



Workshop:

“Ontologies, Standards and Best Practice”

Gent, May 21-23, 2007

June 13, 2007

Compiled by F. Fiorani, M. Kuiper, and P. Hilson

1. SUMMARY AND ACTION POINTS FOR AGRON-OMICS.....	3
2. WORKSHOP PROGRAM	4
3. ABSTRACTS	5
3.1 <i>Session I: Ontologies</i>	5
3.2 <i>Session II: Data Standards and Best Practice</i>	7
3.3 <i>Session III: Web Services for Biological Data.....</i>	9
3.4 <i>Session IV: Biological Pathways</i>	10
4. TRAINING SESSIONS.....	12
4.1 <i>Recording leaf phenotypes: practical session.....</i>	12
4.2 <i>Web-services demo session</i>	12
4.3 <i>Using ontologies and ontology editing tools: introduction and hands-on session on the applicability to capturing plant phenotypes.....</i>	12
4.4 <i>Manual literature encoding tool: introduction and hands-on session</i>	12
5. PARTICIPANTS' LIST.....	13
6. RELEVANT WEBSITES	14
7. SELECTED READINGS.....	16

1. Summary and Action Points for AGRON-OMICS

The AGRON-OMICS workshop “Ontologies, Standards and Best Practice” took place at the Department of Plant Systems Biology, VIB-UGent (Gent, May 21-May 23, 2007). There were thirty-seven participants of which about half are affiliated with partner laboratories and directly contribute to AGRON-OMICS. This event was organized to raise awareness about the need of adopting specific biomedical ontologies and data standards among biologists, and to stimulate discussions about the issues connected with the use of these classification and standard systems in plant research. The program consisted of four seminar sessions with international speakers of different scientific background, and four practical training sessions designed to provide hands-on experience on software for viewing and editing biological ontologies, recording plant phenotypes using a controlled vocabulary, and for building data acquisition pipelines through web-services.

Throughout the workshop the scientific organizing committee has structured part of the discussions to evaluate which of the existing ontologies and biological pathway standards may be adopted and/or enriched by AGRON-OMICS. This effort resulted in the definition of workgroups responsible to follow-up development paths in the consortium context:

- i) Use of ontologies (how to use/enrich PO, GO, PATO, EO)
 - Definition of an Arabidopsis leaf phenotype ontology (Belfield, Chih-Wei Tung, Fiorani, Granier, Hilson, Perez-Perez)
 - Definition of an Arabidopsis genotype ontology (Fiorani, Hilson, Usadel, Chih-Wei Tung, Kusnierczyk, MPI- DB)
- ii) Information exchange and inter-operability of current biological pathways systems, and acquisition of specific analytical results concerning genes included in the AGRON-OMICS molecular scaffold (Usadel, Zimmermann, Walsh, Higgins, Kuiper, Rhee, Schmidt)
- iii) Building dedicated literature mining and plant phenotype annotation tools
 - Coordination of a consortium trial for a software currently developed at PSB (Vercruysse, Kuiper, Hilson, Fiorani, Kusnierczyk)
 - Evaluation of available tools for phenotype annotation (Perez-Perez, Granier, Fiorani)
- iv) Development of dedicated web-services driven by partners’ requests via Wiki style AGRON-OMICS web pages (Schildknecht, Kusnierczyk)

The conclusions of these workgroups will be summarized in working documents that will be accessible to all partners or will be the object of specific communications directed to AGRON-OMICS partners, with particular reference to deliverables of work-package 1 (Leaf Growth), work-package 6 (Data Integration), and work-package 7 (Models).

In conclusion, this workshop was essential for future research within the AGRON-OMICS Consortium, and raised the awareness of the participants of the challenging tasks of adopting and developing resources specifically designed for the plant research community. The participants generally expressed a positive judgment and actively contributed to the discussions after the seminars and during the training sessions.

