ecker lab published 8,577 interactions among Arabidopsis transcription factors (AtTFIN-1) using a next-gen sequencing-based Y2H assay (CrY2H-seq) method they developed (Trigg *et al.*, 2017).

Integrative tools: The ePlant tool for exploring Arabidopsis data from the kilometre- to nanometre-scale in an integrative manner was published by the Provart Lab and collaborators (Waese *et al.*, 2017), and is available at <http://bar.utoronto.ca/eplant> or as an "app" on Araport. A 2018 update to ePlant includes Klepikova *et al.*'s (2016) RNA-seq-based developmental atlas for Arabidopsis and O'Malley *et al.*'s DAP-Seq effort (2016) to map the binding specificity of 529 TFs from Arabidopsis – more than 2.8 million protein-DNA interactions are now viewable in ePlant's Interaction Viewer.

GENIUS is a user-friendly web server that uses a novel machine learning algorithm to infer functional gene networks focused on specific genes and experimental conditions that are relevant to biological functions of interest, developed by the Gutiérrez Lab. Try it for 8 species, including Arabidopsis at <http://networks.bio.puc.cl/genius>.

Overviews/collections: To reiterate from last year's report, a good overview of several plant genomic databases is provided in volume 1533 of *Methods in Molecular Biology (Plant Genomics Databases)*, edited by Aalt-Jan van Dijk, 2016), including chapters on ENSEMBL Plants, the BAR, PLAZA, Plant Promoter DB and many others. A similar collection providing a good overview of plant bioinformatic resources is available in the open-access *Current Plant Biology* special edition on genomic resources and databases (Naithani and Van de Peer - eds., 2016). The 2018 *Nucleic Acids Research* database issue (Rigden and Fernández, 2018) contains updates or reports on several plant databases. Of note here: Plant Chromatin State Database (Liu *et al.*, 2018), Planteome - a portal for plant ontologies (Cooper *et al.*, 2018), updates to PLAZA (Van Bel *et al.*, 2018) and Gramene (Tello-Ruiz *et al.*, 2018), and the aforementioned AraGWAS. AgriGO for Gene Ontology enrichment analyses in several plant species was also updated to version 2.0 by the Su Lab at the Chinese Agricultural University (Tian *et al.*, 2017).

References

1001 Genomes Consortium (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell* 166: 481–491.

Aghamirzaie, D., Raja Velmurugan, K., Wu, S., Altarawy, D., Heath, L.S., and Grene, R. (2017). Expresso: A database and web server for exploring the interaction of transcription factors and their target genes in *Arabidopsis thaliana* using ChIP-Seq peak data. *F1000Research* 6: 372.

Balaguer, M.A. de L., Fisher, A.P., Clark, N.M., Fernandez-Espinosa, M.G., Möller, B.K., Weijers, D., Lohmann, J.U., Williams, C., Lorenzo, O., and Sozzani, R. (2017). Predicting gene regulatory networks by combining spatial and temporal gene expression data in Arabidopsis root stem cells. *Proc. Natl. Acad. Sci.*: 201707566.

Banf, M. and Rhee, S.Y. (2017). Enhancing gene regulatory network inference through data integration with markov random fields. *Sci. Rep.* 7.

Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M., and Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. *Proc. Natl. Acad. Sci. U. S. A.* 114: E10018–E10027.

Becker, M.G., Walker, P.L., Pulgar-Vidal, N.C., and Belmonte, M.F. (2017). SeqEnrich: A tool to predict transcription factor networks from co-expressed Arabidopsis and Brassica napus gene sets. *PLoS One* 12: e0178256.

Cooper, L. *et al.* (2018). The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Res.* 46: D1168–D1180.

David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., Preiss, T., and Searle, I.R. (2017). Transcriptome-Wide Mapping of RNA 5-Methylcytosine in Arabidopsis mRNAs and Noncoding RNAs. *Plant Cell* 29: 445–460.

Ding, R., Boutet, E., Lieberherr, D., Schneider, M., Tognolli, M., Wu, C.H., Vijay-Shanker, K., and Arighi, C.N. (2017). eGenPub, a text mining system for extending computationally mapped bibliography for UniProt Knowledgebase by capturing centrality. *Database* 2017.

Ezer, D., Shepherd, S.J., Brestovitsky, A., Dickinson, P., Cortijo, S., Charoensawan, V., Box, M.S., Biswas, S., Jaeger, K., and Wigge, P.A. (2017). The G-box transcriptional regulatory code in Arabidopsis. *Plant Physiol.*: pp.01086.2017.

