

2020 Vision for Biology Workshop: The role of plants in addressing grand challenges in biology

"Plant Systems 2020: A systems understanding of development and adaptation at the level of cells, tissues, organisms and ecosystems"

Summary and Recommendations

During the last 20 years Arabidopsis has emerged as the primary experimental system for essentially all aspects of plant biology. By focusing on a single tractable system, the international Arabidopsis community has made dramatic advances in nearly every area of plant research. Further, because of the close evolutionary relationships between all flowering plants, discoveries in Arabidopsis have been readily translated to other plant species such as economically important crops. In addition, discoveries made in Arabidopsis have impacted research in animal systems including disease processes in human.

The remarkable success of Arabidopsis research is partly the result of wise investment by the NSF, first through the Arabidopsis Genome sequencing program and attendant technology development and subsequently via the Arabidopsis 2010 Program. This project, now nearing its completion, has funded the generation of a broad range of powerful genetic and genomic resources and technologies. The Arabidopsis toolbox, together with the unique qualities of Arabidopsis and allied species, will now facilitate effective studies at all levels of biological organization including, molecular, cellular, organismal, and ecological. In addition, the Arabidopsis 2010 Program has fostered the development of a vigorous and dynamic international community of researchers, a process that has included the training of many graduate students and postdoctoral researchers, and recruited many scientists not trained initially in plant biology.

Because of this investment, the Arabidopsis research community is ideally, and uniquely, positioned to address the Grand Challenges in biology as described below. A true systems biology approach, encompassing all of life's components from molecules to populations, is now possible using Arabidopsis. To achieve this goal we make the following general recommendations:

- 1) NSF should continue to provide major and specific support to integrate molecular, cellular, organismal, and ecological research on Arabidopsis as a system to understand how a living organism develops, functions and adapts to its environment.
- 2) Funds should be provided for development of additional and new types of large-scale experimental genomics resources that will be required to effectively address the Grand Challenges.
- 3) Efforts should be made to encourage the development of new quantitative approaches to the study of biological systems using Arabidopsis and allied species. This should involve the development of collaborations between biologists, mathematicians, computer scientists, engineers and scientists in other quantitative disciplines.

- 4) Data acquisition should remain a major focus of future programs to fuel iterative cycles of data analysis, integration, hypothesis generation and testing. The emergence of new technologies will enable the collection of new and higher quality data of all types, thus permitting more sophisticated systems analyses.

Rationale for the Workshop

In November 1998, a group of scientists met in Palo Alto, California in an NSF sponsored workshop to discuss the grand challenges facing biology at that time. Their primary aim was to identify areas in which plant science could provide important and lasting contributions to addressing these challenges. A key recommendation from their report was to focus sufficient resources on the reference plant *Arabidopsis* to be able to exploit fully its soon to be completed genome sequence. An action plan subsequently developed from a second meeting at the Salk Institute gave rise to the NSF *Arabidopsis* 2010 program. Nine years later, as we near completion of the 2010 program, whose ultimate goal “is to determine the function of every gene in *Arabidopsis*”, it appeared appropriate to step back, assess progress and once again ask how plants can contribute to addressing the grand challenges in Biology.

What is crystal clear is that there are a host of real-world problems that involve plants: escalating pressures for food, the need for renewable biofuels, and habitat preservation are being driven by the demands of an ever growing and ever more prosperous human population. The ability to face these challenges is dependent on a fundamental understanding of the basic biology and ecology of plants. NSF-funded plant biology research in the past decade has provided the foundations for solving some of these pressing problems. By identifying the next set of grand challenges in biology, and how they can best be met by plant research, we will be preparing for the unforeseen problems of tomorrow. For this purpose, a workshop was held on Jan 3-5, 2008 in Arlington, VA that focused on the future directions for plant research, with special emphasis on the role of the reference species *Arabidopsis* in uncovering fundamental biological principles that will enable us to face future challenges to our wellbeing and that of the global environment.