2. Workshop Program

SEMINARS

Session I: Ontologies (chairperson: Pierre Hilson, PSB)

"Biomedical Ontologies: Foundations and Principles of Design" - [Waclaw Kusnierczyk](#) (University of Trondheim)

"Ontology-driven Data Integration in Life Sciences" - [Vladimir Mironov](#) (PSB)

"Applications of Plant Ontologies for Describing and Comparing Phenotypes and Gene Expression in Angiosperms" - [Katica Ilic](#) (former coordinator of Plant Ontology Consortium)

"Plant Ontology" - [Chih-Wei Tung](#) (present coordinator of Plant Ontology Consortium)

"Leaf Growth Phenotyping : Which Data for Which Objectives, in Which Context ?" - [Christine Granier](#) (INRA-LEPSE)

Session II: Data Standards and Best Practice (chairperson: Martin Kuiper, PSB)

"MIAME and Other Functional Genomics Standards - The State of the Art" [Helen Parkinson](#) (EBI)

"SBML: Systems Biology Markup Language" - [Sarah Keating](#) (University of Hertfordshire, UK)

"Implementation of Best Practices in Large Scale EU Projects" - [Martin Kuiper](#) (PSB)

"Literature Semantic Search and Encoding" - [Steven Vercruysse](#) (PSB)

Session III: Web-services for Biological Data (chairperson: Katica Ilic)

"Data Integration" - [Beatrice Schildknecht](#) (NASC)

Session IV: Biological Pathways (chairperson: Fabio Fiorani, PSB)

"Quality Control of Microarray Data and Pathways in Genevestigator" - [Philip Zimmermann](#) (ETH)

"Integration of Various Multiparallel Data Sources Using the MapMan Software Family" - [Björn Usadel](#) (MPI-MP)

"Reactome - a Knowledgebase of Biological Pathways" - [Esther Schmidt](#) (EBI)

"The Arabidopsis Reactome : Status and Applications" - [Sean Walsh](#) (JIC)

TRAINING SESSIONS

"Recording leaf phenotypes: practical session" - [Daisy Belfield](#) (NASC)

"Web-services demo sessions (NASC, TAVERNA)" - [Beatrice Schildknecht](#) (NASC)

"Using ontologies and ontology editing tools: introduction and hands-on session on the applicability to capturing plant phenotypes" - [Katica Ilic](#), [Jane Lomax](#) (EBI)

"Manual literature encoding tool: introduction and hands-on session" - [Steven Vercruysse](#) (PSB)

Scientific organizing committee: Pierre Hilson, Fabio Fiorani, Martin Kuiper

3. Abstracts

3.1 Session I: Ontologies

Biomedical Ontologies: Foundations and Principles of Design

Waclaw Kusnierczyk

Department of Computer and Information Science, University of Trondheim, Norway.

During the past decade, biomedical ontology development has become one of the most visible research activities within the field of bioinformatics. However, despite a long tradition of knowledge representation in philosophy and computer science, biomedical ontologies are notorious for mistakes made while their authors recurrently face similar problems. In this talk, we outline a number of rules of thumb designed to guide the every-day practice of ontological engineering. The presentation is organized around the issue of explicit commitments -- decisions that have to be made and appropriately documented before an ontology is released to the public. We discuss the need for precisely stating i) the intended use of the ontology, ii) its scope, iii) the philosophical perspective adopted, iv) the terminological conventions employed, and v) the restrictions on the content following from the limitations of the underlying representational formalism.

Ontology driven data integration in Life Sciences

Erick Antezana^{*1}, Mikel Egaña², Vladimir Mironov¹ and Martin Kuiper¹

¹ Department of Plant Systems Biology, Flanders Institute for Biotechnology (VIB)/ Ghent University., Technologiepark 927, Ghent, Belgium

