

2011 MASC Meeting

Date/time: Saturday, 25th June, 12.30-14.00; Beefeaters Room, Madison, WI, USA

Action items:

- Kazuo Shinozaki (Japan) retired as the MASC Chair and the current Co-Chair Mark Estelle (CA, USA) becomes the new MASC Chair (2011 – 2012). Wolfram Weckwerth (Vienna), nominee in absence, was voted to become the new MASC Co-Chair (2011-2012).
- ICAR 2012 (Europe) will be in Vienna, 3-7th July at the Hofburg Palace, organized by Marie-Theres Hauser, Magnus Nordborg and Wolfram Weckwerth.
- ICAR 2013 (Australasia) will be in Australia, Sydney, 24-28th June. It is organized by Barry Pogson.
- ICAR 2014 (American continent) venue is yet to be decided. Proposed locations include Vancouver, San Francisco, Las Vegas, Alaska, Seattle, Madison and Hawaii. It will be organized by with the help of the NAASC Coordinator (Joanna Friesner)
- MASC decided to continue its activities and that it will need a MASC Coordinator after 2012, when the current BBSRC (UK) grant that supports it terminates.
- The MASC 2020 Roadmap document outlining the scientific goals for the next decade is to be finalized in the very near future and should include a timeline. The Committee will provide feedback on the expected timeline of the goals. A MASC progress report will continue to be prepared and distributed at the next ICAR (Vienna, in 2012).
- A MASC Subcommittee Workshop and a MASC session should be held at ICAR 2012 to build a more direct relationship with the community. It is envisaged that a MASC session will be also held at ICAR 2012. The MASC Chair and the MASC Coordinator will liaise with ICAR 2012 to facilitate these events.

Discussion Points:

MASC Roadmap: From Molecules to Fields – A Roadmap for International Arabidopsis Research 2011-2020: At the MASC 2010 Meeting in Yokohama, Japan, MASC decided to continue its activities and this was unanimously reiterated at the MASC 2011 Meeting, where there was a general enthusiasm to continue MASC. Since 2010 marked the end of the Functional Genomics project a new theme for this decade needed to be identified and articulated in a document that could then be used to inform funders and the rest of the scientific community on the developments of Arabidopsis research. The document has been prepared during the year by the MASC Coordinator with help of the MASC Chair and Co-Chair and has been circulated to the rest of MASC for feedback in March 2011.

- The Roadmap document identifies the following scientific goals for the next decade: A) Build a predictive model of an Arabidopsis plant from its molecular parts; B) Exploit the wealth of natural variation that exists in Arabidopsis to further our understanding of adaptation and evolution; C) Establish an effective knowledge exchange pipeline from the laboratory to the field and vice versa; D) Build the International Arabidopsis Informatics Consortium (IAIC), an international informatics and data infrastructure; E) Deepen international Cooperation and Coordination.
- The Committee agreed on the scientific objectives and goals of the Roadmap (see Roadmap document at the end of these Minutes). The committee suggested minor changes and agreed to provide feedback to enable the construction of a timeline to be included in the document.
- The Committee decided it would be most useful to have a MASC session at ICAR 2012, Vienna, Austria to address the specific objectives of the Roadmap. This would provide an excellent opportunity to explain the Roadmap and have talks for each of the objectives.

MASC Coordinator position

- With current BBSRC (UK) funding that supports the MASC Coordinator position ending 30th Nov 2012, arrangements must be made to ensure a smooth transition by the end of 2012 and give the chance to another country to support this position. The Committee discussed different possible funding models, including a distributed system that used profits from ICAR. However, this would not ensure stability as ICAR profits can vary greatly between conferences and in many cases ICARs do not make any profits.
- The Committee agreed that after 2012 the position should be supported via a grant to ensure stability and commitment from a Coordinator and conversations with different funding agencies about the feasibility of supporting this position have started.

MASC Subcommittees

- The committee agreed that a reassessment of the subcommittees of MASC will be needed in response to the new decadal plan, some subcommittees could be merged and some added if necessary. In addition, it was thought that all subcommittees should include a translational aspect.
- A lunchtime short open workshop (organized by Josh Heazlewood) was held at ICAR 2011 to showcase the activities of the MASC subcommittees to the rest of the community. At the MASC 2011 Meeting it was decided that this was very useful to both MASC and the rest of the community and that it would be most desirable to have a MASC Subcommittee workshop next year at ICAR 2012 in a workshop-dedicated slot to encourage attendance and disseminate information on the activities of the subcommittees in the community.

Bioinformatics for Arabidopsis: Nicholas Provart summarised the new developments to work towards the establishment of an International

Arabidopsis Informatics Consortium (IAIC) to fulfil the informatics needs of the community.