Major Accomplishments of the *Arabidopsis* 2010 project

Although still two years from completion, the *Arabidopsis* 2010 Project has done nothing less than revolutionize plant biology. Because of its rapid generation time and ease of genetic manipulations, *Arabidopsis* has attained universal acceptance by the plant science community as the “reference” system for nearly all-biological processes. The term “reference” species indicates that *Arabidopsis* is more than a model for plants; it is itself a plant closely related to most plants of economic and ecological importance, so that inferences are direct. In the larger biological science community, because of the data produced from the 2010 and genome sequencing-projects, *Arabidopsis* is utilized as the reference organism for domain/kingdom comparisons with animal model systems, such as *Drosophila*, *C. elegans* and mouse.

An extensive genetic tool kit and a large variety of genomic technologies have been produced as major outcomes of the 2010 Project. Examples include vast collections of sequence-indexed gene mutations, thousands of RNA, protein and metabolic profiles, dozens of sequenced

accessions/strains and associated nucleotide variation information, as well as tens of thousands of recombination vector-based open-reading-frame clone reagents that are allowing construction of protein interactome networks.

This expansive toolbox has enabled new approaches to address many long-standing questions in plant biology. For example, detailed knowledge of the pathways for biosynthesis of, and responses to all of the major plant hormones has been obtained. Further examples of major breakthroughs include a detailed understanding of the plant immune system, new insights into the mechanisms of plant evolution, adaptation and ecology, and a deeper appreciation of genome organization at all levels, from nucleotides to chromosomes. Cellular systems level understanding of root and flower development and other developmental processes are now emerging. Moreover, we now recognize and strive to understand the unexpected level of complexity and robustness of plant developmental processes in response to genetic perturbation. In several frontier areas of biology, studies resulting from Arabidopsis are particularly notable. For example, in the field of epigenetics, studies of the mechanisms of DNA methylation and small RNAs in plants have led the way to new insights that are organism independent.

Another major accomplishment of the 2010 Project has been the demonstration that groundbreaking science is possible utilizing a reference plant system. Beyond providing an essential reference for nearly all plant physiological processes, researchers studying human disease now routinely come to plant biologists for insights into how genes work. This is because genome sequencing revealed that 40% of human disease-related genes are found in Arabidopsis; the functions of which, in many cases, are best understood in Arabidopsis. For the first time since the rediscovery of Mendel in the early 1900's, researchers from outside of plant biology can no longer ignore findings from plant research.

An unexpected success of the Arabidopsis 2010 project is that this effort has acted as a magnet to attract scientists from other fields (such as computer scientists, as well as biologists trained in microbial, yeast and animal systems) not typically involved in studies of plant biology. For example, work in the 2010 program has fostered modeling approaches not previously attempted in the plant sciences. Importantly, the Arabidopsis 2010 Project has also facilitated the integration of experimentation, theory and modeling, providing a single experimental system to generate, iteratively test and validate new models and theories.

Finally, from an international perspective, the Arabidopsis 2010 Project, with its attendant focus on basic research, has provided a major contribution to international plant sciences, which are largely limited to crop plant species.

What role can plant science play in addressing the grand challenges in biology?

Grand challenges facing biology were identified as:

1. How do cells work and how do they interface with the environment?
2. How do single cells develop into multicellular organisms?
3. How do genomes generate organismal robustness and diversity?
4. What is the molecular basis of evolution?
5. How are levels of biological organization integrated from molecules to ecosystems?
6. How can the environment be made sustainable for future generations?

Answers to these grand challenges will contribute both to a deeper intellectual understanding of life on earth and to addressing current challenges facing mankind. Many experimental systems can and will be employed to meet these challenges. To best understand biological systems, scientists need to be able to compare and contrast independently evolved processes across organisms. Among the three independently evolved multicellular lineages – fungi, plants and animals, experimental systems featuring plants are exquisitely well suited for addressing many of these grand challenges.

