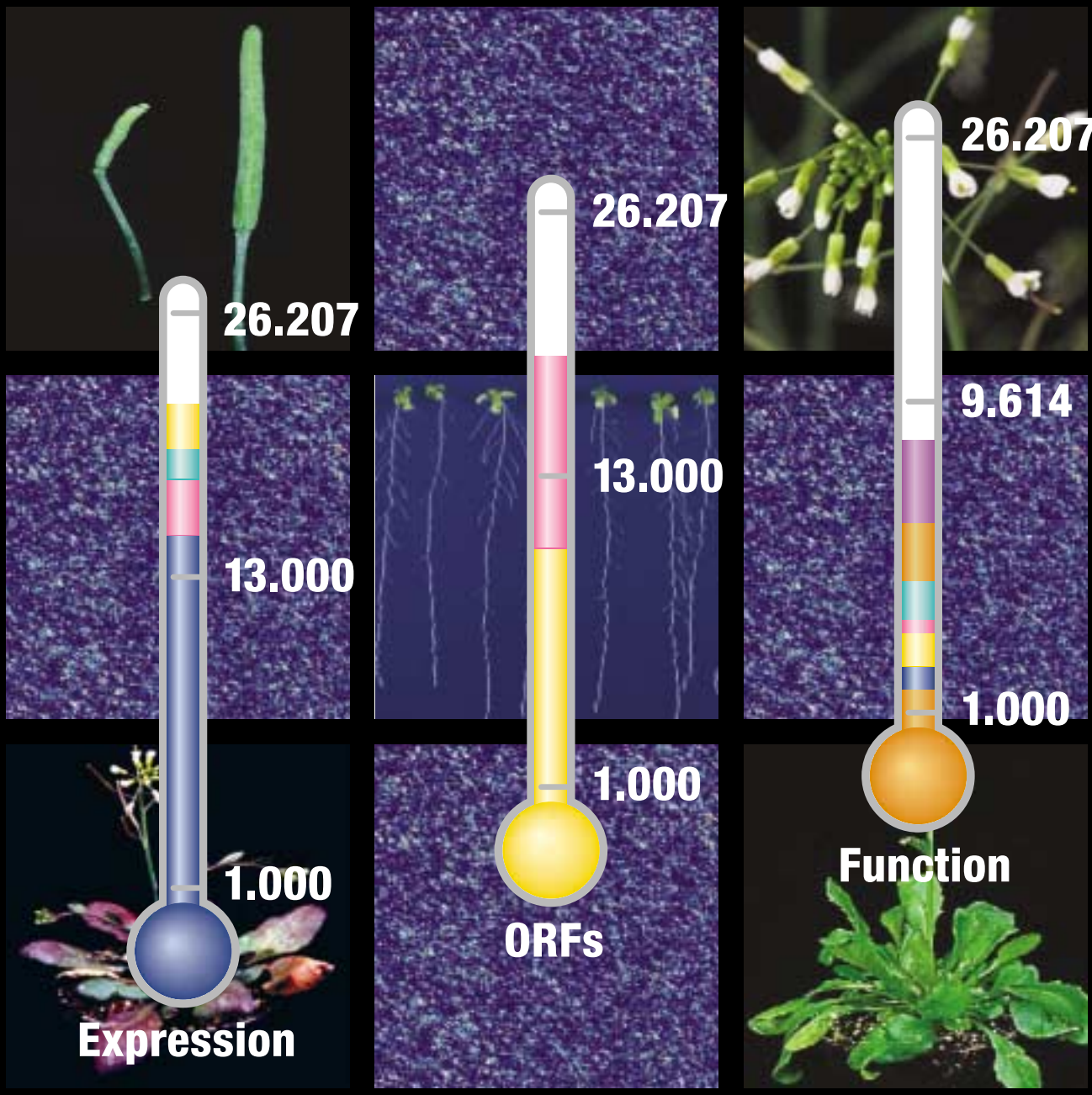
The Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project Annual Report 2004



The Multinational *Arabidopsis* Steering Committee · July 2004

The Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project

Annual Report 2004

Thomas Altmann
Chair

Philip Benfey
Co-chair

Jorge Casal
Bill Crosby
Ian Furner
Jens Freitag
Mary Lou Guerinot
Pierre Hilson
Heribert Hirt
Gerd Jürgens
Jaakko Kangasjärvi
Maarten Koornneef
Ottoline Leyser
Keith Lindsey
Sean May
Peter McCourt
Veronica Ongaro
Javier Paz-Ares
Kazuo Shinozaki
Willem Stiekema
Ian Small
Paola Vittorioso
Geoffrey Wasteneys
Jianru Zuo

Isabell Witt
Executive Secretary and Coordinator

The Multinational *Arabidopsis* Steering Committee · July 2004

Table of Content

Foreword to the Report

This is the third annual report of the **Multinational *Arabidopsis* Steering Committee (MASC) on the status of the Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project**. This 10-year program initiated in 2001 was described in a long-range plan for this new phase of the *Arabidopsis* Genome Research Program in the 2002 MASC report.

The goals of this project are to determine the function of every *Arabidopsis* gene and obtain a detailed and comprehensive understanding of the molecular processes underlying the development, metabolism and interaction with the environment of a flowering plant. The intent is that the knowledge gained on this experimental model organism will serve as the central reference and conceptual framework for all of plant biology. *Arabidopsis* is uniquely situated to play this role for a number of reasons: 1) Its genome is the best characterized among plants, 2) It has the most comprehensive reverse and forward genetic tools and resources, 3) The international research community that uses these tools and resources is among the most active and co-operative, and 4) The solution of most problems in plant biology, whether applied or basic, can be achieved rapidly and in a cost-effective manner through the use of *Arabidopsis*.

The results generated by the *Arabidopsis* Functional Genomics Project, which are made publicly available through central databases, not only provide unprecedented insight into plant function by uncovering basic biological concepts, but also greatly advance our knowledge of the genetic determinants of important traits in crop plants. Studies comparing the physiology, biochemistry,and development of *Arabidopsis* with that of other plant taxa and with economically important agricultural species will be of increasing importance.

The availability of the complete genome sequence of *Arabidopsis thaliana*, the ultimate accomplishment of the previous phase of the Multinational Coordinated *Arabidopsis thaliana* Genome Research Project, provided a "quantum leap" in the information base for plant molecular biological research. On the one hand, this information allowed the *Arabidopsis* research community to develop new approaches and research tools. On the other hand, it highlighted the enormous complexity of the plant biological system and the difficulty of deciphering the function of every gene. Nevertheless, a new goal emerged: "To uncover the mechanisms and processes underlying the spatial, temporal-, and conditional control of the activity of the genes, the identity, function-, and localization of gene products and their interactions in the cellular context, which are the basis of the multitude of cellular, physiological and developmental processes of plant life." To accomplish this goal requires the use of newly developed high throughput technologies, novel experimental tools and comprehensive collections of plant resources as well as powerful procedures for data analysis, storage and display.

Arabidopsis research has provided the cutting edge in generating resources and analytical tools, providing an example for the investigation of other plant species. One of the most important determinants of the success of the *Arabidopsis* Functional Genomics Project is the integration of worldwide efforts. The nature and volume of the proposed work necessitates the marshalling of all resources to attain a maximum level of synergy as well as the avoidance of duplication of effort to enable the *Arabidopsis* community to achieve its ambitious goals. The Multinational *Arabidopsis* Steering Committee plays a key role in supporting this international coordination by collecting and disseminating information from the various initiatives and projects on technology development and functional analyses and by giving specific recommendations for further directions.

As is outlined in this report, the high level of cooperation and widespread willingness to share data throughout the *Arabidopsis* community as well as the support by the funding agencies has yielded important and exciting results. These favorable developments **indicate that the ambitious goal of understanding the function of all *Arabidopsis* genes as a first step toward an in-depth understanding of the biology of higher plants to the benefit of our society can be achieved, if sufficient and sustainable research funding is secured, biological materials and services are made available around the world, and human resources are further developed**.

The Multinational *Arabidopsis* Steering Committee
July 2004

Executive Summary

After completion of the genome sequence of *Arabidopsis thaliana* in the year 2000, the *Arabidopsis* research community put forward the goal of elucidating the function of all genes of the *Arabidopsis thaliana* model organism to provide the basis for achieving full understanding of the biology of a flowering plant. *Arabidopsis* not only serves as a plant model-, but it has, proven to be a key system for uncovering basic biological processes. *Arabidopsis,* therefore, plays an indispensable and unique role as an experimentally tractable system for the advance of basic science. Moreover, due to the dependency on plants of all human and animal life, the results of research on *Arabidopsis* and their application to crop science are of central relevance to our global society.

The major aim of the 10 year Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project is the elucidation of the function of every *Arabidopsis* gene in its cellular, organismal, and evolutionary context, ultimately leading to a comprehensive understanding of the biology of a flowering plant. The research program has been broken down into short-, mid- and long-term goals. Of the short-term goals most have already been achieved, including:

- Genome annotation has been dramatically improved through the use of information on expressed sequences and expert knowledge
- Comprehensive sets of sequence-indexed mutants have been created and are widely used for gene function determination
- Genome-wide sets of gene-specific probes have been created and are in widespread use for expression analysis
- Full-length cDNA sequences have been defined for more than 60% of all *Arabidopsis* genes
- Methods for global metabolic profiling have been established
- Databases have been dramatically expanded providing access to large functional data sets

Of particular note are the availability of full-length cDNA sequences of ca. 16,000 genes, knowledge of expression for ca. 21,000 (80%) of all genes, the identification of T-DNA or transposon insertions in 22,400 genes (ca. 85%), and the establishment of a freely accessible transcriptome reference data set.

The available resources and their widespread use and the strong support of individual projects devoted to gene function analysis has resulted in an **increase of the of the fraction of experimentally studied genes functionally characterized to various extents from 9% in the year 2000 to a current 20%**.

There are a few areas that are lagging behind original plans or are currently underrepresented, particularly proteomics and ORFeome work. Several projects that have recently been initiated to move this field forward should be given further emphasis. Urgently needed by the *Arabidopsis* research community are complete collections of full-length cDNAs / ORF clones and a further development of a comprehensive microarray repository suitable for data mining, as well as several other resources including protein function analysis tools, comparative genomics, improved access to metabolomics technology, and means for facile access to genomic data.

Another area in need of development is the large-scale generation of standardized molecular (transcriptome, proteome, metabolome, etc.) and phenotype information. One way of providing this type of information is through genome technology centers that carry out large-scale, high-throughput analysis programs, place high-quality data freely accessible into the public domain and provide service to individual labs as well as education through courses and workshops.

Several developments, highlighted by multinational co-operation projects such as AtGenExpress and the interaction of funding agencies such as the NSF-DFG co-operation, indicate that the international *Arabidopsis* community is ready to move towards a global research strategy composed of multinational, large-scale projects and numerous individual expert knowledge-driven projects. Given that this transition receives sufficient support, the Multinational *Arabidopsis* Functional Genomics Project will meet its goal of elucidating the function of every *Arabidopsis* gene.

Analysis and Recommendations

After completion of its third year, the Multinational Coordinated Arabidopsis thaliana Functional Genomics Project has moved forward tremendously with unprecedented increase in publicly accessible information. Expression profiling is an example of the dramatic change of paradigms that Arabidopsis research is undergoing. Very few years ago gene expression profiling technology was available in only a few labs and limited to a fraction of the genome. Today, although still costly, genome-wide array analyses are accessible in principle to every researcher! A hallmark of this technique is that the researcher who carries out the experiment will usually use only a minor fraction of the information generated, allowing a large number of colleagues to benefit from access to the complete data set. To make use of this new opportunity of creating synergy and of further accelerating our gain of knowledge, the research community needs to move to a new level of co-operation and further increase its willingness to share data. Building on its history of cultivating a strong spirit of collaboration and free exchange, the Arabidopsis community is well prepared to meet this challenge. This report contains numerous examples of sharing of materials, resources, and data, which are widely used throughout the community. A promising feature of recent years is the growing international cooperation, driven by researchers with common goals and supported by the funding agencies. For example, the recently launched Austrian Platform of Arabidopsis Research (APAR) has been initiated as a further sister program to the US 2010 project and the German AFGN program. The co-ordination among the latter two has also taken another level: This year a common panel of reviewers jointly evaluated proposals submitted to the two programs. This is the first time ever that the two funding agencies, NSF (US) and DFG (Germany) co-ordinated their project proposal reviewing process so tightly.

Current status of the program

This spirit of cooperation has contributed to fulfilling most of the short-term goals of the first phase of the Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project.

In the first three years the major accomplishments include: 1) The release of improved whole genome annotations, the most recent versions supported by full-length cDNA sequences and expert input; 2) Generation of comprehensive sets of sequence-indexed mutants, listed in an integrated database and made available as seed stocks; 3) Implementation of whole genome mapping procedures and development of facile conditional expression systems; 4) The production of genome-wide sets of gene-specific probes for expression analysis; 5) Isolation of full-length cDNAs for more than 60% of the genes, facilitating genome annotation and protein analyses; 6) Establishment of methods for global metabolic profiling and 7) The establishment of MASC with a full-time coordinator to foster information flow, international collaboration and coordination, and to monitor progress of the program.

The improved genome annotation (i.e., identification of genes) is one of the most notable recent achievements. Of critical importance was information from full-length cDNA sequences and from full-genome tiling chip hybridizations. The current release lists a total of 26,207 genes (not including 2,355 transposons and 1,652 non-transposon pseudogenes) of which 16,138 are represented by a full-length cDNA and 20,901 (ca. 80%) have been shown to be expressed. Resources for the characterization of the (biochemical) functions of the gene products have been vastly expanded: Full length cDNA clones of ca. 13,000 genes are being distributed and almost 12,000 ORF clones useful for recombination cloning have been created and deposited at the *Arabidopsis* Biological Resource Center (ABRC). Among the most highly utilized tools for the determination of gene function are vastly expanded resources for reverse genetics. Most notably, the sequence-indexed T-DNA and transposon mutants now cover insertions in 22,400 different genes (ca. 85% of all genes), which probably present null alleles of ca. 70% of all *Arabidopsis* genes. Once an *Arabidopsis* researcher identifies a gene (e.g., by a particular expression pattern) she or he can start to analyze the corresponding knock-out mutant only a few days later. Another boost to the functional analysis of *Arabidopsis* genes is the free access to the rapidly growing genome-wide transcriptome data, which has been started by NASCArrays, the array facility and expression profile repository of the GARNET program. A major milestone in the establishment of community resources in *Arabidopsis* is the recently released AtGenExpress reference transcriptome data set. This data set has been produced and compiled by a multinational initiative and covers genome-wide transcript profiles of all major organs at various developmental stages, responses to diverse environmental stimuli and challenges (abiotic and biotic), and responses to phytohormones. The availability of theses resources and the support of a wide range of individual project on gene function analyses resulted in an increase of the fraction of experimentally studied genes functionally characterized to various extents from 9% in the year 2000 to a current 20%.

As documented in a survey carried out by the North American *Arabidopsis* Steering Committee (NAASC), the *Arabidopsis* community has made extensive use of these resources. Of the resources most frequently cited as key to individual research programs, was the collection of sequence-indexed insertion mutants. Many colleagues also pointed out the fundamental importance of the complete genome sequence and the information and resources provided by the databases and stock centers. Access to and availability of seed and DNA stocks (mutants, accessions/ecotypes, full-length cDNA sequences and gene clones) were also cited as highly important.

In addition to the worldwide access to these resources, the Multinational *Arabidopsis* Functional Genomics Program is characterized by increasing international integration, both at the level of research programs and at the individual project level. For example, the recently launched Austrian Platform of *Arabidopsis* Research (APAR) has been initiated as a sister program to the US AT2010 project and the German AFGN program. The coordination among the latter two has recently become even more integrated: this year's proposals submitted to the two programs were jointly evaluated by a common panel of reviewers. This is the first time ever that the two funding agencies, NSF (US) and DFG (Germany), coordinate so tightly their project proposal reviewing process. The growing international cooperation at the project level is exemplified by AtGenExpress, mentioned above. This initiative has been developed from its inception as an international co-operation, where the participants from Germany, Japan, the UK, and the US used or sought their own funding but agreed to produce a comprehensive, well balanced data set and make it immediately accessible to the research community. It is also an example of the way that biological research in the functional genomics era will be conducted in the future: comprehensive data sets, too large to be produced by an individual lab, will be established by multinational consortia for use by the entire research community. Such global, basic reference data sets will more and more be complemented by the more focused experiments conducted by individual labs. Similar initiatives need to be started for other target areas (see below). Thus, basic global data will more and more become generally accessible for the use by individual researchers, who are experts in certain biological topics and who will use these data sets to deduce novel hypotheses (e.g., on the involvement of certain genes or the contribution of certain pathways) that they test by specific individually designed experiments.