Grimm, D.G., Roqueiro, D., Salomé, P.A., Kleeberger, S., Greshake, B., Zhu, W., Liu, C., Lippert, C., Stegle, O., Schölkopf, B., Weigel, D., and Borgwardt, K.M. (2017). easyGWAS: A Cloud-Based Platform for Comparing the Results of Genome-Wide Association Studies[OPEN]. *Plant Cell* 29: 5–19.

Harkey, A.F., Watkins, J.M., Olex, A.L., DiNapoli, K.T., Lewis, D.R., Fetrow, J.S., Binder, B.M., and Muday, G.K. (2018). Identification of Transcriptional and Receptor Networks That Control Root Responses to Ethylene. *Plant Physiol.* 176: 2095–2118.

Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D., and Penin, A.A. (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.* 88: 1058–1070.

Kulkarni, S.R., Vanechoutte, D., Van de Velde, J., and Vandepoele, K. (2018). TF2Network: predicting transcription factor regulators and gene regulatory networks in Arabidopsis using publicly available binding site information. *Nucleic Acids Res.* 46: e31.

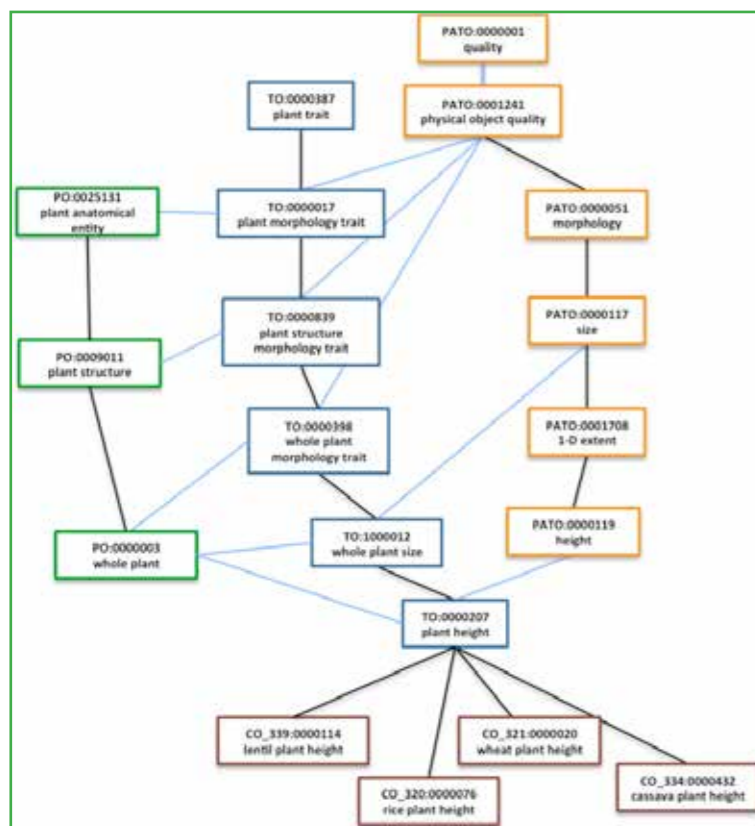


Figure 3: A view of the ontology hierarchy around Trait Ontology term plant height (TO:0000207). Crop Ontology (CO) terms for plant height from the lentil, wheat, rice and cassava ontologies are mapped to the Trait Ontology term for data integration.

Rigden, D.J. and Fernández, X.M. (2018). The 2018 Nucleic Acids Research database issue and the online molecular biology database collection. *Nucleic Acids Res.* 46: D1–D7.

Sijacic, P., Bajic, M., McKinney, E.C., Meagher, R.B., and Deal, R.B. (2018). Changes in chromatin accessibility between *Arabidopsis* stem cells and mesophyll cells illuminate cell type-specific transcription factor networks. *Plant J.* 94: 215–231.

Smet, R.D., Sabaghian, E., Li, Z., Saeys, Y., and Peer, Y.V. de (2017). Coordinated Functional Divergence of Genes after Genome Duplication in *Arabidopsis thaliana*. *Plant Cell: tpc.00531.2017.*

Tello-Ruiz, M.K. *et al.* (2018). Gramene 2018: unifying comparative genomics and pathway resources for plant research. *Nucleic Acids Res.* 46: D1181–D1189.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45: W122–W129.

Togninalli, M., Seren, Ü., Meng, D., Fitz, J., Nordborg, M., Weigel, D., Borgwardt, K., Korte, A., and Grimm, D.G. (2018). The AraGWAS Catalog: a curated and standardized *Arabidopsis thaliana* GWAS catalog. *Nucleic Acids Res.* 46: D1150–D1156.