A summary of the workshops held in 2010 (Nottingham, UK and Washington DC, USA) and the Plant Cell Report was given. In addition updates on the Research Coordinated Network (RCN) application with NSF, which has recently been funded, were given. A Scientific Advisory Board (SAB) and a Scientific Advisory Panel (SAP) should be established and RCN-IAIC is looking for candidates to appoint for the SAB. Each MASC Subcommittee and country representatives will be invited to nominate 1-2 people each.

MASC Meeting Attendees Present:

Thomas	Altmann	IPK Gatersleben- Germany rep.
Ruth	Bastow	GARNet- UK rep. and BBSRC
Katia	Baerenfaller	Proteomics Subcommittee
Jim	Beynon	GARNet- New GARNet/MASC Coord. Grant PI
Danny	Chamovitz	Israel rep
Brian	Dilkes	Purdue University, US – Nat Var and Comp Genomics Subcommittee Chair
Joe	Ecker	Salk Institute, USA
Mark	Estelle	UCSD, CA, USA – incoming MASC chair and ICAR 2011 lead
Joanna	Friesner	USA rep, NAASC Coordinator
Michael	Groszmann	CSIRO, Australia – representing ICAR 2013 team
Erich	Grotewold	ABRC
Josh	Heazlewood	JBEI/Lawrence Berkeley Lab- Proteomics Subcommittee
Pierre	Hilson	Belgium rep.
Masatomo	Kobayashi	RIKEN BRC- ICAR 2010 team
Irene	Lavagi	University of Warwick –UK-- MASC Coordinator/GARNet UK
Minami	Matsui	RIKEN PSC- Phenomics Subcommittee
Sean	May	NASC/European Stock Centre- UK rep ex officio and NASC
Basil	Nikolau	Metabolomics Subcommittee
Scott	Poethig	Pennsylvania University, USA –USA rep and NAASC President
Nicholas	Provar	University of Toronto – Canada, Bioinformatics Subcommittee Chair
Randy	Schöll	ABRC- ex officio
Kazuo	Shinozaki	RIKEN PSC- Japan rep., MASC chair
Stefanie	Wienkoop	ICAR 2012
Viktor	Zarsky	Czech Republic rep
Karen	Cone	National Science Foundation- Observer
Diane	Okamuro	National Science Foundation- Observer

From Molecules to Fields - A Roadmap for International Arabidopsis Research 2011-2020

Executive summary

As a number of projects and funding programmes involving Arabidopsis research came to a close in 2010, the international Arabidopsis community decided to articulate a new road map for Arabidopsis research entitled “From Molecules to Fields” to inform scientists, funding bodies and decisions makers on the future foci of Arabidopsis research.

Drawing from a number of reports, discussions and recommendations that have emerged in recent years including the more recent EU 2020 Vision for Plant Science (6) and An international Model for the Future of Plant Science (7), this document details the future direction of an internationally coordinated road map for Arabidopsis research within the wider Plant Science landscape.

This roadmap aims to continue the substantial international effort on Arabidopsis research to understand the fundamental processes of plant biology with sufficient detail to engineer plant specifically to meet the future needs of society.

To achieve these goals, it will be necessary to integrate vast amounts of heterogeneous quantitative data generated both in the laboratory and in the field. The integration of these data at many levels of biological organisation, including different accessions of Arabidopsis and its close relatives, will allow researchers to build functional networks from the single molecule to the whole plant community, that will in turn provide the knowledge required to address current global challenges.

To establish an effective pipeline from the laboratory to the field and vice versa the Multinational Arabidopsis Steering Committee in consultation with the wider community have formulated a road map for the future of Arabidopsis research which is based on the following objectives:

- A) **Build a predictive model of an Arabidopsis plant from its molecular parts;** this will require embracing novel technologies to collect data that capture the complexity of the biological system both in the laboratory and in the field, thus providing a link to environmental factors enabling the construction of *in silico* models with predictive value.
- B) **Exploit the wealth of natural variation that exists in Arabidopsis to further our understanding of adaptation and evolution;** By utilising emerging technologies, such as next generation sequencing and phenotyping platforms, Arabidopsis ecotypes and Arabidopsis relatives will be analysed to uncover genetic variation thus developing an efficient method to allow prediction of phenotype from the genotype.
- C) **Establish an effective knowledge exchange pipeline from the laboratory to the field and vice versa;** Promoting exchange of information and personnel between the laboratory and the field and engaging with other plant communities will provide an efficient two-ways exchange flow. This will facilitate the translation of basic biology, increase the awareness of the needs and requirements of agriculture and commerce and help us to understand how basic science can best be feed into to practical applications.
- D) **Build the International Arabidopsis Informatics Consortium (IAIC), an international informatics and data infrastructure;**

Developing an integrated system to coordinate, store, visualize and analyze the heterogeneous data produced by objectives A-C, will provide the infrastructure required to enable the accomplishment of the above objectives. This will require strong links with data providers and users from other scientific communities as well and collaborations with computational experts. An international Informatics Consortium will permit the integration of existent databases and develop standards to facilitate access to all public data.