Only a few areas lag behind initial plans or are currently underrepresented. The development of facile technology for heterologous protein expression has not yet been achieved for all proteins. This is an issue that transcends *Arabidopsis* research. However, ongoing efforts to generate ORF clone collections suitable for recombination cloning are a critical first step in addressing this urgent need. Furthermore, major projects on the elucidation of *Arabidopsis* protein structures and functions have been initiated. In addition to transcriptome technologies, proteomics and comparative genomics need to be addressed with more emphasis. Enabling resources for proteomics include: production of antibodies against, or epitope tags on, all deduced proteins and a catalogue of protein profiles at organ, cellular and subcellular levels under a wide range of environmental conditions. For comparative genomics, sequences of related brassica species are needed for the identification of coding and non-coding conserved functional sequences. Another area that has not yet reached the anticipated level is the establishment of genome technology centers that carry out technology development, perform large scale, high-throughput analysis programs, provide service to individual labs, and that offer training and education through courses and workshops.

Recommendations

Based on the analysis outlined above as well as direct community feedback to MASC, comments by the MASC subcommittees, and the aforementioned community survey carried out by NAASC, MASC makes the following specific recommendations:

A number of resources are urgently needed by the *Arabidopsis* community and their establishment as freely accessible materials and information should be given a high priority. These include

- A complete collection of verified homozygous knock-out lines;
- A complete collection of full-length cDNAs/an ORFeome clone collection as an essential foundation and for high-throughput protein analysis and for elucidation of gene/protein functions;
- Further development of a comprehensive microarray repository suitable for data mining. While NASCarray and AtGenExpress are playing major roles in starting to satisfy this need as will CAGE, this has to be expanded and most importantly, ways of access to advanced analysis of these data have to be established;
- Improved capabilities and integration of the *Arabidopsis* database(s), with better ways to locate information and strongly enhanced mechanisms for import of data from individual researchers. In order to do simultaneous queries and analyses on large datasets, these finally have to be merged into in one central public database or a well integrated network of databases.
- Resources for studying protein interaction networks, specifically comprehensive sets of epitope-tagged versions of ORFeome clones;

Further desired resources are (listed here in the order of urgency): Robust RNAi technology for multiple gene "knock-down", a comprehensive set of promoter-reporter lines, more affordable

and complete microarrays including genome tiling arrays, improved metabolomics and proteomics resources including complete antibody collections, high-throughput genotyping methods, ordered overlapping genomic clone collections in versatile shuttle vectors and/or plant autonomously replicating vectors, expanded sets of recombinant inbred lines (RILs) with available suitable molecular markers, and public databases providing access to information such as ecotype and RIL data, insertion mutant, and TILLING resources in different ecotype backgrounds.

Beyond the aforementioned development of resources, novel strategies need to be initiated to be able to achieve the desired level of in-depth knowledge on every *Arabidopsis* gene. Thus, systematic, high-throughput analyses of cellular networks including transcriptional, protein and metabolic networks need to be pursued. A key part of this "systems biology" approach will include subcellular protein localization, protein-protein interactions, protein turnover rates, and protein modifications. To achieve this will probably require, in addition to individual, expert knowledge-driven projects, centralized, large-scale phenotyping projects that apply in a highly standardized fashion (semi-)automated multi-level analytics (such as transcriptome, proteome, metabolome). These analyses will be applied to collections of mutants and genetic variants subjected to various environmental conditions. Like the sequencing projects in the past, these central, large-scale projects can use economics of scale for data generation and have to immediately place extracted, high-quality data freely accessible into the public domain. Due to the high cost of such large-scale projects, they will only be possible by multinational co-operation. This dual research strategy will require tight interaction and collaboration among the researchers as well as co-operation among the funding agencies.

A prerequisite for successful data mining is seamless access to all available information. This can only be achieved via well-connected high-performance databases that contain up-to-date information. *Arabidopsis* databases were improved during the last year and offer access to a rapidly growing set of information, resources, and tools. However, researchers have problems to locate information or to mine data from different sources simultaneously. Furthermore, integration of community input (i.e., improved annotation, localization and interaction of gene products, and functional information) is still slow. Until now, upload into central databases of experimental data provided by individual researchers has been very limited. If at all directly accessible, such information has mostly been deposited in dispersed databases of individual labs. Consequently, the great potential of extracting novel information through mining of existing data will only be exploited to a very limited level if import of data into centralized databases or establishment of seamless links to lab databases is not achieved soon. Some reasons for these limitations could be insufficient communication between the community and the management of databases as well as too little coordination among the various databases. Therefore, MASC will identify the current problems and limitations in detail, and will propose a series of action points to improve the present situation.

A brief history of *Arabidopsis* research and its value as a model system

***How Arabidopsis* became a model**

Genetic and cytogenetic studies on *Arabidopsis* were initiated soon after the re-discovery of the principles of genetics in the beginning of the past century. By 1907, it was already known that *Arabidopsis* contains only n = 5 chromosomes. As early as 1937 *Arabidopsis* was studied for the effects of light on flowering time and seed dormancy. Different ecotypes were also then systematically collected due to the large natural variation in physiological traits among *Arabidopsis* accessions. In 1943, *Arabidopsis* was proposed as a model for genetic and developmental biology given its beneficial characteristics for plant research. Specifically, production of a large number of progeny, rapid development, easy of cultivation in limited space, abundant natural variation, production of fertile hybrids, and a low chromosome number. Later in 1947, it was shown that mutations in *Arabidopsis* could be induced using X-ray irradiation. This discovery prompted many scientists to adopt the weed in their studies. In the sixties, *Arabidopsis*, like other genetically tractable organisms, was employed in the search for chemical mutagens. Consequently, the *Arabidopsis* Information Service, a forum to publish reports on *Arabidopsis*, was established, and a small *Arabidopsis* community came to be. In 1965, the First International *Arabidopsis* Conference was organized. About 25 attendees met in Göttingen, Germany, to discuss *Arabidopsis* research.

However, the widespread adoption of *Arabidopsis* as a model plant really began in the early eighties. Newly developed molecular biology methods encouraged young scientists to address previously intractable problems. Most publications of the time describe mutants affected in various processes. Nevertheless, some key observations were also made such as the estimate of ~70 Mb of nuclear DNA for the *Arabidopsis* genome size, and the discovery that this species has the smallest known genome of any seed plant. Because plant Southern blots and cloning of genes from large genomes was then difficult, *Arabidopsis* small genome size was highly advantageous. In fact, it was claimed that 16,000 λ phage clones could represent the entire genome. In the mid-eighties, large numbers of scientists who previously worked on other organisms turned their attention to *Arabidopsis*. It was then that biologists started to believe that *Arabidopsis* could in fact become a model plant.

An influential review article published in 1987 (Elliot Meyerowitz, *Arabidopsis thaliana*) outlined that the small genome would permit isolation of any *Arabidopsis* gene by positional cloning. Restricted Fragment Length Polymorphism (RFLP) maps, cosmid and Yeast Artificial Clone (YAC) libraries were established and in 1992 the first *Arabidopsis* genes were isolated by map based cloning. From mid- to late-eighties, methods for genetic transformation were developed and the first (small) T-DNA insertion mutant collection was produced. In addition to tissue culture procedures, imbibed seeds with *Agrobacterium tumefaciens* were shown to yield stably transformed plants without the drawbacks of somaclonal variation, caused by *in vitro* culture and regeneration of plant cells. Further improvements of *Agrobacterium* mediated transformation were achieved by the discovery of vacuum infiltration of flowering *Arabidopsis* plants, which resulted in high frequency transformants in the progeny. Refined, this procedure allowed access to easy *Arabidopsis* transformation to almost every laboratory and provided the means to create very large populations of independently transformed lines. In addition to T-DNA insertion mutagenesis, maize transposons were shown to be active in *Arabidopsis* and useful for insertion mutagenesis. These technical advances and early successes solving, until then, intractable problems, highlighted the advantages of *Arabidopsis* as plant model system and triggered biologists around the world to make use of it. The extraordinary usefulness of this model plant was highlighted by the establishment of the ABC model of floral development and isolation of the corresponding genes, the identification of plant hormone receptors and signal transduction factors, plant resistance genes, and the isolation of genes, considered then inaccessible, with metabolic function (e.g., fatty acid desaturases). It became clear that even complex biological processes in plants could be dissected and this resulted in large scale genetic experiments.

In the early nineties, it was evident that *Arabidopsis* was an excellent model system for plant research and the milestones of a typical model organism's career were quickly accomplished. In the last few years, the international conferences on *Arabidopsis* have had increasing numbers of attendees, Cold Spring Harbour Laboratory Courses were implemented, nomenclature guidelines have been produced, and a very active email newsgroup has been established. Also, national and multinational steering committees, whose members are elected via email, were formed to provide organizational support and coordination to the *Arabidopsis* community.

The first report of the multinational steering committee was published in 1990 and stated the goal "to understand the physiology, biochemistry, growth, and development of a flowering plant at the molecular level, using *Arabidopsis* as an experimental model system." For this and other ambitious goals it was necessary to establish stock centres, databases, polymorphism mapping, large insertion libraries and the characterization of mutants by forward genetics. Major milestones during preparation for genome sequenc-

ing were the generation of YAC and BAC libraries and the establishment of physical maps of individual chromosomes or the entire genome. In 1996, the *Arabidopsis* Genome Initiative was established as an international collaboration of scientists from the US, the EU , Japan, and the group's effort resulted in a very successful sequencing programme. The complete *Arabidopsis* genome sequence was published in December 2000, four years earlier than initially expected.

In 1998, the next ambitious goal for *Arabidopsis* as a model organism was formulated during a small workshop in the United States and later validated in 2000 during a National Science Foundation-Sponsored Workshop That is, to determine the function of all *Arabidopsis* genes by 2010. Since the National Institutes of Health (NIH) excluded plant biology from funding, the National Science Foundation (NSF) has taken responsibility for funding *Arabidopsis* research in the United States. NSF adopted important elements of the MASC's proposal that emerged from the workshop, and created the "2010 Project". The *Arabidopsis* Functional Genomics Network (AFGN), funded by the Deutsche Forschungsgemeinschaft – German Research Foundation – (DFG), was subsequently established. The "2010 Project" and AFGN are closely coordinated and share the same goal of determining the function of all *Arabidopsis* genes by 2010. In addition, as described elsewhere in this report, many countries around the world have established national programs focused on *Arabidopsis* functional genomics research.

In a little more than two decades, *Arabidopsis* has established its place in plant research as a model organism and will serve as the blue print of a plant for a long time

Best understood biological processes in *Arabidopsis*

The *Arabidopsis* photoperiod pathway is currently the most completely understood aspect of flowering time control. Although circadian rhythms and light responses in plants have been known for thousands of years, only in the past two decades have the molecular components of light receptors and the circadian clock been identified through genetic approaches. *Arabidopsis long hypocotyls (hy)* mutants were an invaluable source for the identification of light perception and signaling factors. Five phytochromes and their functions have been analyzed in *Arabidopsis* (*hy)* mutants, and the molecular identification of cryptochrome in *Arabidopsis* has led to advances in other organisms in which cryptochromes are responsible for various responses to blue light, including the circadian clock. About 5-6% of the *Arabidopsis* genes are regulated by the circadian clock. The circadian clock is intimately connected with light receptors and with the mechanism by which plants measure day length. This mechanism, in turn, influences many physiological processes, including the time at which flowering occurs.

Genes involved in abiotic stress signaling were investigated in detail in *Arabidopsis* especially, genes involved in heat stress, salt, drought, and cold stress have been described extensively.

Plant-pathogen interactions have been studied in many plant species. The availability of the complete *Arabidopsis* genome sequence, however, made possible the cataloguing of all gene sequences related to known resistance (R) genes. Most notable is the degree of polymorphism observed between *Arabidopsis* accessions. Comparison of R gene loci has shown the deletion and duplication of specific R genes. Extreme divergence of R gene haplotypes was found even between laboratory strains, like *Columbia* and *Landsberg erecta* which on average show less than 0.1% nucleotide sequence divergence. An important step from quantitative trait loci (QTL) to the relevant genes is the separation of an individual QTL from other segregating loci to obtain genotypes with monogenic segregation. Such a "mendeliazation" of a QTL can be achieved by constructing near isogenic lines (NILs) and was first performed in *Arabidopsis* for loci controlling flowering time. Polymorphisms causing single amino acid substitutions in the *Arabidopsis* photoreceptors CRY2 and PHYA could be linked to QTLs of flowering time and to natural variation in light sensitivity, respectively. The biochemical analysis of these natural variants identified properties of the corresponding proteins which explain their different behavior with respect to flowering time and light sensitivity. Studies in this area have shown how the plasticity in plant growth and development among different *Arabidopsis* accessions can be used advantageously to dissect plant biology, and how it will ultimately lead to a better understanding of natural selection. Transcriptional and post transcriptional gene silencing, the two best investigated epigenetic mechanisms, have been studied in many organisms. Molecular studies in plants have been mostly done in tobacco and petunia but the accompanying genetic studies have also been conducted in *Arabidopsis*.

Arabidopsis mutants are a powerful pool to dissect gene functions in many aspects of a plants life.

Pattern formation and cell type specification in leaf epidermis cells were studied on various trichome mutants. In *Arabidopsis*, trichome development and root epidermal patterning use a common mechanism involving closely related cell fate transcription factors and a similar lateral inhibition signaling pathway. Yet the resulting patterns differ substantially, primarily due to the influence of a prepattern derived from subepidermal cortical cells in root epidermal patterning.

Flowering plants have the unique ability to produce new organs continuously from stem cell populations maintained at their actively growing tips. Stem cell maintenance is an active process, requiring constant communication between different regions of the shoot apical meristem to coordinate loss of stem cells from the meristem through differentiation with their replacement through cell division. *Arabidopsis* mutants with altered meristem cell identity or accumulation are viable, allowing dissection of stem cell behavior by using genetic, molecular, and biochemical methods. A range of *Arabidopsis* mutants affecting key stages in meiosis have been identified using a combination of screening for plants exhibiting reduced fertility and, more recently, using a reverse genetics approach. These are now providing the means to identify and characterize the activity of key meiotic genes in flowering plants.

Genetic screens on *Arabidopsis* mutants impaired for hormone perception or signal transduction have been extremely useful in identifying plant hormones receptors and genes involved in hormone signal transduction.

The recent finding that miRNAs in *Arabidopsis* are involved in developmental processes by targeting mainly transcription factors has changed our view on how gene expression is regulated. Dis-

covery of hundreds of putative miRNAs in *Arabidopsis* was possible because the genome is known. Hence, miRNAs serve as a good example for new insights into biological functions we might not have seen without knowing the sequence of the genome.

Testing *Arabidopsis* genes in crops

The central role of *Arabidopsis* in driving innovation in applied plant biology cannot be underestimated. In fact, the July 2004 issue of the journal PLANT PHYSIOLOGY is focused on the importance of this plant in translational biology. The following examples illustrate that genes that have been functionally characterized in *Arabidopsis* bestow the same function to crops upon heterologous expression.

GIGANTEA (GI) is a gene involved in photoperiodic flowering and controlling of circadian rhythms (e.g., gi mutants of *Arabidopsis* exhibit delayed flowering). Early bolting in radish is a problem of agronomic dimension in Asian countries. Because of the taxonomic closeness between *Arabidopsis* and radish, a delay in bolting and flowering in radish were achieved by transferring an *Arabidopsis* antisense *GI* gene fragment into the crop to down-regulate the expression of native *GI*. In *Arabidopsis*, giberellin signalling is mediated via *GAI*, a nuclear member of the *GRAS* transcription factor family. This orthologue of the green revolution gene of wheat and rye is a GA derepressible repressor of plant growth that caused higher yield on the expense of short straw in wheat. Transgenic expression of the *Arabidopsis GAI* and gai confers altered giberellin response in tobacco and rice, and causes dwarfism even in these comparably unrelated plants.

Another hormone, ethylene, requires specific receptors for perception and signal transduction pathways to coordinate downstream responses. *Etr1-1* encodes a mutated receptor that confers dominant ethylene insensitivity in *Arabidopsis* but causes significant delay in fruit ripening in tomato and petunia.

Constitutive expression of the *Arabidopsis* gene *LEAFY (LFY)* is enough to trigger the transition from the vegetative shoot apical mersitem to an inflorescence meristem, and to cause early flowering. Also, the function of *LFY* is highly conserved in unrelated plant species (e.g., the expression of the *Arabidopsis* gene from the same viral promoter results in a similar phenotype in transgenic aspen trees).

In general, tomato is considered a chilling sensitive plant. Like in *Arabidopsis*, constitutive expression in tomato of the *Arabidopsis* C-repeat/dehydration response element binding factor 1 (CBF1), a transcription factor of the AP2/EREBP family, confers elevated tolerance to chilling, oxidative stresses, and water deficit stress in tomato and canola. *Arabidopsis CONSTANS (AtCO)*, a putative transcription factor that accelerates flowering in response to long days, impairs tuberization under short-day inductive conditions when expressed constitutively in potato. Grafting experiments using these lines indicated that *AtCO* exerts its inhibitory effects on tuber formation by acting in the leaves. *CONSTANS* might be involved in generating the elusive and long-distance acting florigen-tuberigen signal(s) in the leaves.

One mechanism by which plants could survive salt stress is to compartmentalize sodium ions away from the cytosol. Overexpression of *AtNHX1* has been correlated with higher vacuolar NA+/H+ antiport activity and confers normal growth and development in plants watered with 200 millimolar sodium chloride.

