Developing common models for molecular mechanisms, crop physiology, and ecological studies: rhythms, stress challenges and growth opportunities

Organizers

Katherine Denby - University of Warwick, Assoc. Prof. Warwick Systems Biology Centre. Tel: +44 (0)24 7657 5097; Fax: +44 (0)24 7657 4500; email: k.j.denby@warwick.ac.uk. Research: Plant signaling mechanisms mediating responses to biotic and abiotic stress. Education: Coordinator of Systems Biology M.Sc. at Warwick. Outreach: Participated in Schools Day at University of Cape Town to raise awareness of science among school pupils.
Stacey Harmer – University of California, Davis; Department of Plant Biology. Tel: (530)752-8101; Fax: (530)752-5410; email: slharmer@ucdavis.edu. Research: Study of the molecular basis of circadian rhythms using the model plant *Arabidopsis*. Outreach: Introducing high school students and teachers and undergraduates from underrepresented minorities to research.
Cynthia Weinig – University of Wyoming; Department of Botany. Tel: (307)766-6378; Fax: (307)766-2851; email: cweinig@uwyo.edu. Research: Mechanisms of adaptation in heterogenous settings; effects of natural variation in circadian rhythms on physiological traits in diverse environments. Outreach: Teacher education; development of evolutionary labs appropriate for use in a secondary school setting.
David Wild – University of Warwick, Prof. Warwick Systems Biology Centre. Tel: +44 (0)24 76150242, Fax: +44 (0)24 76575795, email: d.l.wild@warwick.ac.uk. Research: Bioinformatics, Bayesian statistical machine learning to model gene regulatory networks. Education: Teaches Advanced Bioinformatics for Systems Biology M.Sc. at Warwick. Outreach: Co-organizer of several meetings on computational systems biology.

Designated contact person for group: Stacey Harmer (Tel: (530)752-8101; email: slharmer@ucdavis.edu)

Likely Participants

Reka Albert – Penn State; dynamic modeling of plant signaling networks (Computation)
Tom Brutnell – Cornell; plant/environment responses (Maize, Rice, Sorghum)
Ed Buckler – Cornell; plant/environment responses (Maize)
George Coupland – Max Planck, Cologne; natural variation in flowering time (Arabidopsis)
Raissa D'Souza – UC Davis; modeling complex systems (Computation)
Frank Doyle – UC Santa Barbara; mathematical modeling of circadian clock (Computation)
Jane Glazebrook – U Minnesota; plant/pathogen responses (Arabidopsis)
Stefan Jansson - Umea University; seasonal variation in gene expression/physiology + circadian clock (Poplar)
Tom Juenger - U Texas; ecology; natural variation and drought/stress responses (Ipomopsis and Arabidopsis)
Fumiaki Katagiri – U Minnesota; systems approaches to understanding plant pathogen signaling networks (Arabidopsis)
Dan Kliebenstein – UC Davis; natural variation and plant metabolites (Arabidopsis)
Insuk Lee - Yonsei University; probabilistic functional gene networks (Computation)
Julin Maloof – UC Davis; natural variation and plant responses to light (Arabidopsis)

Rob McClung – Dartmouth; molecular basis of circadian rhythms and natural variation in rhythms (Arabidopsis, Brassica)
Todd Mockler – Oregon State; circadian and diurnal regulation in crop species (Rice, Poplar, Brachypodium)
Tom Oinn - European Bioinformatics Institute; software tools for designing and executing workflow solutions (Computation)
Dave Pink – U of Warwick; crop genetics, crop improvement for food and non-food traits (Brassica and other crops)
Pam Ronald – UC Davis; biotic and abiotic stress responses (Rice)
Herbert Sauro – U Washington; synthetic biology, experimental and computational standards and specifications (Computation)
Mark Stitt – Max Planck, Potsdam; systems approaches to understanding how the environment affects primary carbon and nitrogen metabolism (Arabidopsis and Tomato)
Detlef Weigel – Max Planck, Tubingen; natural variation in flowering time; circadian clock and flowering time, comparative genomics (Arabidopsis)
Xinguang Zhu – U of Illinois; mathematical modeling of photosynthsis (Computation)

Summary of Proposed Workshop

An increasing world population and predicted changes in the global climate demand improvements in plant stress tolerance and crop productivity. These traits are determined by a complex web of interactions between plant genotype and biotic and abiotic factors in the local environment. The ability to understand, predict, and then manipulate plant responses to the environment will be critical for increased food security in the 21st century.

An essential tool to understand complex biological interactions is the construction of computational models. Such models organize information about biological pathways in a systematic manner. Since these models allow scientists to perturb network components and observe the predicted outcomes, they are especially useful when the systems under study are so complex that scientific intuition is no longer an adequate guide for further experimentation. The influx of information produced by genomic and post-genomic technologies is making modeling increasingly attractive to a wide range of scientists.

Biological modeling is being applied to diverse plant datasets, pathways, processes, and organisms. An important goal of such modeling is to develop predictive and easily modified models that can anticipate plant performance under different circumstances. Such models will help breeders, farmers, and agricultural specialists develop agronomic practices that maximize crop yield, help ecologists and conservationists understand population dynamics in complex environments, and allow basic researchers to better understand the signaling networks underlying plant responses to the environment. Currently most models are developed to describe a single biological network, in relative isolation from related networks, and are difficult for non-specialists to access. Combining isolated and divergent models would provide synergism, allowing much larger datasets to be used and achieving greater predictive power. The development of user-friendly applications for model access and manipulation will allow a broad range of citizens and scientists to benefit from this effort.