Tokizawa, M., Kusunoki, K., Koyama, H., Kurotani, A., Sakurai, T., Suzuki, Y., Sakamoto, T., Kurata, T., and Yamamoto, Y.Y. (2017). Identification of *Arabidopsis* genic and non-genic promoters by paired-end sequencing of TSS tags. *Plant J.* 90: 587–605.

Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, Castanon R, Goubil A, Feeney J, O'Malley R, Huang SC, Zhang ZZ, Galli M and Ecker JR (2017) CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nature Methods* 14(8):819-825. doi: 10.1038/nmeth.4343

Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., Coppens, F., and Vandepoele, K. (2018). PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res.* 46: D1190–D1196.

Vanechoutte, D., Estrada, A.R., Lin, Y.-C., Loraine, A.E., and Vandepoele, K. (2017). Genome-wide characterization of differential transcript usage in *Arabidopsis thaliana*. *Plant J.* 92: 1218–1231.

Waese, J. *et al.* (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* 29: 1806–1821.

Liu, Y., Tian, T., Zhang, K., You, Q., Yan, H., Zhao, N., Yi, X., Xu, W., and Su, Z. (2018). PCSD: a plant chromatin state database. *Nucleic Acids Res.* 46: D1157–D1167.

Müller, H.-M., Van Auken, K.M., Li, Y., and Sternberg, P.W. (2018). Textpresso Central: a customizable platform for searching, text mining, viewing, and curating biomedical literature. *BMC Bioinformatics* 19: 94.

Naithani, S. and Van de Peer, Y. (2016). Genomic resources and databases for plant research community. *Curr. Plant Biol.* 7–8: 1.

O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* 165: 1280–1292.

Palovaara, J., Saiga, S., Wendrich, J.R., Hofland, N. van 't W., Schayck, J.P. van, Hater, F., Mutte, S., Sjollem, J., Boekschoten, M., Hooiveld, G.J., and Weijers, D. (2017). Transcriptome dynamics revealed by a gene expression atlas of the early *Arabidopsis* embryo. *Nat. Plants* 3: 894–904.

Plant Genomics Databases - Methods and Protocols | Aalt-Jan van Dijk | Springer.

Puelma T, Araus V, Canales J, Vidal EA, Cabello JM, Soto A, Gutiérrez RA (2017) GENIUS: web server to predict local gene networks and key genes for biological functions. *Bioinformatics*. doi: 10.1093/bioinformatics/btw702

Clone-Based Functional Genomics Resources (ORFeomics)

<http://Arabidopsisresearch.org/index.php/subcommittees/orfeomics>

Motoaki Seki (Chair), motoaki.seki@riken.jp, RIKEN CSRS
 Joe Ecker (Co-Chair), ecker@salk.edu, Salk Institute
 Pascal Braun (Subcommittee Member), pascal.falterbraun@helmholtz-muenchen.de, Helmholtz Zentrum München (HMGU)
 Satoshi Iuchi (Subcommittee Member), satoshi.iuchi@riken.jp, RIKEN BRC
 Debbie Christ, crist.30@osu.edu, ABRC

Open Tools and Resources for Arabidopsis Researchers

We prepared the updated list of Full-length cDNA and ORF clones that are available from Resource Centers (Please see Table 1). The revised ones are shown in red.

Recent activities of Subcommittee members.

Keeping tracking progress made towards the production of full-length cDNAs and open reading frame (ORF) clones for all annotated *Arabidopsis* protein-coding genes.

Future Activities of the Subcommittee.

ORFeomics subcommittee would like to propose a new project to collect all ORF (full-length cDNA) clones from every *Arabidopsis* protein-coding gene so as to test protein-protein, protein-DNA and protein-RNA interactions.

Our recent search showed that now about 23,000 *Arabidopsis* protein-coding genes have been isolated as Full-length cDNA (ORF) clones. One of the last unexplored continents of *Arabidopsis* are the remaining 5,000 protein-coding genes. After that, only the non-coding genes remain to be isolated.

With the completion of isolating all 28,000 *Arabidopsis* protein-coding genes, comprehensive analysis of plant gene function will become possible by various functional analyses using transgenic and protein expression approaches.

The human whole ORFeome project is already ongoing. *Arabidopsis* is a model plant, thus this will represent the first big plant ORFeome project. On completion it might be possible to start synthetic biology using the whole gene set of *Arabidopsis* to allow functional studies of corresponding proteomes.