E) Deepen International Cooperation and Coordination;

International Cooperation and coordination will be necessary to achieve all objectives of this roadmap, to minimise duplication of efforts and maximize efficient use of resources through collaboration.

The next decade will present many challenges for plant science, including the sustainable production of food and energy. To meet these challenges the international plant research community will need to maintain the fundamental research that provides key advances in understanding whilst simultaneously guaranteeing that this augments practical advances through efficient knowledge sharing and beneficial collaborations. For example, research in commercially important species such as wheat is beginning to blossom but technology advances have yet to become available to overcome the obstacles posed by their less tractable genetics. It is therefore of fundamental importance that Arabidopsis continues to evolve to be the key plant for providing the underpinning information on basic plant biology. For this purpose the integration of Arabidopsis research with its close relatives and other plant species and model organisms will be key to the future success of plant research.

To achieve the new decadal vision "From Molecules to Fields" researchers must continue to build on the ever-expanding sets of data, tools and resources so that plant science can develop more vigorously than ever before.

Foreword

In the decade since the Arabidopsis genome has been sequenced (1) research on this reference plant has flourished. Publically available shared genomic tools and resources such as structural and functional genome annotation (2,3) and collections of genome wide insertion mutations (4) have allowed researchers to make major advances in plant and general biology. The outstanding progress that has been achieved over the past decade results from the substantial investment and support from funding bodies across the globe and a collaborative international community with the Multinational Arabidopsis Steering Committee (MASC) at its centre.

In 2010 a number of projects and funding programmes for Arabidopsis research that began a decade earlier came to a close. The international Arabidopsis community therefore decided to come together to articulate a new vision for Arabidopsis research entitled '**From Molecules to Fields**' to help inform scientists, funding bodies and decisions makers as to the future direction of Arabidopsis research. This vision is based on the discussions and recommendations of a set of reports that have been commissioned in recent years to evaluate the past success of Arabidopsis and plant biology and the role plants will play in our future including, '2020 Vision for Plant Biology' (5), 'EU 2020 Vision for Plant Science' (6) and 'An International Model for the Future of Plant Science' (7).

The goal of this new decadal vision 'From Molecules to Fields' is to understand the

fundamental processes of plant biology with sufficient detail to be able to accurately predict both outputs of the system and the effect of perturbations. Only by obtaining an in-depth of knowledge of how a genome sequence is translated into a continuum of processes from the single molecule to cells and tissues, the whole plant, populations of plant, and finally the field, will researchers be able to solve the perplexing question of how genotype maps to phenotype. This information is essential in order to engineer plants specifically to meet societies future needs.

Achieving the objectives of the roadmap will require precise quantitative data obtained at many levels that can be used to build functional networks of interactions from the single molecule to the whole plant community. In addition the parallel development of informatics and computational platforms to collect, store, visualize and interpret the diverse array of biological data will be essential. These ambitious goals will require considerable effort and coordination. It will be a significant challenge to combine heterogeneous data types into *in silico* models with predictive value. However the outputs and breakthroughs generated by this effort will drive plant science forward in the future.

As we progress into the next decade a major challenge faced by plant sciences is how to maintain curiosity-based research that generates fundamental advances whilst simultaneously ensuring that this new knowledge flows into practical programmes such as plant breeding. To help achieve this balance it will be essential that Arabidopsis is not viewed as separate from the broader plant science community. Rather, Arabidopsis research should be an essential component in a continuum of discovery from the basic to the applied in which plant scientists collaborate with each other and scientists from other disciplines to ensure that underpinning knowledge generated at the bench can be exploited in the field.

Introduction

Humans depend on plants for almost every aspect of life, from the food we eat to the medicine we rely upon. However this interdependent relationship is being stretched to its limits as a result of an ever-increasing world population and the effects of climate instability. Understanding how plants grow, develop and interact with their environment is therefore more important now than ever. Only by continuously expanding our knowledge of plant processes can we hope to meet the pressing need for food, feed, fuel, shelter and to mitigate the negative effects of climate change.

Arabidopsis thaliana is a member of the Brassicaceae family that includes mustard, cabbage and oilseed rape. As a result of its small size, rapid generation time, small genome and ability to generate a large amount of seed by self-pollination Arabidopsis has emerged as the organism of choice to study plant biology. This has led to a significant investment in Arabidopsis research over past decades that have generated a fully annotated genome and a myriad of tools and resources, which, in turn, have led to a flourishing international community of researchers. Combined, these efforts have made Arabidopsis the most studied of all flowering plants and a reference for all biology; in 2010 over 3,400 peer-reviewed papers were published on Arabidopsis (8). By focusing on a single tractable system, the international community has made dramatic advances in our understanding of plants and basic biology including the identification of the major plant hormone receptors (9-17), the elucidation of florigen (18-19) and the identification of small RNAs (20-21). As a result of this outstanding progress, the Arabidopsis research community is now ideally placed to help address current and future global challenges.