² University of Manchester Computer Science, Oxford Road, M13 9PL, UK

New biological discoveries are being reported at an extremely rapid rate. This new information is found in diverse resources that encompass a broad array of journal articles and public databases associated with different sub-disciplines within biology and medicine. The integration of biological knowledge and information is recognized as a critical knowledge gap in science and as essential for the future of the field because dissemination and subsequent deployment of the knowledge by automated applications and by researchers who need to access and connect the diverse information is also recognized as critical. The integration of data in biomedicine is not a trivial task. On one hand, on-line bioinformatics databases are especially difficult to integrate because they are complex, highly heterogeneous, dispersed and incessantly evolving. On the other hand, on-line biological data are often described only in human-readable formats (most commonly free text) that are difficult for computers to analyze, due to the lack of standardized structures. An ontology (sensu computer science) is a computer-processable method for encoding knowledge which specifies and defines not only the concepts in a given field but also the relationships among those concepts. Ontologies provide insight into the nature of information produced by that field and are an essential ingredient in any attempt to arrive at a shared understanding of the subject domain. Thus, the development of ontologies for biological information and the sharing of those ontologies within the bio-informatics community are pivotal for life scientists, as they heavily rely on on-line data. In our presentation we will demonstrate the use of the ontology driven approach to data integration in one particular knowledge domain of high scientific interest and social-economic importance. The domain of discourse covers the knowledge over cell (division) cycle - the process underlying cell proliferation in eukaryotic organisms. The resulting ontology, Cell Cycle Ontology www.cellcycleontology.org/, integrated data from various sources for four selected organisms.

Applications of Plant Ontologies for Describing and Comparing Phenotypes and Gene Expression in Angiosperms

Katica Ilic & Chih-Wei Tung
Plant Ontology Consortium

The Plant Ontology Consortium (POC) is a collaborative effort of several plant databases and experts in plant systematics, botany and genomics, sharing a goal of developing controlled vocabularies that accurately reflect morphology, anatomy, growth and developmental stages of a flowering plant. Plant Ontology (PO) has two aspects: 1) Plant Structure Ontology, with terms describing organs, organ systems, tissues, and cell types; and 2) Plant Growth and Developmental Stage Ontology, with terms describing spatial-temporal growth stages of a whole plant and developmental stages of organs and organ systems. After integration of existing ontologies for *Arabidopsis*, maize and rice, the PO was expanded to encompass Fabaceae, Solanaceae and other cereal crops. Consistent use of these ontologies potentially reduces the problem of heterogeneity of terminology used to describe comparable object types in plant databases - an obvious obstacle for conducting queries for multiple species across plant genomic databases. As a part of ongoing functional annotation efforts, participating databases, such as TAIR, NASC, Gramene, MaizeGDB and SOL Genomics Network, have been using PO to describe expression patterns of genes and phenotypes of mutants and natural variants. Over 10,000 gene annotations and phenotype descriptions from several species-specific databases can now be queried and retrieved using the Plant Ontology browser. A short summary of the organizing principles and rules followed in developing PO will be presented. Standards and methods for using PO to annotate gene expression patterns and to describe phenotypes of mutants and natural variants will be addressed with examples from member databases, demonstrating how the PO supports searching and retrieval of genetic and genomic information as well as mutant stocks and germplasm data. Active maintenance and further development of the PO will be outlined. Ontologies are work in progress - we envision the PO as a continuously expanding ontology that will gradually encompass many angiosperms and become a robust resource available for plant genomic databases as well as bench scientists.

Leaf growth phenotyping : which data for which objectives, in which context ?

Christine Granier,
INRA-SUPAGRO, LEPSE Montpellier-FRANCE

Two automated phenotyping platforms were developed in our group for capturing data describing leaf growth in hundreds of genotypes grown in rigorously reproducible environmental conditions. The first one, PHENODYN, is dedicated to monocotyledonous plants such as maize and rice, the other, PHENOPSIS, is dedicated to *Arabidopsis thaliana*. Both platforms allow capturing with different time resolution: leaf growth stages, leaf size, underlying dynamic leaf growth variables, underlying cellular leaf growth variables and set of micro-meteorological data. The sets of leaf growth variables recorded in both platforms allowed to identify leaf growth phenotypes altered by environmental conditions and/or genetic modifications. Response curves between leaf growth variables and environmental conditions were built and they were used to model the genotype x environment interaction. This approach allows characterising each genotype by a set of parameters which predict its leaf growth in a range of environmental conditions. Results presented here set in evidence that there is a challenge in the future to combine the development of Plant Ontology with biological and modelling approaches of the genotype x environment interaction.

3.2 Session II: Data Standards and Best Practice

MIAME and other functional genomics standards - The State of the Art

Helen Parkinson

Head of Production, ArrayExpress Database. Microarray Informatics Team, European Bioinformatics Institute, Hinxton. UK.