Our proposed Grand Challenge Workshop aims to identify the barriers to the generation of user-friendly models and to develop a plan to overcome these challenges. Using the integration of circadian, diurnal, and environmental response models as a biological framework, we will generate models that are broadly useful, easy to manipulate, and yield predictions that can be tested in the lab and the field. Important points we will address include data exchange and database integration, integration of models across biological and computational levels, and how to make advanced informatics technologies widely accessible both to scientists and the general public. We anticipate 25 – 35 participants with expertise in topics ranging from engineering to ecology.

Katherine Denby

Present Address: Warwick HRI and Warwick Systems Biology Centre,
University of Warwick, Coventry CV4 7AL, UK

Education: University of Bristol, UK. Microbiology, B.Sc. (Hons) 1st Class, 1991

University of Oxford, UK.D.Phil Plant Science. 1995
"Metabolic Regulation of Glyoxylate Cycle Genes in Higher Plants",

Boyce Thompson Institute, Cornell University.
Regulation of Tryptophan Biosynthesis. 1995-1998

Appointments:

2006 - Assoc. Prof. University of Warwick, UK.
Warwick HRI and Warwick Systems Biology Centre

2005 – 2006 Senior Lecturer, University of Cape Town, South Africa
Department of Molecular and Cell Biology

1999 – 2004 Lecturer, University of Cape Town, South Africa
Department of Molecular and Cell Biology

Relevant Publications:

Murray, S., Ingle, R., Petersen, L. and **Denby, K.J.** (2007) "Basal resistance against *Pseudomonas syringae* in Arabidopsis involves WRKY53 and a protein with homology to a nematode resistance protein" Molecular Plant Microbe Interactions20(11):1431-8

Daniel J. Kliebenstein, Heather C. Rowe and **Katherine J. Denby**(2005) "Secondary metabolites influence *Arabidopsis/Botrytis* interactions: variation in host production and pathogen sensitivity" Plant Journal 44: 25-36

Katherine J. Denby,Laure J.M. Jason, Shane L. Murray, and Robert L. Last (2005) "*ups1*, an *Arabidopsis thaliana*camalexin accumulation mutant defective in multiple defence signalling pathways" Plant Journal 41:673-684

Shane L. Murray, Nicolette Adams, Daniel J. Kliebenstein, Gary J. Loake and **Katherine J. Denby** (2005) "A constitutive *PR-1::luciferase* expression screen identifies *Arabidopsis* mutants with differential disease resistance to both biotrophic and necrotrophic pathogens" Molecular Plant Pathology 6(1): 31-41

Denby, K.J., Kumar, P. and Kliebenstein, D.J. (2004) "Identification of *Botrytis cinerea* susceptibility loci in *Arabidopsis thaliana*" Plant Journal. 38(3):473-486

Other Publications:

Katherine Denby and Chris Gehring (2005) “Engineering drought and salinity tolerance in plants: lessons from genome-wide expression profiling in Arabidopsis” Trends in Biotechnology 23(11): 547-552

Donaldson, L.E., Ludidi, N., Knight, M.R., Gehring, C. and **Denby, K.J.** (2004) “Salt and osmotic stress cause rapid increases in *Arabidopsis thaliana*cGMP levels.” FEBS Letters.Vol 569/1-3 pp 317-320

Graham, I.A., **Denby K.J.** and Leaver, C.J. (1994) “Carbon Catabolite Repression Regulates Glyoxylate Cycle Gene Expression in Cucumber” The Plant Cell, 6: 761-772

Synergistic Activities:

i. Establishment of Microarray Facility at University of Cape Town and development and running of courses for a wide range of scientists on microarray expression profiling, experimental design and data analysis (2000-2006)

ii. Coordination of Systems Biology M.Sc. course at Warwick University (2007-)

iii. Committee member (1999-2002) and Chair (2002) of the Experimental Biology Group, South Africa (scientific networking group with quarterly meetings)

Collaborators and Other Affiliations:

David Berger, University of Pretoria, South Africa
James Beynon University of Warwick, UK
Vicky Buchanan-Wollaston, University of Warwick, UK
Thomas Eulgem, UC Riverside, CA
Rob Ingle, University of Cape Town, South Africa
Nicola Illing, University of Cape Town, South Africa
Marc Knight, University of Durham, UK
Paul Tudzynski, University of Münster, Germany
Jan van Kan, Wageningen University, The Netherlands
MelaneVivier, University of Stellenbosch, South Africa
David Wild, University of Warwick, UK

Graduate/Postdoctoral Advisors: Chris Leaver, University of Oxford
Rob Last, University of Michigan

Thesis Advisor: Total number of graduate students = 8

Joseph Mulema, Makerere University, Uganda
SanushkaNaidoo, University of Pretoria, South Africa
MarykeCarstens, University of Pretoria, South Africa

CURRICULUM VITAE

Name: STACEY LYNN HARMER
Assistant Professor
Section of Plant Biology
University of California, Davis

Professional Preparation:

University of California, Berkeley	Biochemistry	B.A.(Highest Honors), 1991.
University of California, San Francisco	Biochemistry	Ph.D. 1998.
The Scripps Research Institute	Chronobiology	10/1998 – 10/2002

Professional Positions Held:

University of California, Davis	Assistant Professor	10/02 – present

Publications Most Pertinent to Proposed Research (5)