Plant science is currently undergoing an era of great change that is being driven by genome scale technologies, theoretical based approaches and informatics. As we look to

the future, new technologies such as next generation sequencing will inevitably lower the barriers to research into previously intractable economically important plants, as well as lesser-studied wild plant species. As a consequence of this progress, genomic approaches will be applied to a wider range of plant species, which will in turn allow the fundamental knowledge, accrued in Arabidopsis to be more readily translated into plants of agricultural and commercial importance. The broad availability of emerging technologies is also driving a convergence of disciplines and fields across plant biology that have been historically separate, e.g., plant breeding and ecology.

Arabidopsis research played a central role in bringing about revolutionary changes. The wealth of information and resources available for Arabidopsis, in addition to the inherent strengths that led to its initial adoption as a reference system, will increase the importance of Arabidopsis as the nodal experimental model system for plant science for the foreseeable future. The vast toolbox that is available to Arabidopsis researchers is likely to result in this being the plant species in which novel forms of data are first accrued and exploited, and new techniques tested and perfected. Arabidopsis is also uniquely placed to contribute to our understanding of ecology and evolution. The abundant natural variation that exists amongst Arabidopsis accessions and relatives coupled with a rich knowledge base make Arabidopsis ideally suited for understanding adaptation and evolution at the mechanistic level.

Despite these and other advantages, Arabidopsis is not the best model system to study a number of plant processes, such as C4 metabolism or nitrogen fixation. At the same time, there are examples in which Arabidopsis research has contributed to our understanding of plants in some very unexpected areas. Arabidopsis is not a halophyte yet a clever screen for salt hypersensitivity (22) provided a plethora of information on a signalling pathway that is not only central to salt tolerance and sodium homeostasis in Arabidopsis but that also exists in crops and halophytes (23-28). It will therefore be essential that Arabidopsis along with a small number of other key plant species continue to be utilised for focused studies of plant biology.

Research in Arabidopsis is central to our ability to generate fundamental knowledge that can inform a broad research base impacting on medical, plant, environmental and evolutionary sciences, and to provide plant breeders with the knowledge and resources (i.e. techniques and tools) to achieve sustainable crop production. If Arabidopsis research is to help address the current challenges facing the planet, we must therefore ensure that Arabidopsis is the first step in a pipeline that begins in the laboratory and ends in the field.

Molecules to Fields - An International Coordinated Roadmap for Arabidopsis Research

In the last decades of the 20th Century, Arabidopsis and plant biology research largely focused on a reductionist approach to answering biological questions. Studies often centred on understanding a single gene/group of genes in specific pathways or processes. This approach had been extremely successful, and built a solid foundation from which Arabidopsis research community was able to move onto the era of genomics toward the end of the 1990's. Completion of the whole genome sequence of Arabidopsis in 2000 opened up a new way to study life processes in plants. New methodologies such as systems biology are revealing that genotype to phenotype relationships are much more complex than previously thought; mutations don't always generate distinct phenotypes and/or a phenotype is often dependent on environmental conditions. We therefore have to re-evaluate concepts such as a distinct genetic blueprint in which

there is a simple linear relationship between a gene and its biological output. A major challenge for the future is to understand this level of complexity so that genomic sequences can be linked to physiological outcomes. This will provide the fundamental knowledge base required for the rational improvement of plants for traditional agriculture, for the vast array of plant-based products, and for efforts to improve our environment.

We therefore propose a new decadal vision for International Arabidopsis research “From Molecules to Fields”. The goal of this effort will be to understand the fundamental processes of plant biology in sufficient detail to accurately predict both outputs of the system and the effect of perturbations. This will require an in-depth knowledge of how the genome is translated into a continuum of processes from the single molecule, to cells and tissues, the whole plant, plant populations, and finally the field. Such a goal will present a substantial challenge for the next decade as it will require accurate and quantifiable data obtained at many levels including transcripts, methylation patterns, chromatin patterns, comparative genomics, proteomics, metabolomics, protein interactions, protein location and temporal data. This quantitative data must then be translated into functional networks at various scales, linked to mathematical models, compared across thousands of accessions, and associated with meteorological and geographical information so that we can eventually provide real time data and information as a plant grows or an epidemic unfolds. Such a vision will also demand new research tools, the development of new experimental methods and an internationally coordinated data infrastructure to collect, store, visualise and interpret this diverse array of biological and theoretical data.

The investments made in Arabidopsis over the past decades have created a wealth of data, tools and resources, which in turn have led to the creation of a flourishing and collaborative international body of researchers. As a result of this investment the Arabidopsis research community is now uniquely positioned to address the goals outlined in this document.