The gene expression standard MIAME has been available and supported by public databases (including ArrayExpress) and funding bodies now since 2001. In this period numerous other standards have been published with varying degrees of adoption. This talk will provide an overview of the MIAME standard development process, provide a retrospective look at the MIAME strengths and weaknesses and examine the need for allied data exchange formats and ontology development to support standards. The future of MIAME will also be discussed.

SBML - Systems Biology Markup Language

Sarah Keating

University of Hertfordshire

SBML is a computer-readable format for representing models of biochemical reaction networks and has become the de-facto standard format in this field. This allows development of software tools to focus on their functionality rather than the model format and to date there are over 100 known software tools that support SBML, providing users with a wide range of functionality and the facility to exchange models between both tools and researchers. The most recent definition of SBML (Level 2 Version 3) incorporates support for SBO (the Systems Biology Ontology) and an RDF-based annotation scheme that facilitates compliance with MIRIAM (the Minimal Information Requested In the Annotation of biochemical Models).

SBML development is driven by an international community of modellers and developers.

Implementation of Best Practices in Large Scale EU Projects

Martin Kuiper

Computational Biology, VIB Department of Plant Systems Biology

Coordinator of DIAMONDS, Coordinator of EMERALD.

The objectives of the EU Coordination Action EMERALD (<http://www.microarray-quality.org/>) include the development and implementation of quality metrics (QM, done by EBI) to assess the quality of a microarray experiment. Such QM will provide the means for assessing the quality of new data, but also the means to filter out bad data when re-analysing repositories. In addition, EBI will also develop a Normalization and Transformation ontology to be added to MGED. EMERALD will link with ongoing efforts on standard development, most importantly MAQC (MicroArray Quality Control project) and ERCC (External RNA Control Consortium). The results will be communicated to the microarray user community at a series of workshops. MIBBI (Minimum Information for Biological and Biomedical Investigations, <http://mibbi.sourceforge.net/>) is an initiative to develop a repository for checklists that describe the necessary information for a variety of technologies. Such a repository should help to channel community efforts to build, refine and use such checklists to ensure that biological data will be amply complemented with experimental metadata.

Literature Semantic Search and Encoding

Steven Vercruysse
VIB Department of Plant Systems Biology

Biological information is accumulating to enormous proportions, and keeping an overview is becoming a substantial problem. Based on new high-throughput experiments and a high interest in biotechnology, the information in this field is growing at an alarming rate. Although some information can flow to existing databases, most of the knowledge is still reported in publications, as unstructured text. Automated text-mining techniques try to understand this biological prose, but unfortunately, the intricacies of natural language are still a major obstacle to let them comprehend most of the relations between biological entities. Currently only human intelligence can reliably understand biological essays. As an alternative, we studied the requirements for a community effort in biology, where many biologists can read articles of their choice, and translate the information into a computer-readable standard format. We believe that a joint effort in a Wikipedia-like 'Web 2.0' style, is the only possible way to oversee and to meaningfully conserve the information that we have built together as a gigantic literature corpus. Key aspects to enable this community organisation are: 1) a controlled language that can capture very heterogeneous information coming from various biological disciplines; 2) a general appreciation of the necessity of computerized information management; and 3) a substantial reward for an achieved effort, such the automatic integration of your own and of colleagues' gathered information, into an intuitively browsable, powerful graphical overview.

3.3 Session III: Web Services for Biological Data

Data Integration

Beatrice Schildknecht
Nottingham Arabidopsis Stock Centre

The World Wide Web has brought a whole new world of databases and tools within easy reach of scientists. With this wealth of information, however, scientists and bioinformaticians can quickly become overwhelmed with the volume, diversity and complexity of data. Many bioinformatics resources are available to scientists through web interfaces, which require no programming knowledge. In this post-genomic era, however, web interfaces are limiting in many ways, such as when performing complex or very large analyses. An increasing number of resources are now becoming available as 'Webservices', which provide a programmatic interface for web-based services and are gradually being used more and more in biology and bioinformatics to automate the execution of services, circumventing many of these limitations. These Webservices can be chained together to create complex 'Workflows'. Taverna is a sophisticated, but easy to use application, developed for biologists to organize, automate and visualize these workflows.

