The Computable Plant: Annual Report, 9/2003-5/2004.

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

Overview

Major accomplishments of the NSF FIBR Computable Plant project for the partial year from 9/2003 through 5/2004 can be categorized as experimental, computational, and outreach-related. Experimental accomplishments include the purchase and installation of a dedicated confocal laser scanning microscope for live plant imaging (Figure 1), the development of improved imaging techniques including GFP (Green Flourescent Protein) variants and floral meristem primordium markers, the creation of 3D and 4D image data sets bearing on Shoot Apical Meristem (SAM) morphogenesis and cell division, and the discovery of coordinated patterns of cell division in the floral primorida of the SAM. Computing accomplishments include the creation and maintenance of the web site www.computableplant.org, the development of models and simulations exploring mechanisms for phyllotaxis compatible with auxin/PIN1 regulatory interactions and with expression data, the design and prototyping of new simulation software which will in the future allow exchange of developmental models between research groups, and progress towards the formulation of a next-generation "beveled polyhedron" modeling framework for the interaction of biomechanics and regulatory networks. We have also initiated an outreach program (www.outreach.caltech.edu/computableplant) that will host its first teacher workshop at the Huntington Botanical Gardens in August 2004.

Figure 1. Confocal laser scanning microscope setup dedicated to plant developmental imaging in the Meyerowitz laboratory at Caltech, immediately after installation (photo taken 4/26/04). This is a centerpiece and a major investment of the FIBR Computable Plant project and is now producing useful imagery.

Imaging technology

Marcus G. Heisler

In order to understand the temporal and spatial relationship between various markers it is necessary to make full use of the available GFP variants. Previously we have successfully used ER localized GFP, Histone 2B fused YFP and GFP fused to proteins of interest. However we have discovered a number of limitations so far including:

1) Tissue damage due to bleaching and excessive laser exposure.

2) Excessive detector noise from the Meta spectral analyzer when using suboptimal confocal pinhole apertures

To address these issues we have been testing new GFP variants and multimerized GFPs in an effort to increase signal strength. Data so far suggests this approach can indeed increase signal to noise ratios very significantly (see Figure 2). Another strategy is to use GFP variants that do not require the Meta for signal separation. Thus we have tested a GFP/dsRED express combination and shown that this works well even for time lapse (see Figure 3).

These approaches are also being adapted for use with novel ubiquitous promoters and our basic set of nuclear and plasma membrane localization tags in order to significantly improve our ability to geometrically map cell positions and sizes over time.

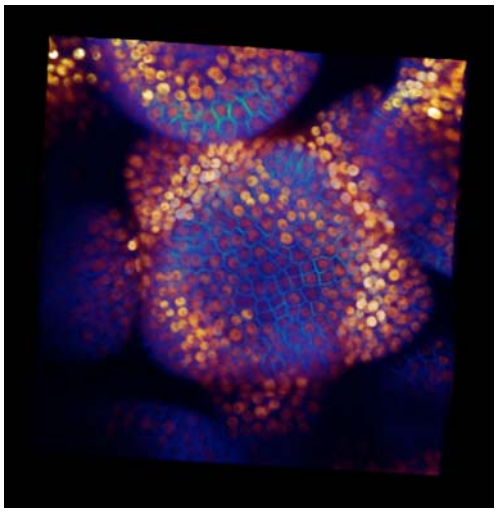

Figure 2. PIN1GFP in combination with pCUC:3XVENUS N7. The pCUC:VENUS N7 signal is the strongest reporter generated so far suggesting that multimerization of GFPs increases signal strength.