Objectives

To achieve the vision of ‘Molecules to Fields’ the international Arabidopsis community have identified five main objectives

- A. Build a predictive model of an Arabidopsis plant from its molecular parts
- B. Exploit the wealth of natural variation that exists in Arabidopsis to further our understanding of adaptation and evolution
- C. Establish an effective knowledge exchange pipeline from the laboratory to the field and vice versa
- D. Build the International Arabidopsis Informatics Consortium (IAIC), an international informatics and data infrastructure
- E. Deepen International Cooperation and Coordination

A. Build a Predictive Model of Arabidopsis from its Molecular Parts

The profusion of genome scale datasets that have emerged over the past years has allowed researchers to address how pathways function and interact globally. This type of ‘systems based’ approach has illustrated that a plant is not just a sum of its constituent parts but a complex set of interactions that occur on various levels and scales to generate phenotypes at the whole plant and population level.

Systems biology approaches have begun to transform our understanding of how genes function as part of a whole system. By exploiting and combining 'omics' data such as transcriptomics, transcription factor binding, proteomic and metabolomic data researchers have already begun to reconstruct networks and pathways, infer functional interactions and uncover emergent properties. However, this is just the beginning. As technology advances, larger and more diverse datasets will become available. When integrated together it will become possible to build and combine computational and theoretical models that enhance our understanding of the complex and dynamic networks that underpin plant growth and development.

Arabidopsis provides an ideal plant system in which to undertake a system-based approach to build a computable plant. In common with other plants, Arabidopsis has a simple body plan of repeating units, a small set of stem cells that continually develop organs throughout the whole life cycle, and unlike animals, all plant growth is postembryonic. Plant development and growth is also very robust as a result of the sessile nature of plants allowing researchers to easily impose internal/external perturbations and measure the effects from the molecular level to the whole plant. However, it is the unparalleled breadth and depth of Arabidopsis data that are available to researchers that sets Arabidopsis apart from other plants. These data, coupled with the ease of genetic manipulation and myriad of genetic resources, allow researchers to easily and consistently test hypothesis generated by modelling approaches. Thus Arabidopsis is a very amenable organism in which to generate the systems level understanding of biological organisation that will be required for the rational engineering of plants to meet the challenges we face.

Goals:

1. Collect and collate accurate and quantifiable data obtained at multiple levels of abstraction including a wide range of genomic transcripts, methylation patterns, chromatin modification patterns, comparative genomics, proteomics, metabolomics, DNA-protein interactions, protein-protein interactions, protein location, cell type specific, phenomic, spatial and temporal data. Collected data must also describe the variation amongst Arabidopsis accessions and relatives. The scale of this task should not be underestimated and will need to be repeated as new concepts, tools and technologies arise.
2. Continue to develop collections of mutants and mutant lines for researchers to test hypotheses generated from modelling approaches. This needs to be accompanied by support of public and accessible stock centres that enable all researchers to access this material
3. Develop new research tools and experimental methods to address the lack of global assays for a number of plant processes. (Include barriers in proteomics, metabolomics, imaging and informatics; examples needed from the MASC community).
4. Fully exploit emerging technologies. For example Next Generation Sequencing is providing unparalleled opportunities for novel applications (29). Any cellular process or assay for which the output is nucleic acid based can in theory be exploited by next generation sequencing to generate a toolkit that is specific, sensitive, quantitative and flexible enough to deal with the unknown novelty of sequence regulation that has yet to be discovered

5. Build the International Arabidopsis Informatics Consortium (IAIC), an international informatics infrastructure that is able to handle, store, visualise and interpret a diverse array of biological and theoretical data that has already been generated and data that will be created in the future.
6. Translate quantitative data into functional network models and maps that describe plant processes at various levels of organisation and temporal and spatial scales.
7. Develop new methods for modelling across scales to build a computation/mathematical model that stretches from the molecule to the field.
8. Compare and integrate models to generate an *in silico* Arabidopsis plant.
9. Combine this *in silico* model with meteorological and geographical data to provide data and information from the field as a plant grows or an epidemic unfolds.
10. Provide training for plant scientists in computational science and theoretical approaches.
11. Encourage further collaboration between plant biologists and theoretical scientists.

B. Exploit the wealth of natural variation that exists in Arabidopsis to further our understanding of adaptation and evolution

One of the long-standing goals of plant biology has been to link the variations that can be observed in natural/wild populations to the underlying genotypic variation, with the aim of exploiting this variation to engineer beneficial traits into plants of agricultural and commercial importance.

Arabidopsis has proven to be the most efficient plant model for the genetic analysis of natural variation because of the ease of manipulation, breadth and depth of understanding of genetic and biochemical pathways and most importantly its abundance of natural variation. Combined, these attributes have allowed the largest number of genes and nucleotide polymorphisms underlying natural variation to be uncovered in Arabidopsis compared to any other plant species (30). Despite this success, it still remains a considerable challenge to map genotype to phenotype. A great deal of the natural variation we observe results from complex, quantitative, multi-gene traits, which are greatly influenced by environmental factors. Recent advances and cost reduction in high throughput sequencing combined with progress in statistical techniques indicate that it will be possible to identify the causal genetic polymorphism underlying natural variation via genome wide association (GWA). GWA exploits variation in a collection of genotypes/individuals to identify genotype-phenotype associations. Projects to explore this emerging area are already underway and the most ambitious of these is the 1001 Genomes project (31), which aims to indentify the whole-genome variation in 1001 Arabidopsis accessions. By providing access to this unexploited wealth of variation, Arabidopsis researchers can begin to link phenotypic differences with genotypic variations on a global scale.

