The Computable Plant: Annual Report, 6/2004-5/2005

Eric Mjolsness, Principal Investigator Institute for Genomics and Bioinformatics University of California, Irvine <u>emj@uci.edu</u>

A substantial escalation in activity in the FIBR Computable Plant project has taken place during this year, with the full operation of the microscope, the arrival of new personnel, successful computer analysis of the imagery, and the development of a first round of models for meristem maintenance and phyllotaxis. A first round of results have been disseminated. In this report we will describe the logistical setting of the research (major activities, personnel, and space); then highlight selected sub-projects including the WUSCHEL model accepted for presentation at the Intelligent Systems in Molecular Biology conference; then list more comprehensively the achievements of the funded work for the year; then list the publications and presentations that resulted, and conclude.

Setting of the Research

Major Activities, 2004-2004

Research was managed through twice weekly meetings, augmented with special occasions, as follows.

Weekly meetings - Caltech

Usual attendees: Agrawal, Gor, Heisler, Jönsson (biweekly videoconference), Meyerowitz, Mjolsness, Reddy, Shapiro Occasional guests: Bacarian, Bhan, Sadovsky, Wong

Weekly meetings – UCI

Usual attendees: Bacarian, Baldi, Bhan, Fang, Liang, Mjolsness, Sadovsky, Wong

Computable Plant teachers' workshop: "The ABCs of Developmental Botany: Integrating Plants into the Classroom", August 23-27 2004.

Computable Plant FIBR Project retreat, Kerckhoff Marine Laboratory, Corona del Mar, CA January 22, 2005

Visit of ICG collaborators January 2005 (two weeks).

NSF Site Visit, Chris Greer, PhD. Caltech, Huntington, and UCI sites. February 28-March 1, 2005.

Numerous papers and presentations (see Dissemination, below).

Personnel

A strong group of people are now involved either as project employees, part time or full time, or as collaborating scientists whose research is synergistic with the project. In addition to the PI's (Mjolsness, Meyerowitz) and co-I's (Baldi, Folsom) the cast includes:

Postdocs: Marcus Heisler, Venu Reddy (Caltech Biology), Tigran Bacarian, Ashish Bhan (UCI Bioinformatics).

Staff: Martha Kirouac (Huntington); Alexey Vorobyov (1/4 time programmer, UCI); Victoria Gor, Bruce Shapiro (Caltech Jet Propulsion Laboratory).

Graduate Students: Elaine Wong (UCI Math), Fang Fang, Guy Yosiphon (UCI Computer Science).

Collaborators: Henrik Jönsson, research scientist, Lund University, Sweden (former Caltech and UCI postdoc); Alex Sadovsky, postdoc, UCI; Gang Liang, UCI statistics faculty; collaborators from the Institute for Cytology and Genetics, Novosibirsk, Russia (Nikolay Kolchanov, Nadezda Omelianchuk, Vitali Likoshvai, Nikolay Podkolodny, Sergey Nikolaev).

Space. New space at UCI was occupied in May 2005 at the UCI branch of the California Institute for Telecommunications and Information Technology <u>http://www.calit2.net/</u>.

Selected Highlights of this Year's Work

Data acquisition and image analysis

Quantification of growth. (Heisler, Bacarian, Gor)

After several rounds of experiments, conditions for imaging apical meristems have been worked out such that the density of time points is sufficient for nuclei tracking with minimal affect on plant growth rates. Furthermore procedures have been established that allow us to correct for growth deformations during the collection of confocal data. These advances have resulted in the generation of a first set of data that is accurate enough to enable us to obtain some insight into meristem growth dynamics. Improvements in progress include the generation of nuclear markers using the dsRED protein which when combined with gene expression markers based on GFP should allow for the simultaneous measurement of growth and gene expression dynamics.

Image analysis (Bacarian, Heisler, Mjolsness)

One of the goals of the year was to develop necessary tools to match and track similar objects through time series of 2D as well as 3D stacks of real images to obtain descriptive numerical and visualizing data. The application of the matching code was one part of more general efforts in tracking plant cells from 3D fluorescent image stacks of growing

Arabidopsis apical shoot meristem (SAM). The general target goal of these efforts was to receive correct descriptive scalar and/or vector parameter fields for digital illustration of the growth patterns of the SAM and possible numeric comparison of those patterns with the gene expression maps.

Among the challenging problems to be resolved were:

- 1. Originally distorted state of the data images due to non-negligible growth speed of the plant during the scanning procedure;
- 2. Substantial range of error while extracting positions of the cell nuclei centers due to their natural jitter, as well as shape, brightness and contrast variation of relatively large image of a separate nuclei. The latter lead to extraction mistakes, comparable or larger than the tracked growth displacements between consecutive time-points.
- 3. Existing lower limit for the time span between consecutive scans of the plant due to the plant reaction on disturbance.
- 4. Errors of the extraction program itself (the code based on contrast gradient driven algorithm by H. Jöhnsonwas used for this purpose).