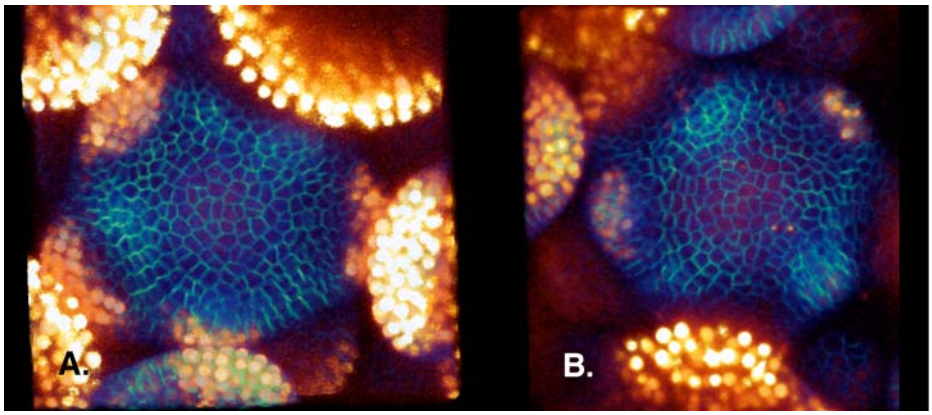

Figure 3. pPIN1:PIN1GFP in combination with pFIL:dsRED N7. (A) Expression pattern at 0 hrs. (B) Expression patterns at 33 hrs with imaging every 3 hrs in between. Note two new primordia are expressing the dsRED marker.

Spatio-temporal dynamics of primordial gene expression and growth

Currently a large number of genes are known to be expressed in primordial regions and many of their functions are at least partially understood. However their relative expression domains are poorly characterized both spatially and temporally. The aim of this project is to determine the spatial and temporal relationships between such genes in an effort to understand how primordial cells are progressively specified. These markers will also be combined with a ubiquitous nuclear marker in order to correlate gene expression dynamics with cell behavior.

So far we have completed the construction of primordial makers including:

pFIL:dsRED Express N7 (N7 is nuclear localized) (see Figure 3)
pPID:dsRED Express N7
pREV:REV:VENUS translational fusion
pCUC:3XVENUS N7 (see Fig. 1)

Markers in progress include:
pZLL:dsRED N7
pANT:dsRED N7
pAS1:dsRED n7
pSTM:3XVENUS N7

One of the consequences of this work will be to enable the testing of auxin/PIN1 phyllotaxis models (see Modeling Phyllotaxis section below) by monitoring the polarity of pPIN1:PIN1GFP in response to localized auxin sources.

Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana*

G. Venugopala Reddy, Marcus G. Heisler, David W. Ehrhardt and Elliot M. Meyerowitz

Precise knowledge of spatial and temporal patterns of cell division, including number and orientation of divisions, and knowledge of cell expansion, is central to understanding morphogenesis. Our current knowledge of cell division patterns during plant and animal morphogenesis is largely deduced from analysis of clonal shapes and sizes. But such an analysis can only reveal the number, not the orientation or exact rate of cell divisions. In this study, we have analyzed growth in real time by monitoring individual cell divisions in the shoot apical meristems (SAMs) of *Arabidopsis thaliana*. The live imaging technique has led to the development of a spatial and temporal map of cell division patterns. We have integrated cell behavior over time to visualize growth. We have analyzed the cell behavior associated with flower primordium development starting from a stage in which the future flower is comprised of four cells in the L1 epidermal layer. Our results show that primordium development is a sequential process linked to distinct cellular behavior. Our analysis reveals that the amount of cell division is comparable in regions of the SAM where successive primordia arise. Changes in cell division orientation are associated with initial outgrowth of a flower primordium. These changes are followed by a rapid burst of cell expansion and cell division, which transforms a flower primordium into a three dimensional flower bud.

One of the major limitations in understanding growth in both plants and animals has been the inability to monitor cell behavior in real time. Several studies have tried to address this issue starting from inference of cell behavior from clonal analysis, to generative modeling of growth through computer simulations. Our analysis of growth in real time circumvents the requirements for inference in studies of clonal growth, or for theoretical growth simulations. Once cell positions can be extracted by cell-finding algorithms, it should be possible to integrate cell coordinates in time-lapse observations. Such efforts are currently in progress. The challenge for the future is to superimpose models of gene regulatory networks on such models of growth, and to integrate with these models the cell-cell interactions involved in meristem maintenance and morphogenesis.

Work listed above has been accepted for publication in DEVELOPMENT (Reddy *et al.* 2004).