Publication

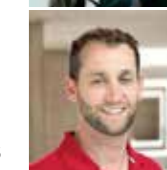
Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, Castanon R, Goubil A, Feeney J, O'Malley R, Huang SC, Zhang ZZ, Galli M and Ecker JR (2017) CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nature Methods* 14(8):819-825. doi: 10.1038/nmeth.4343.

Epigenetics and Epigenomics

Xuehua Zhong
xuehua.zhong@wisc.edu
 University of Wisconsin-Madison



Robert Schmitz
schmitz@uga.edu
 University of Georgia



Open Tools and Resources for Arabidopsis Researchers

Enhanced Jbrowse Plug-ins for Epigenomic Data Visualization.
<https://github.com/bhofmei/bhofmei-jbplugins>

CoGe browsers for Arabidopsis, maize and soybean - online servers for plant epigenomic data.

Conferences and Workshops

2017
 International Conference on Arabidopsis Research, St. Louis, MO, June 2017 (Session on Epigenetics).

Gordon Research Conference, Epigenetics. Holderness, NH, August 2017

Plant & Animal Genomes Conference, San Diego, CA, January 2017 (Session on Epigenomics of Plants International Consortium)

40th New Phytologist Symposium – Plant Epigenetics, Vienna, Austria, June 2017

2018
 Plant & Animal Genomes Conference, San Diego, CA, January 2018 (Session on Epigenomics of Plants International Consortium)

Cold Spring Harbor Asia meeting on Chromatin, Epigenetics & Transcription, Suzhou, China, April 2018 (Section on Epigenetic inheritance and Plant Epigenetics)

Gordon Research Conference, Plant Molecular Biology. Holderness, NH, June 2018 (Session on Epigenetics)
 Midwest Chromatin & Epigenetics meeting, West Lafayette, IN, June 2018 (Session on Plant Epigenetics)

Reports from MASC Subcommittees

| Creator | Format | Focus | Validation | Scale | URL | Stock center |
|--------------------------|---|-------------------------------------|--------------------|--------|--|----------------|
| ORF clones | | | | | | |
| SSP/RIKEN/Salk Institute | Univector pUNI51 | | Full sequence | 14,398 | signal.salk.edu/cdnastatus.html http://methylo.me.salk.edu/cgi-bin/clones.cgi | ABRC |
| Salk/Invitrogen | Gateway entry | | Full sequence | 12,114 | signal.salk.edu/cdnastatus.html http://methylo.me.salk.edu/cgi-bin/clones.cgi | ABRC |
| CCSB/Salk | Y2H clones | Plant Interactome Network Map | Full sequence | 18,258 | http://interactome.dfci.harvard.edu/A_thaliana/host.php | ABRC |
| TIGR | Gateway entry | Hypothetical genes | Full sequence | 3,041 | www.tigr.org/tdb/hypos/ | ABRC |
| Peking-Yale Joint Center | Gateway entry | Transcription factors | 5' and 3' end seq. | 1,282 | | ABRC |
| Dinesh-Kumar et al. | Gateway expression | TAP-tagged transcription factor | 5' and 3' end seq. | 15,543 | | ABRC |
| REGIA | Gateway entry | Transcription factors | 5' and 3' end seq. | 982 | gabi.rzpd.de/materials/ | GABI/ RZPD |
| Dinesh-Kumar et al. | Gateway entry, no stop pLIC-CTAP | Plant protein chips | 5' and 3' end seq. | 7,300 | plants.gersteinlab.org/ | ABRC |
| ATOME collection | Gateway entry | | 5' and 3' end seq. | 6,448 | http://urgv.evry.inra.fr/ATOMEdb | ABRC, CNRGV |
| Doonan et al. | Gateway Expression | GFP fusion for subcellular location | | 155 | | ABRC |
| Callis et al. | Gateway entry | Protein ubiquitination | Full sequence | 111 | plantsubq.genomics.purdue.edu | ABRC |

Table 1. List of available ORF Resources. Revised numbers are shown in Red

American Society for Plant Biologist, Montreal, Canada, July 2018 (Session on Plant Epigenetics)

28th International Conference on Arabidopsis Research, Turku Finland, June 2018 (Session on Epigenetics)

2019

Impact of Nuclear Domains On Gene Expression and Plant Traits, France, December 2019
Plant Epigenetics, Vienna, Austria, (hosted by Magnus Nordborg) Date:TBD, 2019

2020

Plant Epigenetics, Japan (hosted by Keiko Sugimoto and Toshiro Ito) Date:TBD 2020