In addition to variation in DNA sequence, large scale data sets from other 'omic' technologies such as proteomics, metabolomics ionomics and epigenomics are also

beginning to be utilised to assess the variation that exists at different levels of biological organisation and molecular regulation. By combining large scale 'omic' data sets obtained across populations it is possible to bring the power of systems biology to quantitative genetics. This approach can be used to construct and infer biologically meaningful regulatory networks (32-34) that can reveal emergent properties that play a role in systems robustness and adaptability, and are crucial to the survival and adaptation of species.

Natural variation and comparative genomics are also helping to provide insights into another fundamental question in biology. Why does the genome sequence of Arabidopsis make an Arabidopsis plant and not one that looks more like Capsella, Brassica, Cleome or cotton? Or to put it another way, how do specific genetic differences result in the diverse phenotypes of different plant species? Arabidopsis and its sequenced relatives are providing researchers with an unsurpassed set of tools to analyse the genetic basis of developmental, metabolic or physiological differences. This will help to deliver a genetics based view of evolution and ecology and ultimately provide an understanding of how the diversity within the plant kingdom is encoded.

By combining the power of emerging technologies with the extensive knowledge base that has built up in previous decades, the Arabidopsis community is perfectly poised to lead the way in the utilization of natural variation to understand how sequence variation affects biological and evolutionary processes.

Goals:

1. Continue to develop genomic sequencing and computational resources in order to understand and utilize the natural variation of Arabidopsis and related species. This will include close interactions with the Brassicaceae Map Alignment Project (BMAP) and the 1001 Arabidopsis Genomes project.
2. Survey the variation in Arabidopsis and related species at numerous levels of biological organisation and regulation including genomic, proteomic, metabolomic, and epigenomic to infer biological networks.
3. Develop tools and techniques to facilitate the identification of QTLs that have subtle effects on plant phenotypes
4. Utilise the information gathered in Arabidopsis and related species to undertake comparative genomics/ comparative evolution/ comparative ecological genomics
5. Develop appropriate open access informatics and data infrastructure for storage, retrieval and analysis of variation and QTL data
6. Develop accessible statistical and computational methods for the analysis of natural variation and QTL data.
7. Analyse ecotypes and Arabidopsis relatives in association with their rhizosphere, endophyte and epiphyte communities in various ecological settings.
8. Develop high-throughput methods for measuring phenotypes in the lab and the field

9. Extend systems monitoring from the defined conditions of the laboratory to the variable conditions of the field.

C. Establish an effective knowledge exchange pipeline from the laboratory to the field and vice versa

Plant science is currently undergoing an era of great change that is being driven by genome scale technologies, theoretical approaches and informatics. For example, the advances we have witnessed in genome sequencing will inevitably lower the barriers to research using previously intractable economically important plants, as well as lesser-studied and wild plant species. As a consequence, the fundamental knowledge that has been accrued in *Arabidopsis* can now more readily inform the development of tools and techniques, such as marker assisted breeding, for the generation of improved crops for agriculture and commerce. The broad availability of emerging technologies is also driving a convergence of disciplines and fields within plant biology that have been historically separate. This period of revolutionary change provides plant biologists with an opportunity to work together to construct a pipeline from basic to applied science that is capable of tackling global problems.

Although not of economical importance per se, *Arabidopsis* will have a major role to play in providing the first step in this pipeline. The investments that have been made in *Arabidopsis* research over past decades have generated a fully annotated genome and multitude of publicly available tools and resources, as well as a highly trained and creative international body of researchers. The *Arabidopsis* community is therefore ideally positioned to deliver underpinning knowledge in key areas such as the regulation of carbon fixation, growth, development, nutrient uptake, source-sink relationships and biotic and abiotic stresses. A detailed knowledge of these fundamental processes will provide the basis for the rational improvement of plants to meet our needs for food, fuel, shelter, plant based products and to mitigate the effects of climate change.

As we progress into the next decade and beyond, a major challenge that plant researchers will face is how to maintain curiosity-based research that generates the fundamental advances, whilst simultaneously ensuring that this new knowledge flows into plant breeding programmes. To help achieve this delicate balance it will be essential that *Arabidopsis* is not viewed as separate from the wider plant science community but rather as an essential component in a continuum of discovery from the basic to the applied.