These were resolved in one line of work as follows: A correctional procedure for the growth-caused distortion of the original data images was developed, based on additional reference scan of the plant at different speed. A Matlab preprocessing toolbox was developed for extrapolating, registering, and correctively transforming the original image data to achieve undistorted 3D time-series of the growing SAM images. A separate processing package consisting of an infrastructure of software tools, utilizing the extraction code, and the above described matching codes has been developed. A problem of correction for the nuclei center-positional errors was approached with the development of a new algorithm involving matched data from all time points simultaneously in one treatment, iteratively utilizing Thin Plate Spline (TPS) interpolation, polynomial smoothing, concurrent, statistically based, rejection of erroneous trajectories, and fine tuning for image registration mistakes. A separate post-processing Matlab toolbox, for obtaining resulting differential growth fields (velocity field, volumetric growth rate, linear growth components) was written, based on the described above procedure. The data were optionally reformatted for subsequent use in Amira software based enhanced visualization. The resulting movies showed for the first time with this level of accuracy the pattern of development of a plant SAM primordium. Particularly, it showed increased volumetric growth rate in the regions of the future primordium formation, followed by change in primary linear growth direction from tangential to the surface to normal to the surface of the SAM. The results have been reported at the 7th Plant Science Institute Symposium, Meristems 2005 (Iowa, June 2-5, speaker M. Heisler). Further reports both on the image processing and biological analysis of the SAM growth will follow.

Auxin modeling (Jönsson, Heisler, et al.)

Our auxin concentration model is now close to publication and we are starting to develop ideas for new improved models. The experimental investigation into the regulation of PIN1 by auxin is ongoing. An important new insight has been provided by examining the PINOID (PID) protein fused to YFP in combination with PIN1GFP. PID encodes a protein kinase with no obvious transmembrane domain and yet surprisingly appears membrane localized in a polar manner similar to PIN1. PID activity is necessary for primordium development and it promotes the apical localization of PIN1 in the epidermal cells of the meristem. By co-visualizing PIN1GFP and PIDYFP we have been able to determine that PID is expressed strongly in cells in which PIN1GFP undergoes polarity changes towards the apex. Hence PID is a likely candidate for mediating this reversal. With these data we can now include PID in our phyllotactic modelling.

Modeling WUSCHEL (excerpted and edited from Jönsson et al, Bioinformatics, 2005)

We address the question of how the WUS expression pattern is activated and confined to a small region of cells located centrally in the SAM. We do this by combining *in vivo* confocal microscopy data with computer simulations of different models for WUS activation. An important step in the methodology is the attempt to quantitatively measure WUS expression. Using a set of image processing tools, we compartmentalize and quantify the confocal microscopy data, resulting in an approximate cell-based quantitative template for WUS expression (Figure 1).

We use spatial compartmentalization at a cellular level and do not account for cellular growth or proliferation. Molecular levels are described by concentrations, and deterministic ordinary differential equations (ODEs) are used for describing the dynamics. The models presented here are built from three main components:

(1) *Molecular reactions*. Basic mass action kinetics is used, where reaction rates are proportional to the concentration of reactants.

(2) *Gene regulation*. We use sigmoidal functions to describe the interactions in a gene regulatory network (GRN) (Mjolsness *et al.*, 1991; Mjolsness, 2001).

(3) *Molecular transport*. Transport between neighboring cells is modeled by a passive diffusion.

Activator model The main assumption in this model is that WUS expression is induced by an activator produced by a pattern-forming, reaction-diffusion model (Meinhardt, 1982). Inherent in the activator model dynamics is the ability to create regular patterns of activator concentrations at distinct spatial distances, even from close-to-homogeneous initial concentrations. The distances between concentration peaks can be tuned by model parameters to allow for only a single peak within the SAM. To ensure that the activator peak is positioned at the center of the SAM, a small repression of the activator from an L1 originating signal is included.

Repressor model We have also implemented a model with a simpler hypothesis for the activation of WUS expression. The basic idea of this model is that WUS is normally expressed everywhere, unless a repressor is present. In the model, an L1 originating signal represses WUS expression.



Figure 1. Quantification of the WUS expression by computer image analysis. (A) A doubly labeled confocal image showing a horizontal section $\sim 17 \mu m$ into the plant shoot. A cell membrane dye marker is shown in red, and fluorescence resulting from expression of a WUS::GFP construct is shown in green. (B) Pixels extracted as background. As can be seen, older primordia are not included in the SAM template. (C)Walls (defined by the borders between cell compartments) extracted by the watershed algorithm. (D) Average GFP::WUS intensity for individual cells. These numbers are interpreted as relative concentrations. The color coding is defined as black (min)-blue-red-yellow (max).



Figure 2. WUSCHEL equilibrium concentration for both models simulated on the extracted template. (A) Activator model. (B) Repressor model. Color coding as in previous figure.

Wild-type expression domain. We first set out to determine how well the models recreate the wild-type WUSCHEL expression domain. We chose to perform these simulations on the extracted template in order to directly compare the result with the quantified experimental data (Fig. 3). The equilibrium WUS concentration is presented in Figure 6. As can be seen in the figure, the equilibrium expression for both models defines a small, fairly distinct region of WUS expression in the central part of the meristem, similar to the experimental data. The spatial position of the WUS expression domains in the two models are slightly different from the experimental position (cf. Figs 1 and 2). The positioning in the models depends on the L1 originating signal Y, which in turn depends on the topology of the L1 layer. ... The shift in position might be a consequence of implementing the models in two dimensions, since the actual peak position is determined by the complete three-dimensional contribution of the repressing signal.



Figure 3. WUSCHEL equilibrium concentrations in an unperturbed lattice simulation. (A) Activator model. (B) Repressor model. Color coding as in previous figure.