Training

EpiDiverse Research & Training Network, Wageningen, Netherlands, June 5-8, 2018

Reports from MASC Subcommittees

| Creator | Format | Focus | Validation | Scale | URL | Stock center |
|-----------------------------------|---------------------------------------|------------------------------------|-----------------------------------|--------|--|--------------|
| ORF clones | | | | | | |
| | | | | | sheenweb/category_genes.html | |
| Steve Clouse | Gateway expression | N-terminal Flag/His tagged kinases | | 855 | http://www4.ncsu.edu/~sclouse/Clouse2010.htm | ABRC |
| Frommer et al. | Gateway entry, no stop | Membrane and signaling proteins | 5' and 3' end seq. | 2,712 | http://associomics.org | ABRC |
| Frommer et al. | Gateway Expression (mbSUS clones) | Membrane and signaling proteins | 5' and 3' end seq. | 5,414 | http://associomics.org | ABRC |
| AIST/RIKEN | Gateway entry, no stop | Transcription factor | Full sequence | 1,998 | | BRC |
| RIKEN | Gateway entry, no stop | Transcription factor | Full sequence | 399 | | BRC |
| Allie Gaudinier and Siobhan Brady | Y1H, AD vector | Transcription factor | Full sequence | 635 | | ABRC |
| SALK/Promega | pIX-HALO vector | | | 12,069 | | ABRC |
| O'Malley et al. | pIX-HALO vector | Transcription factor | Full sequence | 1,315 | | ABRC |
| Guillaume Pilot | pDONRZEO vector | Membrane protein | 5' and 3' end seq. | 192 | | ABRC |
| Pruneda-Paz et al. | Gateway Entry, no stop | Transcription factor | Full sequence | 1,956 | | ABRC |
| Pruneda-Paz et al. | Gateway Destination, no stop, pDEST22 | Transcription factor | Full sequence | 1,956 | | ABRC |
| Lao et al. | Gateway DONR, no stop, pDONR223 | glycosyltransferases | Full sequence | 429 | http://gt.jbei.org/ | ABRC |
| Amita Kaundal et al. | Gateway DONR, no stop, pDONR201 | | | 26 | | ABRC |
| cDNA clones | | | | | | |
| RIKEN/SSP/ | λ ZAP or λ PS | | Full sequence/ 5' and 3' end seq. | 22,671 | www.brc.riken.go.jp/lab/ | BRC |
| Salk Insitute | | | | | epd/Eng/order/order.shtml | |
| MPI-MG | Gateway expression | | 5' end seq. | 4,500 | gabi.rzpd.de/materials/ | GABI/RZPD |
| Génoscope/LT I | Gateway entry | | Full single pass seq. | 28,866 | www.genoscope.cns.fr/Arabidopsis | CNRGV |

Table 1a. List of available ORF Resources. Revised numbers are shown in Red

Selected Publications

Biochimica et Biophysica Acta Special issue on Plant Gene Regulatory Mechanisms and Networks included 16 articles edited by Nathan Springer and Erich Grotewold.
<https://www.sciencedirect.com/journal/biochimica-et-biophysica-acta-bba-gene-regulatory-mechanisms/vol/1860/issue/1>

Genome Biology Special issue on Plant Epigenomics included 30 articles edited by Claudia Kohler and Nathan Springer
<https://www.biomedcentral.com/collections/plantepigenomics>

Metabolomics

<http://Arabidopsisresearch.org/index.php/subcommittees/metabolomics>

Kazuki Saito (chair) and Wolfram Weckwerth (co-chair) with contributions from subcommittee members and the wider Arabidopsis community



Open Tools and Resources for Arabidopsis Researchers

www.masc-metabolomics.org
Metabolomics subcommittee website

<https://www.arabidopsis.org/portals/metabolome/index.jsp>
TAIR metabolomics Resources

<http://prime.psc.riken.jp/>
PRIME, a web-based service for metabolomics, containing computational metabolomics toolbox (MS-DIAL, MS-FINDER, etc), metabolomic datasets (DROP Met), metabolic modeling system (PASMet), metabolomic characterization of 50 Arabidopsis mutants and the database as a functional genomics tool (MeKO), standard compounds (MetBoard), Arabidopsis metabolome expression databases (AtMetExpress development, AtMetExpress 20 ecotypes) and a collection of literature and in-house MS spectra data (ReSpect for Phytochemicals).