Plant biology research profoundly, and increasingly, informs fundamental human activities. Hence, detailed knowledge of plant biology, and the ability to manipulate plant growth in a predictive manner, is required to improve agricultural productivity while simultaneously minimizing its environmental impact. Plant biology is, therefore, essential to human health and wellbeing. Plant biology however has been woefully under-funded compared to biomedical research focused on human disease in both the public and private sector. We are now at a crossroads where this discrepancy continues at our peril. The incongruity between our ability to manipulate and control plant growth and development, contrasted, for example, to our ability to detect and treat many human diseases, will soon have negative societal consequences. The need for detailed understanding of the basic biology and ecology of plants has never been greater. In a globally heating world, a renewed and enlarged focus on plant biology to meet these challenges must be combined with sound basic science to guide ecological stewardship. These goals require a deeper understanding of plant growth and development as shaped by interactions of plants with biotic and abiotic factors across various populations and in varied ecosystems.

In the next 10 years, plant biology will move from a descriptive to a predictive science. It will have its foundation in a detailed understanding of the systems that control the behavior of individual cells and the interactions between them, resulting in outputs at the whole plant level. To achieve this will require transformation of molecular maps into networks of functional interactions within living cells, the integration of the levels of biological organization from molecules to ecosystems, and the parallel development of mathematical platforms to handle, quantify, integrate and interpret biological data obtained from diverse experimental platforms. This will require a detailed knowledge of what happens when we alter networks of genetic components, through mutations induced in the laboratory or mined from natural variation and adaptation in ecologically relevant settings. Plants are very well suited to experimental studies and theoretical modeling to drive new rounds of experiment across all levels of biological organization.

Almost all major research breakthroughs in plant science in the last 20 years have relied on development of *Arabidopsis thaliana* as a reference system. The large and vibrant Arabidopsis research community is furthermore a model for international collaboration. All aspects of basic plant biology will continue to be informed by studies in Arabidopsis and allied species and genera, since they collectively provide both the deepest and broadest experimental tools and features to rapidly enable new discoveries relevant to all of the grand challenges listed above. Paradigms established using Arabidopsis have, and will continue, to be applied to crop species (see for example the 2008 NRC report on the National Plant Genome Initiative). This paves the way for rational improvements in a variety of agricultural traits. It is essential that major, dedicated support continue for studies of basic biology in Arabidopsis, as the breakthroughs in this reference plant will assuredly blaze the most efficient path to the future optimization of species of economic importance.

Biological grand challenges and how to address them

1. How do cells work and how do they interface with the environment?

The complex functions performed by multicellular organisms ultimately depend on processes that occur within cells. The cell theory, formally articulated in 1839 by Schleiden & Schwann was based on the observation of cells initially in plants, the discovery of which predates other great paradigms of biology discovered in plants including Mendel's laws of inheritance (1865). A deep understanding of how cells work is essential for understanding how organisms develop and function within their environment. This requires dynamic and quantitative information about the control of gene expression, protein modifications, protein transport and protein localization, protein-protein interactions, protein-nucleic acid interactions and the metabolic status of cells in the context of the tissues and organs in which they reside. Advances in genomic and proteomic technology have made it possible to establish maps of the genetic and protein interaction networks in cells. A major challenge is to transform these molecular maps into networks of dynamic interactions within living cells and to develop mathematical models that accurately describe and predict cellular phenomena. Plants are particularly well suited for these studies because they undergo processes common to all eukaryotic cells and, in addition, perform specialized functions essential to life on earth. These include capturing energy from sunlight and assimilating essential nutrients from the air and soil. Furthermore, as they are sessile, the exquisite responsiveness of plant cells to environmental conditions is an important feature of their biology that must be understood if we are to predict the long-term consequences of environmental change.

The reference plant *Arabidopsis* will continue to be used extensively for the analysis of cellular processes because of its rich genetic and genomic resources in addition to its rapid generation time. Advances in imaging technologies coupled with genome-wide knowledge of cellular components will provide the starting points for understanding the dynamics of regulatory networks controlling cellular states.