The genes mentioned above were already known before *Arabidopsis* was fully sequenced, and before high throughput functional genomics was applied to *Arabidopsis*. Still, they give us a taste of the tremendous potential of pool of genes from model plants, characterized during the global functional genomics era, which may be used in the near future to alter plant properties.

Arabidopsis continues to play an important role in the discovery of genes that can favorably impact the nutritional quality of crop plants. In a recently published example, map-based cloning in *Arabidopsis* identified the *VTE3* gene, encoding the Vitamin E biosynthetic enzyme 2-methyl-6-phytylbenzoquinol methyltransferases. This enzyme was long sought after because it was proposed to be able to convert delta-tocopherol, which accumulates to relatively high levels in soy and other oilseed crops, to the more biologically active gamma-tocopherol. Transgenic soy plants that expressed the *VTE3* gene during seed development were found to no longer accumulate delta-tocopherol, thus improving the nutritional value of the soy oil. Co-expression of *VTE3* and *Arabidopsis* gamma-tocopherol methyltransferases *(VTE4)* in the seed caused virtually 100% alpha-tocopherol accumulation, which has the highest biological activity of the naturally occurring tocopherols. Thus, introduction of two *Arabidopsis* genes created soy oil with a much higher nutritional quality than in standard varieties, a result with nutritional and food quality implications.

Synteny and colinearity between *Arabidopsis* and other plants

The use of model species in biological research is based on the assumption that many of their features are shared among a wide range of related taxa. Consequently, it is expected that many of the genes associated with important traits in crop plants can be identified via homology to their counterparts in *Arabidopsis*. In addition to a high degree of conservation of individual gene sequences throughout the plant kingdom, comparative genomics has revealed a high degree of conservation in genome structure (synteny) among closely related taxonomic groups. Our current knowledge on synteny indicates that, despite plasticity contributing to the diversity of plant genomes, the organization of genes is conserved within large sections of chromosomes. This fact constitutes another validation of the considerable efforts made on model plants. Our understanding of plant genomes gained through model plants has fostered a massive surge in plant biotechnology, which is currently changing our vision of crop production and protection. Indeed, such technological progress presently enables the insertion of useful genes into crop plants, at a fast rate and in a much more precise manner than with conventional genetic methods.

Colinearity refers to a certain degree of conservation of gene content, order and orientation between chromosomes of different species or non-homologous chromosomes of a single species. The investigation of microsynteny requires sequencing and annotation of genomic DNA, enabling direct comparison of the sequences using

various computational tools. The completed *Arabidopsis* genome sequence and growing lists of genomic resources for other plants have been of incredible benefit to comparative genomics research. *Arabidopsis* genomic segments exhibit extensive colinearity with genomic regions of the closely related genera Brassica and Capsella. However, even when compared to more distantly related species, *Arabidopsis* has shown some degree of conserved synteny.

Transfer of knowledge from *Arabidopsis* to crops

Day length provides an important environmental cue by signaling conditions favorable for flowering. Knowledge of flowering in *Arabidopsis* was relevant for understanding how flowering is controlled in rice. That is, while *Arabidopsis* promotes flowering in response to long days, rice promotes flowering in response to short days. Distinct photoperiod responses in these two plants involve related components acting in the same sequence. The key difference is that the activity of the transcription factor CONSTANS is reversed under long day conditions in rice compared to *Arabidopsis*. This breakthrough result was built on two significant areas of work. First, the elucidation of the photoperiod pathway in *Arabidopsis* and, second, the genetic analysis of rice quantitative trait loci (QTL) affecting flowering time. The realization that the photoperiodic pathway is highly conserved in rice provides a new grasp on the strategic control of rice flowering.

The adaptive value of flowering time control reveals an agriculturally important trait at the heart of breeding programs. The modification of flowering time is one of the most important properties governing geographic distribution of crop plants. It is possible that, in some parts of the world, the acceleration of rice flowering time could enable the growing of two crops per year on the same piece of land where, currently, only one is cultivated. The exploitation of allelic variation in genes of this pathway provides one way to manipulate this process through marker-assisted breeding programs while the use of predictable transgenic approaches provides another. Therefore, the genetic study of photoperiodism in the model plant *Arabidopsis* provides a meaningful knowledge base for a crop that contributes significantly to world food security.

Other important factors targeted by biotechnological approaches are dispersion and seed dormancy. For example, in oilseed rape, the seeds that fall to the ground during dehiscence represent a loss of almost 20% of the harvest for the farmer. To address this problem, the genes responsible for silique opening during dehiscence have been identified in *Arabidopsis*. Recent research has shown that it is possible to genetically control dormancy, with regard to hormone biosynthesis (i.e., abscisic acid and gibberellins) and to transcription factors involved in the phenomenon.

An omega-3 desaturase has first been identified by map based cloning and characterization in *Arabidopsis*. The gene was used to isolate homologous genes from other organisms that were exploited for genetic engineering of multiple unsaturated fatty acids in plants.

The examples given here are not reflecting an exhaustive overview but rather serve to give an impression of how important it is to have easy access to genes and their characterization. This is only possible with an expanded tool kit as it is available for *Arabidopsis*.

Why are model plants essential for our survival

Plants are essential for our existence on Earth. Photosynthesis provides the biological energy that fuels our world and is responsible for the oxygen and carbon dioxide cycling that makes our life possible. However, plants do a lot more than photosynthesis. They provide essential nutrients and vitamins, they are an invaluable source of medicinal compounds (or lead structures), and they provide fibers and wood for clothing and constructing. Because of their sessile life style plants have developed unique properties. Instead of a skeleton, structural components such as fibers give plants stability. Many aspects of signal transduction are different in sessile multicellular, photo-synthetically active organisms as compared to animals. Differences in defense and detoxification strategies led to an invaluable variety of secondary metabolites. These active compounds constitute a large pool for pharmaceuticals, now and in future. In addition, healthy food is found in many plants.

Arabidopsis is an ideal dicotyledonous model to gain a principal understanding of pathways for primary and secondary metabolites. Based on the knowledge acquired through this model plant, the investigation of compounds in more complex plants of interest, gains a lot of speed. Genetic engineering provides the means to produce these compounds in different plants that are adapted to certain climates, grow fast and produce high biomass.

Plants constitute the essential and main food resource for animals and humans. This is why the foremost mission of agriculture is to produce plants in sufficient quantities and at high quality to respond to the absolute necessity of feeding the world. Today, this problem has become acute in the face of demographic developments, erosion of arable land, intensive farming which environmental damage, and increasing climate changes. There is substantial evidence to indicate that significant global warming will occur during the 21st century. Climate change could lead to harsher winter weather conditions, strongly reduce water availability, and intensify winds in regions that currently provide a significant fraction of the world's food production, including Europe and the United States. With inadequate preparation, the result could be a significant drop in the human carrying capacity of the Earth's environment. Deep understanding of at least one or two model plants, enriched with specific knowledge about crops and coupled to enhanced breeding procedures, including plant genetic engineering, is most likely the only option to address these problems with the necessary speed.

Note.: Because we were not able to mention here all references that were used to support this article, the reference list will be available on the MASC web page shortly after the annual Arabidopsis conference in Berlin.

Progress and Activities of the Multinational *Arabidopsis* Steering Committee (MASC)

The *Arabidopsis* community continues to grow. Approximately 13,000 researchers around the world, affiliated with over 4,900 laboratories, are currently engaged in unraveling the function of the *Arabidopsis* genome and applying the knowledge gained to other plants. Currently, the *Arabidopsis* Information Resource (TAIR) (http://*arabidopsis*.org) has over 13,000 registered users.

In 2003, the Multinational *Arabidopsis* Steering Committee (MASC) was very active in establishing and maintaining communication within the large *Arabidopsis* functional genomics community, ensuring information availability to the *Arabidopsis* research community and to the biological research community at large, and coordinating *Arabidopsis* functional genomics activities around the world. In fact, the MASC has met all short-term goals identified at the onset of the Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project with positive results for the community.

A full-time coordinator has worked for MASC since 2002. The coordinator's work was supported in 2002 and 2003 by a NSF grant and through supplemental support for traveling by several MASC member countries. In 2004, the MASC coordinator is being supported by DFG and is located in Germany. Isabell Witt is the present coordinator. Her tasks include the organization of the 15th International *Arabidopsis* Conference in Berlin and communication and coordination of efforts within the MASC, and between the MASC and the *Arabidopsis* research community. She has also overseen the publication of this Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project report and maintains the internet homepage for MASC at TAIR (http://www.*arabidopsis*.org/info/2010_projects/).

The MASC internet site contains project and resource information for scientists actively engaged in *Arabidopsis* functional genomics research as well as information for those seeking to learn about the progress of the *Arabidopsis* Functional Genomics Project, the MASC, and the *Arabidopsis* research community in general. Users of the MASC homepage are able to search for genes under investigation by individual functional genomics project, or for the projects themselves by name, principal investigator or gene. The MASC website was last updated and extended in December 2003. It will be updated again in September 2004. The expert staff at TAIR maintains the search capability and has been actively involved in getting the MASC site launched.

One of the many positive results of MASC's work in 2003 for the *Arabidopsis* research community was the growth in communication within the community. Awareness is increasing continuously among *Arabidopsis* researchers and other biologists, not only about the activities of the *Arabidopsis* functional genomics community but also about the multinational and inclusive nature of the MASC and its efforts. Representatives of each of the subcommittees and multinational community are encouraged to be in frequent contact with *Arabidopsis* colleagues from their country or region who are contributing to research in their fields. In turn, members of the *Arabidopsis* community are encouraged to be in contact with their MASC or subcommittee representatives and communicate where their research fits into our community efforts and what they identify as needs or new opportunities. As in the years before, also in 2004 a letter was prepared by the MASC coordinator and distributed to every registered TAIR user and via the *Arabidopsis* News Group, explaining the purpose of the MASC and inviting everyone to make use of the resources being made available for *Arabidopsis* research.

There is growing interest from scientists around the world to participate in the MASC and to establish *Arabidopsis* functional genomics research in countries which currently lack active involvement in this field. New contributors to the MASC in 2004 include, for example, representatives from Eastern European countries. They have organized themselves in the Eastern European *Arabidopsis* Activity (EEAA), composed of 18 scientists and their groups from seven countries: Czech Republic, Hungary, Lithuania, Poland, Russia, Uzbekistan and Yugoslavia. Similarly, South Africa has initiated contact with other African countries in order to build an African *Arabidopsis* community.

Despite new forms of electronic communication, we are all aware of the many social, cultural and political forces that strive to divide and separate us. *Arabidopsis* has provided a means for unifying plant scientists all over the world and continues to offer an important motivation to ensure the free exchange of information and materials across borders. Examples of such exchanges are present throughout this report, including database links and the public availability of large expression profiling data sets such as AtGenExpress.

Highlights of the past year

The Arabidopsis transcriptome, a new community resource

AtGenExpress is a multinational project that established a reference transcriptome data set covering gene expression profiles of *Arabidopsis* plants and their organs at different developmental stages, of plants subjected to various different environmental conditions, and of responses to phytohormones. This initiative was started in the frame of the AFGN program through recognition that every project devoted to the elucidation of gene function(s) needs basic, global information on the gene(s) expression profile(s). With the availability of a worldwide accessible, highly standardized genome-wide expression profiling technology provided by the Affymetrix Ath1 gene chip methodology, the essential prerequisites for the establishment for a widely usable reference data set were given. According to the enormous volume of the experiments needed to cover a basic set of expression profiles, it was immediately obvious that such a project had to use all resources that could be mobilized worldwide for this common goal and had to avoid duplication of efforts. Thus, an international consortium of groups agreed to cooperate towards the common goal of setting up such a freely available transcriptome data resource within less than one year using already available or newly raised funds. This consortium is co-ordinated by Lutz Nover, Detlef Weigel and Thomas Altmann, who received funding from the DFG to conduct a total of ca. 500 chip hybridizations and for an additional ca. 100 chip hybridizations to be performed at NASCArray (with substantial support through a GARNET grant by BBSRC). The corresponding RNAs and experiment descriptions (MIAME) are provided by the groups of D. Weigel, B. Weisshaar, and D. Twell (developmental series), J. Kudla, H. Puchta, D. Bartels, K. Harter, P. v. Koskull-Döring (abiotic stress responses), T. Kretsch (light responses), and T. Nürnberger (responses to selected pathogen infections). Chip hybridizations are carried out by the Weigel lab (Tübingen), the Deutsches Ressourcenzentrum für Genomforschung GmbH (RZPD, Berlin), and NASCArray (Nottingham). Further members of the AtGenExpress consortium are the NSF funded 2010 project "Expression profiling of plant disease resistance pathways" led by X. Dong, F.M. Ausubel, and S. Somerville (responses to a broad range of pathogen infections), and the RIKEN plant science center groups represented by S. Yoshida with data contributed by Y. Shimada, E. Nambara, I. Yamaguchi, and H. Takahashi (phytohormone responses, seed germination and nutrient starvation). Supported by NSF, all collected data are integrated and displayed for public access at TAIR led by S. Rhee. In addition to this core of AtGenExpress groups, further contributions are provided by the labs of C. Somerville (generated as part of the NSF 2010 project "Identification of the function of a family of putative glycosyltransferases"), M. Stitt (diurnal cycle and nutrient conditions), and another a large data set is supplied by the group of W. Gruissem. These additional contributions are excellent demonstrations of how AtGenExpress should further develop: Starting out from the core (reference) data set, this repository should grow rapidly through submission and integration of compatible data collected throughout the entire *Arabidopsis* community.

NASCarray (http://affy.*arabidopsis*.info) offers already free public access to 1000 Affymetrix chips experiments covering more than 60 different biological associations. Another large project, CAGE http://www.psb.ugent.be/CAGE/objectives.htm, carries out hybridizations of 2000 biological samples in twofold repetition on 4000 arrays based on CATMA-GSTs. These data will be available by 2005.

AtGenExpress reference transcriptome data (core project supported by DFG, NSF, BBSRC, RIKEN and Max-Planck-Society)

Type of experiments	Number of experiments	Number of replicates	Total Number of Affymetrix Ath1 chips	Principal Investigator or Institution
Developmental series	64	3	192*	D. Weigel, B. Weisshaar, Germany, D. Twell, United Kingdom
Abiotic stress	151	2	302*	J. Kudla, H. Puchta, D. Bartels, K. Harter, P. v. Koskull-Döring, RZPD, Germany
Light regimes responses, pathogen infections	46	3	138*	T. Kretsch, T. Nürnberger, Germany, NASCArray (GARNet), United Kingdom
Pathogen infections	80	3	240**	X. Dong, F. Ausubel, S. Somerville, USA
Phytohormones, seed germination, and nutrient starvation		2-3	Several hundred**	S. Yoshida, Y. Shimada, E. Nambara, I. Yamaguchi, H. Takahashi, RIKEN, Japan

The AtGenExpress data become publicly and freely available at TAIR (http://*arabidopsis*.org).
*Data set(s) fully completed / **partially completed at the time of publication of this report.

Uncovering the hidden transcriptome

Functional analysis of a genome requires accurate gene structure information and a complete gene inventory. A novel strategy was used to verify and correct the initial genome sequence annotation of the reference plant *Arabidopsis* and identify thousands of new transcription units (Yamada *et al.*, 2003). A set of 12 Affymetrix genome tiling arrays that contain nearly the entire genome sequence was hybridization with RNA populations from various tissues. This study allowed correction of the annotation of thousands of gene structures. In addition, 5817 novel transcription units were identified, including a substantial amount of antisense gene transcription (~30% of all genes show anti-sense transcription), and 40 genes within the genetically defined centromeres. The use of unbiased whole genome tiling arrays has revealed the presence of a "hidden transcriptome" where approximately 25% of transcription is derived from the previously unannotated intergenic regions. The study also resulted in completion of 30% of the *Arabidopsis* ORFeome as a resource for global functional experimentation of the plant proteome. This approach is now being adopted for genome annotation and novel transcription unit identification for the *Drosophila* and human genomes.

Measuring the gene function knowledge

In the summer of 2003, the 14th International Conference on *Arabidopsis* Research took place in Madison, Wisconsin, USA. During the MASC annual meeting held during that conference, it was agreed that a better update on gene functions and quantification for how many genes the function is known for would be established. Hence, functional categories have been defined for easier quantification.

1. For genes that encode a protein
- Protein activity/ Molecular function (catalytic or otherwise/e.g., kinase, chaperone, phosphatase, proteinase) We should be using the GO ontology, trait ontologies for functional categorizations
- Tertiary structure
- Post-translational modification data
- Expression pattern of protein at cell and tissue level
- Subcellular localization
- Protein interaction data
- Phenotype of genetic knockout/ other loss-of-function alleles
- Biological processes (e.g., photosynthesis, amino acid metabolism, cell wall biosynthesis)

2. For genes that do not code for a protein
- Activity of RNA/ gene product
- Expression pattern at cell and tissue level
- Structure of RNA/ gene product
- Subcellular location for RNA/ gene product
- Phenotype of genetic knockout/ other loss-of-function alleles
- Interaction data

The Gold Standard of a genes functional characterization is reached when we have full information for each of these categories.