To achieve this objective we will need to meet the following goals:

1. Undertake the approaches outlined section in sections A and B to help us understand important complex traits such as: yield, cell wall biogenesis, plant disease resistance, water-use efficiency, nutrient-use efficiency, photosynthesis, generation of biomass and growth. This is needed to generate a knowledge base that can help develop plants to meet our future needs.
2. Promote active dialogue, knowledge and data exchange between plant communities and various fields of expertise. This will increase awareness of the richness of plant research, bring together previously separate communities, and help identify potential bottlenecks and barriers to translation.

3. Develop a data and informatics infrastructure in which underpinning knowledge generated in Arabidopsis can flow easily to plant breeding.
4. Showcase examples of the role of Arabidopsis in rational improvement of plant species for agriculture and other plant-based industries through the annual MASC report and annual International Conference on Arabidopsis (ICAR).
5. Promote exchanges of information and personnel between Arabidopsis groups and those working on other plant species and vice versa to increase the flow of knowledge and expertise between different communities and to help develop an understanding of how translation from basic to applied research could be improved. Combined with workshops, conference sessions and representation with the other plant communities, this would promote an efficient two-way knowledge exchange.
6. Promote knowledge exchange with data providers/users of other model organism communities and facilitate interactions with computational/theoretical researchers.

D. Build an International Informatics and Data Infrastructure

Over the past decade there has been an unprecedented explosion of data, tools and resources. In virtually all areas of experimental biology, databases and repositories have evolved to become a central tool in the modern laboratory, with *in silico* experiments now regularly playing a role in the processes of scientific discovery.

In common with many other communities, Arabidopsis researchers have to deal with a range of new data types, informatics approaches and tools, whilst simultaneously trying to cope with the mountain of data that they are faced with on a daily basis. In order to provide solutions to these challenges members of the International Arabidopsis community recently came together to assess its future informatics needs. The major conclusion of the meetings was that the Arabidopsis community was at a critical junction and it was recommended that it should move beyond current structures to develop novel approaches to data management access and integration via the initiation of the International Arabidopsis Informatics Consortium (IAIC) (35). The IAIC would consist of globally distributed system of data and resources that would be managed by a single international consortium capable of leveraging expertise and funding on a global scale. Existing data and databases will be the basis for the new infrastructure. The development of such an international consortium would be essential to achieving the goals outlined in this vision document, as they are heavily reliant upon the development of an appropriate cyber-infrastructure that is able to handle, store, visualize and interpret this diverse array of biological and theoretical data.

Only by developing a coordinated infrastructure will we be able to take full advantage of the current 'data deluge' and exploit future technological advances. This will allow researchers to effectively translate data and information into scientific understanding and improve our ability to address global challenges.

Goals:

1. Develop an international coordinated informatics and data infrastructure such as the IAIC, that is able to store, integrate, visualise and analyze a plethora of heterogeneous experimental and theoretical data, which would be generated by objectives A-C of this vision in addition to those previously generated by the Arabidopsis community
2. Develop an infrastructure that is flexible enough to respond to new/ future approaches in data management, access and integration.
3. Develop appropriate data standards and establish their widespread use within the community
4. Generate an infrastructure that promotes data exchange and collaboration. For example, to ensure that integration of data allows users to move vertically between Arabidopsis associated data as well as horizontally to other plant species and model organisms.
5. Ensure all data and resources generated are available via the appropriate public data repositories.
6. Ensure there is interoperability between the data and resources generated by the Arabidopsis community and those generated by other communities.
7. Establish strong links with other data providers/ users and computational experts to allow exchange of information and best practice.

E. Deepen International Cooperation and Coordination

The challenges that plant biologists must address are too large for any single country to tackle alone. International coordination of both research and investment will therefore be essential to the future success of plant science research.

The Arabidopsis community has a long tradition of international collaboration and sharing of data and resources that is exemplified in the sequencing of the Arabidopsis genome (35) and the Multinational Coordinated *Arabidopsis thaliana* Functional Genomics Project. In each of these initiatives MASC undertook a pivotal role of community coordination to ensure success. Originally established in 1990 to promote international cooperation and assist in the free exchange of ideas and information, MASC has continued to provide a unifying voice for the international community in subsequent decades. In the current period of financial uncertainty the role of MASC will be more important than ever if we are to achieve the ambitions of the 'Molecules to Field' vision outlined above, leverage national investment to support this vision, and synergize outcomes.

Goals

1. Continue to represent each country that is undertaking Arabidopsis research around the globe.

2. Increase awareness of the richness of international Arabidopsis research via the production and distribution of the annual MASC report and the International Conference on Arabidopsis Research.
3. Help coordinate international Arabidopsis research, to minimize duplication of efforts and maximize efficient use of resources through collaboration.
4. Promote open communication and free exchange of data, materials, resources and ideas among the Arabidopsis research community.
5. Liaise with funding agencies supporting Arabidopsis research.
6. Provide coordination for the 'Molecules to Field' roadmap
7. Periodically assess the status of the 'Molecules to Field' roadmap and adjust goals as required.