<http://gmd.mpimp-golm.mpg.de/>
Golm Metabolome Database facilitates the search for and dissemination of reference mass spectra from biologically active metabolites quantified using gas chromatography (GC) coupled to mass spectrometry (MS).
www.plantmetabolomics.org A web portal of Arabidopsis Metabolomics Consortium that contains data from an NSF-2010 funded project concerning metabolite profiling of a set of metabolic mutants. <http://mmcd.nmrfam.wisc.edu/>
The Madison-Qingdao metabolomics consortium database has emphasis on Arabidopsis and contains both NMR and MS data of metabolites.

<http://www.ebi.ac.uk/metabolights>
MetaboLights, a database for metabolomics experiments and derived information. The database is cross-species, cross-technique and covers metabolite structures and their reference spectra as well as their biological roles, locations and concentrations, and experimental data from metabolic experiments and is a collaborative multi-laboratory effort including groups specialising in plant metabolism.

Recent activities of Subcommittee members.

Since metabolomics is an important component of Arabidopsis omics, a continuous goal of this subcommittee will be to promote metabolomics research of Arabidopsis leading to functional genomics and systems biology. Full integration of Arabidopsis-based metabolomics research with the activity of the Metabolomics Society <<http://www.metabolomicssociety.org/>> is also an important goal of this subcommittee. Several members of the subcommittee are involved in drawing up the plant biology specific documentation for the Metabolomics Society. In addition this committee will aim to establish a mechanism that allows the dissemination of metabolomics datasets to the wider Arabidopsis community and encourage and facilitate initiatives for the integration of metabolomic datasets with other omic datasets. This will involve depositing metabolomic data in a usable format for data integration.

Future Activities of the Subcommittee.

To realize the goals, we aimed to establish the subcommittee website www.masc-metabolomics.org for more efficient exchange of information and dissemination of the subcommittee's activity. The subcommittee discussion will be taken not only in the occasion of ICAR annual meeting but also in the occasions of several other metabolomics-related meetings, where the subcommittee members can join. The webinterface will provide user with a user-friendly tool to search for *Arabidopsis thaliana* metabolomics data in available databases. In addition, the people in plant metabolomics community actively provide open tools and resources useful for Arabidopsis researchers as indicated above.

Conferences, Workshops and Training events

2017/6/26-29 Metabolomics 2017, Brisbane, Australia
2017/7/9-14 Gordon Research Conference, Plant Metabolic Engineering, Waterville Valley, NH, US
2017/7/16-20 The 4th International Conference on Plant Metabolism (ICPM 2017), Dalian, China
2018/6/18-22
MU Metabolomics Workshop, Metabolomics Center, University of Missouri, US
2018/6/24-28
Metabolomics 2018 (In partnership with The Plant Metabolomics Forum), Seattle, Washington, US
2018/7/8-13
The 23rd International Symposium on Plant Lipids, Yokohama, Japan

Proteomics

<http://Arabidopsisresearch.org/index.php/subcommittees/proteomics>

<http://www.masc-proteomics.org/>

Joshua Heazlewood (chair)

Open Tools and Resources for Arabidopsis Researchers

Pep2Pro - proteo-genomic resource
<http://fgcz-pep2pro.uzh.ch/>
by Baerenfaller

ProMEX - mass spectral resource
<http://promex.pph.univie.ac.at/promex/>
by Wienkoop and Weckwerth

GelMap - 2-DE proteomic repository
<https://gelmap.de/projects-arabidopsis/>
by Braun

APP - targeted proteomics (SRM) tool <http://www.plantenergy.uwa.edu.au/APP/>
by Taylor and Millar

SUBA - subcellular database
<http://suba.live/>
by Millar

PPDB - plant protein database
<http://ppdb.tc.cornell.edu/>
by van Wijk

AT_CHLORO - plastid proteomic resource
http://at-chloro.prabi.fr/at_chloro/
by Rolland and Ferro

Plant PTS1 Protein Prediction-
<http://ppp.gobics.de/>
by Reumann

PhosPhAt - phosphoproteome database
<http://phosphat.uni-hohenheim.de/>
by Schulze

P3DB - plant phosphoproteome database
<http://p3db.org/>
by Thelen

MASCP Gator - proteomics aggregation portal
<http://gator.masc-proteomics.org/>
by MASCP

1001 Proteomes - non-synonymous SNPs from Arabidopsis natural variants
<http://1001proteomes.masc-proteomics.org/>
by Heazlewood



MRMaid - MRM design tool
<http://elvis.misc.cranfield.ac.uk/mrmaid/>
 by Jones

Multiple Marker Abundance Profiling
<http://suba.live/toolbox-app.html>
 by Heazlewood and Millar

Recent activities of Subcommittee members.