2. How do single cells develop into multicellular organisms?

How the fertilized zygote develops into a multicellular organism has intrigued biologists for over 150 years. Plants are excellent systems for investigating how single cells develop into multicellular organisms, as they possess a relatively small number of cell types, distinct, positionally cued cell lineages and no cell migration. During development, all multicellular organisms face a similar set of challenges: they must create a diverse set of specialized cell types, organize these cells into functional tissues and organs, and maintain pools of stem cells throughout their lifetimes. Organisms must have developmental networks robust enough to ensure that they correctly assemble their parts, but flexible enough to modify development in the face of injury or environmental change. The availability of nutrients, presence of toxins, and changing climate have all been linked to altered development and viability of both plants and animals. In sessile organisms like plants, the extent to which the environment affects development is magnified.

The ability to grow large populations of *Arabidopsis* in the lab or in the field in defined environmental conditions makes *Arabidopsis* a powerful experimental system to dissect the contribution of specific environmental conditions to development. The availability of genomic

sequences for multiple ecotypes also allows *Arabidopsis* researchers to uncover the genetic changes underlying specific responses to the environment in natural populations. One mechanism that contributes to developmental flexibility is the production of stem cells that can renew and replenish organs during the life of the organism. How stem cells are created and maintained (currently a subject of great interest for human health) can be easily studied in plants. Plant stem cells are found in discrete and known locations, they are readily accessible, and they are maintained for exceedingly long periods of time. For example, stem cells at the apex of trees can provide the materials to produce new organs for hundreds of years. These attributes make plants an attractive model to study all aspects of stem cell biology. Live imaging of stem cell populations and their differentiating progeny in *Arabidopsis* has already started to provide answers about how stem cells are formed and maintained at a cellular resolution not currently possible in animal models. Finally, as biologists attempt to account for, measure and model the emergent properties of complex systems, the advances in single tissue genome-wide profiling allow *Arabidopsis* researchers the unique ability to monitor how individual cell types change RNA, protein and metabolite levels in response to perturbations of the whole organism.

3. How do genomes generate organismal robustness and diversity?

A critical question underlying our understanding of the diversity of life on earth is how highly related or even identical genomes can yield distinctly different individuals. Plant science has led the way in elucidating connections between the genetic blueprint and its biological output. For example, plants provide exquisitely sensitive reporter systems such as flower and seed pigmentation that have driven key discoveries in gene activation and silencing, including seminal work on genome instability caused by transposable elements, elucidation of telomere function, and the discovery of small RNAs. *Arabidopsis* has provided the proof-of-concept system showing that we can now map the dynamics of the genome at every level from DNA sequence to DNA methylation patterning to genomic association of histones and transcription factors.

A related question is how genomes balance beneficial and deleterious changes to drive overall robustness of the organism. Plant genomes are remarkably plastic, and thus illustrate important mechanisms to achieve this balance. For example, plant genomes typically maintain long segmental duplications that provide opportunities to diversify gene function without compromising fitness. Even more strikingly, plants frequently use polyploidization - the co-existence of multiple related genomes in the cell nucleus - as a means to evolve highly robust individuals. Notably, many important crop plants such as wheat and cotton are polyploid.

Robustness also depends on the ability to make rapid changes in gene expression in response to a broad range of changing environments. The dynamics of accessing genomic information through changes in chromatin state is especially critical in plants, which due to their sessile life style have no choice but to adapt rapidly to environmental challenges including extremes of light, temperature, and nutrition. Plants thus offer a rich source of gene regulation strategies, especially higher-order networks, the nature and structure of which we are just beginning to understand.

4. What is the molecular basis of evolutionary change?

Elucidating the molecular basis of evolutionary change is one of the major challenges for modern biology. Understanding how genotype maps into phenotype is a critical first step toward

understanding the molecular basis of evolution. Although, itself a major challenge, a complete understanding of adaptation is only possible if we understand the genetic differences that cause adaptively important phenotypic differences, and a complete understanding of speciation must include knowledge of the genetic differences that underlie species differences.