1. Covington, M.F., Maloof, J.N., Straume, M., Kay, S.A., and **Harmer, S.L**. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. (submitted)
2. Covington, M.F. and **Harmer, S.L** (2007) The circadian clock regulates auxin signaling and responses in Arabidopsis. PLoS Biology, 5(8): e227.
3. Nozue, K., Covington, M.F., Duek, P.D., Lorrain, S., Fankhauser, C., **Harmer, S.L**., and Maloof, J.N. (2007) Rhythmic growth explained by coincidence between internal and external cues. Nature, *448*:358-61.
4. Walley, J.W., Coughlan, S., Hudson, M.E., Covington, M.F., Kaspi, R., Banu, G., **Harmer, S.L.**, and Dehesh, K. (2007) Mechanical stress induces biotic and abiotic stress responses via a novel *cis*-element. PLoS Genetics, 3(10): e172.
5. **Harmer, S. L.**, Hogenesch, J. B., Straume, M., Chang, H. S., Han, B., Zhu, T., Wang, X., Kreps, J. A., and Kay, S. A. (2000). Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. Science 290:2110-3.

Other Significant Publications (5)

1. Martin-Tryon, E.L. and **Harmer, S.L**. *XAP5 CIRCADIAN TIMEKEEPER* coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. Plant Cell (in press).
2. **Harmer, S. L.**, Covington, M. F., Bläsing, O., and Stitt, M. (2005). Circadian regulation of global gene expression and metabolism. In: Endogenous Plant Rhythms, ed. Hall, A., and McWatters, H., Blackwell Publishing, p. 133-165.
3. Martin-Tryon, E.L., Kreps, J.A., **Harmer, S.L**. (2007). *GIGANTEA* acts in blue light signaling and has biochemically separable roles in circadian clock and flowering time regulation. Plant Physiology 143:473-486.
4. **Harmer, S. L.** and Kay, S. A. (2005). Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*. Plant Cell 17:1926-1940.
5. **Harmer, S. L.**, Panda, S., and Kay, S. A. (2001). Molecular bases of circadian rhythms. Annual Review of Cell and Developmental Biology 17:215-53.

<u>Synergistic Activities</u>

Science education

Participated in the Biophotonics Teacher Research Academy, a program that helps high school teachers integrate current techniques and research findings into their curricula.

Lectured at the Cold Spring Harbor Plant Sciences post-graduate course.

Have supervised 15 undergraduate and high school students carrying out independent research since arriving at UC Davis in 2002.

Scientific reviewer

Reviewed manuscripts for the journals:

Biochimica et Biophysica Acta, BMC Genomics, EMBO Journal, Genes & Development, Molecular Biology & Evolution, Nature Genetics, Plant Cell, Plant Journal, Plant Physiology, Planta, PLoS Biology, PLoS One, and PNAS.

Reviewed grant proposals for: BARD, Systems Approaches to Biological Research (SABR), Israel Science Foundation, NIH, NSF, USDA NRI, and NWO Council for Earth and Life Sciences (Netherlands Organization for Scientific Research).

<u>Academic Collaborators (past 48 months)</u>

Bläsing, O.	Max Planck Institute, Potsdam, Germany
Casal, J.	Universidad de Buenos Aires, Argentina
Dehesh, K.	University of California, Davis
Fankhauser, C.	University of Lausanne, Switzerland
Farre, E. M.	The Scripps Research Institute
Harmon, F.G.	Plant Gene Expression Center, USDA
Kay, S.A.	University of California, San Diego
Maloof, J.N.	University of California, Davis
Más, P.	IBMB-CSIC, Barcelona, Spain
Putterill, J.	University of Auckland, New Zealand
Stitt, M.	Max Planck Institute, Potsdam, Germany
Straume, M.	The University of Virginia
Yanovsky, M. J.	IFEVA, Universidad de Buenos Aires, Argentina

<u>Advisors</u>

Graduate: A.L. DeFranco	University of California, San Francisco
Postdoctoral: S.A. Kay	The Scripps Research Institute

<u>Students & Postdoctoral Fellows Trained</u>

Students trained (2)	Martin-Tryon, E.	University of California, Davis
	Ellison, C.	University of California, Davis
Postdocs trained (4)	Covington, M.F.	University of California, Davis
	Jones, Matthew	University of California, Davis
	Rawat, R.	University of California, Davis
	Schwartz, J.	University of California, Davis

CYNTHIA WEINIG

PROFESSIONAL PREPARATION

Brown University Majors: Biology B.A. 1991
Mainz University Fulbright Fellow 1991-1993
Indiana University Area: Evolutionary Biology Ph.D. 1998
Brown University Post-doctoral fellow
Area: Molecular evolutionary ecology of plant developmental plasticity 1999-2002

APPOINTMENTS

Asst. Professor, Dept. of Plant Biology, Univ. of Minnesota-St. Paul 2002
Assoc. Professor, Department of Botany, Univ. of Wyoming-2007-current

SELECTED PUBLICATIONS

(from a total of 30 in peer-reviewed journals)

1. Weinig, C., L. A. Dorn, N. C. Kane, Z. German, S. S. Halldorsdottir, M. C. Ungerer, T. F. C. Mackay, M. D. Purugganan, and J. Schmitt. 2003. Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. Genetics 165: 321-329.
2. Sequence diversity and haplotype associations with phenotypic responses to crowding: *GIGANTEA* affects fruit set in *Arabidopsis thaliana*. Molecular Ecology 16: 3050-3062.
3. Weinig, C. 2002. Phytochrome photoreceptors mediate plasticity to light quality in flowers of the Brassicaceae. American Journal of Botany 89: 230-235.
4. Weinig, C., and M. T. Brock. Association analyses of candidate genes underlying complex ecologically relevant traits: *PHYTOCHROME INTERACTING FACTOR 4* influences stem elongation in *Arabidopsis thaliana.* Functional Ecology. Accepted pending minor revision
5. Willis, C. G., M. T. Brock, and C. Weinig. The genetic architecture of competitive response and effect: each strategy is genotype-specific. Evolution. In review