- A) The subcommittee members maintain an array of Arabidopsis specific proteomic repositories. These resources currently contribute to information in the Arabidopsis Information Portal (AIP).
- B) The 1001 Proteomes portal provides pre-computed nsSNP data from the sequenced accessions. Data from this portal are now available in ePlant.
- C) Members have ongoing interests in applying proteomic approaches in Arabidopsis to important crop species. A number of the on-line proteomic resources also provide proteomic datasets for agricultural relevant species.
- D) Subcommittee members maintain extensive resources in the area of protein function in Arabidopsis, and it is expected that this information will be coordinated with the IAIC.

E) Subcommittee members are involved with the initiative on Multi-Organism Proteomes (iMOP) as part of the Human Proteome Organization (HUPO) and are active members of the International Plant Proteomics Organization (INPPO). As well as many serving as members of their national plant societies.

Future Activities of the Subcommittee.

- (A) Maintain and develop new resources for Arabidopsis focusing on protein function and proteomics.
- (B) Demonstrate applicability of 1001 proteome data to the research community. Update the data for the recently released accessions.
- (D) Continue to develop proteomic Apps for the Arabidopsis Information Portal.
- (E) Ensure attendance and involvement in future INPPO and HUPO activities.
- (F) Maintain an active Arabidopsis proteomics subcommittee (<http://www.masc-proteomics.org/>)

Publications

Hooper CM, Stevens TJ, Saukkonen A, Castleden IR, Singh P, Mann GW, Fabre B, Ito J, Deery MJ, Lilley KS, Petzold CJ, Millar AH, Heazlewood JL, Parsons HT (2017) Multiple

marker abundance profiling: combining selected reaction monitoring and data-dependent acquisition for rapid estimation of organelle abundance in subcellular samples. *Plant J* 92: 1202-1217

Kraner ME, Muller C, Sonnewald U (2017) Comparative proteomic profiling of the choline transporter-like1 (CHER1) mutant provides insights into plasmodesmata composition of fully developed *Arabidopsis thaliana* leaves. *Plant J* 92: 696-709

Willems P, Ndah E, Jonckheere V, Stael S, Sticker A, Martens L, Van Breusegem F, Gevaert K, Van Damme P (2017) N-terminal Proteomics Assisted Profiling of the Unexplored Translation Initiation Landscape in *Arabidopsis thaliana*. *Mol Cell Proteomics* 16: 1064-1080

Rantala M, Tikkanen M, Aro EM (2017) Proteomic characterization of hierarchical megacomplex formation in Arabidopsis thylakoid membrane. *Plant J* 92: 951-962

Carrera DA, Oddsson S, Grossmann J, Trachsel C, Streb S (2018) Comparative Proteomic Analysis of Plant Acclimation to Six Different Long-Term Environmental Changes. *Plant Cell Physiol* 59: 510-526

Systems and Synthetic Biology

<http://Arabidopsisresearch.org/index.php/subcommittees/systems-biology>

Siobhan Brady, UC Davis
sbrady@ucdavis.edu
 Malcolm Bennett, University of Nottingham
Malcolm.Bennett@nottingham.ac.uk,
 Gabriel Krouk, CNRS
gabriel.krouk@cnrs.fr
 Nicola Patron, Earlham Institute
Nicola.Patron@earlham.ac.uk,
 Rodrigo Gutierrez, P. Universidad Catolica de Chile
rgutierrez@bio.uc.cl
 Pascal Falter-Braun, Hemholtz Zentrum Munchen
pascal.falter-braun@helmholtz-muenchen.de



Open Tools and Resources for Arabidopsis Researchers

GB3.0 A Digital Toolbox for Plant Synthetic Biology
<https://gbcloning.upv.es>
 LOOP A simple, open-source system for recursive fabrication of DNA circuits
<https://www.biorxiv.org/content/early/2018/01/15/247593>

OpenMTA. A simple, standardized legal tool that enables individuals and organizations to share their materials on an open basis. Developed as a collaborative effort between the BioBricks Foundation and the OpenPlant Initiative, with input from researchers, technology transfer professionals,

social scientists, lawyers, and other stakeholders from across the globe, the OpenMTA reflects the values of open communities and the practical realities of technology transfer.
<https://www.openplant.org/openmta/>

OpenSimRoot - Postma, J.A., Kuppe, C., Owen, M.R., Mellor, N., Griffiths, M., Bennett, M.J., Lynch J.P., Watt, M. (2017) OpenSimRoot: Widening the scope and application of root architectural models *New Phytologist* 215 (3) 1274-1286. This paper describes how researchers from 3 centres in the US, UK and Germany worked together to create an open source version of SimRoot, termed OpenSimRoot, to make this FSPM (functional-structural plant model) available to the wider scientific community