The opposite of the Gold Standard is we "don't know anything"
- Sequence has no homology to any sequence that we know the function of
- ORF has no expression
- no cDNA has been isolated, just predicted

Although it is evident that the Gold Standard is a very high standard, the final goal is that it will be applied to all genes. In the next seven years, it should be possible to collect at least one category for every gene in the genome. As shown in the thermometers below, expression patterns for more than 80% of the genes can be extracted by the various expression profiling experiments, although not many at the cellular level. The progress made in each of the categories will be measured and illustrated with thermometers in the subsequent annual MASC reports as well.

For this year's report, Chris Town and Hank Wu from TIGR provided the actual numbers of genes that fall into different evidence codes, genes for which there are fullength c-DNAs, genes that have been detected in various expression profilings, and genes for which there was experimental evidence in the literature for a function. Information for the thermometers was also obtained through the *Arabidopsis* community. A questionnaire was sent to 2010/AFGN researchers by the MASC coordinator about the categories listed under 1.: functional categories for genes that code for proteins. Fifty three 2010/AFGN projects supplied data which was forwarded to TIGR to be filtered for redundancy with other data. New non-redundant information was integrated into the thermometer and called community input (CIP). TIGR provided numbers of genes for which "we don't know anything about".

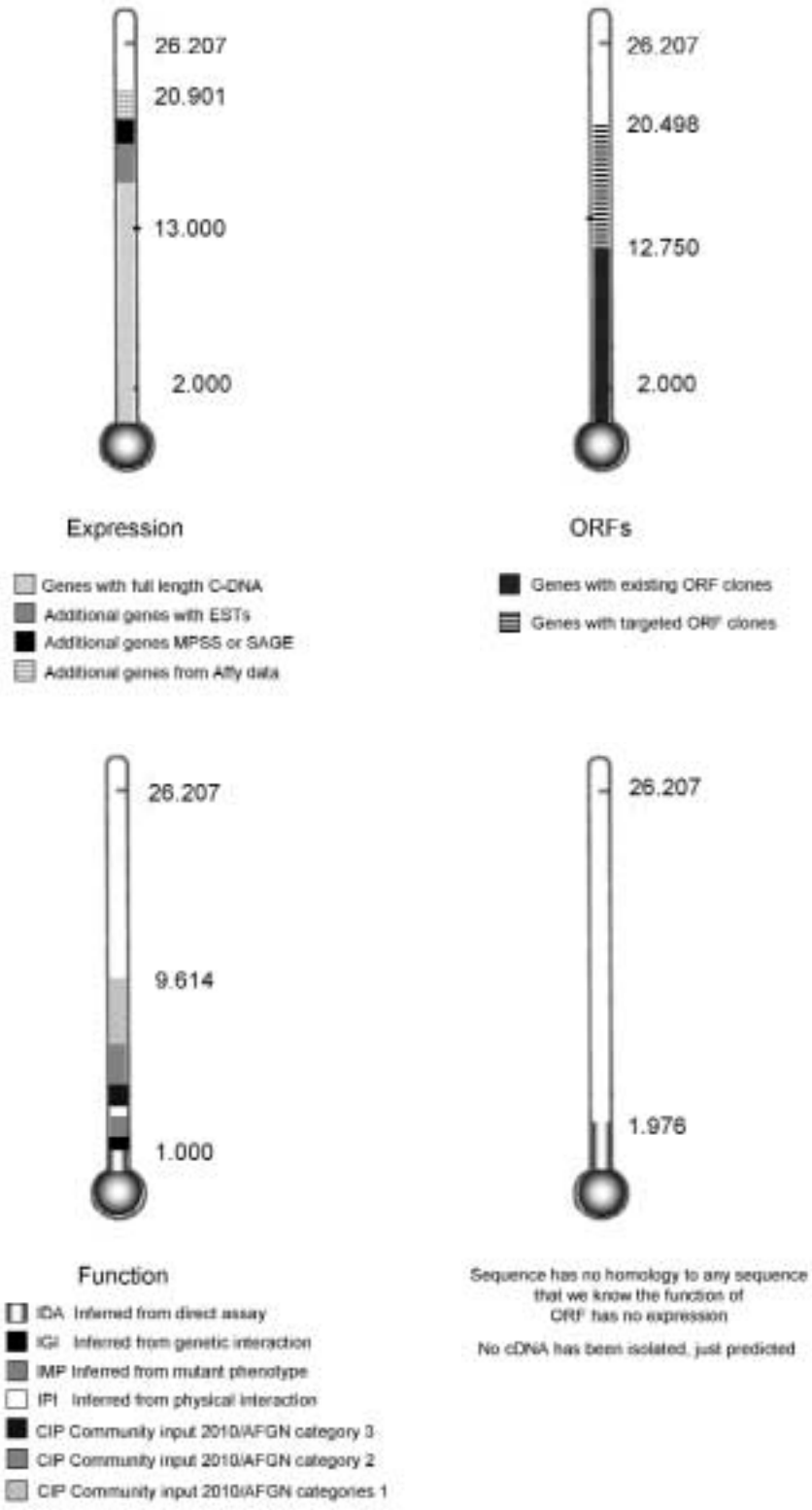


Figure 1: Measures of knowledge on *Arabidopsis* genes. Exact numbers for the different categories are as follows: Genes with full-length cDNA (16,138), additional genes with ESTs (2,701), additional genes from MPSS or SAGE (397), additional genes from Affymetrix data (1,665). Genes with existing ORF clones (12,750), genes with targeted ORFs for cloning* (7,748) Gene function IDA (889), IGI (135), IMP (342), IPI (45), CIP class 3 is when a characterization is almost finished (656), CIP class 2, genes have been partially (3,110) and CIP class 1, genes were selected for characterisation but have not been characterized yet (4,437). Please note that gene accessions were compared for redundancies. The numbers in each thermometer refer to non-redundant gene accessions. The total number of genes "we don't know anything about" and that have only been predicted by computational methods is 1,976.

* genes targeted by Ecker's group (list provided to TIGR), CESG, Wisconsin (list provided to TIGR), ORPHEUS Group (data extracted from ORPHEUS DB, Ghent), Atome (taken from their web site), TIGR, and a few other minor sources.

Reports of the MASC Subcommittees

Bioinformatics

Prepared by Chris Town, Chair
MASC Bioinformatics Subcommittee

The last meeting of the Bioinformatics subcommittee took place in 2003 during the annual *Arabidopsis* conference in Madison. Issues discussed at that occasion related to ways in which major bioinformatics outlets can benefit from and be a service to the functional genomics community. Specifically,

1. The need for a centralized information source (a - what is the scope of the activities and datasets at the major centers? and b - how should disparate pieces of functional genomics information be gathered into a "one-stop" resource for community?).
2. Agreement both on criteria for assignment of gene function and on a more standardized nomenclature to minimize or eliminate the use of ambiguous and/or overlapping terms (e.g., "unknown protein," "hypothetical protein," and "putative protein") at different bioinformatics outlets.
3. How can individual “expert” databases (or the information therein) in functional and other genomics projects be best captured by the major community databases.

The European view on this third point, materialized by the PLANET-project (http://mips.gsf.de/proj/planet/), is to promote expert databases by ‘keeping the experts with their database’ and to develop an exchange protocol in order to generate a federated system accessible to everyone that can be integrated through a central protocol and website for viewing data. This protocol, called BioMoby, enables data and services offered by a group or institution to be transparently used and displayed by others on their website.

4. Develop mechanisms for discovering the types of bioinformatics capacity and tools needed by the *Arabidopsis* community while maintaining the bioinformatics community abreast of tools and capacity being developed by the functional genomics projects.

Therefore, the following on-going goals have been derived for the *Arabidopsis* bioinformatics community:

- Establish and maintain stronger interactions/connectivity between the major databases (see note on PLANET above).
- Survey and compile a comprehensive list of functional genomic resources that is web-accessible, searchable and extensively hyperlinked.
- Define "unknown function" and establish parameters for assigning a function to a gene.
- Establish parameters for minimal data content for submission of expert data sets into backbone databases.
- Create "exit strategies" for functional genomics projects to plan for successful integration of expert database contents into backbone databases.
- Make community and funding agencies aware of the need for exit strategies and encourage funding agencies to require an explicit exit strategy.
- Utilize backbone genomics sites as a repository for output by projects.
- Keep the *Arabidopsis* community at large informed about the activities of the bioinformatics community.

Some additional issues deserve our consideration in the near future. For example, how to encourage people to adopt and use easy formats to exchange various data. Technically, this means enhancing the use of XML format for data exchange. Also, standards for describing data should be promoted. Standards would solve the problem of having various terms to describe "unknown protein", including GO annotations which are not perfect at the moment but will soon become a standard. To move one step further, we should discuss as well the use of "web services" which would allow for decentralized services appearing through one web portal. It is the subcommittee’s opinion that one centralized big database may not be the best option for the future (see for example the European initiative PLANET).

The MASC Bioinformatics subcommittee will meet again during the next *Arabidopsis* annual conference in Berlin to address the issues mentioned above and formulate an action plan. The *Arabidopsis* community is invited to participate and contribute to the discussion.

Currently, these are the bioinformatics resources specifically serving the *Arabidopsis* research community:

- AGR at UK Cropnet (http://ukcrop.net/agr/) and at NASC have been replaced by http://atensembl.*arabidopsis*.info. AtEnsembl is an integrated genome browser that displays both MIPS and TIGR annotations in-line with NASCstocks, NASCarrays and CATMA amplicons. AtEnsembl is based on EBI ENSEMBL software.

- GABI-Info and GABI-Primary Database (links from http://www.gabi.de)
- Genoplante-Info (http://genoplante-info.infobiogen.fr/)
- Kazusa Department of Plant Gene Research (http://www.kazusa.or.jp/en/plant/) (plus ESTs)
- MAtDB at MIPS (http://mips.gsf.de/proj/thal/db/index.html)
- NCBI (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/map_search?chr=arabid.inf)
- RIKEN Genomic Sciences Center (http://pfgweb.gsc.riken.go.jp/) RIKEN *Arabidopsis* Genome Encyclopedia (RARGE) (http://rarge.gsc.riken.go.jp/)
- SIGNAL (http://signal.salk.edu)
- TAIR (http://*arabidopsis*.org)
- TIGR (http://www.tigr.org/tdb/e2k1/ath1/)
- VIB Department of Plant Systems Biology, Bioinformatics and Evolutionary Genomics (http://bioinformatics.psb.ugent.be/bioinformatics/)
- MIAMEplant - a Swiss/UK/US consortium deriving controlled vocabularies for Plant microarray data. Developmental/morphological ontologies are being standardized through /http://www.plantontology.org /
- PLANET has been started to bring together the main European *Arabidopsis* Data providers: http://www.eu-plant-genome.net/partners.html

Representatives of MIPS, TIGR and TAIR also met in Madison to discuss ongoing coordination of AGI identifier assignments. It was agreed that TAIR would become the curator of AGI identifiers in the future. Other centers such as TIGR and MIPS would submit lists of new genes with proposed identifiers for approval before incorporating them into their annotation and making them available to the public. Later in 2003, VIB released its 'own' structural annotation, generated with the software package EuGene, performed on the TiGRv4 pseudomolecules within PLANET. In January 2004, TIGR made its fifth whole genome annotation release and has now turned over the primary responsibility for maintaining all *Arabidopsis* genome annotation to TAIR. The VIB Department of Plant Systems Biology is currently working on a whole genome structural annotation on the TiGRv5 pseudomolecules. This release will be made available as soon as possible through TAIR for assignment of AGI-codes to new genes.

cDNAs and Clone-Based Functional Proteomics (ORFeomics)

Prepared by Pierre Hilson, Chair
MASC cDNAs and Clone-Based Functional Proteomics (ORFeomics) Subcommittee

Collectively, the *Arabidopsis* community has now gathered full-length (fl) cDNA sequence information for about 16,000 of the 26,207 protein-encoding genes, excluding transposable elements and pseudogenes, identified in the latest TIGR nuclear genome annotation (January, 2004; release 5.0). This experimental confirmation of gene models is crucial for a high quality annotation because, in many cases, the predicted models are corrected by the actual transcript sequences and, in other cases, some transcription units are simply not predicted at all. However, the isolation of novel fl cDNA clones becomes more laborious as it focuses increasingly on genes expressed at low level, in particular conditions or in specific cell types. Consequently, alternative approaches are welcome at this stage of the genome structural annotation. Such an example is the use of transcript profiling tiling arrays or the systematic RT-PCR amplification and sequencing of cDNAs based on predicted gene models. New methods to intentionally capture cDNAs originating from uncharacterized transcription units will soon be necessary as the fraction of genes lacking experimental expressed sequence data narrows down.

Fl cDNA clones are not only important for genome annotation. They also constitute crucial reagents for the functional analysis of protein-encoding genes. In this respect, major projects have already resulted in the construction of open reading frame (ORF) collections that can be transferred at large-scale via recombinational cloning techniques from a reference clone to a wide-variety of expression vectors, each designed for a specific functional assay. These ORF collections are, or soon will be, publicly available. They will undoubtedly foster research projects that either focus on the analysis of selected gene subsets with various methods or on the systematic genome-scale characterization of certain protein properties. Because different applications dictate incompatible sequence constraints (ORF with or without stop codon including or not terminal tags) and because each recombinational cloning technique has its own pros and cons (at this stage either the Gateway or CRE-lox systems), the ORF collection format cannot be unique and settled once and for all. However, the community would greatly benefit from a centralized database that would inform all potential users of the status of the cDNA/ORF cloning and sequencing progress for their genes of interest in any of the ORFeome projects, together with the restrictions that may or may not apply to their use. Obviously, less or no restriction is preferable to boost the *Arabidopsis* research community.

Ongoing large-scale projects

RIKEN Centers (Motoaki Seki and Kazuo Shinozaki)

The Genomic Sciences Center (GSC) is involved in the systematic isolation and characterization of fl cDNA clones. About 18,000 independent fl cDNA clones, called the RAFL clones, have been collected so far, from which about 15,000 fl sequences have been deposited to GenBank. Over 13,000 cDNAs are now distributed by the Bioresource Center (BRC). In the framework of the National Bioresource Project, GSC and BRC are continuing the sequencing of the remaining 8,000 RAFL cDNAs. Once characterized and functionally annotated, these will be available from the BRC.
http://pfgweb.gsc.riken.go.jp/projects/raflcdna.html.
http://rarge.gsc.riken.go.jp/.
Seki et al. (2002). Functional annotation of a full-length Arabidopsis cDNA collection. Science, 296, 141-5.

SSP consortium (Joe Ecker, Ron Davis, Sakis Theologis)

The SSP contribution includes fl cDNA clone sequences (11,794 RAFL clones in GenBank), ORF cloning as well as transcript unit mapping with genome tiling arrays. This project completed in September 2003 resulted in the production of 10,556 fully sequenced ORF clones tailored for recombinational cloning. All match the genome sequence. Approximately 9,000 clones were derived from RAFL cDNA inserts and 1,556 from RT- PCR.
http://signal.salk.edu/SSP/
Yamada et al. (2003). Empirical analysis of transcriptional activity in the Arabidopsis genome. Science, 302, 842-846.

Salk 2010 cDNA project (Joe Ecker)

The Ecker lab is continuing the SSP effort to experimentally verify the annotation of an additional 3,330 genes by fl cDNA sequencing. It is constructing and sequencing the corresponding ORF clones. This project focuses on genes that are known to be transcribed, so-called "annotated expressed genes", but for which fl ORF clones are not available. So far, 1235 fl cDNA sequences have been submitted to Genbank. As of May 2004, from the SSP contribution listed above and this project combined, 11,734 ORF clones had been deposited and arrayed for distribution by the ABRC. The majority of clones are pUNI vector derivatives designed for CRE-lox mediated subcloning, although 1,799 are Gateway entry clones. All these ORFs are in the closed configuration (with stop codon).
http://signal.salk.edu/csummary.html.
http://signal.salk.edu/cdnastatus.html.

TIGR 2010 cDNA project (Chris Town)

The Town lab is focusing on the isolation (RT-PCR) and characterization of about 2,000 transcripts for which no experimental cDNA sequences are available. This project includes validation of predicted gene structures by 5' and 3' RACE and cloning of the corresponding ORFs in the Gateway pDONR221 entry vector, with the original stop codon. http://www.tigr.org/tdb/hypos/
Xiao et al (2002) Cloning and sequencing of cDNAs for hypothetical genes from chromosome 2 of Arabidopsis. Plant Physiology, 130, 2118-28.

Génoscope, Unité de Recherche en Génomique Végétale (URGV) and Invitrogen (Marcel Salanoubat)

A novel collection of full-length cDNA clones has been analyzed that matches at least partially 11,500 genes. It provides information on about 2,000 genes covered by new fl cDNA sequences.
Castelli et al (2004) Whole genome sequence comparisons and "full-length" cDNA sequences: a combined approach to evaluate and improve Arabidopsis genome annotation. Genome Research, 14, 406-413.