Laser ablation experiment An important feature of the SAM, including the WUS expression domain, is the capability of reorganization, which we address by modeling the laser ablation experiment of Reinhardt *et al.* (2003). To do this, the central cells are removed from the SAM, and the models are simulated on the new lattice (Fig. 10). In the laser ablation simulation, the activator model creates two new, spatially distinct WUS domains at opposite sides of the ablated region. This is in full agreement with the experiment. Furthermore, this simulation recreates the dynamics of the experiment, since it first induces a circular domain of weakly expressing WUS cells surrounding the ablated cells. The repressor model, on the other hand, does not induce any spatially distinct WUS domain. The equilibrium WUS expression is in a ring-shaped domain surrounding the ablated cells with fairly low expression. Hence, the activator model exhibits the very important ability to recreate a spatially distinct WUS expression domain from a physically perturbed system. Since this is a feature of the actual biological mechanism, it is a required property for any model trying to address the mechanism underlying WUS expression.



Figure 4. WUS concentrations in the simulation of the laser ablation experiment. (A) Early time point for the activator model. (B) Equilibrium concentration for the activator model. (C) Equilibrium concentration for the repressor model. Color coding as in Figure 3, where the maximal value is set to the maximal value in the unperturbed situation.

Modeling CLV3/1 (Shapiro, Heisler, Agrawal)

We have extended and enhanced our previously reported developmental simulations of the shoot apical meristem (SAM) (Jönsson et al 2003, 2005; Mjolsness et al 2004). Our working hypothesis is that SAM development can be described by the differential expression of key regulatory proteins such as CLV1 (a receptor kinase), CLV3 (thought to be the CLV1 ligand), WUS (a transcription factor negatively regulated by CLV1) and a layer-1 specific protein (L1SP). For example, activation of CLV1 might be described by the reactions



$$CLV1 + CLV3 \rightleftharpoons CLV1 *$$
$$CLV3 \rightarrow \emptyset$$

The dependence of CLV1 and CLV3 on WUS, perhaps through a hypothetical diffusible intermediary (CLV3I1), has been inferred from experiments. A second diffusive signal is postulated to originate from L1SP and diffuses into the rest of the meristem via messenger CLV3I2. CLV3 is turned on only if the sum CLV3I1+CLV3I2 exceeds threshold.

$$WUSI \mapsto WUS$$
$$WUS \mapsto CLV 3I1$$
$$L1SP \mapsto CLV 3I2$$
$$CLV 3I1 + CLV 3I2 \mapsto CLV 3$$

where WUSI is a hypothetical WUS inducer that originates either in or below the corpus. The expression $A \mapsto B$ is the Cellerator notation for genetic regulation (see Table 1). Inhibitory feedback is provided by a proposed entity Z that sequesters activated CLV1, and when activated, sequesters or removes WUSI:

$$Z + CLV1^* \rightleftharpoons Z1$$
$$Z1 + WUSI \rightleftharpoons Z2$$

Additionally, an unknown diffusible messenger Y creates a surface specific expression pattern for L1SP, which is itself inhibited by STEM, a hypothetical gene expressed only in the lowest meristem layer:

$$\begin{array}{c} Y \mapsto L1SP \\ STEM \mapsto Y \end{array}$$

Here the first reaction is activating, and the second is inhibitory; both genetic regulation and inhibition are modeled by Hill functions with different parameters. To maintain homeostasis we include the reactions

$$CLV3, CLV3I1, WUS \rightarrow \emptyset$$

and describe diffusion using a simple compartmental approach. A two-dimensional 133cell Cellerator implementation has 5422 reactions and 1596 differential equations.

Polarity Model (Shapiro, Heisler)

The polarity model is based on HD-ZIP genes, miRNAs, KANADI and other players, and is derived from the model in [Engstrom, Izhaki & Bowman, Plant Physiology 135:685-694 (2004)] (see figure) using Auxin as an inducing factor. The preliminary model (see figure) has not been calibrated. We are currently investigating the ability of Auxin (or any other factor) to induce a channel of increased protein concentration. (see figure). The



preliminary 130 cell model has 1740 reactions and 780 differential equations.

Figure 6. Polarity Model from Engstrom et al (Left) and Cellerator single-cell implementation (Right).



Results of a preliminary simulation (see text).

Modeling Software (Shapiro)

We are using Cellerator to generate models of simulated two dimensional meristems. The cellerator plugin Cellaray generates pseudo-meristems with Cell wall locations based on Voronoi Cells and cell-cell interactions derived from Delaunay triangulation. Horizontal and vertical cross sections as well as three-dimensional models can be developed. Because of the difficulty in parameter fitting with large models we are currently concentrating on the two-dimensional simulations. A typical 50-cell vertical cross section is illustrated below:



Because of the difficulty of simulating large models, we have implemented algorithms to improve the processing speed of Cellerator. Simulation in Cellerator is composed of two steps: (1) conversion of reactions to differential equations; and (2) numerical solution of the differential equations. The vast bulk of CPU time is utilized in the first, or *interpret* step. The second, or *run* step, is highly efficient because it utilizes NDSolve which is based on the LSODx family of numerical solvers. We have developed an improved interpret algorithm that improves processing speed by as much as two orders of magnitude. This algorithm has been implemented in a new extensible version of Cellerator. Because the fundamental algorithm is completely different, the new program has a different name, xlr8r. We have implemented all standard Cellerator mass-action based and Michaelis-Menten reactions (and cascades thereof) in xlr8r. (Not all of the cascades have been implemented in Cellerator). The new algorithm is much easier to extend and we will add additional reactions during the upcoming year. The timing improvement is illustrated in the following figure (Blue: xlr8r, Red: cellerator).