Plants are ideally suited for studying the genetics of both adaptation and speciation because they are experimentally tractable and exhibit a broad range of adaptations to both the biotic and abiotic environment. The canonical experiments for studying local adaptation, reciprocal transplantation and “common garden” studies, are perfectly suited for plants, but difficult or impossible in many other organisms. With respect to speciation, morphologically and ecologically distinct plant species are often interfertile, making it possible to study the genetics of species differences directly.

The long-term goal of evolutionary biology is to discover general principles and laws governing evolution. Some of the questions have long histories, including “Are the kinds of genetic changes responsible for differences between species the same as the kinds of genetic changes responsible for within-species variability?” and “Is adaptation generally limited by mutation, or is sufficient genetic variability usually present?” Other questions such as: “Are there constraints or fundamental laws underlying the evolution of network states and structures?” can’t be addressed without access to many examples from different kinds of organisms. At the same time, detailed elucidation of the genetics of adaptation and speciation is sufficiently challenging that no progress will be made without focusing on reference systems like *Arabidopsis*. Genomics-based answers to each of the questions posed above are within reach using plant reference systems as study organisms.

Plants are a major part of life on earth, and it is clear that a general understanding of evolutionary processes will have to encompass them. Indeed, plants are likely to lead the way, both because of the ease with which adaptation and speciation can be studied, and because they are well suited to defining the dynamics of networks through applying the insights of systems biology to natural variation from the laboratory to field.

5. How are biological systems integrated from molecules to ecosystems?

The genomic revolution has paved the way for understanding the integrated function of biological systems across molecular, organismal, and ecological scales. *Arabidopsis* is ideal for this endeavor because of the wealth of genomic and functional data and systems biology tools available, and because it is well suited to manipulative ecological experiments. For example, systems biology and gene network modeling provide new tools for examining how perturbation of a gene’s function within a signaling network will affect ecologically important organismal phenotypes, such as flowering time, adaptation to soil nutrients, disease resistance, stress tolerance, and herbivore resistance. These studies can be scaled up to the ecological level to ask how the phenotypic effects of network perturbation, in turn affect fitness and population dynamics in different environments.

This integrated approach promises to shed new light on fundamental ecological problems, such as the factors limiting the distribution and abundance of species. Although factors limiting species ranges or population sizes have sometimes been identified at the ecological level, almost nothing is known about the underlying genetic mechanisms or the genetic changes that might lead to ecological release from such limits. Plants are ideal for investigating the molecular basis of ecological limits because ecological experiments such as transplants and environmental

manipulations are feasible. Arabidopsis is particularly valuable for these studies because the genetic pathways underlying key ecological traits such as stress tolerance, flowering time, disease resistance, or herbivore resistance are rapidly being elucidated, and mutants or natural variants in these pathways can be used for ecological experiments. Population genomic analyses can identify which genes in these pathways are subject to strong selection.

Arabidopsis and its relatives will also be invaluable tools for identifying and understanding the genes controlling ecological interactions and the genetic mechanisms of co-evolution with ecological partners. Rapid progress is being made in elucidating the function of the interacting pathways involved in resistance to pathogens and herbivores at the molecular level as well as the pathogen effectors that promote disease by interfering with plant resistance mechanisms. Extensive genomic data on natural sequence variation has made it possible to test hypotheses about evolutionary dynamics of host resistance and pathogen virulence, linking studies of molecular mechanisms with molecular evolution. At the same time, ecological field experiments have revealed natural selection on resistance in Arabidopsis and its relatives. The opportunity now exists to measure natural selection on specific genes and pathways controlling species interactions in the real world in real time, and to dissect the dynamics of natural selection and co-evolution of host species with pathogens, microbial community, and herbivores on a broad ecological and evolutionary scale.