Atome URGV (Claire Lurin, Ian Small)

ORFs identified in the cDNA clones described in Castelli *et al* (2004) are being transferred into Gateway vectors. Already, 2000 cDNAs from the Genoscope/INRA/Invitrogen collection were transferred into the Gateway pDONR207 vector. These cDNAs can be used for native protein expression in plants and in other eukaryotes. They are not optimal for bacterial expression of native proteins because they lack the Shine/Dalgarno sequence before the ATG and they carry 5' and 3' UTRs. Therefore, ORF entry clones are now being generated without UTRs for subsequent transfer to destination vectors designed for the expression of fusion proteins in any system. Two versions of the clipped ORFs (with and without stop codon) are being cloned. About 500 ORFs have been transferred to entry clones. Some 2,000 more should be transferred before the end of the year 2004. The project's goal is to generate 6,000 to 10,000 end-sequenced ORF clones for about 3,000 to 5,000 individual genes.
http://genoplanteinfo.infobiogen.fr/Databases/CT_Nouveaux_Outils/NO2001054/

Table 1

Salk, Stanford, PGEC (SSP) full-length cDNAs	11,737	04/14/2004, completed
Salk Stanford, PGEC full length, error-free ORFs	10,568	04/14/2004
Salk *Arabidopsis* Gene Collection/ORFome project full length cDNA sequence	487	05/13/2004
Salk *Arabidopsis* Gene Collection/ORFome 2010 project full length error-free ORFs	840	05/13/2004
Arabidopsis research community cDNAs	12,523	05/13/2004
Peking Yale transcription factor read	1,501	05/28/2004
RIKEN *Arabidopsis* full-length (RAFL) cDNA single reads	154,640	06/2004
RIKEN *Arabidopsis* full-length (RAFL) cDNA full-length sequences	2055	04/14/2004
GSLT cDNA single reads	28,816	04/07/2004

Peking-Yale consortium (Peking-Yale Joint Center of Plant Molecular Genetics and Agrobiotechnology, Peking University and Yale University, coordinators Xing Wang Deng and Yuxian Zhu)

A collection of cDNA clones containing the precise ORFs of 1282 *Arabidopsis* transcription factors has been generated and donated to ABRC. All ORFs were individually cloned as Gateway entry vectors and end-sequence validated.

Gong et al. (2004) Genome wide ORFeome cloning and analysis of Arabidopsis transcription factor genes. Plant Physiology, 135, in press.

As listed above, most efforts so far have been devoted to the generation of cDNA and ORFeomics resources. Yet, a few projects are gradually stepping up the systematic functional characterization of proteins. Notably, important structural genomics initiatives are developing the technologies needed for high-throughput structure determination of eukaryotic proteins by X-ray crystallography and NMR spectroscopy. Those focusing on *Arabidopsis* include the Center for Eukaryotic Structural Genomics
(CESG; http://www.uwstructuralgenomics.org/) and the
RIKEN Structural Genomics project (http://protein.gsc.riken.go.jp/).

Tracking ORF cloning projects

Because multiple *Arabidopsis* ORFeome projects are currently underway, TIGR and the ORFEUS consortium (www.orfeome.org) are coordinating their effort to create an online database that will track the progress made in the framework of these initiatives and provide the information to the community with regular updates. In addition, a standard format called the Minimum Information about an ORF (MIAO) has been proposed to exchange all relevant ORFeome information. Tools are being developed on this basic structure to support its implementation including a simple mark-up language as well as a conversion and visualization interface
(see www.orfeome.org/miao).

Tools for functional assays

Parallel to the construction of comprehensive cDNA and ORF collections, several laboratories are developing vectors designed specifically for functional assays in plant cells and compatible with the systematic recombinational cloning of fl cDNAs and ORFs.
http://www.psb.ugent.be/gateway/.
http://signal.salk.edu/pHOST.html
Karimi et al (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Science, 7:193-195.
Curtis and Grossniklaus (2003) A Gateway cloning vector set for high-throughput functional analysis of genes in planta.
Plant Physiology, 133:462-469.
Guo and Ecker (2003) Plant responses to ethylene gas are mediated by SCF (EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor. Cell, 115, 667-677.

Multiparallel Analytical Tools & Phenotype Analyses

Prepared by Mike Beale, Chair - Multiparallel Analytical Tools Subcommittee and by Mary Lou Guerinot, Chair - MASC Functional Proteomics, Metabolomics, and Phenotype Analysis Subcommittee

The Multiparallel Analysis Tools and Phenotype Analyses subcommittee met to review progress at last year's *Arabidopsis* meeting in Madison. The good news is there is more and more data available to be analyzed. For example, NASC arrays now have over 1,000 Affymetrix chip experiments open to public use which we believe is the largest set of public access Affymetrix data for any single organism. Also, AtGenExpress is a co-ordinated international program to generate more than 1000 new Affymetrix baseline data sets for a number of critical developmental stages, tissue types and stereotyped challenges. NASCArrays microarray data are freely available on the web http://affy.*arabidopsis*.info. It offers spot histories, two-gene scatterplots across all experiments, gene swinger, subset bulk gene downloader and other tools including ExpressionProfiler friendly files for clustering. John Ward has taken a first step with the NASC data by converting the microarray information into electronic northerns for each gene that is present on the ATH1 chip. This information is available at his *Arabidopsis* Membrane Transport Library database website [http://www.cbs.umn.edu/*arabidopsis*/] under "Search Expression." GABI MapMan site [https://gabi.rzpd.de/projects/MapMan/data.shtml] already uses NASCarray transcriptomic data presented online linked within pathways. It is clear that if TAIR is to be the main repository of *Arabidopsis* data, more bioinformatic support needs to be devoted to this effort.

True metabolomics – The goal proving to be the most difficult to achieve is to simultaneously quantify all of the metabolites at the cell, organ or plant level. Traditional analytical chemistry based on chromatographic separation of metabolites and subsequent identification by techniques such as GC-MS and LC-MS has played an important role in opening up this area. Most work published so far utilizes these techniques to profile crude plant extracts or to home in on particular classes of compounds in purified extracts. Recently, the application of 'fingerprinting' to unchromatographed extracts by NMR and direct injection ESI-MS or FT-ICR-MS have proved to be promising techniques. They are perceived as a way forward for high-throughput mass screening of mutants and natural variants.

Much *Arabidopsis* metabolomics is being pursued in the private sector. Service and large-scale activities in publicly funded *Arabidopsis* metabolomics are less prevalent than the other 'omics'. Nevertheless, the UK GARNet project contains an *Arabidopsis* metabolomics service and activities in *Arabidopsis* metabolomics are also beginning to emerge in the Netherlands (http://www.biosystemsgenomics.nl/) and in Sweden (http://wcn.ntech.se/platforms/Metabolomics.htm). In addition, the MeT-RO project (Metabolomics at Rothamsted) is a newly funded initiative in the UK which has built on the GARNet project to establish a National Centre for Plant and Microbial Metabolomics. (http://www.metabolomics.bbsrc.ac.uk)

The plant metabolomic community holds an annual International Congress (Potsdam, 2003; Iowa State, 2004) and has formed a platform to further international discussions (www.metabolomics.nl). The integration of metabolomics data with other functional genomic data is a difficult goal to achieve and concerns the community. Pathway databases (http://www.*arabidopsis*.org/tools/aracyc) and (http://www.genome.ad.jp/kegg/pathway.html) are being developed but they do not yet contain metabolomic datasets. Problems of alignment of datasets, databases and effective query tools are still being researched. However, software solutions to some of the problems are emerging (see for example metAlign at http://www.plant.wageningen-ur.nl/default.asp?section=products).

Reverse and Forward Genetic Stocks

Prepared by Bernd Weisshaar, Chair
MASC Reverse and Forward Genetic Stocks Subcommittee

Fast and reliable access to mutants in selected genes is crucial for systematic reverse genetic approaches. The MASC Reverse Genetic Stocks subcommittee addresses issues of coordination and communication among the existing projects in this field. The next meeting will take place during the international conference on *Arabidopsis* research in 2004.

The integration and data exchange between the various projects has progressed well. Most providers of flanking sequence tag (FST)-based mutant collections do allow access to the primary FST sequence information, either from their web sites or via GenBank. In addition, the discussion on, for example, what constitutes a "potential FST gene hit" has resulted in more detailed evaluation and annotation of FSTs in terms of the location of the insertion within a given gene. Analysis of the current resources has shown that coverage of the *Arabidopsis thaliana* gene inventory with knockout mutations is already impressive (see Table 2), but it also demonstrates that the number of really useful insertion lines, namely those that are likely to be a null mutation, is still not saturating. If one considers insertions in coding exons as good candidates for NULL alleles, then there are about 22.400 different genes covered with insertions. However, it is clear that alleles with insertions at the end of the ORF may still result in a (partially) functional protein. On the other hand, insertions in introns are also often good candidates for NULL alleles. In conclusion, we can argue that for about 70% of all *Arabidopsis thaliana* genes useful NULL alleles are available.

The Salk Institute Genomic Analysis Laboratory (SIGnAL, http://signal.salk.edu/cgi-bin/tdnaexpress) has integrated FST data from GABI-Kat, SAIL and FLAGdb as well as data from RIKEN, Wisconsin and several other FST resources (Table 1) into their T-DNA express database. As a result, a quite comprehensive collection of sequence-indexed T-DNA insertion mutants can be searched at a single location on the basis of FST sequence information. This allows users of reverse genetic resources a "one-stop" access to almost all available information on T-DNA insertions in a given gene. The "*Arabidopsis* Knockout Facility" at the University of Wisconsin-Madison has announced the availability of a new collection of T-DNA lines containing Ds-Lox launching pads and Cre/Lox recombination sites (see http://www.hort.wisc.edu/krysan/DS-Lox/). The specific features of these lines can be used to delete tandemly duplicated gene family members, or to generate insertion mutants at flanking loci that are not covered by currently available T-DNA collections. FSTs from an initial set of ca. 10,000 lines have been registered with the SIGnAL.

To complement the efforts to saturate the A. thaliana genome with addressable insertion mutations, other projects are under way to systematically set up collections of RNAi lines that cover the genome (e.g., the EU-funded AGRIKOLA project; see http://www.agrikola.org/). So far, more than 5000 hairpin plasmids have been created and more than 1000 of these have been introduced into *Arabidopsis thaliana*. Preliminary analysis of the transformants indicates that (i) phenocopies of previously described knockout mutants can be obtained, (ii) viable 'knockdown' mutants of genes known to be essential can be obtained, and (iii) the project will reveal many informative phenotypes by inhibition of genes of currently unknown function. Also, projects based on TILLING (see http://www.*arabidopsis*.org/abrc/henikoff.jsp) allow access to additional mutations, including change-of-function alleles of a given gene.

Table 2. Modified version of a table found at http://signal.salk.edu/cgi-bin/tdnaexpress. Basis for the numbers is the TIGR/AGI genome annotation version 5 that contains 30,700 genes. Numbers as of May 13, 2004.

T-DNA population	SALK	SAIL	GABI	FLAG	SMa	Wisc	RIKEN	Totalb
Total Mapped	145,417	51,706	59,441	24,594	23,411	10,459	18,551	333,479
Coding Exon	14,259	5,721	9,324	3,121	3,559	1,954	3,488	22,423
Intron	7,260	2,572	4,534	1,671	1,020	954	957	11,649
5' UTR	5,048	1,912	2,478	957	627	459	818	9,611
Promoter (1st 500bp)	9,879	4,455	5,695	2,652	1,058	1,196	1,368	16,61
Unique At Genes Identified	21,858	11,444	16,177	7,360	5,230	4,201	5,493	27,723

a) FSTs from transposon insertions.

b) The number of "Total mapped" lines is given in row 1. The numbers below refer to the total number of genes in the Arabiodpsis thaliana genome covered by the mentioned T-DNA populations and not to the simple sum of the row.

A comprehensive summary of forward genetic stocks, recombinant inbred line (RIL) populations and other such resources is available at http://www.inra.fr/qtlat/NaturalVar/RILSummary.htm. At the moment, seven RIL populations are available as seed stocks from the public stock centers, but more than 56 different RIL populations and two populations of genetic substitution lines (nearly isogenic lines, NILs) are presently being established (see Table 3). Single Nucleotide Polymorphisms (SNPs) detection is of increasing importance in the forward genetics tool kit. Several large SNP collections are available through TAIR, including those of Cereon/Monsanto (approximately 37,500 SNPs), the Stanford Genome Technology Center (at least 11,000 SNPs) and GABI-MASC ("MASC" stands for the Max-Planck *Arabidopsis* SNP Consortium; over 8,000 SNPs; these are also available via http://www.mpiz-koeln.mpg.de/masc/). With the exception of the GABI-MASC SNPs, which were obtained by re-sequencing between 6 and 12 accessions, the SNPs above were identified as a difference between a single accession (typically Ler) and the reference genome. Thus, little is known about their frequency in other accessions. In contrast, over 17,000 polymorphisms obtained through a re-sequencing study of 96 accessions is available through http://walnut.usc.edu/2010, and will shortly be available through TAIR as well. The 96 accessions, which include many of those being used to generate RILs (see Table 3) are available as a set from the stock centers.

Table 3. Modified version of a table found at http://www.inra.fr/qtlat/NaturalVar/RILSummary.htm

Accession	Stock Center	JIC RIL (1)	Natural RIL (2)	Salk RIL (3)	UTexas RIL (4)	Versailles RIL (5)	Other RIL
Abd-0	CS0932					x Col-0	
Ag-0	CS0936	x Cvi					
Ak-1	CS0938		x C24				
An-1	CS0944		x Ler				
Bay-0	CS0954					x Shahdara	
Bch-1	CS0956		x C24				
Bla-1	CS0970					x Col-0	
Blh-1	CS1030					x Col-0	
Br-0	CS6626	x Kondara					
Bur-0	CS1028					x Col-0	
C24	CS0906		x Col-0 ;				
Can-0	CS1064					x Col-0	
Can-0	CS6660				x Sav-0		
Co-4	CS1090					x Col-0	
Ct-1	CS1094	x Wt-5				x Col-0	
Cvi	CS0902	x Ag-0	CS8580 x Ler			x Col-0	
Da(1)	CS0917		x Ei-2				
Db-1	CS1102					x Col-0	
Dijon-G	CS0910					x Col-0	
Ei-2	CS6689		x Da(1)				
Eri-1	CS22548		x Ler				
Es-0	CS6699				x Pa-3		
Est-0	CS1148					x Col-0	
Est-1	CS6701			x Col-0			
Fei-0	CS???		x Ler				
Ga-0	CS6714	x Nok-3					
Ge-0	CS1186					x Col-0	
Gr-3	CS1202					x Col-0	
Gy-0	CS6732	x Sorbo					
Kas-1	CS0903					x Col-0	xTsu-1 (Mc Kay)
Kas'-2'	CS1264		x Ler				
Kas-1	CS3880						x Col-gl1 (Somerville)
Kin-0	CS6755			x Col-0			
Ko-2	CS1288					x Col-0	
Kondara	CS0916	x Br-0					
Kondara	CS6175		x Ler				
Kyo-1 (JW1)	CS???		x Ler				
Ler	CS0020	x Col-0					
Ler	CS8581				x No-0		
Ll-0	CS6781		x Ler				
Lu-1	CS1352					x Col-0	
Mh-0	CS0904					x Col-0	
Mh-0	CS6792				x Sf-2		
Mr-0	CS6795			x Col-0			
Mz-0	CS6800	x Ts-5					
Nd-1	CS1636						x Col-3/5 (Holub)
Nd-1	CS6922		x C24				
No-0	CS6805				x Ler		
Nok-0	CS6807				x Uk-3		
Nok-1	CS1400					x Col-0	
Nok-3	CS6810	x Ga-0					
Pa-3	CS6827				x Es-0		
Ri-0	CS1492					x Col-0	

Sav-0	CS6856				x Can-0		
Sf-2	CS6857				x Mh-0		
Shahdara	CS0929		x Ler			x Bay-0 ; x Col-0	
Sorbo	CS0931	x Gy-0					
Ts-5	CS6871	x Mz-0					
Tsu-0	CS1564					x Col-0	
Tsu-1	CS1640						x Kas-1 (Mc Kay)
Tul-0	CS1570					x Col-0	
Uk-3	CS6880				x Nok-0		
Van-0	CS6884			x Col-0			
Ws	CS2223						x W100F (Scolnik)
Wt-5	CS6896	x Ct-1					
Yo-0	CS1622					x Col-0	
Number RIL Pop		7	14	4	5	24	of 4

Populations indicated in bold are already available through the stock centers:
1) http://www.jic.bbsrc.ac.uk/corporate/Science_Departments/crop_gen.html
2) http://www.dpw.wau.nl/natural/
3) http://www.naturalvariation.org/
4) http://www.biosci.utexas.edu/MCDB/lloyd.html
5) http://dbsgap.versailles.inra.fr/vnat/

In conclusion, the combination of all existing resources significantly increases our chances to obtain plants containing the mutation(s) and alleles we need to find answers to the biological questions of interest to us.

The International *Arabidopsis* Functional Genomics Community

Argentina

http://www.*arabidopsis*.org/info/2010_projects/Argentina.html
Contact: Jorge Casal
Universidad de Buenos Aires
casal@ifeva.edu.ar

The first Symposium on Arabidopsis functional genomics in Argentina was held in Buenos Aires on October 27, 2003 with speakers from Argentina and Chile. Currently, several university-associated groups are actively engaged in Arabidopsis research in Argentina. Funding for Arabidopsis research is available from the organizations listed below:

Analysis of transcriptome in plant-pathogen interactions: Plant genes required for susceptibility to fungal infection. Malena Alvarez, malena@dqb.fcq.unc.edu.ar CIQUIBIC-CONICET, Facultad Ciencias Quimicas, Universidad Nacional de Cordoba Province of Córdoba, http://www.fcq.unc.edu.ar/ciquibic.