Continuum Mechanics models of meristems (Sadovsky, Baldi, Heisler)

Formulated and justified continuum-mechanical models for Arabidopsis thaliana meristems without primordia for the following static cases: A. Isotropic medium; B. Isotropic medium simplified for computational purposes: velocity field outside meristem stele assumed irrotational, stele regarded as an obstacle; C. Anisotropic medium, with longitudinal preferred cell relaxation; D. Anisotropic medium, with transverse preferred cell relaxation. Solved the above models numerically. Formulated a dynamic, anisotropic model of a meristem with primordia, and with a growth source term. (Numerical solutions in preparation.) Formulated and solved a 1-dimensional stochastic model of the stele of a root meristem. Obtained a qualitative agreement with experimental data. Made progress in formulating a viscoplastic model of an individual cell.

Outreach Modeling (Wong)

Developed a Cellerator/Mathematica notebook that uses reaction notation (e.g. $A+B\rightarrow C$ with rate k) to represent mathematical models of biological growth, starting from exponential growth, in an "easy to interact" format for students. In other words, the more advanced math (differential equations) is hidden from the user, but the graphs are displayed, and engaged students could do new experiments. This notebook and activities (in the form of worksheets) will be made accessible through the web for teachers and students. In development: An updated version of the lesson plan that will include the use of a Cellerator notebook to explore mathematical models of biological growth.

Predator-Prey-Renewable Resource Chain

species8 = {{T + S + 2 T, 1}, {S + R + 2 S, 1}, { $\phi \rightarrow R$, 0.2}, {S + ϕ , 0.2}, {T + ϕ , 0.1}} {{S + T + 2 T, 1}, {R + S + 2 S, 1}, { $\phi \rightarrow R$, 0.2}, {S + ϕ , 0.2}, {T + ϕ , 0.1}} s8 = run[interpret[species8], initialConditions + {R[0] == 1, S[0] == 1, T[0] == 1}, timeSpan + {0, 25}, plotVariables + All]



 $\{ \{0, 25, \{ \{R \rightarrow InterpolatingFunction[\{ \{0., 25.\} \}, <> \}, \\ S \rightarrow InterpolatingFunction[\{ \{0., 25.\} \}, <>], \\ T \rightarrow InterpolatingFunction[\{ \{0., 25.\} \}, <>] \} \} \}$

 $\begin{array}{l} run[interpret[species8], \ initialConditions \rightarrow \{R[0] == 1, \ S[0] == 1, \ T[0] == 1\}, \\ timeSpan \rightarrow \{0, \ 25\}, \ samePlotVariables \rightarrow \{R, \ S, \ T\}] \end{array}$



Dissemination of the work

Web site maintained and updated: <u>www.computableplant.org</u> . (Shapiro, et al.)

Journal publications

"Modeling the Organization of the WUSCHEL Expression Domain in the Shoot Apical Meristem", Henrik Jönsson, Marcus Heisler, G. Venugopala Reddy, Vikas Agrawal, Victoria Gor, Bruce E. Shapiro, Eric Mjolsness, Elliot M. Meyerowitz. Accepted for Bioinformatics; also Intelligent Systems in Molecular Biology 2005.

Reddy, G.V., Heisler, M.G., Ehrhardt, D.W. and Meyerowitz, E.M. (2004) "Real-time analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana." Development 131:4225-4237.

Book chapters

"A Software Architecture for Developmental Modeling in Plants: The Computable Plant Project", Victoria Gor, Bruce E. Shapiro, Henrik Jönsson, Marcus Heisler, G. Venugopola Reddy, Elliot M. Meyerowitz and Eric Mjolsness. In Bioinformatics of Genome Regulation and Structure, Second Biennial Volume", Nikolay Kolchanov, editor. Kluwer, in press, 2005.

"AGNS—A Database on expression of *Arabidopsis* Genes". Nadezhda Omelyanchuk, Viktoria Mironova, Aleksander Poplavsky, Nikolay Podkoldny, Nikolay Kolchanov, Eric Mjolsness, Elliot Meyerowitz. In Bioinformatics of Genome Regulation and Structure, Second Biennial Volume", Nikolay Kolchanov, editor. Kluwer, in press, 2005.

Conference papers

"Tracking Cell Signals in Fluorescent Images", Victoria Gor, Tigran Bacarian, Michael Elowitz, Eric Mjolsness. Accepted for the Computer Vision Methods for Bioinformatics (CVMB) workshop at Computer Vision and Pattern Recognition (CVPR). June 2005.

"Modeling the Organization of the WUSCHEL Expression Domain in the Shoot Apical Meristem", Henrik Jönsson, Marcus Heisler, G. Venugopala Reddy, Vikas Agrawal, Victoria Gor, Bruce E. Shapiro, Eric Mjolsness, Elliot M. Meyerowitz. Accepted for Intelligent Systems in Molecular Biology 2005.

"A Multicellular Model of a Feedback Network Regulating Spatial Gene Expression Domains in the Shoot Apical Meristem", Henrik Jönsson, Bruce E. Shapiro, Victoria Gor, Marcus Heisler, G. Venugopala Reddy, Elliot M. Meyerowitz, and Eric Mjolsness, ICSB 2004.

Talks

Talks by Elliot Meyerowitz, featuring Computable Plant work:

Royal Society, London, July 15, 2004.

Temasek Life Sciences Laboratory, Singapore, July 19, 2004.

NSF FIBR Outreach Teacher's Workshop, Huntington Gardens, San Marino, CA.

September 9, 2004, Salk Institute, La Jolla, CA.