An integrated approach to plant biology is also essential for understanding how biological systems will respond to changing environments in an era of unprecedented rapid global change. Plants are the basis of terrestrial ecosystem function and their responses to global change such as increasing atmospheric carbon dioxide, global warming, and anthropogenic nitrogen fertilization will determine the ultimate sustainability of life as we know it. Such questions have thus far largely been addressed by ecosystem ecologists with short-term ecological experiments and coarse-grained models, and very little is known about mechanisms for plant acclimation and adaptation to rapid environmental change. Understanding the genetic networks involved in plant responses to environmental drivers will improve our ability to predict the response of plants and ecosystems to different scenarios of global change. Although as an ephemeral annual Arabidopsis may not be a “keystone species,” it nevertheless provides a wealth of genomic tools and resources to identify “keystone genes” critical for the resilience of ecosystem function in response to global change.

6. How can the environment be made sustainable for future generations?

As primary producers, plants are the ultimate source of food, fiber, and fuel for all forms of terrestrial life, including humans. They are the foundation of ecosystem function and the global biogeochemical cycles that sustain life on earth. Around 40% of the earth’s land area is now used for agriculture, with major impacts on global carbon and nitrogen budgets. An integrative systems approach to plant biology will be necessary both to predict the response of biological systems to human perturbations, and to design adaptive strategies to enhance system resilience in the face of rapid environmental change. Designing sustainable agricultural practices and land use strategies will be essential both for ensuring human food and energy security in the face of catastrophic global change and for mitigating further damage to biological systems and human life from such change. Systems biology approaches founded on model systems such as Arabidopsis, united with parallel efforts by ecosystem modelers, can provide the necessary tools to ask: “What are the principles of biological resilience as local and global environmental limits

are reached?” “How do these processes scale from molecular systems to ecosystems?” “Are there universal rules for system resilience?” Identification of such rules will allow us to evaluate different models of extended ecosystem, community, or population disruption to design perturbations that preserve, sustain, or restore system function. By adjusting flexible regulatory modules it may be possible to direct change toward a self renewing, human-coupled natural system.

An integrative systems approach to plant biology will also allow the development of new tools to effect such change through assessment, prediction, and adaptive management of biological systems. For example, the era of genomics has opened the door to quantifying biodiversity. Genomic DNA sequences of organisms can be ordered into lineages that have coevolved by sharing genetic information vertically and horizontally. This knowledge allows diversity to be catalogued and prioritized for conservation. Moreover, existing genomic diversity can be re-assorted into new combinations through genome-assisted plant breeding. Genomes with new functions can be created through experimentally directed evolution to meet new environmental challenges. Alternatively, biotechnology allows novel genomes to be created in-silico and printed via synthetic biology. Plants are remarkably amenable to such synthetic biology based on the knowledge of systems processes at different levels of organization.

Integrative Biology: From Molecules to Systems

The overarching goal of biology is to fully understand life and the operation of biological processes. The achievement of this goal requires detailed knowledge of life's component parts at every level from molecules to ecosystems, as well as a holistic view of these components and their dynamic relationships over space and time. Arabidopsis is uniquely suited to research encompassing these diverse approaches. The development of powerful genetic and genomic technologies will continue to energize ongoing efforts to understand the molecular basis for cellular and physiological processes. These technologies will also facilitate major advances in our understanding of evolution. New analytical tools promise to provide near complete catalogs of all the constituents of every cell type as well as detailed knowledge of macromolecular function. Further, dramatic advances in cell imaging will produce unprecedented views of the cellular environment. Finally, the integration of this diverse knowledge to produce a unified view of life will require the development and application of novel “systems” approaches. These approaches, collectively called Systems Biology can be defined as “the exercise of integrating the existing knowledge about biological components, building a model of the system *as a whole* and extracting the unifying organizational principles that explain the form and function of living organisms” (von Bertalanffy, 1968). Systems biology is poised to play a major role in defining and advancing biology as we plan the course of the next decade of research.