The genetic network involved in plant responses to the light environment, analysis of transcriptome in phytochrome and cryptochrome mutants. Jorge J. Casal, casal@ifeva.edu.ar, IFEVA, Facultad de Agronomía, Universidad de Buenos Aires. Buenos Aires http://www.ifeva.edu.ar/staff/perpages/casal.htm.

Cytochrome c, cytochrome oxidase subunit 5b and other genes involved in respiration. Daniel H. Gonzalez, dhgonza@fbcb.unl.edu.ar, Facultad de Bioquímica y Ciencias. Biológicas Universidad Nacional del Litoral, Province of Santa Fe.

Role of senescence associated genes in the formation of lytic vacuoles during senescence. Juan José Guiamet, jguiamet@museo.fcnym.unlp.edu.ar, Instituto de Fisiología Vegetal, Universidad de La Plata. Province of Buenos Aires.

Genes involved in Potassium and Sodium transport. Guillermo E. Santa-Maria, gsantama@pop.unsam.edu.ar, Instituto de Investigaciones Bioteconológicas, Universidad Nacional de San Martin. Province of Buenos Aires.

Regulatory genes involved in the control of transcription of genes of the photosynthetic antenna. Roberto J. Staneloni, RStaneloni@Leloir.org.ar, Instituto Leloir, Buenos Aires.

Functional analysis of oxidative stress-regulated genes Estela M. Valle, evalle@arnet.com.ar, Instituto de Biología Molecular y Celular de Rosario, Facultad Ciencias Bioquimicas y Farmaceuticas, Universidad Nacional de Rosario. Province of Santa Fe.

Identification of key components for retrograde signalling between mitochondria and nucleus in higher plants by transcriptomic, proteomic and functional analyses of respiratory complex mutants in Arabidopsis. Eduardo Zabaleta, ezabalet@mdp.edu.ar, and Diego Gómez-Casati, diego.gomezcasati@intech.gov.ar, Universidad de Mar del Plata and Instituto de Investigaciones Bioteconológicas, Universidad Nacional de San Martin Province of Buenos Aires.

The main sources of financial support are the Agencia Nacional de Promoción Científica y Técnológica (ANPCYT; functional genomics has been one of the priority subjects in recent calls for proposals), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and the FUNDACION ANTORCHAS (Argentina).

Australia & New Zealand

http://www.*arabidopsis*.org/info/2010_projects/Australia.html
Contact: Geoffrey Wasteneys
The Australian National University, Canberra
geoffw@rsbs.anu.edu.au

Australia has a strong tradition in plant scientific research. Many institutions are engaged in *Arabidopsis* Functional Genomics work including the Plant Industry Division of the Commonwealth Scientific and Industrial Research Organization (CSIRO), the major Universities and private enterprise. Their work ranges from individual projects to international collaborations and major resource development. Funding is mainly available through the Australian Research Council's (ARC's) Discovery and Linkage Grant Schemes and the Grains Research and Development Corporation of Australia (GRDC).

Researchers in all Australian States and the Capital Territory now use *Arabidopsis* functional genomics approaches. Projects are generally highly focused but increasingly involve international collaborators. Canberra, Australia's capital city, remains a major node for *Arabidopsis* research activity. Together, CSIRO's Division of Plant Industry, the Australian National University (ANU) and the Center for the Application of Molecular Biology to International Agriculture (CAMBIA) form a remarkable unit of fundamental, industrial and application-driven research.

The Australian Center for Plant Functional Genomics is a major initiative announced in 2001, and it is now underway at the University of Adelaide. Established jointly by the ARC and the GRDC, the center's objective is to contribute to ensuring that Australia remains internationally competitive in plant science research. However, its current focus on major crop plants with little emphasis on *Arabidopsis*.

New Zealand has a small population but is nevertheless home to several *Arabidopsis* research programs. Increasing numbers of New Zealand plant scientists are incorporating *Arabidopsis thaliana* into their research, and at least six groups are using functional genomics approaches. Funding is principally available through the Royal Society of New Zealand's Marsden Fund and the New Zealand Foundation for Research, Science and Technology. Geographically, *Arabidopsis* research seems to be concentrated in three regions: in the North Island cities of Auckland and Palmerston-North and at the University of Otago in Dunedin, on the South Island. In addition to the projects being conducted at the universities, research programs are carried out at the Government-owned Crown Research Institutes, including Horticulture and Food Research Institute of New Zealand (HortResearch) and the New Zealand Institute for Crop & Food Research Limited (Crop & Food Research).

The horticultural industry is a big part of the New Zealand economy and, reflecting this, much of the *Arabidopsis* research impinges on reproductive development and fruiting. Other functional genomics programs include work on a magnesium transporter gene family and a recently initiated study on the role and function of carboxylesterases.

Austria

http://www.*arabidopsis*.org/info/2010_projects/Austria.html
Contact: Heribert Hirt
Vienna Biocenter, Vienna
HEHI@gem.univie.ac.at

In recent years, major changes have taken place in the development of molecular biology research facilities in Austria. One of the hot spots of constant change is the Vienna BioCenter, a newly established science campus close to the city center. In addition to several smaller biomedical companies, the Vienna BioCenter has become home of various research institutes from the University of Vienna, the Academy of Sciences and the pharmaceutical company Boehringer-Ingelheim. These developments prompted the government, local authorities and the University of Vienna to concentrate plant molecular research groups from Botany, Microbiology and Genetics, Biochemistry and Molecular Cell Biology, and Medical Biochemistry institutes at the Pflanzen Molekularbiologie Zentrum (PMZ). The PMZ facilities are already constructed and the center is expected to open in early 2005.

Adjacent to the PMZ, the Austrian Academy of Sciences is establishing two new institutes: the Gregor-Mendel-Institute of Molecular Plant Sciences (GMI) and the Institute of Molecular Biotechnology (IMBA). Whereas the IMBA will concentrate on generating knowledge that aims ultimately at curing major human diseases, the goal of the GMI is a basic understanding of how plants work. Construction of both institutes has just begun and their opening is scheduled for 2005. Also, the Gregor-Mendel-Institute has accepted to buy an Affymetrix workstation so that Austrian researchers can process Affychip microarray data of the various *Arabidopsis* genomics consortia. These new developments add considerable value to Austria's research potential and provide the necessary critical mass for starting a coordinated thematic program on *Arabidopsis* biology.

It is the intention of the Austrian Platform of *Arabidopsis* Research (APAR) consortium to function as a research platform coordinating and promoting *Arabidopsis* research in Austria. The activities of APAR are tightly linked to several programs of the European Union and to the worldwide coordination efforts by MASC. Additional Austrian project partners will be incorporated into APAR in the future. APAR currently comprises several projects. For example, (i) molecular regulation of cytokinesis during plant development, (ii) molecular analysis of MAPK-mediated ethylene signaling in *Arabidopsis thaliana*, (iii) analysis of glycogen synthase kinase/shaggy-like kinases, (iv) novel aspects of salt stress signaling in plants, (v) specificity and functional analysis of a PP2C protein phosphatase gene subfamily, (vi) calcium-dependent protein kinases in *Arabidopsis* signal transduction, and (vii) the functional study of the Ku complex at *Arabidopsis* telomeres. One hundred and fifty participants joined the trilateral (Austrian, German and Swiss) APAR meeting held in Vienna, 15-17 April, 2004. Additional activities on *Arabidopsis* research in Austria include projects examining structure-function relationships of ribonucleoproteins, signal transduction and cell cycle regulation, auxin and cytokinin, transport and cell differentiation, epigenetics, chromosome biology, genes involved in the reprogramming of microspores, and MAP kinase signal transduction in plants.

Funding for *Arabidopsis* research in Austria is available from Fonds zur Förderung der wissenschaftlichen Forschung (FWF; basic research only) http://www.fwf.ac.at, Wiener Wissenschafts-, Forschungs- und Technologiefonds (Vienna region) http://www.wwtf. at, Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) http://www.bmbwk.gv.at/, and the Austrian Industrial Research Promotion Fund (FFF; applied research),
http:// www.fff.co.at/.

Canada

http://www.*arabidopsis*.org/info/2010_projects/Canada.html
Contacts:
Bill Crosby
University of Saskatchewan, Saskatoon, Saskatchewan
bcrosby@cs.usask.ca
Peter McCourt
University of Toronto, Toronto, Ontario
mccourt@botany.utoronto.ca

Arabidopsis functional genomics efforts are ongoing at several major institutions in Canada. The *Arabidopsis* Research Group (ARG) at the University of Toronto, which includes eight research groups housed out of the Department of Botany, was originally established to provide resources and expertise for the *Arabidopsis* community in Canada. Programs sponsored by ARG are jointly funded through the Ontario Genomics Initiative (OGI), Genome Canada, the National Science and Engineering Research Council (NSERC) and by private industry. All resources and data will be made publicly available through various databases and international stock centers. Contacts for each program are listed at http://www.genomecanada.ca/GCprogrammesRecherche/projets/index.asp?l=e or the ARG program director, John Coleman, can be reached directly at coleman@botany.utoronto.ca.

The functional genomics program at the University of British Columbia includes participants from the Biotechnology Laboratory, Botany and Plant Science Departments, along others. The program has recently received diverse funding input to support its projects, including CFI, NSERC, OTIP, FRBC, HFSP, Genome BC, and Genome Canada. Select program elements include the exploitation of *Arabidopsis* as a model system for studying development and the development of TILLing resources.

The recently implemented University of Saskatchewan program derives from activities initiated in late 1999, under the auspices of the National Research Council Genomics in Health and Agriculture Initiative (NRC - GHI). The program was additionally funded by Genome Canada, the Saskatchewan-Canada Agriculture-Food Innovation Fund and, more recently, it has been linked to an NSF 2010 project concerned with the functional genomics of the Ubiquitin-Protein Ligase (E3) families in *Arabidopsis*. In addition, the United States have supported a new Bioinformatics group that includes a research emphasis involving plant genomics and Systems Biology.

The ongoing program at the NRC Plant Biotechnology Institute continues to explore the interface between *Arabidopsis* functional genomics for its implication to Brassica crop improvement with a new emphasis on food quality and secondary metabolism.

The Saskatoon Research Center of Agriculture Canada is conducting an active program designed to exploit *Arabidopsis* model system in support of genomics approaches to Brassica crop development. The program is funded by the Agriculture Canada Genomics Program and is supplemented by recent support from Genome Canada. Program elements include genetic, physical and bioinformatics approaches to defining the relationship between the *Arabidopsis* and Brassica genomes and the development of an *Arabidopsis* activation-tagged T-DNA insert population.

China

http://www.*arabidopsis*.org/info/2010_projects/China.html
Contact: Jianru Zuo
Institute of Genetics and Developmental Biology
Chinese Academy of Sciences, Beijing
jrzuo@genetics.ac.cn

The *Arabidopsis* community has rapidly expanded in China these past few years. More than 250 participants attended the Annual Workshop on *Arabidopsis* Research, held in Shanghai on November 30, 2003. The workshop was organized by Zhihong Xu, President of Peking University, and featured eighteen oral presentations.

In 2002, the National Science Foundation of China (NSFC) provided a grant of US$1.5 millions for a major international collaborative project aimed at the proteomic characterization and functional studies of approximate 1,600 *Arabidopsis* transcription factors. The project involves multiple leading academic institutions in China including Peking University, the Institute of Genetics and Developmental Biology of the Chinese Academy of Sciences (CAS), Fudan University, Wuhan University, Shanghai Jiao Tong University, and Shanghai Institute of Plant Physiology and Ecology of CAS. The coordinators of the project are Xing-Wang Deng (Peking University/Yale University/CAS Center for Plant Molecular Genetics and Agrobiotechnology and Yale University) and Yuxian Zhu (Peking University). During the first phase of the project, an ORFeome collection for the *Arabidopsis* transcription factor genes has been generated in a Gateway high-throughput cloning vector (Gong *et al.*, Plant Physiology, in press). Currently, 1,282 clones containing full-length ORF regions have been deposited at ABRC and will be available by May 2004 (http://www.*arabidopsis*.org/news/news.jsp#orf).

In a separate effort, funded by the Ministry of Science and Technology of China (MOST; US$ 350,000), an inducible enhancer/promoter vector was used to generate activation tagging lines (Jianru Zuo, Institute of Genetics and Developmental Biology, CAS). More than 55,000 T1 transgenic lines had been collected by the end of 2003, 35,000 of which were generated in Zuo's lab and 20,000 lines generated in Yingtang Lu's lab at Wuhan University.

Funding for *Arabidopsis* functional genomic research is available from the Ministry of Science and Technology of China (www.most.gov.cn), National Science Foundation of China (NSFC - www.nsfc.gov.cn), CAS (www.cashq.gov.cn), and other sources on a competitive basis.

Eastern European *Arabidopsis* Activity

Contact: Martin Fellner
Palacky University
Department of Cell Biology and Genetics and
Institute of Experimental Botany ASCR
Olomouc, Czech Republic
emfee@prfholnt.upol.cz

Arabidopsis research in former communist countries is relatively new, small and often isolated. The goal of the Eastern European *Arabidopsis* Activity (EEAA) is to integrate the *Arabidopsis* community in Eastern Europe and incorporate its program into the international *Arabidopsis* effort. The purpose of the EEAA is to initiate a joint research project, potentially in collaboration with some of the already established *Arabidopsis* laboratories around the world. EEAA's long-term objective is to grow the *Arabidopsis* community and boost the prestige of plant science in Eastern Europe. During 2003, laboratories from six countries demonstrated repeated interest in the EEAA and are currently investigating various topics.

Czech Republic

Interaction between blue light signaling and abiotic stress
(supported by Academy of Sciences of the Czech Republic).
Identification of genes integrating hormone and light signaling.
Martin Fellner, emfee@prfholnt.upol.cz, http://genetika.upol.cz/.

Analysis of 1500 *Arabidopsis* insertional lines (containing insert of T-DNA with tetramer of enhancer from 35S promoter) with respect to flower and root mutations, mutations in responses to elevated boron concentrations and in response to Plasmodiophora brassiceae infection.
Tomá_ Kocábek, kocabek@umbr.cas.cz,
http://www.umbr.cas.cz/805_www/Kocabek/index.htm.

Hungary

Function of phosphoprotein phosphatases
(supported by the Hungarian Scientific Research Fund).
Ilona Farkas, farkas@jaguar.dote.hu,
http://www.dote.hu/tudomany/whoiswho98/119.htm

Collecting cDNA clones and insertion mutants representing protein family designated as "Receptor-like cytoplasmic kinases VI" involved in Rop GTPase-dependent signaling cascades (supported by the National Grant Agency).
Attila Fehér, fehera@nucleus.szbk.u-szeged.hu.

Lithuania

Functional activity of the plasmatic and vacuolar membrane.
Plant responses to salt stress.
Gemir Maksimov, gemir@botanika.lt,
http://ml.lms.lt/200203/disertacijos.htm.

Poland

Identification and characterization of enzymatic activity of all *Arabidopsis* ORF's containing Nudix/MutT domain and *Arabidopsis* protein Ku70.
Marta Dobrza_ska, martad@ibb.ww.pl, http://www.ibb.waw.pl/

Investigation of plant genes transcriptional activation and repression mechanisms through remodeling of chromatin structure.
Biological functions of linker (H1) histones.
Andrzej Jerzmanowski, andyj@ibb.waw.pl, http://www.ibb.waw.pl/.

Russia

Characterization of the state of phytochromes and (proto) chlorophylls in their native state in the cell (supported by state funding and by the Russian Foundation).
Vitaly Sineshchekov, V.Sineshchekov@mtu-net.ru.

Uzbekistan

Thionines-cystein rich peptides and Isolation and physico-chemical characterization of hormone-binding proteins.
O. Veshkurova, G. Mavlonov, ali@ibchem.ccc.uz.