PUBLICATIONS (5 others)

1. Weinig, C. 2000. Differing selection in alternative competitive environments: shade-avoidance responses and germination timing. Evolution 54: 124-136.
2. Weinig, C. 2000. Plasticity versus canalization: population differences in the timing of shade-avoidance responses. Evolution 54: 441-451.
3. Weinig, C. 2000. Limits to adaptive plasticity: temperature and photoperiod influence plant competitive responses. American Journal of Botany 87(11): 1660-1668.
4. Weinig, C. and L.F. Delph. 2001. Phenotypic plasticity early in life constrains developmental responses later. Evolution 55: 930-936.
5. Weinig, C., and J. Schmitt. 2004. Environmental effects on the expression of quantitative trait loci and implications for phenotypic evolution. BioScience 54: 627-635.

SYNERGISTIC ACTIVITIES

Advisor to 13 undergraduate students on independent research (since arriving at UMN). All of these students now have jobs or are attending graduate school in science-related fields. Two attended the National Conference on Undergraduate Research this spring, and 2 have manuscripts published or in review from their research.

Advisor in UMN SEPGM program.

SEPGM trains students for teaching careers in greater Minnesota. We have had two students from this program rotate through the lab, and have developed evolutionary labs appropriate for use in a secondary school setting.

Lecturer for undergraduate non-majors General Botany

I have developed several laboratory exercises for use in the General Botany lab that illustrate principles of plant growth, genetics, and evolution.

Affinity group mentor for Undergraduate Women in Science and Engineering

COLLABORATORS AND OTHER AFFILIATIONS

Thesis advisor: Dr. Lynda Delph (Indiana University)
Collaborators during Ph.D.: Dr. Roger Hangarter (Indiana University)
Post-doctoral collaborators: Trudy Mackay (North Carolina State University), Michael Purugganan (North Carolina State University), Johanna Schmitt (Brown University)
Other collaborators: Lisa Dorn (UW-Oshkosh), Candace Galen (University of Missouri), Julin Maloof (UC Davis), Ken Olsen (Washington University), John Stinchcombe (University of Toronto), Peter Tiffin (University of Minnesota), Mark Ungerer (Kansas State University), Stephen Welch (KSU)

BIOGRAPHY – DAVID L. WILD

A. Education: University of York, UK, B.A., Physics (1974)
Council for National Academic Awards, UK, M.Phil., Biophysics (1977)
University of Oxford, UK, D. Phil., Molecular Biophysics (1982)

Present Address: Systems Biology Centre, University of Warwick, Coventry, CV4 7AL., UK.

B. Academic/Professional Appointments

August 2006 Professor of Bioinformatics, Systems Biology Centre, University of Warwick, UK

1999 – Present Associate Professor and Director of Computing, Keck Graduate Institute of Applied Life Sciences, Claremont, CA

1999- 2006 Plenary Faculty, Claremont Graduate University, Claremont, CA

1997 – 1999	Senior Bioinformatics Analyst, Oxford Molecular and Bioinformatics Group, GlaxoWellcome Medicines Research Centre, Stevenage, UK
1994 –1997	Senior Staff Scientist, Structural Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA
1989 –1994	Staff Scientist, Biological Structures and Biocomputing Programme, European Molecular Biology Laboratory, Heidelberg, Germany
1988 – 1989	Senior Research Scientist, Allelix Biopharmaceuticals Inc., Mississauga, Canada
1987 – 1988	Research Scientist, Allelix Biopharmaceuticals Inc.
1984 – 1987	Staff Scientist, European Molecular Biology Laboratory, Grenoble, France
1982 – 1984	Research Fellow, Department of Applied Biochemistry and Food Science, University of Nottingham, UK
1980 – 1982	Part-time Computing Assistant, Department of Metallurgy and Science Materials, University of Oxford, UK
1974 – 1976	Research Assistant, Manchester Polytechnic, UK

Visiting Appointments

2001	Visitor, European Bioinformatics Institute, Hinxton, UK
2002-2006	Visitor, Gatsby Unit for Computational Neuroscience, University College, London, UK
2004	Senior Fellow, Institute of Pure and Applied Mathematics, UCLA

C. Relevant Publications

1. Beal M.J., Li, J. Ghahramani Z. and Wild, D.L. **Reconstructing Transcriptional Networks using Gene Expression Profiling and Bayesian State Space Models** in *Introduction to Systems Biology* (Ed: Choi, S.), Humana Press, Totowa (2007), pp 217-241.
2. Beal, M.J., Falciani, F., Ghahramani, Z., Rangel C. and Wild, D.L. **A Bayesian approach to reconstructing genetic regulatory networks with hidden factors.** *Bioinformatics* (2005), 21: 349--356.
3. Rangel, C.,Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., A., Gaiba, A.,.Wild, D.L. and Falciani, F. **Modeling T-cell activation using gene expression profiling and state space models.** *Bioinformatics* (2004)*,* 20(9):1361-1372.
4. Rangel, C. Angus, J., Ghahramani, Z. and Wild, D.L. **Modeling genetic regulatory networks using gene expression profiling and state space models.** In Husmeier, D., Roberts, S. and Dybowski, R. (Eds.) , ***Applications of Probabilistic Modelling in Medical Informatics and Bioinformatics.*** Springer Verlag, (2004), Springer Verlag, (2004), pp. 269-293.
5. Rangel, C., Wild, D. L. Falciani, F., Ghahramani, Z., and Gaiba, A., **Modeling biological responses using gene expression profiling and linear dynamical systems.** *Proceedings of the 2nd International Conference on Systems Biology*. Madison, WI: OmniPress, pp 248-256 (2001).