Conferences, Workshops and Training events

- Conferences and Workshops
 Practical Synthetic Biology. Pretoria, South Africa, February 2017
- Synthetic Biology for Natural Products. Cancun, Mexico, May 2017
- SB7.0 Singapore, June 2017
- Synthetic Biology: Engineering, Evolution & Design (SEED). Vancouver, July 2017
- Gordon Research Conference, Plant Metabolic Engineering – Plant Engineering in the Synthetic Biology Era. Waterville Valley, NH, July 2017
- Open Plant. Cambridge, UK, July 2017
- Genome Science/Genome10K. Norwich, UK, August 2017
- Natural Products and Synthetic Biology: Parts and Pathways. Olympic Valley, California USA, January 2018
- Molecular Biosystems, Puerto Varas, Chile, September 23-26, 2017
- 1st Latin American Workshop and Conference on Systems Biology, Mexico City, Mexico, May 2-5, 2017
- Systems Biology: Networks, Cold Spring Harbor, New York, USA, March 14-18, 2017
- International Conference on Arabidopsis Research, St. Louis, Missouri, USA, June 19-23, 2017, session on Modeling, Gene Regulation, Systems, Quantitative and Computational Biology.
- Synbiosys Summer School, Copenhagen Plant Sciences Centre, Denmark, August 2017
- The Synthetic and Systems Biology Summer School (SSBSS) Robinson College, University of Cambridge, UK, July 2017
- Second International Plant Biology Course (Santiago, Chile), 2017.

Future Conferences or Workshops:

- EMBO Workshop: Integrating Systems Biology, April 15-17, 2018; with prominent plant talks and Sorina Popescu and as organizers Pascal Falter-Braun, Luis Serrano and Sandra Orchard.
- iPSB - First International Plant Systems Biology Meeting; September 10-14, 2018; organized by Gabriel Krouk, Pascal Falter-Braun, Siobhan Brady, Gloria Coruzzi and Rodrigo Gutierrez
- SEB Interdisciplinary Sessions: The 2018 Annual Meeting of the Society for Experimental Biology will be held in Florence on the 3rd-6th July 2018. This will feature many parallel sessions including interdisciplinary sessions on "Shaping root architecture - from nutrient sensing and tropisms to systemic signals and decision making" organised by Stefan Kepinski and Julia Davis,
- Systems analyses of multicellularity complexity and organ biology organised by George Bassel, Leah Band and Mark Fricker,
- Quantitative synthetic biology organised by Christian Fleck, Mustafa Khammash and Robert Smith.

Selected Publications

- Robaina-Estévez S, Daloso DM, Zhang Y, Fernie AR, Nikoloski Z.; Resolving the central metabolism of Arabidopsis guard cells. *Sci Rep.* 2017 Aug 16;7(1):8307. doi: 10.1038/s41598-017-07132-9.
- Kautsar SA, Suarez Duran HG, Blin K, Osbourn A, Medema MH (2017) plantiSMASH: automated identification, annotation and expression analysis of plant biosynthetic gene clusters *Nucleic Acids Research.* 45(W1):W55-W63. doi: 10.1093/nar/gkx305.
- Di Mambro R, De Ruvo M, Pacific E, Salvi, E, Sozzani R, Benfey PN, Busch W, Novak O, Ljung K, Di Paola L, Maree AFM, Costantino P, Grieneisen VA, Sabatini S. Auxin minimum triggers the developmental switch from cell division to cell differentiation in the Arabidopsis root. *PNAS.* (2017) 114(36):E7641-E7649.
- Li B, Tang M, Nelson A, Caligagan H, Zhou X, Clark-Wiest C, Ngo R, Brady SM, Kliebenstein, DJ. Network-Guided Discovery of Extensive Epistasis between Transcription Factors Involved in Aliphatic Glucosinolate Biosynthesis. *Plant Cell.* 2018 Jan;30(1):178-195. doi: 10.1105/tpc.17.00805. Epub 2018 Jan 9.
- Jackson MD, Xu H, Duran-Nebreda S, Stamm P, Bassel GW. Topological analysis of multicellular complexity in the plant hypocotyl. *Elife.* 2017. pii: e26023. doi: 10.7554/eLife.26023

Additional Information

iPSB was a direct result of activities and leadership efforts on behalf of the Systems and Synthetic Biology Subcommittee.

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/>