European Union

http://www.*arabidopsis*.org/info/2010_projects/EU.html
Contact: Bernard Mulligan
bernard.mulligan@cec.eu.int

Opportunities for functional genomics research on all organisms can be found in several areas of the current 6th Framework Program "FP6"(2002-2006), the European Union's research funding program. To be eligible for FP6 funding, research programs must involve laboratories from several European Member States. However, many opportunities also exist for researchers from countries outside Europe to be involved in programs funded through FP6. In fact, in certain cases, researchers from countries outside Europe can receive FP6 funding. FP6 funds large scale "networks of excellence" and "integrated projects" with grants of Euro 10 million or more as well as smaller targeted projects and individual research fellowships. Funding opportunities for coordination projects and for activities (e.g., conferences and workshops) to support the development of European Union science policy (e.g., in areas relating to functional genomics research) are also available. Details about all these opportunities can be found at http://fp6.cordis.lu/fp6/home.cfm and http://europa.eu.int/comm/research/fp6/index_en.html

The large-scale projects are often very multidisciplinary in nature. A good example is the integrated project "Grain Legumes". This highly multidisciplinary project will develop new genetic, genomic, post-genomic, and bioinformatics tools to improve and sustain grain legume seed production and quality. Notably, the project will contribute to the complete sequencing, within an international project, of the gene-rich regions of the Medicago truncatula genome which is a relevant model system for European grain legumes. "Grain Legumes" fully recognizes the value of the model plant *Arabidopsis* and consequently will fully integrate *Arabidopsis* research or data derived from this model system in several of its activities. With 54 partners in 18 countries, this project is expected to build a European area for Grain Legumes research. Further information about the project can be found at http://www.eugrain-legumes.org/

Another example is the network of excellence "Epigenetics" which includes a joint research program in the field of epigenetics. Epigenetics involves 25 research teams of top European scientists with a proven track record as leaders in their field. They will constitute the 'virtual core center' by combining their expertise and resources. The 25 core research teams are geographically clustered around eight established centers of epigenetic research and in some cases benefit from established collaborations and synergies that have emerged from previous European Union programs (e.g., 5th Framework Program). The research program of the core teams addresses the functional analysis of epigenetic control in many different organisms (e.g., S.cerevisiae, S.pombe, plants, Drosophila, Xenopus, mouse, human) and applies varied and wide ranging genetic, biochemical and cytological approaches. The strength of such a core program lies in its focus on the molecular mechanisms underlying epigenetic control rather than on purely descriptive and phenomenological analyses. For further details please see http://www.epigenome.imp.ac.at/

The ERA-NET grant scheme is a novel feature of the 6th Framework Program. It provides support for transnational networking and coordination of national research programs. Therefore, the scheme's participants are program managers working in national ministries and funding agencies. The "European Research Area – Plant Genomics", with a grant of 2.2 million Euros, focuses on networking of national programs to help maximize the return on the Euro 80 million invested in plant genomics across Europe each year. The network will formulate long-term research goals and objectives for plant genomics in Europe and identifying areas in which Europe should contribute to international programs (see http://www. cordis.lu/coordination/publications.htm and
http:// www.genomics.nl/homepage/research/
funding_opportunities/eranet_(fp6)_projects/)
In addition, a project database is being set up for projects funded under FP6 (http://www.cordis.lu/ fp6/projects.htm). A database of previously funded European Union projects is available at http://www.cordis.lu/en/home.html.

France

http://www.*arabidopsis*.org/info/2010_projects/France.html
Contact: Ian Small
Plant Genomics Research Unit-URGV
Institut National de la Recherche Agronomique (INRA), Evry
small@evry.inra.fr

The major source of funding in France for the *Arabidopsis* Functional Genomics project is Génoplante (http://www.genoplante.com/), a joint venture created by public funding agencies (INRA, CNRS, CIRAD, IRD) and several French ag-biotechnology companies such as Biogemma, Aventis CropScience, and Bioplante). Génoplante has joined forces with GABI, a similar German initiative, and several joint projects are being funded.

Génoplante funded programs

FLAGdb++, an *Arabidopsis* genomics database including an inventory of flanking sequence tags from the Versailles *Arabidopsis* T-DNA collection (http://genoplante-info.infobiogen.fr/FLAGdb/)

CATMA, a complete *Arabidopsis thaliana* microarray (http://www.catma.org/). This is a program involving several EU countries. The URGV is now printing CATMA arrays for a number of collaborative projects. The gene-specific tags used to print the arrays have been cloned in the AGRIKOLA program (see below) and will soon be available from NASC.

ATOME: An *Arabidopsis* ORFeome (http://genoplante-info.infobiogen.fr/Databases/CT_Nouveaux_Outils/NO2001054/index.html).
Analysis of the proteome of *Arabidopsis* (contacts: Jacques Joyard, jjoyard@cea.fr and Michel Rossignol, rossignol@ensam.inra.fr)
Metabolomics – several projects are analyzing levels of various metabolites or protein co-factors in *Arabidopsis* mutants. Some examples are:

- The *Arabidopsis* metabolome by NMR and mass spectroscopy. R. Bligny, CEA, Grenoble
- Cytochrome P450s. D. Werck, IBMP, Strasbourg
- Glycoproteins. V. Gomord, U. de Rouen
- Cell wall polysaccharides. H. Höfte, INRA, Versailles, and
- Numerous other projects aimed at functional analysis of specific genes or gene families.

The Génoplante-Info database (http://genoplante-info.infobiogen.fr/) contains data from several *Arabidopsis* projects including:

- FLAGdb, the FST database
- PlantGene and GeneFarm, *Arabidopsis* gene annotation projects
- ATOME, an ORFeome project
- EST, SAGE, and microarray data on transcription profiles.

Major Generic Non-Génoplante Programs

A panel of sequenced *Arabidopsis thaliana* full-length cDNAs (contact: Marcel Salanoubat, salanou@genoscope.cns.fr).

AraCORE: Analysis of genetic variability between *Arabidopsis thaliana* ecotypes. Several hundred accessions, thousands of recombinant inbred lines, constitution of an *Arabidopsis* core collection based on SNP genotyping (contacts: David Bouchez, bouchez@versailles.inra.fr, Dominique Brunel, brunel@versailles.inra.fr, and Georges Pelletier, pelletie@versailles.inra.fr).

AGRIKOLA: *Arabidopsis* Genomic RNAi Knock-out Line Analysis and Construction of resources for systematic RNAi in *Arabidopsis*. http://www.agrikola.org/ (contact: Ian Small, small@evry.inra.fr).

Germany

http://www.*arabidopsis*.org/info/2010_projects/Germany.html
Contacts:
Thomas Altmann
Max Planck Institute for Molecular Plant Physiology, Golm
Altmann@mpimp-golm.mpg.de
Gerd Jürgens
University of Tübingen, Tübingen
gerd.juergens@uni-tuebingen.de

Arabidopsis functional genomics research has received strong support in Germany through the implementation of two major research programs supported by the Federal Ministry for Education and Research (BMBF) and the Deutsche Forschungsgemeinschaft - German Research Foundation (DFG).

The first of these programs is ***G****enom****a****nalyse im* ***bi****ologischen* System Pflanze (GABI), genome analysis in the plant biological system (http://www.gabi.de/). GABI was initiated in 1999 aiming at strengthening plant genome research in Germany, establishing a network of competence including public, private research groups and corporations, and enhancing international collaboration and transfer of knowledge into application. The second phase of the program has recently been started and will last until the end of 2007 with a budget of Euro 10 million per year. GABI is funded by the Federal Ministry of Education and Research and private business companies, a public-private partnership *par excellence*. The support of private partners involved in the program has increased from 10% in the first program phase to 20% in the second phase. This is a clear indicator of the importance of plant genomics for our societies and economies now and in the future.

About 50% of GABI's funding in the first program phase was devoted to work on the model system *Arabidopsis thaliana*. In the second program phase, support for the model organism will be somewhat reduced. However, the interlocking of research on a model organism and the transfer of these results to crops plants is a fundamental principle of GABI. Therefore, the so-called "bridging projects" embed research on the model *Arabidopsis* with crops within single research consortia. Established rules regulate disclosure and use of research results obtained though GABI activities.

Several GABI projects provided major recent contributions to the international efforts on *Arabidopsis* functional genomics: a large collection of sequence-indexed T-DNA insertion lines (GABI-KAT; http://www.mpiz-koeln.mpg.de/GABI-Kat/), a database of membrane proteins (Aramemnon; http://crombec.botanik.uni-koeln.de/index.html), and extensive SNP information for 13 different *Arabidopsis* accessions (MASC-DB; http://www.mpiz-koeln.mpg.de/masc/). Maintenance and further development of MAtDB at MIPS http://mips.gsf.de/proj/thal/ are also being supported by GABI.

One of GABI's major targets is the establishment and support of international collaborations. A first step towards setting up direct collaborative efforts in Europe has been the establishment of joint research projects between the French plant genome program, Génoplante, and the German GABI initiative. This bilateral interaction is currently being expanded to a trilateral co-operation including the Spanish genome program. Starting this year, nine trilateral research projects plus five bilateral projects between France and Germany will change research structures in Europe. Once more, Plant Genomics will become an excellent example of how research will be organized in the 21st century. A recently funded European Research Area Network Plant Genomics (ERA Net PG; http://www.erapg.org) is another example of

plant genomics as a front runner. Both German programs, GABI and AFGN (see below), played an important role during the establishment of this network and, consequently, performed a significant function in the creation of the European Research Area.

The second major funding initiative for *Arabidopsis* functional genomics research is the *Arabidopsis* Functional Genomics Network (AFGN), funded since 2001 by the DFG. AFGN was founded in close coordination with the 2010 Project of the United States National Science Foundation. Both programs were established with the goal of elucidating the function of all *Arabidopsis* genes by the year 2010. Eleven AFGN projects started in 2001 and 20 more projects were added in 2002. All currently active AFGN projects run until 2004 when the second phase of the program will begin. As a result of the increasing interaction between these two funding agencies, in 2004, AFGN proposals submitted to the DFG and 2010 Project proposals submitted to NSF were co-reviewed by a joint AFGN-NSF panel. Transnational co-operative projects, including Germany-America or multiple country partnerships, were especially encouraged. Information about the AFGN project can be found at http://www.uni-frankfurt.de/fb15/botanik/mcb/AFGN/AFGNHome.html and information about individual AFGN-funded projects at http://www.uni-frankfurt.de/fb15/botanik/mcb/AFGN/Members.html or at the functional genomics web page http://www.*arabidopsis*.org/info/2010_projects/AFGN_Abstracts.jsp.

AFGN has taken the lead in setting-up an international joint effort to establish a comprehensive genome-wide *Arabidopsis* transcriptome reference database. AtGenExpress is a multinational coordinated effort to uncover the transcriptome of the multicellular model organism *Arabidopsis* thaliana coordinated by Detlef Weigel, Thomas Altmann and Lutz Nover. The overall database derived from about 1300 microarrays (i.e., more than 30 million data points) will be accessible via TAIR and will be released to the public Gene Expression Omnibus (GEO) and ArrayExpress databases. Data processing and publication already started to be managed by the NSF-supported *Arabidopsis* database TAIR, in the United States.

Italy

http://www.*arabidopsis*.org/info/2010_projects/Italy.html
Contact: Paola Vittorioso
University of Rome "La Sapienza", Rome
paola.vittorioso@uniroma1.it

Several Italian groups have been actively engaged in *Arabidopsis* research in recent years. Most of these groups are involved in national and international plant functional genomics network projects. In 2003, a common technological platform was developed creating a network among groups of the highest qualification active in Italian universities, public research institutes and the most relevant plant biotechnology companies. This national network, funded by the Italian Ministry of Research (MIUR; www.miur.it), could represent a first step towards the establishment of a National Plant Biotechnology Center (From *Arabidopsis* to tomato: A scientific network and a technological platform for the functional genomics of plant development).

This network intends to exploit a functional genomics approach to analyze selected regulatory aspects of *Arabidopsis* development through the analysis of the function and interactions of members of different families of regulatory and structural genes. On these genes, laboratories involved in the network have achieved results and know-how of the highest international standards. The scope of this project is to gain knowledge on the function of individual genes involved in the different developmental processes analyzed and to identify regulatory networks and interactions between different genes and different processes. It has become increasingly evident that in higher organisms, individual genes influence several processes and, therefore, a satisfactory comprehension of developmental processes can only be achieved through a functional genomics approach.

Analyzed in this research are genes from the:

- Dof family (Costantino) involved in auxin-dependent meristem formation, in the control of seed germination and in the response to gibberellins and stress;
- rol family (Costantino) that influence meristem formation, floral transition and sexual organ formation;
- HD-Zip family involved in the regulation of primary and secondary meristem activity (Morelli) and in developmental processes as a response to the environment, such as shade-avoidance response (Ruberti);
- MYB family (Tonelli) involved in morphogenesis, stress response and in the biosynthesis of nutritionally relevant polymers;
- NF-Y family which interact with several families of transcription factors crucial in differentiation and development in eukaryotes (Tonelli);
- MADS family (Colombo) involved in vegetative and reproductive development, HMG and TAF (Colombo) known as important factors in modulating transcription; and from the
- E2F family (Cella, Albani) involved in cell-cycle regulation and development.

This research analyzed as well genes involved in response to red and far-red light (PHY; Bowler), response to blue light (CRY; Bowler, Benvenuto), in signalosome assembly (DET; Bowler), in the biosynthesis of carotenoids (UR Benvenuto, Cellini), and genes important for photosynthetic activities and nutritional quality. Included were also genes of proteins involved in the response to pathogens and development (PG, PGIP; Cervone), members of the 14.3.3 protein class (Aducci) involved in cell cycle control and in several signal-transduction pathways, and members interacting with 14.33 (Soave) and genes involved in iron homeostasis and in detoxification of ROS (Soave).

The different lines of research on these genes are coordinated. New post-genomic technologies will be set up and the use of existing technologies will be made available to all partners of the network. The network will develop and utilize technologies for the functional analysis of the genes (i.e., RNA interference, negative and positive dominant, chemical gene-machine/Tilling), technologies for the analysis of interactions between genes (i.e., *Arabidopsis* macro- and micro-arrays, real-time PCR) as well as technologies for the identification of protein partners and targets (i.e., Surface Plasmon Resonance, two hybrid in yeast and plant, stable antibodies phage display libraries). In addition, Mariotti, Marmiroli, Migliaccio, and Perata groups are involved in *Arabidopsis* projects funded by the Italian Space Agency, the European Space Agency, and the Institut Pasteur.

Japan

http://www.*arabidopsis*.org/info/2010_projects/Japan.html
Contact: Kazuo Shinozaki
Plant Functional Genomics Research Group, RIKEN GSC
Lab of Plant Molecular Biology, RIKEN Tsukuba Institute
sinozaki@rtc.riken.go.jp

Japan has been a worldwide leader in *Arabidopsis* research and is continuing that tradition by moving forward into the world of functional genomics. In Japan, ongoing programs for *Arabidopsis* functional genomics are found at RIKEN Genomic Sciences Center Plant Functional Genomics Research Group (http://pfgweb.gsc.riken. go.jp/), RIKEN Plant Science Center (http://www.psc.riken.go.jp/indexE.html), Kazusa DNA Research Institute (http://www.kazusa. or.jp/en/plant/), the CREST program of the Japan Science & Technology Corporation, and NEDO project. Both the RIKEN Genomic Sciences Center Plant Functional Genomics Research Group and the Kazusa DNA Research Institute have ongoing bioinformatics programs as well.

Arabidopsis functional genomics research at RIKEN Genomic Sciences Center (GSC – Kazuo Shinozaki and Minami Matsui) includes (i) collection and phenotype analysis of Ds-tagged lines (Takashi Kuromori), (ii) collection of full-length cDNAs (Motoaki Seki), (iii) collection and phenotype analysis of activation tagging lines (Miki Nakazawa), (iv) full-length-cDNA-overexpressing transgenic lines (Takanari Ichikawa), (v) structural proteomics of plant regulatory proteins with novel structures in collaboration with Protein Research Group of RIKEN GSC (PI: Dr. Shigeyuki Yokoyama) and (vi) transcriptome analysis of genes expression in response to both abiotic and biotic stress using RAFL full-length cDNA microarray analysis (Motoaki Seki). Further work on reverse proteomics for functional analysis of in vitro expressed proteins using the wheat germ cell-free protein synthesis system is taking place at RIKEN GSC, in collaboration with a group at Ehime University (Yaeta Endo, Principal Investigator). The RIKEN Plant Science Center (Takuji Wada and Kiyotaka Okada) is active in phenotype analysis of Ds-tagged lines in collaboration with RIKEN GSC (Takashi Kuromori). At the Kazusa DNA Research Institute (Satoshi Tabata), ongoing projects include the collection of T-DNA tagged lines and *Arabidopsis* and Lotus japonicas ESTs. A major project is the genomic sequencing of Lotus japonicas.. In addition, *Arabidopsis* T87 cultured cells have been transformed with RAFL cDNAs and other cDNAs for metabolic profiling of primary and secondary metabolites (Daisuke Shibata and Kazuki Saito).

Several groups at other centers and universities are also involved in *Arabidopsis* functional genomics. The projects involve metabolic profiling in *Arabidopsis* (Chiba University - Kazuki Saito), genome-wide analyses of the two-component system (Takeshi Mizuno), cell wall genes in *Arabidopsis* and rice (Tohoku University - Kazuhiro Nishitani), small G proteins (RIKEN - Akihiko Nakano), P450 genes (RIKEN PSC - Yuji Kamiya), and transcription factor function using repressor domain and overexpressors (Agency of Industrial Science & Technology in Tsukuba - Ohme-Takagi and Kaoru Suzuki).

RIKEN BRC (http://www.brc.riken.jp/lab/epd/Eng/) is funded by the National Bioresource Project of Japan and collects various plant resources from Japanese research institutes and universities. The RAFL clones, Ds-tagged lines and Activation tagging lines mentioned above are distributed from tRIKEN BRC. This year, RIKEN BRC takes over the distribution service of the ecotypes and mutants of *Arabidopsis* from the Sendai *Arabidopsis* Seed Stock Center (SASSC; Nobuharu Goto). Since established in 2001, RIKEN BRC has already distributed approximately 7,000 *Arabidopsis* resources to the world. Masatomo Kobayashi (kobayasi@rtc.riken.jp) is in charge of *Arabidopsis* resources distribution at RIKEN BRC.