D. Significant Publications

1. Rasmussen, C.E., de la Cruz, B.J., Ghahramani, Z.., Wild, D.L. **Modeling and Visualizing Uncertainty in Gene Expression Clusters using Dirichlet Process Mixtures**. *IEEE/ACM Transactions on Computational Biology & Bioinformatics (2008)* http://doi.ieeecomputersociety.org/10.1109/TCBB.2007.70269
2. Chu W., Ghahramani, Z., Falciani, F. and Wild, D.L. **Biomarker Discovery in Microarray Gene Expression Data with Gaussian Processes.** *Bioinformatics*, 21: 3385-3393 (2005).
3. Wild, D. L., Tucker, P. A. & Choe, S., **A Visual Data Flow Environment for Macromolecular Crystallographic Computing.** *Journal of Molecular Graphics*, **13:** 291-298 (1995).
4. Timmins, P. A., Wild, D. & Witz, J., **The three-dimensional distribution of RNA and protein in the interior of tomato bushy stunt virus: a neutron low-resolution single-crystal diffraction study.** *Structure*, **2(12):** 1191-1201 (1994).
5. Weber, I. T., Johnson, L. N., Wilson, K. S., Wild, D. L., Yeates, D. G. R. & Jenkins, J. A., **Crystallographic studies on the activity of glycogen phosphorylase b.** ***Nature*** **274:** 433-437 (1978).

E. Synergistic Activities

1. Co-organizer, Symposium on Complexity and Systems Biology, Warwick Mathematics Institute, 2009-10
2. Organizing Committee, Program on Proteomics, Institute of Pure and Applied Mathematics, UCLA (2004).
3. NIH Study Section SSSH-90 (Computational Biology), 2003 and BDMA (Biodata Management and Analysis), 2004, NSF Grant Review Panel, Computational Biology, 2006
4. Reviewing Committee, 3rd International Conference on Systems Biology, 2002, Stockholm.
5. Associate Guest Editor IEEE Intelligent Systems. Special Issue, "Intelligent Systems in Biology," November/December 2001; March/April 2002.

E. Collaborators & Other Affiliations

i. *Collaborators*
Katherine Denby University of Warwick, UK
James Beynon University of Warwick, UK
Vicky Buchanan-Wollaston University of Warwick, UK
Elizabeth Wellington University of Warwick, UK
Zoubin Ghahramani University of Cambridge, UK
James Griffin University of Kent, UK
Mansoor Saqi, Rothamsted Research, UK
Francesco Falciani, University of Birmingham, UK
Philip Bourne, University of California, San Diego

ii. *Graduate Advisors*
L. Johnson, Oxford, UK
K. Wilson, York, UK

Postdoctoral Advisor
J. Blanshard, Nottingham, UK

iii. *Thesis Advisor and Postdoctoral Scholar Sponsor*
Claudia Rangel, INMEGEN, Mexico
Alpan Raval, Keck Graduate Institute
Seungwoo Hwang, Korean BioInformation Center (KOBIC)
Alexei Podtelezhnikov, Michigan Technological University

Scientific problem

Recently soaring food prices have dramatically illustrated the precarious nature of the global food supply[1]. It is estimated that crop productivity will need to double by the end of the century to keep pace with the world's increasing population[2]. Global climate change will complicate the process of crop improvement, as changing local environments will make it difficult to predict which genotypes will perform best in a particular locale. Important ways to increase yields include increasing the photosynthetic efficiency of plants and reducing crop losses due to abiotic and biotic stresses. It is estimated that the maximum conversion efficiency of solar energy to biomass is only 4.6% for C3 and 6% for C4 plants[2]. Abiotic and biotic stresses significantly reduce yields from these theoretical maxima. Within the United States, it is thought that drought causes an annual average crop loss of between 9 and 12 billion USD (2008 dollars)[3] while drought and other abiotic stresses such as salinity cause even greater productivity losses worldwide. Furthermore, biotic stresses cause an estimated 30% reduction in crop yield; these losses can reach close to 50% in underdeveloped agricultural settings[4].

Many environmental changes that affect plant productivity, such as light availability and biotic and abiotic stresses, recur with predictable daily or seasonal patterns. Plants have evolved an internal timer, the circadian clock, to anticipate these predictable challenges and opportunities. The circadian system provides plants with an adaptive advantage, presumably through its influence on processes including the timing of the transition to flowering, photosynthetic efficiency, responses to biotic and abiotic stresses, and hormone signaling[5-10]. In addition, plant responses to these stimuli can be directly modulated by cyclic changes in the environment, such as alterations in light or temperature[11-13]. Anticipation of and direct response to daily and seasonal growth opportunities and stresses are thus central to plant physiology and performance. Although some pathways such as vernalization differ between monocots and dicots, the circadian clock mechanism and many connections with biological response pathways are widely conserved across higher plants[14-16]. The intricate connections between the clock, plant physiology, and the environment strongly suggest that understanding these systems and their interactions is essential for food security.