It has long been apparent that macromolecules form complex networks of functionally interacting components, and that the molecular mechanisms underlying most biological processes correspond to particular steady states adopted by such cellular networks (Delbruck, 1949; Novick and Weiner, 1957; Waddington, 1957; Monod and Jacob, 1961; Kauffman, 1969; Thomas, 1973). However, until recently, systems-level theoretical models were based on untestable hypotheses in biological systems because of a lack of supporting experimental data. With the advent of a first generation of genome-wide datasets, and the promise of even larger and more robust datasets to be generated in the near future, analysis of interactions within gene regulatory networks can reveal emergent properties of networks that play a role in systems

robustness and adaptability, which are crucial to the survival and adaptation of species. An example of an emergent property identified through systems analysis is the oscillating properties of interconnected networks in the circadian clock (Locke et al. 2006, Zelinger et al. 2006).

Practically speaking, a systems approach to understanding biological processes can be described as an iterative process that includes (1) data collection and integration of all available information (ideally all components and their relationships in the organism), (2) system modeling (3) experimentation at a global level and (4) generation and testing of new hypotheses (modified from Ideker et al., 2001). With the full genome of several plant species in hand, the challenge ahead is to decode this information into the components (genes and their products), identify their relationships (e.g. regulatory, physical interactions) and build a dynamic and predictive *in-silico* model of a plant cell, tissue, and organism that encompasses all levels of organization.

In Arabidopsis and other model organisms, integration of genomic information via systems biology approaches has provided the first glimpses of genome-wide network interactions between genes, proteins and RNA molecules. A first major goal is to elucidate the organization and interactions among the various network modules that regulate the complex biological processes underlying growth and development. This understanding will entail analysis of the dynamic interactions among genes, RNAs, proteins and metabolites (i.e. the genome, the transcriptome, the proteome and the metabolome). A second major goal is not simply to describe and model what we know, but to predict how the system will react upon perturbation, either by changing the genome and/or epigenome, or by changing the external environment.

The exciting prospect of the next 10 years in biology is for the first time to be able to integrate knowledge across different levels of biological organization and to anchor this at the level of molecular interactions. To accomplish this will require plant biologists to (1) determine the biochemical and physiological function of all cellular components (genes, proteins etc) (2) generate much more detailed genome-wide experimental data of the system states and molecular interactions (3) integrate and analyze these data originating from heterogeneous technologies to make it biologically coherent and build predictive models of the system. Plant sciences in general, and Arabidopsis research in particular, are leading the way in many areas of biology and will play a major role in meeting these challenges.

Broader Impacts

Plant research plays a central role in broadening the appreciation and understanding of biological sciences. Beyond the fundamental appeal of plants, the focus on a single reference organism, Arabidopsis has transformed biology over the past two decades. Arabidopsis provides a uniquely tractable multi-cellular experimental system. Its ease of culture and extensive toolkit make it an entry point to scientific discovery at every level, from the classroom to cross-disciplinary and international collaborations.

Perhaps the most significant way in which Arabidopsis has shaped science today is by drawing in researchers trained in other fields who would not otherwise have considered working with a plant system. For example, many of the current leaders in the Arabidopsis community come from microbiology or animal biology backgrounds. Furthermore, Arabidopsis is leading the way in projects that bring together biologists with chemists, mathematicians, physicists, statisticians, engineers and computer scientists. Arabidopsis is now widely accepted as one of the

premier laboratory organisms for studying basic eukaryotic processes, and it has the added value of serving as a direct reference for other plants.