References

1. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana* *Nature* **409** 796-815
2. <http://www.arabidopsis.org>
3. <http://arabidopsis.info>
4. Alonso, J.M. et al (2003) Genome-wide Insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**:653-57
5. 2020 Vision for Biology Workshop: The role of plants in addressing grand challenges in Biology, 3-5 January 2008, Arlington, VA, USA <http://www.arabidopsis.org/portals/masc/workshop2020.pdf>
6. Beynon, J. Coupland, G. Graham, I. Harter, K. 2020 European Vision for Plant Science, 2-3 June 2008, Bonn, Germany
7. http://www.arabidopsis.org/portals/masc/2020_European_Vision.pdf
8. Dennis, L. Ecker, J. Gale, M. Harter, K. Kersey, P. Lee. S> Matsui, M. Millar, A. Parkin, I. van Sluys, M.A. Traas, J (2011) An international Model for the Future of Plant Science, March 2009, Banbury Center, USA
9. The Multinational Arabidopsis Steering Committee (2010) The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project http://www.arabidopsis.org/portals/masc/2010_MASC_Report.pdf
10. Dharmasiri, N. Dharmasiri, S. Estelle, M. (2005) The F-box protein TIR1 is an auxin receptor. *Nature* **435**, 441–445
11. Dharmasiri, N. et al. (2005) Plant development is regulated by a family of auxin receptor F box proteins. *Dev. Cell* **9**, 109–119
12. Kepinski, S. & Leyser, O. (2005) The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature* **435**, 446–451
13. Ueguchi-Tanaka, M. et al. (2005) GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* **437**, 693–698
14. Chini, A. et al. (2007) The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* **448**, 666–671
15. Thines, B. et al. (2007) JAZ repressor proteins are targets of the SCF(CO11) complex during jasmonate signalling. *Nature* **448**, 661–665
16. Melotto, M. et al. (2008) A critical role of two positively charged amino acids in the Jas motif of Arabidopsis JAZ proteins in mediating coronatine- and jasmonoyl isoleucine-dependent interactions with the CO11 F-box protein. *Plant J.* **55**, 979–988
17. Ma, Y et al. (2009) Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* **324**, 1064-8
18. Park, SY et al. (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* **324**, 1068-71
19. Tamaki, S. et al. (2007) Hd3a Protein is a Mobile Flowering Signal in Rice. *Science* **316**, 1033-36
20. Corbesier, L. et al. (2007) FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. *Science* **316**, 1030-33
21. Hamilton, A. Baulcombe, D. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–2
22. Baulcombe, D. (2004) RNA silencing in plants. *Nature* **431**, 356-363.
23. Zhu, J.K. (2000) Genetic Analysis of Plant Salt Tolerance Using Arabidopsis. *Plant Physiol* **124**, 941-8
24. Weinl, S. Kudla, J. (2009) The CBL-CIPK Ca(2+)-decoding signaling network: function and perspectives. *New Phytol* **184**, 517-528
25. Ho, C.H. et al. (2009) CHL1 functions as nitrate sensor in plants *Cell* **138**, 1184-1194
26. Li, L. et al. (2006) A Ca(2)+ signaling pathway regulates a K(+) channel for

- low-K response in Arabidopsis. *PNAS* **103**, 12625-30
27. Xu, J. *et al.* (2006) A Protein Kinase, Interacting with Two Calcineurin B-like Proteins, Regulates K⁺ Transporter AKT1 in Arabidopsis. *Cell* **125**, 1347-60
 28. Remans, T. *et al.* (2006) The Arabidopsis NRT1.1 transporter participates in the signaling pathway triggering root colonization of nitrate-rich patches. *PNAS* **103**, 19206-19211
 29. Amtmann, A. *et al.* (2008) The effect of potassium nutrition on pest and disease resistance in plants. *Physiol Plant* **133**, 682-691
 30. Jiao and Meyerwitz (2010) Cell-type specific analysis of translating RNAs in developing flowers reveal new levels of control. *Mol. Sys. Bio* **6** 149
 31. Alonso-Blanco C. *et al* (2009) What has natural variation taught us about plant development, physiology, and adaption? *Plant Cell* **21** 1877-96
 32. Weigel D and Mott R (2009) The 1001 genomes project for Arabidopsis thaliana. *Genome Biol* **10** 107
 33. Keurentjes J *et al* (2006) The genetics of plant metabolism *Nature Gen* **38** 842-9
 34. West M *et al* (2007) Global eQTL mapping reveals the complex architecture of transcript-level variation in Arabidopsis *Genetics* **175** 1441-1500
 35. Keurentjes *et al* (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci *PNAS* **108** 1708-1713
 36. International Arabidopsis Consortium (2010) An international bioinformatics infrastructure to underpin the Arabidopsis community *Plant Cell* **22** 2530-2536