Given the need to improve plant productivity, it is critical that we identify both genetic traits and agronomic practices that will increase performance in a wide range of environments. A critical means to achieve these improvements is the development of accurate, modifiable and testable computational models of plants and their responses to abiotic and environmental perturbations. Such models would allow crop breeders or physiologists to rapidly simulate the effects of changes to a plant's genome or environment on performance before devoting resources to testing these changes in the field. Numerous models are currently being developed in a wide range of plant sciences, including physiology, breeding, agronomy, and fundamental science[17-24]. However, for the potential power and utility of modeling approaches to be realized, information from disparate datasets, processes, disciplines and organisms must be combined and made accessible to scientists with a range of expertise. Thus there is a need to develop a simple computational platform for the integration of such models and their application across diverse plant sciences. The workshop we propose aims to explore ways in which an iPlant Grand Challenge project could meet this need.

The likely participants in this workshop possess a broad range of expertise. We have recruited: computational scientists with expertise in the generation of mathematical and network models (Albert, Doyle, D'Souza, Lee, Sauro, Wild, Zhu) and development of standards for systems biology and software integration (Oinn, Sauro); fundamental biologists studying the mechanisms underlying circadian clock function (Harmer, McClung, Mockler), biotic and abiotic stress responses (Denby, Glazebrook, Katagiri, Kliebenstein, Maloof, Ronald), central metabolism (Stitt), and flowering time regulation (Coupland, Maloof, Weigel); biologists working towards crop improvement (Brutnell, Buckler, Jansson, Mockler, Pink, Ronald); and ecologists studying plant/environment interactions (Juenger, Weinig). Together, we will be able to decide upon common language and data exchange standards, identify computational and modeling tools that will be useful to a broad range of biologists and breeders for basic and applied research, and plan the optimization, robust software engineering, and dissemination of these tools.

Need for new tools and approaches

A wide range of detailed studies of plant/environment interactions, many of which also explicitly examine the effects of genotype and the time of day or season on responses, have been completed and more are underway. Parameters recorded in these studies include meteorological data, high resolution images, gene, protein and metabolite profiles, flowering time, growth, and yield. Although potentially complementary, it is currently difficult for an individual researcher to find and coordinate data generated by groups working at different biological levels or using different species. When this data is located, it can be hard for those working in a slightly different field to understand what is being reported. Therefore one pressing need is for a central clearinghouse for these diverse datasets, with data deposited using the format (such as MIAME for microarray data) considered the standard by each community. We must also develop seamless methods for data exchange between these formats. All types of data will need to be incorporated in future modeling efforts, and a standardized way to deposit and access diverse datasets will greatly facilitate this process. Life scientists and software engineers will need to determine whether this will best be accomplished by a single centralized data repository or by a workflow environment that links information from different databases[25].

Another challenge is that computational scientists and biologists are generating predictive and descriptive models of various plant response networks using diverse tools and approaches, which make these potentially complementary models difficult (if not impossible) to compare and integrate with each other. At a computational level, it is very difficult to integrate existing models with each other, especially when they are based upon different techniques such as ordinary differential equations (ODE), linear dynamical systems (LDS), or Boolean models. At a biological level, it is very challenging to integrate discrete models that describe separate but related processes, such as circadian clock function, light signaling, and pathogen response. Both types of model integration must be achieved, however, if we are to generate useful models that will predict plant performance in a variety of environments. Even after models have been generated, it is then necessary to translate them into standard repository formats such as SBML or CELLML. While these formats are appropriate for ODE models, similar formats have not been developed for other types of models such as LDS or Boolean. Thus an important challenge is to devise new ways to integrate and store predictive models.

Most current models have been generated using software that is so specialized that only a limited number of scientists, conversant with software such as MATLAB, are able to effectively use them. To leverage full value from the community investment into data generation and model development, we need to translate prototype specialist academic software into robust engineered production software that can be used by many types of plant scientists, from breeder to molecular biologists. The goal would be to generate flexible interfaces that would allow the non-specialist user to modify the model depending upon their needs and then generate testable hypotheses. Scientists working on crop improvement might wish to determine the predicted effects on performance when a particular genotype was grown in a different environment, or after specific genetic changes were introduced. Scientists studying fundamental processes might wish to determine the predicted effects on phenotype after adding or removing model components or changing connections between them. Such applications by non-specialists are currently quite difficult; for example, although quantitative models of the plant circadian clock have been published[17,18], these models cannot be easily manipulated even using relatively standard programs such as Cell Designer and MATLAB Systems Biology Toolbox. Accessible models with "Lego set"-type flexibility would allow a wide range of biologists to apply different types of data to the improvement and application of models, resulting in a deeper understanding of fundamental plant biology and more rational strategies to improve productivity.

A final challenge will be to integrate such user-friendly models and software into research, education, and public outreach programs. It is vital to educate biologists, breeders, farmers, and the general public about the value of computational approaches to complex biological problems and to engage them in the scientific process. One outreach project we are planning is to provide 100 maize accessions to science classes and volunteer "citizen science" groups located across a wide range of environments. Similar citizen science projects have been successful in the fields of ornithology, botany, and astronomy[26]. We would then provide these students and volunteers with a user-friendly model that will predict traits (such as flowering time and yield) based upon the local environment. Students and volunteers would then measure these traits, compare them to the predictions, and upload their data to the project organizers. This distributed phenotyping would be useful for model improvement and introduce students and members of the general public to modeling and quantitative approaches to biology.