3.4 Session IV: Biological Pathways

Quality control of Microarray Data and Pathways in Genevestigator

Philip Zimmermann
ETH Zurich

Genevestigator is a web-based analysis tool to study gene expression in a variety of biological contexts. The establishment of its microarray database required developments in quality control (QC), ontology annotation, and pathway curation. This presentation will show which QC measures were applied to ensure data comparability, how ontologies were built based on community standards, and challenges faced during pathway and network building. Several of the methodologies developed and used while creating this large database can be applied to small-scale microarray projects within single laboratories in order to improve data analysis capacity.

Integration of Various Multiparallel Data Sources Using the MapMan Software Family

Usadel B¹, Gibon Y¹, Nagel A², Blaesing OE^{1,3}, Piques MC¹, Thimm O^{1,3}, Hannemann J¹, Weicht D¹, Meyer S², Stitt M¹.

¹Max-Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Golm

²RZPD Deutsches Ressourcenzentrum für Genomforschung GmbH, Heubnerweg 6, D-14059 Berlin

³Present Address: Metanomics, Tegeler Weg 33, 10589 Berlin

The MapMan data visualization software has been developed for the visualization of large datasets, such as transcript, metabolite and/or protein profiling experiments. The software displays data in the context of pre-existing biological knowledge such as biological pathways, regulative processes or biological overviews (currently 70 pathways are preloaded). These pathways use an easily understandable, human-readable ontology which can be applied to any plant species. Thus, it is possible to get an idea of the biology behind one's experiments by simply loading the experiment into MapMan and displaying it. The different entities like metabolites, transcripts and enzymes are represented by different glyphs and summary statistics are employed to point the user to the most affected pathway or super-pathway. However, due to the current development to use multiple condition experiments or time courses we introduced a new software, PageMan, which condenses experiments based on ontology based statistics. The resulting summary is then displayed graphically and can be adapted interactively before final export. These tools were employed to investigate into the changes of the metabolism, during diurnal cycles and after transfer into darkness of Arabidopsis plants, where metabolites, enzymes and transcripts were measured. Many metabolite-transcript correlations could be identified, and for several these could be explained by investigating into the respective pathway. Moreover, it could be shown that complex experiments can be visualized efficiently using PageMan.

Reactome - a Knowledgebase of Biological Pathways

Esther Schmidt
European Bioinformatics Institute

Reactome (www.reactome.org) is a manually curated pathway knowledgebase that describes biological processes in a computationally accessible format. Data is provided and peer reviewed by experts in the field. The Reactome data model can accommodate a wide variety of biological pathways, ranging from carbohydrate metabolism to signal transduction and the cell cycle. Pathways serve to group reactions, and reactions form a network of biological interactions through entities that are consumed and produced, or act as catalysts. Annotation in Reactome also keeps track of modifications, e.g. phosphorylations, and cellular location. Cross-references are supplied to a wide selection of publicly available databases (like UniProt, Ensembl, GO, PubMed) in order to facilitate overall integration of biological data as well as to provide multiple entry points into the database. While the main focus of manual curation is on human data, electronic inference to 22 other species, based on orthology data, is performed at each data release. The Skypainter allows visualization of user-supplied data, overlaid onto the Reactome reaction framework, as well as a statistical analysis to assess overrepresentation of entities in Reactome pathways. Reactome data are freely available and can be downloaded in a number of formats (e.g. SBML).

The Arabidopsis Reactome : status and applications

Sean Walsh
John Innes Centre

Knowledge of biological networks is widely distributed in the literature, specialised databases and within the domain of specialist research groups. Having this knowledge in computationally amenable data structures would allow better analysis of functional genomics experiments and provide ready reference to what is known. Furthermore, this knowledge provides the framework for Systems Biology applications. Recognizing that the software and data model from the Reactome team could be usefully applied to the plant reference system *Arabidopsis thaliana*, we have embarked on the construction of the Arabidopsis Reactome. The current status and potential applications will be discussed.

4. Training Sessions

4.1 Recording leaf phenotypes: practical session

Coordinated by: [Daisy Belfield](#) (NASC)

This session focused on recording leaf phenotypes using a controlled vocabulary by the choice of an entity (noun) in Plant Structure Ontology and a quality (description) from PATO. The participants were asked to complete a score sheet after careful observation of a selected set of pictures of Arabidopsis rosette leaves.