We see **four general areas** where Arabidopsis research can continue to make major contributions to the infrastructure of biology. **1)** Arabidopsis provides the “how-to” guide for the development of other plants as facile research systems. Expanding the toolkit in Arabidopsis serves as a test bed for possible expansions into additional plants species. The inevitable outward expansion of paradigms established using Arabidopsis will increase our understanding of species-specific processes including interactions with the environment and secondary metabolism. **2)** The early success of multi-disciplinary projects involving Arabidopsis, notably systems biology and chemical genomics, illustrates that Arabidopsis will be centrally important in training a new breed of researchers with broadened expertise. For example, Arabidopsis could provide a major teaching tool in new graduate curricula for computational and systems biology. **3)** Large-scale Arabidopsis projects such as genome sequencing and mutant collections have involved highly successful international collaborations. “Big science” with Arabidopsis can continue to forge scientific strength beyond the US. In particular, since developing countries place such a high value on agriculture-related research, we foresee expanding future collaborative efforts beyond Europe and Japan. **4)** In this age of genome sequences, heightened concern about the environment and dwindling energy and food resources, it is imperative to improve the scientific literacy of the general public so that we can make informed decisions regarding current and future generations.

Arabidopsis research has brought plants back into Biology departments and into non-traditional settings for plant research. Arabidopsis can readily be grown in a classroom due to its small size and rapid generation time, and used to illustrate fundamental concepts such as Mendelian genetics and the connection between DNA sequence and morphological variation. In addition, Arabidopsis provides the ideal experimental organism to study adaptations to a changing environment. We foresee a new paradigm to maximize the impact of bridging university research programs with the community: organizing educational outreach at an institutional or regional level, rather than at the level of the individual PI, to allow coordination of the most effective efforts and provide a framework for monitoring the effectiveness and adjusting the content of educational programs.

Workshop participants:

Judith Bender (Brown University)
Philip Benfey (Duke University) Co-Organizer
Dominique Bergmann (Stanford University)
Justin Borevitz (University of Chicago)
Gloria Coruzzi (New York University) Co-Organizer
Jeff Dangl (University of North Carolina, Chapel Hill)
Caroline Dean (John Innes Centre, Norwich, England)
Joseph Ecker (Salk Institute)
Mark Estelle (Indiana University) Co-Organizer
Jane Glazebrook (University of Minnesota)
Sarah Grant (University of North Carolina, Chapel Hill)
Mary Lou Guerinot (Dartmouth University)
Rodrigo Gutierrez (Catolica U. of Chile, Santiago, Chile)

Jeff Long (Salk Institute)
Magnus Nordborg (University of Southern California)
Scott Poethig (University of Pennsylvania)
Natasha Raikhel (University of California, Riverside)
Johanna Schmitt (Brown University)
Arp Schnittger (University of Cologne, Germany)
Marc Vidal (Dana Farber Cancer Institute, Harvard)

Cited literature:

- Delbruck M** (1949) Discussion. *In* EdCndIR Scientifique, ed, Unités biologiques douées de continuité génétique. Editions Paris, pp 33-35
- Gutierrez RA, Shasha DE, Coruzzi GM** (2005) Systems Biology for the Virtual Plant. *Plant Physiol.* **138**: 550-554
- Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L** (2001) Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network. *Science* **292**: 929-934
- Kauffman SA** (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* **22**: 437-467
- Kitano H** (2001) Foundations of systems biology. The MIT Press, Cambridge, MA
- Locke JCW, Kozma-Bognar L, Gould PD, Feher B, Kevei E, Nagy F, Turner MS, Hall A Millar AJ.** (2006) Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Mol Syst Biol* **2**
- Monod J, Jacob F** (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* **26**: 389-401
- Novick A, Weiner M** (1957) Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci U S A* **43**: 553-567
- Palsson BO** (2004) In silico biotechnology: Era of reconstruction and interrogation. *Current Opinion in Biotechnology* **15**: 50
- Thomas R** (1973) Boolean formalization of genetic control circuits. *J Theor Biol* **42**: 563-585
- Von Bertalanffy L** (1968) *General Systems Theory: Foundations, Development, Applications.* George Braziller, New York.
- Waddington CH** (1957) *The strategy of the genes* Geo. Allen & Unwin, London
- Zeilinger MN, Farre EM, Taylor SR, Kay SA, Doyle FJ.** (2006) A novel computational model of the circadian clock in *Arabidopsis* that incorporates PRR7 and PRR9. *Mol Syst Biol* **2**