November 10, 2004, Yale University Department of Molecular and Cellular Biology.

December 10, 2004, Cold Spring Harbor Laboratories Plant Genomes meeting.

December 13, 2004, National Science Foundation, Arlington, VA.

January 22, 2005, Computable Plant FIBR Project retreat, Kerckhoff Marine Laboratory, Corona del Mar, CA.

May 25, 2005, Harvard University Plant Science Symposium, Cambridge, MA.

Talks by Eric Mjolsness, featuring Computable Plant work:

"Physics-inspired models of regulatory networks for cellular and developmental biology". Invited seminar Center for Theoretical Biological Physics (CTBP) in San Diego, a joint effort between research groups at UCSD, the Salk Institute, the Scripps Research Institute and the San Diego Supercomputer Center. http://ctbp.ucsd.edu/seminars.html . May 20, 2005.

'Image understanding and modeling for biological development: the case of a plant shoot meristem", Eighth Workshop on Mining Scientific and Engineering Datasets, in conjunction with the Fifth SIAM International Conference on Data Mining (SDM 2005), http://www-aig. jpl.nasa.gov /public/MSD05/. Newport Beach, CA. Saturday. Invited keynote. Apr 23, 2005.

"Variable-structure systems from dependency diagrams and dynamical grammars" E. Mjolsness. The Learning Workshop, http://snowbird.djvuzone.org/ . Snowbird, Utah, April 5-8 2005.

'Modeling and image understanding for biological development: the case of a plant shoot meristem", Eric Mjolsness. Computer Modeling in Biology group, Cambridge University, http://www.anat. cam.ac. uk/~comp-cell/CMBGroup /previous.htm. March 29 2005.

Some Mathematical Problems in Computational Systems Biology", Eric Mjolsness. UC Irvine Applied Mathematics seminar. Invited talk. March 14, 2005.

"Measuring and modeling the growth of a plant shoot apical meristem", Eric Mjolsness. Dynamics Days, 24th Annual International Conference. http://www. physics. uci. edu/dynamicsdays2005/. Long Beach CA, January 8, 2005.

"The Computable Plant: Software and Science", Eric Mjolsness, (Dual poster presentation), Caltech Biology Division Annual Retreat, Redondo Beach CA, October 16 2004.

"The Computable Plant: A Software Architecture for Developmental Modeling in Plants" by Victoria Gor, Bruce Shapiro, H. Jonsson, M. Heisler, GV.Reddy, E Meyerowitz and Eric Mjolsness. Fourth International Conference on Bioinformatics of Genome Regulation and Structure. http://www.bionet. nsc.ru/meeting/ bgrs2004/. July 2004.

Talk by Marcus Heisler

Meristems 2005 7th Plant Science Institute Symposium, (June 2-5 2005).

Talks by Bruce Shapiro

A Computational Modeling Framework for the Arabidopsis Shoot Apical Meristem. (Abstract: Minisymposium Presentation and Poster, *To Be Presented*) Abstract accepted for the Annual Meeting of the American Society of Plant Biologits, July 16-20, 2005, Seattle, WA.

Developmental Modeling of the Arabidopsis Shoot Apical Meristem by the Computable Plant Project. 35th Biological Systems Simulation Conference, April 19-21, 2005, Phoenix Arizona. (Invited Keynote)

Computable Plant: A mathematical framework for developmental modeling in plants. Presented at the 2004 *SIAM Life Sciences Meeting*, Portland, OR, USA, July 2004.

Developmental Modeling of the Arabidopsis Shoot Apical Meristem. 44th Annual Meeting of the American Society for Cell Biology, Washington, DC, USA, Dec 4-8, 2004. (Poster)

Talk by Ashish Bhan

"Static and dynamic models of biological networks", UIUC conference on complex systems, May 17, 2005.

Abstracts

Shapiro BE, Agrawal V, Gor V, Heisler M, Jönsson H, Meyerowitz EM, Mjolsness E, Reddy GV, *A Computational Modeling Framework for the Arabidopsis Shoot Apical Meristem*. Plant Biology 2005, Annual Meeting of the American Society of Plant Biologists, Seattle, WA USA, 16-20 July 2005.

Shapiro BE, Agrawal V, Gor V, Heisler M, Jönsson H, Meyerowitz EM, Mjolsness E, Reddy GV, *Developmental Modeling of the Arabidopsis Shoot Apical Meristem by the Computable Plant Project*. 35th Biological Systems Simulation Conference, US Water Conservation Laboratory, Agriculture Research Service, Phoenix, AZ USA 19-21 April 2005.

Jönsson H, Heisler M, Shapiro BE, Gor V, Venugopala Reddy G, Meyerowitz EM, Mjolsness M, *Modeling and in vivo live imaging of the Arabidopsis shoot apical meristem*. Presented at the 15th International Conference on Arabidopsis Research, Berlin, Germany, July 11-14, 2004.

Shapiro BE, Heisler M, Reddy GV, Jönsson, Gor V, Meyerowitz EM, Mjolsness M, *Developmental Modeling of the Arabidopsis Shoot Apical Meristem. Molecular Biology of the Cell*, 15(Supplement):315A, November, 2004, Abstract Number 1204, 44th Annual Meeting of the American Society for Cell Biology, Washington, DC, USA, Dec 4-8, 2004.

Manuscripts under review

"Dependency Diagrams for Variable-Structure Systems", Eric Mjolsness. Submitted to Neural Information Processing Systems, June 2005.