4.2 Web-services demo session

Coordinated by: [Beatrice Schildknecht](#) (NASC)

This session provided a detailed guide and practical demonstrations of how biologists can use Taverna software to build workflows based on web services. This tutorial was structured in five parts: software installation; creation of a simple workflow; scavenge NASC SOAPlab web services; invoke a single SOAPlab web service; invoke a BioMoby workflow.

4.3 Using ontologies and ontology editing tools: introduction and hands-on session on the applicability to capturing plant phenotypes

Coordinated by: [Katica Ilic](#) (Plant Ontology) and [Jane Lomax](#) (EBI)

This session was divided in two parts, each including a hands-on time slot. The first part was designed to introduce relevant ontologies (GO, PO, PATO) with specific focus on ontology editing tools, in particular the OBO-Edit software. A second part focused on specific examples of how to use these ontologies to specify plant phenotypes using controlled vocabulary syntax.

4.4 Manual literature encoding tool: introduction and hands-on session

Coordinated by Steven Vercruysse (PSB)

The last session was based on a demo of software being developed by Steven Vercruysse and Martin Kuiper at VIB Department of Plant Systems Biology, Ghent University. This software allows the annotation of literature findings based on a new parsable syntax, and the visualization of the annotator statements. Complementing two other training sessions, the focus of this demo was to highlight the potential application of this software to encode ontology-based, entity-quality statement annotation types for leaf phenotypes.

5. Participants' List

NAME	INSTITUTION	CONTACT
Antezana, Erick	PSB-VIB-UGent	erant@psb.ugent.be
Beemster, Gerrit	PSB-VIB-UGent	gebee@psb.ugent.be
Belfield, Daisy	NASC	daisy@arabidopsis.info
Beynon, Jim	WSB	jim.beynon@warwick.ac.uk
Blonde, Ward	PSB-VIB-UGent	wablo@psb.ugent.be
Buchanan-Wollaston, Vicky	WSB	vicky.b-wollaston@warwick.ac.uk
Dhondt, Stijn	PSB-VIB-UGent	stdho@psb.ugent.be
Fiorani, Fabio	PSB-VIB-UGent	fafio@psb.ugent.be
Granier, Christine	INRA-LEPSE	granier@supagro.inra.fr
Grønlund, Jesper	WSB	J.T.Gronlund@warwick.ac.uk
Hernández Romero, Diana	UMH	dhernandez@umh.es
Higgins, Janet	JIC	janet.higgins@bbsrc.ac.uk
Hilson, Pierre	PSB-VIB-UGent	pihil@psb.ugent.be
Ilic, Katica	Ontario, Canada	kat23ica@gmail.com
Keating, Sarah	University of Hertfordshire	skeating@caltech.edu
Kuiper, Martin	PSB-VIB-UGent	makui@psb.ugent.be
Kusnierczyk, Waclaw	NTNU	waku@idi.ntnu.no
Leonelli, Sabina	London School of Economics	sabinaleo@hotmail.com
Lomax, Jane	EMBL-EBI	jane@ebi.ac.uk
Massonnet, Catherine	INRA-LEPSE	massonne@supagro.inra.fr
Messerli, Gäelle	ETHZ	mgaelle@ethz.ch
May, Sean	NASC	sean@arabidopsis.info
Mironov, Vladimir	PSB-VIB-UGent	vlmir@psb.ugent.be
Parkinson, Helen	EMBL-EBI	helen@ebi.ac.uk
Perez-Perez, José	UMH	jmperez@umh.es
Pervent, Marjorie	INRA-LEPSE	prevent@supagro.inra.fr
Rubio-Díaz, Silvia	UMH	silvia.rubio@umh.es
Rymen, Bart	PSB-VIB-UGent	barym@psb.ugent.be
Schildknecht, Beatrice	NASC	Beatrice@arabidopsis.info
Schmidt, Esther	EMBL-EBI	eschmidt@ebi.ac.uk
Tisné, Sébastien	INRA-LEPSE	tisne@supagro.inra.fr
Tung, Chih-Wei	Cornell University	cwt6@cornell.edu
Usadel, Björn	MPI-MP	usadel@mpimp-golm.mpg.de
Van Lijsebettens, Mieke	PSB-VIB-UGent	milij@psb.ugent.be
Vercruysse, Steven	PSB-VIB-UGent	stcru@psb.ugent.be
Walsh, Sean	JIC	sean.walsh@bbsrc.ac.uk
Zimmermann, Philip	ETHZ	phz@ethz.ch