Other funding opportunities for *Arabidopsis* functional genomics in Japan include CREST of Japan Science and Technology Corporation (http://www.jst.go.jp/EN/), the Program of Promotion of Basic Research Activities for Innovative Biosciences (http://www.brain. go.jp/welcome-e.html), NEDO, and Grants-in Aid for Science from the Ministry of Education, Science, Culture and Sports.

The Netherlands

http://www.*arabidopsis*.org/info/2010_projects/Netherlands.html

Contacts:
Maarten Koornneef
Wageningen University, Wageningen
maarten.koornneef@wur.nl
Willem Stiekema
Plant Research International, Wageningen
w.j.stiekema@plant.wag-r.nl

In 2003, many previously established research groups continued active research using *Arabidopsis* studying a wide variety of topics from signal transduction to ecological questions. Groups conducting *Arabidopsis* research are located in all Dutch universities working on plants and in research institutes such as Plant Research International. *Arabidopsis* was the major research object during the national experimental plant science days in Lunteren, attended by more than 300 participants.

The Dutch genome program Centre for Biosystems Genomics (CBSG), headed by Willem Stiekema (www.biosystemsgenomics.nl), was started in 2003. *Arabidopsis* groups received funding for research on 'quality' (metabolic content), protein interactions, plant/pathogen interactions, and chromatin studies. The Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), the Dutch National Science Foundation, continued support for a rather limited number of *Arabidopsis* projects and only a few new projects have been funded. However, two young professors, Marcel Dicke and Corné Pieterse, have received large personal grants which have a major contribution for *Arabidopsis*. In additional, several ´personal´ projects have been awarded to young researchers who make use of *Arabidopsis*.

Very recently, a large national program on proteomics has been initiated: The Netherlands Proteomics Centre (NPC), http://www. netherlandsproteomicscentre.nl/. Medical, animal, microbial, and plant researchers collaborate in this program. The plant projects within NPC will focus mainly on *Arabidopsis* and comprise research on protein interactions, -complexes, ligand-receptors, glycosylation, and novel MS-techniques (coordinator: Gerco Angenent). The NWO has funded the Wageningen Phytoinformatics group, also headed by Willem Stiekema. Their research on bioinformatics issues related to plants is being successfully continued. This group is also involved in the EU PLANET project (www.eu-plant-genome.net) that aims at developing and delivering a high level plant genome database for the systematic exploration of *Arabidopsis* and other plants. Furthermore, the participation of Dutch groups in various European Union projects (e.g., NATURAL, CATMA and CAGE projects, APOTOOL, REGIA, PLANTREC EXOTIC) has either been continued or ended.

Nordic *Arabidopsis* Network

http://www.*arabidopsis*.org/info/2010_projects/Nordic.html
Contact: Jaakko Kangasjärvi
University of Helsinki
Helsinki, Finland
jaakko.kangasjarvi@helsinki.fi

All Scandinavian countries have their own national research funding system. The Nordisk Forskerutdanningsakademi (NorFA), Nordic Research Academy, is funding a 5-year (2001-2005) Nordic Network for groups that are involved in research with *Arabidopsis*. The Nordic *Arabidopsis* Network aims at keeping the groups in regular contact with each other, and it also offers small mobility grants for graduate students and post docs for short-time exchange between groups.

Norway has initiated a national functional genomics program, FUGE. Also, a Norwegian *Arabidopsis* Research Centre has been created. Proteomics is performed in Oslo (UIO, Aalen lab), mutant/clone-collection at the Agricultural University (NLH, Rognli lab) and genomics in Trondheim (NTNU, Bones lab). The intention is that these three labs will serve the plant community in Norway (coordinated by Atle Bones, University of Trondheim).

In Sweden, the Umeå Plant Science Center (UPSC) has been created by moving plant groups from the Umeå University and Swedish University of Agricultural Sciences (Umeå) to the same building. UPSC groups have also received National Center of Excellence status and funding for functional genomics. Their activities are mainly concentrated in trees (hybrid poplar). However, *Arabidopsis* functional genomics is heavily utilized for the determination of the function of poplar genes that have a well-conserved counterpart in *Arabidopsis*. The UPSC is also a partner in the European CATMA-project. Groups from the Uppsala University are involved in two EU-projects that aim at the elucidation of several transcription factors groups in *Arabidopsis*.

The Finnish groups involved in *Arabidopsis* research are concentrating on stress-physiology and functional genomics of plant stress responses, developmental and hormone biology, and in photosynthesis. They are using genomics and proteomics to determine plant defense and adaptation to biotic and abiotic stresses and the function for the proteins in chloroplast thylakoid membranes. In the Spring of 2003, the Finnish Plant Functional Genomics Research Program was created in order to increase Finnish participation in the European functional genomics activities.

The Icelandic investigators involved in *Arabidopsis* research have promoted *Arabidopsis thaliana* as a model research plant within the Icelandic research community. The Danish activities in *Arabidopsis* functional genomics are primarily concentrated on plant-pathogen interactions and plant defense responses, and in photosynthesis.

United Kingdom

www.*arabidopsis*.org/info/2010_projects/United_Kingdom.html
Contact: Ottoline Leyser
University of York, Heslington, York
hmol1@york.ac.u

The major funding agency for plant science in the UK is the Biotechnology and Biological Science Research Council, BBSRC. The BBSRC is encouraging applications that use genomic technologies and has launched several initiatives to stimulate research in this area. The BBSRC Exploiting Genomics Initiative now funds several *Arabidopsis* functional genomics projects. Other initiatives include proteomics, metabolomics and systems biology. More information about these can be found at http://www.bbsrc.ac.uk/science/initiatives/.

GARNet, the Genomic *Arabidopsis* Resource Network, has established infrastructure and expertise to provide reliable and efficient user-driven and publicly available functional genomics resources for *Arabidopsis* research. GARNet started in January 2000 with funding from the UK BBRSC for a three-year period. Funding has recently been extended for further three years to allow establishment of cost recovery systems from GARNet users. Information on GARNet is available via the GARNet web pages http://garnet. *Arabidopsis*.org.uk. GARNet Resources include transcriptome, proteome and metabolite analysis services. Insert clone libraries and a screening service are available from GeTCID and additional insertional mutagenesis populations generated in the first funding period are now available at the Nottingham *Arabidopsis* Stock Center, NASC. Also available at NASC is a large database with results from GARNet Affymetrix experiments and a database containing proteomics results. A database for metabolomics results is being developed. In addition to the GARNet program, many leading universities and institutes in the UK have established their own functional genomics resource centers.

NASC (http://*Arabidopsis*.info/) makes a wide range of material available to the research community such as seeds, DNA and database information. NASC has an agreement with the ABRC in that they both stock the same lines as safety copies and the onus of acquiring, curating, bulking, and distributing is shared by both centers. Distribution from NASC alone is about 30,000 tubes of seed per year worldwide. Data resources made available from NASC include http://atensembl.*arabidopsis*.info, a comprehensive genome browser bringing together a variety of resources including MIPS and TIGR annotation linked to germplasm information, and an extensive database of Affymetrix GeneChip* data. NASC also provides an international (not-for-profit) genechip hybridization service with the sole purpose of increasing public availability of high quality plant gene chip data.

Sir Henry Wellcome Functional Genomics Facility (SHWF) provides a number of technical resources for the functional analysis of genomes and proteomes. These facilities include microarray, proteomics and bioinformatics services. SHWF resources are available to a wide range of UK scientific researchers including the *Arabidopsis* community http://www.gla.ac.uk/functionalgenomics.

Arabidopsis research groups in the UK are involved in several Europe-wide research initiatives, including European Union Framework Program 5 research projects (i.e., REGIA, EXOTIC, CONFAB, EDEN, GVE, PLANET, NATURAL, NONEMA, AGRIKOLA and CATMA). In addition, several genome-related applications have been submitted for Framework Program 6 research projects. Finally, GARNet has teamed with GABI, the German plant functional genomics initiative, and Génoplante, the French functional genomics program, to organize an annual international functional genomics meeting called Plant GEMs.

United States

http://www.*arabidopsis*.org/info/2010_projects/United_States.html

Contacts:
Philip Benfey
Duke University, Durham, North Carolina
philip.benfey@duke.edu
Mary Lou Guerinot
Dartmouth College, Hanover, New Hampshire
Mary.Lou.Guerinot@Dartmouth.edu

The *Arabidopsis* research community in the United States is coordinated by the North American *Arabidopsis* Steering Committee which consists of six elected members who serve four-year terms. Two members rotate off every year. Two members of the Committee represent the U.S. on the Multinational *Arabidopsis* Steering Committee.

The National Science Foundation (NSF) initiated the *Arabidopsis* 2010 Project in fiscal year 2001. The program's goal is to determine the function of 25,000 genes in *Arabidopsis* by the year 2010. The current foci of the Project are to determine the function of a network of genes and to develop research tools and resources that enable the entire research community to participate in the 2010 activities. NSF requires that the 2010 awards be coordinated with similar activities worldwide, that the investigators post publicly the identity of genes under investigation, and that the outcome of the awards (data, information and materials) be made available to the public according to the timetable approved by NSF. Twenty-seven projects were funded under this program in 2001, a further twenty projects in 2002 and 20 more projects were funded in 2003. In May 2004, for the 2010 fiscal year 4, grant proposals were co-reviewed with the AFGN grant proposals at the NSF in order to avoid unnecessary duplications and to support further collaboration between the two projects. Abstracts can be found at http://www.*arabidopsis*.org/info/2010_projects/2010_Abstracts.html. The NSF expects to continue the *Arabidopsis* 2010 Project for 10 years, although the focus of the project may change.

In addition to the *Arabidopsis* 2010 Project, other activities related to *Arabidopsis* research are supported by various programs at NSF, including individual research projects, workshops/meetings, information resources and informatics tools development, and the biological resource center, ABRC. The Center for Eukaryotic Structural Genomics (http://www.uwstructuralgenomics.org/) has been funded by the National Institutes of Health (NIH) to solve three-dimensional structures for many of the proteins of the *Arabidopsis* proteome.

NSF award information can be found at https://www. fastlane.nsf.gov/a6/A6AwardSearch.htm. The U.S. Department of Agriculture, the U.S. Department of Energy and the NIH, especially the National Institutes of General Medical Sciences, support many research projects involving *Arabidopsis*, although they do not have a funding program specifically targeted to *Arabidopsis* research. NIH awards can be searched at http://commons.cit.nih.gov/crisp3/Crisp_Query.Generate_Screen

In early 2004, the North American *Arabidopsis* Steering Committee (NAASC) sent out a questionnaire to individuals in the *Arabidopsis* research community using the electronic *Arabidopsis* news group. The survey was initiated in response to concerns about funding priorities and strategic decisions that are impacting the future direction of plant research. Many of the compiled answers, concerns, suggestions and recommendations were integrated into this MASC report where appropriate.

Members of the Multinational *Arabidopsis* Steering Committee

Thomas Altmann, Chair
Representing Germany
altmann@mpimp-golm.mpg.de
University of Potsdam
Max Planck Institute of
Molecular Plant Physiology
Golm, Germany

Philip Benfey, Co-chair
Representing the United States
of America
philip.benfey@duke.edu
Duke University

Jorge Casal
Representing Argentina
and Brazil
casal@ifeva.edu.ar
University of Buenos Aires
Buenos Aires, Argentina

Bill Crosby
Representing Canada
bcrosby@cs.usask.ca
University of Saskatchewan
Saskatoon, Saskatchewan,
Canada

Jens Freitag
GABI managing office
c/o Max Planck Institute of
Molecular Plant Physiology
GABI managing office
freitag at mpimp-golm.mpg.de

Ian Furner
Representing the
United Kingdom
ijf@mole.bio.cam.ac.uk
University of Cambridge
Cambridge, United Kingdom

Mary Lou Guerinot
Representing the United States
Mary.Lou.Guerinot@
Dartmouth.edu
Dartmouth College
Hanover, New Hampshire, USA

Pierre Hilson
Representing Belgium
pierre.hilson@gengenp.rug.ac.be
Ghent University

Heribert Hirt
Representing Austria
HEHI@gem.univie.ac.at
Vienna Biocenter,
Vienna, Austria

Gerd Jürgens
Representing Germany
gerd.juergens@uni-tuebingen.de
University of Tübingen
Center for Plant
Molecular Biology
Tübingen, Germany

Jaakko Kangasjärvi
Representing the Nordic
Arabidopsis Network
jaakko.kangasjarvi@helsinki.fi
University of Helsinki
Helsinki, Finland

Maarten Koornneef
Representing the Netherlands
maarten.koornneef@wur.nl
Wageningen University
Wageningen, The Netherlands

Ottoline Leyser
Representing the United
Kingdom
hmol1@york.ac.u
University of York
Heslington, York, United
Kingdom

Keith Lindsey
Representing the
United Kingdom
keith.lindsey@durham.ac.uk
University of Durham
Durham, United Kingdom

Sean May
Representing the
United Kingdom
arabidopsis@nottingham.ac.uk
Nottingham *Arabidopsis*
Stock Center
Loughborough, United Kingdom

Peter McCourt
Representing Canada
mccourt@botany.utoronto.ca
University of Toronto
Toronto, Ontario, Canada

Veronica Ongaro
GARNET Coordinator
vo100@york.ac.uk
University of York
Heslington, York, United
Kingdom

Javier Paz-Ares
Representing Spain
jpazares@samba.cnb.uam.es
Centro National de
Biotechnoligia
Madrid, Spain

Kazuo Shinozaki
Representing Japan
sinozaki@rtc.riken.go.jp
RIKEN Genomic Sciences Center
RIKEN Tsukuba Institute
Tsukuba, Ibaraki, Japan

Ian Small
Representing France
small@evry.inra.fr
INRA-Evry
Evry, France

Willem Stiekema
Representing The Netherlands
w.j.stiekema@plant.wag-ur.nl
Plant Research International
Wageningen, The Netherlands

Paola Vittorioso
Representing Italy
paola.vittorioso@uniroma1.it
University of Rome
"La Sapienza"
Rome, Italy

Geoffrey Wasteneys
Representing Australia
and New Zealand
geoffw@rsbs.anu.edu.au
Australian National University
Canberra, Australia

Jianru Zuo
Representing China
jrzuo@genetics.ac.cn
Chinese Academy of Sciences
Beijing, China

Isabell Witt
Executive Secretary
and Coordinator
witt@mpimp-golm.mpg.de
University of Potsdam c/o Max
Planck Institute of Molecular
Plant Physiology
Golm, Germany

Members of the MASC Subcommittees

Bioinformatics

Fabrice Legeai
flegeai@infobiogen.fr

Sean May
arabidopsis@nottingham.ac.uk

Klaus Mayer
Kmayer@gsf.de

Yasukazu Nakamura
yn@kazusa.or.jp

Chris Town (chair)
cdtown@tigr.org

Yves van de Peer
yvdp@gengenp.rug.ac.be

cDNAs and Clone-Based Functional Proteomics (ORFeomics)

Pierre Hilson (chair)
pihil@gengenp.rug.ac.be

Kazuo Shinozaki
sinozaki@rtc.riken.go.jp

Ian Small
small@evry.inra.fr

Chris Town
cdtown@tigr.org

Multiparallel Analytical Tools & Phenotype Analyses

Gerco Angenent
g.c.angenent@plant.wag-ur.nl

Mike Beale (chair)
mike.beale@bbsrc.ac.uk

Jerry Cohen
cohen047@umn.edu

Paul DuPree
p.dupree@bioc.cam.ac.uk

Oliver Fiehn
fiehn@mpimp-golm.mpg.de

Mary Lou Guerinot
mary.lou.guerinot@dartmouth.edu

Heribert Hirt
hehi@gem.univie.ac.at

Martin Kuiper
makui@gengenp.rug.ac.be

Keith Lindsey
keith.lindsey@durham.ac.uk

Sean May
arabidopsis@nottingham.ac.uk

Lutz Nover
nover@cellbiology.uni-frankfurt.de

Kiyotaka Okada
Kiyo@ok-lab.bot.Kyoto-u.ac.jp

Göran Sandberg
goran.sandberg@genfys.slu.se

Motoaki Seki
mseki@rtc.riken.go.jp

Mike Sussman
msussman@facstaff.wisc.edu

Reverse and Forward Genetic Stocks

Thomas Altmann
altmann@mpimp-golm.mpg.de

Steve Henikoff
steveh@howard.fhcrc.org

Maarten Koornneef
maarten.koornneef@wur.nl

Minami Matsui
minami@postman.riken.go.jp

Magnus Nordborg
magnus@usc.edu

Randy Scholl
scholl.1@osu.edu

Ian Small
small@evry.inra.fr

Mike Sussman
msussman@facstaff.wisc.edu

Satoshi Tabata
tabata@kazusa.or.jp

Paola Vittorioso
paola.vittorioso@uniroma1.it

Bernd Weisshaar (chair)
bernd.weisshaar@uni-bielefeld.de