"Dynamical Grammars Diagrams for Variable-Structure Systems", Eric Mjolsness. Submitted to Neural Information Processing Systems, June 2005.

Manuscripts in preparation, with tentative titles

"An auxin-driven polarized transport model for phyllotaxis", Jönsson, H., M. Heisler, E. Meyerowitz, E. Mjolsness. In preparation for journal submission.

"Biodynamics of *Arabidopsis thaliana* Meristems: Empirical Evidence and Mathematical Models", A. Sadovski, M. Heisler, E. Mjolsness, P. Baldi. In preparation for journal submission.

"Static and Dynamic Models of Biological Networks", A. Bhan and E. Mjolsness. In preparation for journal submission to Complexity.

Additional creative works

Arrays proposal for SBML – revision. Finney, Gor, Mjolsness, Shapiro.

"Evaluation of the Workshop, 'The ABCs of Developmental Botany: Integrating Plants in the Classroom' . June K. Hilton, Ph.D., David E. Drew, Ph.D.)

Specific Accomplishments

Image analysis and data acquisition

Novel additions to the Softassign matching algorithm have been developed to extend the algorithm to the case of possible splitting of the objects' tracks, as well as non-isotropic distance measure, generally individual for each matched object from each time-point. (Bacarian)

Correctional procedure for the growth-caused distortion of the original data images was developed, based on additional reference scan of the plant at different speed. (Bacarian)

A problem of correction for the nuclei center-positional errors was approached with the development of a new algorithm involving matched data from all time points simultaneously in one treatment, iteratively utilizing Thin Plate Spline (TPS) interpolation, polynomial smoothing, concurrent, statistically based, rejection of erroneous trajectories, and fine tuning for image registration mistakes. (Bacarian)

The resulting movies showed for the first time with this level of accuracy the pattern of development of a plant SAM primordium. Particularly, it showed increased volumetric growth rate in the regions of the future primordium formation, followed by change in primary linear growth direction from tangential to the surface to normal to the surface of the SAM. (Bacarian, Heisler)

Reconstruction of meristem geometry using 3-D Voronoi tessellation; also support for the use of Amira related and basic computing issues. (Agrawal)

Differentiation patterns: New markers have been generated for monitoring gene expression dynamics during primordial development including protein fusions between GFPs and *SHOOTMERISTEMLESS* (*STM*) and *CUPSHAPED COTYLEDONS2* (*CUC2*). Plant lines have also been generated containing triple combinations of markers such as PIN1GFP, REVYFP and pFIL:dsRED as well as pLFY:GFPER, pCUC2:VENUS and pFILdsRED. These markers are also starting to be examined in *Arabidopsis* mutants such as *filamentous flower-1* and *revoluta-1*. (Heisler)

Modeling

A number of models are described in detail in previous sections, notably models of meristem maintenance (CLV3/CLV1/WUSCHEL) and phyllotaxis (auxin/PIN1).

Evaluated ellipsoid mechanical model for meristem cells and found it wanting. (Sadovsky, Bhan)

Simplified 1-dimensional stochastic model for cell "lanes" in root meristems. (Sadovsky, Baldi)

Network inference

Implemented the Levenberg-Marquardt algorithm for parameter estimation (network inference) in MATLAB for a system of metabolic reactions and compared the performance with the simulated annealing algorithm. (Bhan)

Implemented the fugacity distribution calculations to construct static networks with desired degree distributions for some test cases. (Bhan)

Applied Mathematics

Developed a fundamental theory of "dynamical grammars" in terms of operator algebras, for use in multiscale models in systems biology. Outlined its application at molecular, cellular, and tissue level in image analysis and predictive dynamical modeling of phyllotaxis in SAM. (Mjolsness)

Implemented dynamic grammar algorithms to generate lineage trees (constrained and unconstrained) and compare with theoretical predictions. (Bhan)

Software Tools

A new C++ based general purpose implementation has been developed for the Softassign matching algorithm extended to possible splitting of the objects' tracks, as well as non-isotropic distance measure. (Bacarian)

Designed, implemented and tested pairwise matching algorithm (based on optimization of energy function) for cell data. The algorithm is successful, producing on average 5% error. It is currently being evaluated for matching meristem cell data. (Gor)

Matlab preprocessing toolbox was developed for extrapolating, registering, and correctively transforming the original image data to achieve undistorted 3D time-series of the growing SAM images. (Bacarian)

Designed, implemented and tested heuristic cell tracking algorithm, in current use. (Gor)

Implemented and tested cell tracking algorithm based on optimization of comprehensive energy function model (Gor).

A separate processing package consisting of infrastructure of the tools (C++ code and shell scripts), utilizing the extraction and the described above matching codes has been developed. (Bacarian)

A separate post-processing Matlab toolbox, for obtaining resulting differential growth fields (velocity field, volumetric growth rate, linear growth components) was written,

based on the described above procedure. The data were optionally reformatted for subsequent use in Amira software based enhanced visualization. (Bacarian)

The pre- and post- processing Matlab toolboxes have been posted with the project's main server at computableplant.caltech.edu. (Bacarian)

Designed and Implemented (in C++) SBML dynamic array language extension, for developmental modeling, to new SBML parser written in C++. (Gor)