6. Relevant websites

BIO - ONTOLOGIES

Barry Smith's webpage

<http://ontology.buffalo.edu/smith/>

Institute for Formal Ontology and Medical Information Science - IFOMIS

<http://www.ifomis.uni-saarland.de/>

Open Biomedical Ontology project

<http://obofoundry.org/>

Gene Ontology

<http://www.geneontology.org/>

Plant Ontology Consortium

<http://www.plantontology.org/>

Sequence Ontology

<http://www.sequenceontology.org/>

Cell Cycle Ontology

www.cellcycleontology.org/

Ontology Look-up Service

<http://www.ebi.ac.uk/ontology-lookup/>

Web-ontology browser

<http://www.brenda.uni-koeln.de/ontology/>

STANDARDS

links to Arabidopsis microarray public repositories at TAIR

<http://www.arabidopsis.org/portals/expression/microarray/microarrayDatasetsV2.jsp>

Microarray Gene Expression Data Society

<http://www.mged.org/>

Web portal of EMERALD, an EU project on Standards and Standardization of Microarray Technology and Data Analysis

<http://www.microarray-quality.org/>

Minimum Information for Biological and Biomedical Investigations

<http://mibbi.sourceforge.net/>

MIAME: Minimum Information About a Microarray Experiment

<http://www.mged.org/Workgroups/MIAME/miame.html>

MIAPE: Minimum Information About a Proteomics Experiment

<http://www.psidev.info/index.php?q=node/91>

Systems Biology Markup Language portal

<http://sbml.org/index.psp>

BIOLOGICAL PATHWAYS

Genevestigator software for microarray experiments mining: requires subscription
<https://www.genevestigator.ethz.ch/>

MapMan software for the mapping of gene expression onto biological pathways
<http://gabi.rzpd.de/projects/MapMan/>

software for quick overview of multi-parallel experiments
<http://mapman.mpimp-golm.mpg.de/pageman/>

Curated database of biological processes
<http://www.reactome.org/>

Arabidopsis reactome
<http://arabidopsisreactome.org/>

Kyoto Encyclopedia of Genes and Genomes - pathway database
<http://www.genome.jp/kegg/pathway.html>

Arabidopsis pathways and metabolic maps database at TAIR
<http://www.arabidopsis.org/biocyc/index.jsp>

SOFTWARE

OBO-Edit software for viewing and edit bio-ontologies
<http://oboedit.org/>

Protégé ontology editor
<http://protege.stanford.edu/>

COBrA, an ontology browser and editor for anatomy allowing mapping between ontologies
<http://www.xspan.org/cobra/index.html>

BioMoby system for inter-operability between biological data hosts and analytical services
<http://biomoby.open-bio.org/index.php/what-is-moby/>

Taverna software to create workflows and facilitate data integration in biological sciences
<http://taverna.sourceforge.net/>

Phenote software toolkit for the annotation of phenotypes
<http://www.phenote.org/>

AGRON-OMICS

AGRON-OMICS project portal
<http://www.agron-omics.eu/>

AGRON-OMICS molecular scaffold: access reserved to AGRON-OMICS partners
<http://arabidopsis.info/agronomics/home.html>

AGRON-OMICS Wiki pages: registration required
http://arabidopsis.info/agronomicswiki/index.php/Main_Page