Available datasets and models

Diverse large-scale datasets and models investigating the role of predictable changes in the environment on plant physiology have been generated. Computational models describing the circadian clock (Doyle, Millar), biotic and abiotic stress pathways (Albert, Katagiri, Ronald), photosynthesis (Zhu), and carbon and nitrogen metabolism (Stitt) have already been published. Many microarray experiments examining the effects of the time of day and the season on gene expression have been or will soon be completed in *Arabidopsis* (Chory, Harmer, Maloof, Millar, Stitt), rice (Mockler), poplar (Jansson, Mockler), and *Brachypodium* (Mockler). High resolution time series gene expression data, and network models inferred from this data, are available for leaf development and infection of *Arabidopsis* with the pathogen *Botrytis cinerea* (monitored over two diurnal cycles). Additional gene expression experiments and models for *Pseudomonas syringae* infection and abiotic stresses (drought, high light) in *Arabidopsis* are currently being generated and should be completed within the next 12 – 24 months (Denby, Pink, Wild). In addition, transcript, enzyme activity, metabolite and polysome loading levels have been monitored in one *Arabidopsis* accession over diurnal cycles and in over 100 accessions using a reduced number of time points (Stitt). A comparative study of photosynthetic gene expression in maize, rice and sorghum is currently underway to characterize transcriptomes, proteomes and metabolites associated with C3 and C4 photosynthetic differentiation (Brutnell), providing a platform for future studies to examine the effects of abiotic stresses such as shade and nutrient deficiencies on photosynthetic differentiation. Gene expression studies of responses to biotic and abiotic stresses in rice are also available (Ronald).

The effects of genotype on plant physiology in different environments have been examined by many workers. Large data sets are available in *Arabidopsis* relating growth, enzyme activities and metabolites when carbon and nitrogen levels are limiting (Stitt). Extensive studies in *Arabidopsis* (Coupland, Kliebenstein, Maloof, Weigel, Weinig) and maize (Buckler) have examined effects of genotype and environment on flowering time or biotic interactions using different natural accessions, nested association mapping (NAM), recombinant inbred line (RIL), and F2 mapping populations. In some cases, circadian parameters were also measured for these populations (Coupland, Weigel, Weinig). Growth, performance, and circadian parameters (period, phase, amplitude) have been examined in *Arabidopsis* field-grown RILs (Weinig) and metabolite levels, growth, and yield have been examined in 5000 maize RILs grown in up to ten different environments (Buckler). Detailed genotypic information is already available for many of the *Arabidopsis* lines used in the above studies[27]. Another 1,300 *Arabidopsis* accessions are being genotyped at 250,000 single nucleotide polymorphisms (http://walnut.usc.edu/2010/SNPs) and complete genome sequences are being generated for 1,001 accessions (http://1001genomes.org/). The maize NAMs have been genotyped and initial maps will be publicly available within the next 12 months and Solexa sequencing data will be released within the next 12 to 24 months.

The above studies thus range from those performed with a single genotype grown in a controlled environment facility to those carried out using diverse populations in a variety of natural environments, and measure parameters ranging from gene expression to plant fitness and performance.

Goals for the workshop

The primary goal of the workshop will be to develop an action plan to generate the new tools and approaches needed to make computational models more broadly useful. It will first be necessary for workers in diverse fields to develop a common language and educate each other about appropriate reporting standards for models and data exchange. This will facilitate data exchange and model generation and integration. To this end, participants will first describe the state of the art in different disciplines, discuss the types of standards required for both data exchange and models, and solicit ideas on how to integrate molecular, genetic, metabolic, phenotypic, and environmental datasets. We will also discuss what will be required to integrate models at various levels: this includes models generated using diverse techniques (such as ODE, Boolean, and LDS); models of related biological pathways (such as the circadian clock and stress responses); and models describing different types of events (such as molecular, metabolic, physiological, and meteorological processes). Existing datasets that should be incorporated into new and existing models will be identified.

Another important topic will be a discussion of ways to make advanced informatics technologies usable by a wide range of biologists. We will discuss current computational models and tools that would merit development from specialized academic prototypes into highly accessible and robust software. In addition to providing useful tools for basic scientists and breeders, this will allow us to introduce school-age children and citizen scientists to quantitative approaches to biology via outreach efforts. Ways in which these tools can be disseminated among target audiences will also be discussed. Finally, we will plan how to develop a full Grand Challenge proposal and identify core leadership for this effort.

Related efforts

Several Grand Challenge Workshop proposals are related to this proposal. Katherine Denby, David Pink, David Wild, and Xinguang Zhu are likely participants in the "climate change" iPlant proposal organized by Steve Howell while Pam Ronald is a likely participant in the "climate change" iPlant proposal organized by Ruth Grene. Ed Buckler is an organizer and Stacey Harmer and Julin Maloof are likely participants in the "plant adaptation" iPlant proposal. Data resulting from an NSF-funded FIBR initiative investigating the role of genotype and the environment in the regulation of flowering time (Award Abstract #0425759; PIs: Schmitt, Welch, Amasino, Purugganan) would prove useful in our modeling efforts.

Meeting format

We anticipate that the proposed workshop would take place over two days. It would consist of a few formal presentations in a specific area followed by group discussion of that topic. Each speaker would spend most of his or her time on background information followed by a brief description of what results he or she would like to see come from the Grand Challenge proposal. We envision the following topics, with the indicated participants playing key roles in each session:

1. Database integration/exchange (Oinn, Sauro)
2. Model integration I: different mathematical approaches (Albert, Doyle, D'Souza, Lee, Wild)
3. Model integration II: related biological pathways (Brutnell, Denby, Katagiri, Ronald, Stitt)
4. Model integration III: across biological levels (Buckler, Pink, Weinig)
5. Model translation: academic prototypes → robust engineered production software (Oinn)
6. Education and outreach tools (Katagiri, Oinn)
7. Summary session: identification of key goals and formation of working groups (All)

Role for iPlant Collaborative

We would appreciate the assistance of the iPlant Collaborative in identifying contributors with expertise in software engineering, plant breeding for the developed and developing world, undergraduate biology education, and citizen science outreach.