Finite elements elastic modeling libraries in Mathematica. 2D model: Developed 2D tissue data structure. Hierarchy: vertices, edges, faces. Developed library to access elements of 2D tissue data structure. Developed library to orient cells (needed to correctly determine direction of forces). Developed library to find different centers of cells: as center of vertices, center of edges, center of faces. Developed library to display 2D cells with specification of vertices, edges, and faces. 1D Finite Elements (Bars) Library: Element Stiffness Matrix: Assemble Master Stiffness Matrix: Calculation of forces acting on edges of cells through turgor pressure; Main simulation function (iteratively calculate element stiffness matrices, assembles master stiffness matrix, calculate forces, solve system of overdetermined linear equations). 3D Finite Elements Library: all of the above and more. 3D model: all of the above, plus: Developed library to get description of tissue on different levels: vertices, edges, faces, cells. Developed library to calculate volumes and areas of cells and cell elements. Developed library to orient cells (needed to correctly determine direction of forces). Developed an additional 3D tissue data structure that included description of faces as polygons as well as sets of edges. Developed library to display 3D tissue from different view points with ability to select cells to display. Main simulation function (iteratively calculate element stiffness matrices, assembles master stiffness matrix, calculate forces, solve system of overdetermined linear equations) . Voronoi tesselation (method to build cells from coordinate of nuclei with assumption of equal distance from neighbor nuclei to the wall) using QHull. (Vorobyov)

Synergistic activity: encoded CLV3/WUSCHEL model for implementation in the "Sigmoid" pathway modeling database, as a prototype of a future web-executable developmental modeling facility. (Agrawal; Ben Compani, at no charge to FIBR project).

Outreach

Web site for 2004 and 2005 Computable Plant workshops. <u>http://www.outreach.caltech.edu/computableplant/</u> (Kirouac)

Computable Plant workshop for high school teachers at the Huntington Botanical Gardens: "The ABCs of Developmental Botany: Integrating Plants into the Classroom", August 23-27, 2004. The course is geared to high school level biology teachers who are looking to enrich their classroom with inquiry-based lessons using plants. These lessons not only incorporate state and federal standards for biological sciences, but expand on them, including state-of-the-art techniques and information. This course is meant to help teachers bridge the gap between high school material and the research world. The course

focuses on giving teachers the confidence, skills, desire, and means to use plants as an experimental system in their classroom. As a demonstration of the teachers' expanded knowledge of botany, they create lesson plans that are appropriate to their classroom. In addition, we are trying to create both a support community for bringing plants back into the classroom, and a sense of ownership of the material. The learning opportunities available through this course have proven to be so attractive that several teachers from out of the area have even found their own support for travel and housing. (Kirouac, Folsom, Reichmann, Meyerowitz, Wong) An external evaluation report on the 2004 teachers' workshop was given to Chris Greer on his site visit to the Huntington. (Kirouac, Folsom, June K. Hilton, Ph.D., David E. Drew, Ph.D.)

Preparations for the 2005 workshop, July 11-15th and July 28-29th, are underway. Based on feedback and and project evaluation from 2004, we have further developed both the notebook for teachers and the material for the course. One of the major focuses this year will be the extension of the idea of the hypothesis, which is well incorporated into science education and into the idea of a model.

Developed a basic 2-day lesson plan for high school students consisting of the mathematics behind biological modeling, focusing on topics such as: graphs, slopes, functions (in particular exponential and logarithmic), characteristics of a mathematical model, best fit lines, and predictive analysis. (Wong)

Developed a Mathematica notebook that uses Cellerator notation to represent basic exponential growth models in an "easy to interact" format for students. In other words, the complicated math (differential equations) is hidden from the user, but the graphs are displayed. This notebook and activities (in the form of worksheets) will be made accessible through the web for teachers and students. (Wong)

KML Computable Plant meeting 1/22/05

One snapshot of the technical work of the project for the first half of the year emerged from the one-day retreat conducted at Kerckhoff Marine Laboratory in January. These minutes were recorded by Meyerowitz and edited by Mjolsness.

20 attendees – from CIT Elliot Meyerowitz, Venu Reddy, Marcus Heisler and Barbara Wold (arr noon); from JPL Victoria Gor; from Novosibirsk [Russian Academy of Sciences Institute of Cytology and Genetics] Nadya Omelianchuk, Nikolay L. Podkolodny, Nicolay Kolchanov and Vitaly A. Likhoshvai; from UCI Pierre Baldi (<u>http://www.ics.uci.edu/~pfbaldi/</u>), Eric Mjolsness, Gang Liang (a new statistics prof. at UCI <u>http://www.ics.uci.edu/~liang/</u>), three of Eric's postdocs, Alex Sadovsky, Ashish Bhan, Tigran Bacarian; several Eric grad students Li Zhang, Fang Fang, Kiril Petrov, and two programmers from Eric's group, Alexey Vorobyov and Ben Compani.

1) Meyerowitz, general intro to Computable Plant, SAM and CLV system

2) Reddy: dynamic analysis of SAM – including new data on rate of cell div in CLV3 RNAi – in CLV3-RNAi, central zone as indicated by CLV3 expression enlarges right away, then meristem gets bigger. While it is growing the former peripheral zone still divides faster that the former CZ, though both are now faster than they were before CLV3 was shut off. So – overall rate of cell division changed, but the distinction between the meristem's center and margin in cell division rate is still present even in the absence of CLV signaling.

3) Heisler: growth quantification and shape change in SAM, then auxin models of phyllotaxis.

4) Nadya Omelianchuk: AGNS (Arabidopsis GeneNet Supplementary database), internet-available resource on gene expression during development of Arabidopsis organs, derived from literature. 6 modules: sequence database with genes and mutations (like a Genbank entry with additional info on mutant alleles), expression database (including stage of development looked at, mutant backgrounds, references to each paper), phenotype database (each allele listed, with reference, and place of phenotype in controlled stage and organ/tissue vocabulary. Each datum accompanied by quote of relevant sentence from paper!). Developmental stages vocabulary, reference database, organs tissues and cells vocabulary.