7. Selected Readings

- Aitken S, Korf R, Webber B, Bard J (2005) COBrA: a bio-ontology editor. *Bioinformatics* 21: 825-826. PMID: 15513995
- Bard JB, Rhee SY (2004) Ontologies in biology: design, applications and future challenges. *Nat Rev Genet* 5(3):213-22 PMID: 14970823
- Blake JA, Bult CJ (2006) Beyond the data deluge: data integration and bio-ontologies. *J Biomed Inform* 39: 314-320 PMID: 16564748
- Brazma A, Krestyaninova M, Sarkans U (2006) Standards for systems biology. *Nature Rev Genet* 7(8):593-605 PMID: 16847461
- Chandrasekaran, Josephson JR, Benjamins VR (1999) What are ontologies, and why do we need them? *IEEE Intelligent Systems*, 14: 20-26
- Cote RG, Jones P, Apweiler R, Hermjakob H (2006) The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries. *BMC Bioinformatics*. 28;7:97. PMID: 16507094
- Day-Richter J, Harris MA, Haendel M, Lewis S (2007) OBO-Edit - An Ontology Editor for Biologists. *Bioinformatics*. 2007 Jun 1; [Epub ahead of print] PMID: 17545183
- Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr J-H, Hunter PJ, Juty ES, Kasberger JL, Kremling A, Kummer U, Le Novere N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J (2003) The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* 19(4): 524-531.
- Ilic K, Kellogg EA, Jaiswal P, Zapata F, Stevens PF, Vincent LP, Avraham S, Reiser L, Pujar A, Sachs MM, Whitman NT, McCouch SR, Schaeffer ML, Ware DH, Stein LD, Rhee SY (2007) The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant. *Plant Physiol*. 143(2):587-99. PMID: 17142475
- Kawas E, Senger M, Wilkinson MD (2006) BioMoby extensions to the Taverna workflow management and enactment software. *BMC Bioinformatics*. 30;7:523. PMID: 17137515
- Kersey P, Apweiler R (2006) Linking publication, gene and protein data. *Nature Cell Biol* 8(11):1183-1189 PMID: 17060904
- Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, Holloway E, Kolesnykov N, Lilja P, Lukk M, Mani R, Rayner T, Sharma A, William E, Sarkans U, Brazma A (2007) ArrayExpress--a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res*. 35:D747-50. PMID: 17132828
- Pujar A, Jaiswal P, Kellogg EA, Ilic K, Vincent L, Avraham S, Stevens P, Zapata F, Reiser L, Rhee SY, Sachs MM, Schaeffer M, Stein L, Ware D, McCouch S (2006) Whole-plant growth stage ontology for angiosperms and its application in plant biology. *Plant Physiol*. 142(2):414-28 PMID: 16905665
- Rhee SY, Dickerson J, Xu D (2006) Bioinformatics and its applications in plant biology, *Annu Rev Plant Biol* 57:335-360 PMID: 16669765
- Rosse C, Mejino JL Jr (2003) A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J Biomed Inform* 36(6):478-500 PMID: 14759820

Schulz S, Kumar A, Bittner T (2006) Biomedical ontologies: what *part-of* is and isn't. J Biomed Inform. 39(3):350-61 PMID: 16442850

Soldatova LN, King RD (2005) Are the current ontologies in biology good ontologies? Nat Biotechnol. 23(9):1095-8. PMID: 16151402

Spear AD Ontology for the twenty first century: An introduction with recommendations.
<http://www.ifomis.uni-saarland.de/bfo/manual/manual.pdf>

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37(6):914-39. PMID: 14996223

Usadel B, Nagel A, Steinhauser D, Gibon Y, Blasing OE, Redestig H, Sreenivasulu N, Krall L, Hannah MA, Poree F, Fernie AR, Stitt M. (2006) PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. BMC Bioinformatics 18;7: 535. PMID: 17176458

Vastrik I, D'Eustachio P, Schmidt E, Joshi-Tope G, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, Lewis S, Matthews L, Wu G, Birney E, Stein L (2007) Reactome: a knowledge base of biologic pathways and processes. Genome Biol. 2007;8(3):R39. PMID: 17367534

Wilkinson M, Schoof H, Ernst R, Haase D (2005) BioMOBY successfully integrates distributed heterogeneous bioinformatics Web Services. The PlaNet exemplar case. Plant Physiol. 2005 138(1):5-17 PMID: 15888673

Yu AC (2006) Methods in biomedical ontology. J Biomed Inform. 39(3):252-66. PMID: 16387553

Zimmermann P, Hennig L, Grisse W (2005) Gene-expression analysis and network discovery using Genevestigator. Trends Plant Sci 10: 407-409