References

1. Fedoroff, N. Seeds of a perfect storm. *Science* **320**, 425 (2008).
2. Zhu, X.G., Long, S.P. & Ort, D.R. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr Opin Biotechnol* **19**, 153-9 (2008).
3. Hayes, M.J., Svoboda, M.D., Knutson, C.L. & Wilhite, D.A. Esimating the economic impacts of drought. in *Fourteenth Conference on Applied Climatology* (Seattle, WA, 2004).
4. Oerke, E.-C. & Dehne, H.-W. Global crop production and the efficacy of crop protection – current situation and future trends. *European Journal of Plant Pathology* **103**, 203 – 215 (1997).
5. Covington, M.F. & Harmer, S.L. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS Biol* **5**, e222 (2007).
6. Sauerbrunn, N. & Schlaich, N.L. PCC1: a merging point for pathogen defence and circadian signalling in Arabidopsis. *Planta* **218**, 552-61 (2004).
7. Dodd, A.N. et al. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**, 630-3 (2005).
8. Garner, W.W. & Allard, H.A. Photoperiodism, the response of the plant to relative length of day and night. *Science* **55**, 582-583 (1922).
9. Fowler, S.G., Cook, D. & Thomashow, M.F. Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiol* **137**, 961-8 (2005).
10. Walley, J.W. et al. Mechanical stress induces biotic and abiotic stress responses via a novel *cis*-element. *PLoS Genet* **3**, 1800-12 (2007).
11. Griebel, T. & Zeier, J. Light regulation and daytime dependency of inducible plant defenses in Arabidopsis: phytochrome signaling controls systemic acquired resistance rather than local defense. *Plant Physiol* **147**, 790-801 (2008).
12. Nozue, K. et al. Rhythmic growth explained by coincidence between internal and external cues. *Nature* **448**, 358-61 (2007).
13. Yan, L. et al. The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. *Science* **303**, 1640-4 (2004).
14. Serikawa, M., Miwa, K., Kondo, T. & Oyama, T. Functional conservation of clock-related genes in flowering plants: overexpression and RNA interference analyses of the circadian rhythm in the monocotyledon Lemna gibba. *Plant Physiol* **146**, 1952-63 (2008).
15. Ramos, A. et al. Winter disruption of the circadian clock in chestnut. *Proc Natl Acad Sci U S A* **102**, 7037-42 (2005).
16. Pino, M.T. et al. Ectopic AtCBF1 over-expression enhances freezing tolerance and induces cold acclimation-associated physiological modifications in potato. *Plant Cell Environ* **31**, 393-406 (2008).
17. Zeilinger, M.N., Farre, E.M., Taylor, S.R., Kay, S.A. & Doyle, F.J., 3rd. A novel computational model of the circadian clock in Arabidopsis that incorporates PRR7 and PRR9. *Mol Syst Biol* **2**, 58 (2006).
18. Locke, J.C. et al. Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. *Mol Syst Biol* **2**, 59 (2006).
19. Zhu, X.G., de Sturler, E. & Long, S.P. Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiol* **145**, 513-26 (2007).
20. Gutierrez, R.A. et al. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proc Natl Acad Sci U S A* **105**, 4939-44 (2008).
21. Hammer, G. et al. Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci* **11**, 587-93 (2006).

22. Espinosa-Soto, C., Padilla-Longoria, P. & Alvarez-Buylla, E.R. A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* **16**, 2923-39 (2004).
23. Smith, R.S. et al. A plausible model of phyllotaxis. *Proc Natl Acad Sci U S A* **103**, 1301-6 (2006).
24. Welch, S., Roe, J.L. & Dong, Z. A genetic neural network model of flowering time control in *Arabidopsis thaliana*. *Agron. J.* **95**, 71–81 (2003).
25. Oinn, T. et al. Taverna: lessons in creating a workflow environment for the life sciences. *Concurrency Computat.: Pract. Exper.* **18**, 1067–1100 (2006).
26. Clark, F. & Illman, D.I. Dimensions of civic science: introductory essay. *Science Communication* **23**, 5-27 (2001).
27. Nordborg, M. et al. The pattern of polymorphism in Arabidopsis thaliana. *PLoS Biol* **3**, e196 (2005).

Suggested reviewers

Phil Benfey, Duke University. Email: philip.benfey@duke.edu
Systems biology, transcriptional networks, and root development.
Steve Briggs, University of California, San Diego. Email: sbriggs@ucsd.edu
Plant pathogen interactions, proteomics, and systems biology.
Dirk Inze, University of Ghent. Email: diinz@psb.ugent.be
Systems biology, plant cell cycle, and yield.
Andrew Millar, University of Edinburgh. Email: Andrew.Millar@ed.ac.uk
Systems biology and plant circadian rhythms.
Thomas Mitchell-Olds, Duke University. Email: tmo1@duke.edu
Evolutionary, ecological, and agricultural functional genomics.
Julian Schroeder, University of California, San Diego. Email: julian@biomail.ucsd.edu
Abiotic stress signaling pathways.