Can query on gene, organ, stage, and on abnormal expression in mutant or transgenic plant. Organ descriptions are hierarchical – e.g. 2.1 embryo apical region, 2.1.1. emb apical domain, 2.1.1.1 emb presumptive SAM, 2.1.1.2 emb apical domain inner cells (not protoderm).

Lunch

5) Alexei Vorobyov: mechanical simulate cell as two nested polyhedra, give each wall a stiffness and then expand and get expansion related to stiffness. Also working on Voronai diagram representation of cells, to eliminate the L1 cells projecting to infinity.

6) Vitaly Likhosvai: Mathematical modeling of gene networks and dynamics of gene function using simple differential equations, in a gene-network frame model.

7) Ashish Bhan: network inference. Boolean networks (Kaufman), Bayesian networks, Nonlinear ODEs (the Cellerator method), piecewise differential equations, PDEs and other spatially distributed models. Levenberg-Marquardt optimization method – minimize a general function assumed to be quadratic, in this case minimize the x square error – difference between the model's estimate of a parameter and the data.

8) Li Zhang: Simulated annealing. Given model and guessed parameters, infer real parameters.

9) Victoria Gor: Cell tracking algorithms that follow cell divisions.

10) Tigran Bacarian: growth mapping of the SAM, velocity vector fields for each 2D slice

11) Alex Sadovsky: Hejnowicz growth map models - conjectures are that if v is a velocity field in a growing SAM in a special coordinate system, in which there is no shearing – deformations are always tensile stretching, compression or rotation.

12) Eric Mjolsness: Bio experiments, image data analysis, modeling software (Sigmoid and Cellerator), model construction, bio application, then mathematical philosophizing and dissemination of results. Mathematical philosophy: dynamical grammars (grammar because it has syntax and semantics) modeled objects change their number (cell birth and death) and change their relations to each other. So it needs to have elementary creation and annihilation operators for cell death and birth, reaction operators.

Potential plans for future:

Computational Modeling: Workforce on modeling, Derived growth fields as template, new generation of mechanical models, self-assembly in callus.

Computational tools: revive autogenerated code effort, GUI for cellerator with cell grids or Sigmoid standard shapes, develop AGNS further.

Bio experiments: SAM CLV, SAM phyllotaxis, SAM growth, callus generation of selfassembled new SAMs; measure physical properties of single cells, modeling on flower gene networks, patterning organs in terms of polarities, growth modeling of flower development, phyllotaxis in flowers and the various PINs on in them.

Discussion

Beneficial effects of large-scale project funding

The FIBR program may be unusual at NSF in allowing larger-scale groups of scientists to work within a common collaborative project in biology. In our experience this has several major advantages.

The unique effect of large-scale funding in the Meyerowitz lab was the purchase of a dedicated Zeiss LSM 510 Meta upright laser scanning confocal microscope, which allows long timecourse real-time dynamic imaging of the shoot apical meristem – so far a capacity not matched in any other laboratory, and the basis for the data acquisition for the meristem and phyllotaxis models.

We have also been able to undertake multiple overlapping approaches to both image analysis and computational modeling, in order to compare and recombine them. Image analysis includes feature matching algorithms and codes by Bacarian and Gor, both starting from Softassign and Thin Plate Splines but diverging substantially afterwards, and both evaluated on meristem and bacterial imagery; a nascent generative approach to 3D plant image modeling and object recovery by Liang, Fang and Mjolsness; and detailed image registration improvements that have proved essential to recovering the 3D growth field by Bacarian and Heisler. Modeling approaches include the originally proposed GSRN/spring model pursued to satisfactory state by Jonsson, Heisler, et al (ISMB 2005); a software project to autogenerate the same kinds of models from biologists' input by Gor; an alternative coarse-scale continuum model pursued by Sadovsky and Baldi; a fine-scale cell-compartmental elastic modeling code developed by Vorobyov and Mjolsness. These three approaches may develop into a combined multiscale spatial modeling framework for plant development. We think this kind of coordinated attack on core problems is succeeding, and would not be possible under ordinary NSF grants.

Another salutary effect of larger-scale project funding is its effect on the development of foundational intellectual tools. As an example, it's not easy to fund "dynamical grammars" through ordinary means, since few reviewers have both the multidisciplinary background and the time to invest in understanding early-stage theory even of demonstrable importance. However it is the intellectual core for Prof. Mjolsness's research program in biological modeling. This year important conceptual progress of relevance to developmental modeling has been made, as recorded in the submissions to the Neural Information Processing Systems conference, and they have been applied to guide the development of image analysis algorithms and multiscale developmental models.

Summary

A substantial escalation in activity has taken this year place in the Computable Plant project, as detailed above. Major scientific results include quantification of meristem growth, and our first round of predictive models of meristem maintainance and phyllotaxis. Numerous supporting developments have occurred in biological experiments, microscopy, image analysis, mathematical modeling frameworks, simulation codes, and network inference algorithms and codes. These continuing developments will enable the construction of more comprehensive models and stronger experimental results. Dissemination efforts represented the project at many conferences (with strong positive response including interest by agronomists) and in publications. A unique outreach program to high school teachers was launched, to encourage the use and understanding of plants in the science classroom. Strong synergies were observed that would be very difficult to reproduce in a smaller scale project.