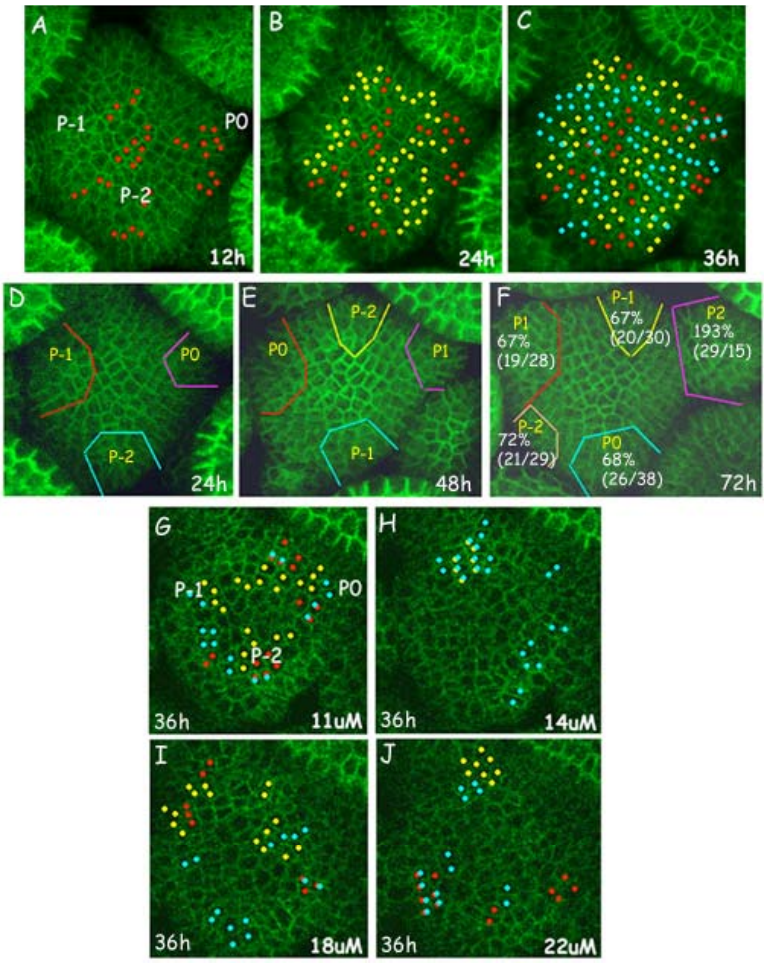

Figure 4. Spatial distribution of mitotic activity over a period of time. (A), (B) and (C) show reconstructed L1 layer of the same plant separated by 12hours interval. Cells divided in each of the 12hours window are differentially color-coded. Red dots represent cells divided in first 12hours window, yellow dots the following 12hours and the blue dots the next 12hours. (G), (H), (I) and (J) are individual optical sections from the same plant depicting cells located in L2 and corpus from the same time point as in (C) and the color code remains the same. (D), (E) and (F) represent reconstructed views of L1 layer of the same SAM followed over 72hours. Elapsed time in individual panels is marked. Different colored sectors represent regions of primordium development marked as P-2, P-1, P0, P1 and P2. The number expressed as percentages in (F) represent averaged mitotic index calculated for every 24hours interval over 72hours, in sectors representing same stage primordium. The numbers in parenthesis indicate the number of cells divided over total number of cells. From (Reddy et al., 2004).

Fluorescent cell type specific markers for Shoot apical meristems (SAMs)
G. Venugopala Reddy and Elliot M. Meyerowitz

The shoot apical meristem (SAM) is a collection of distinct cell types located in specific positions with specialized functions. The central zone (CZ) cells harbor initials or stem-cells, while the cells at specialized regions in the peripheral zone (PZ) differentiate into organ primordia. The cells in Rib-zone are incorporated into the developing stem. All these cell types are located in a dynamic environment and hence they have to constantly assess their position in order to retain their identity. The dynamics of cell fate specification can be best studied by combining cell type specific markers and cell behavior in a single study. In this context, we are developing a set of promoter constructs to direct the expression of fluorescent proteins in distinct cell types. We have designed promoter-fluorescent protein constructs for *CLAVATA3 (CLV3), UNUSUAL FLORAL ORGANS (UFO), LEAFY (LFY)* and *WUSCHEL(WUS)* to mark cells in CZ, PZ, incipient primordia and part of the Rib-zone respectively. The plants carrying these constructs have been obtained and their expression patterns are being characterized in detail, in time-lapse experiments.

Cell behavior, cell types and meristem maintenance
G. Venugopala Reddy

The genetic studies have revealed the signaling mechanisms involved in meristem maintenance. Mutations in *CLAVATA* genes (*CLV1, CLV2* and *CLV3*) result in bigger meristem, while mutations in *WUSCHEL (WUS)* results in a failure to maintain a functional meristem. The nature of gene products, expression domains, genetic interactions and mis-expression studies have contributed to an elegant model involving positive and negative feedback loops operating on each other to maintain meristem size. The function of WUS is required to maintain a constant stem-cell pool in CZ and at the same time CLV genes function to repress WUS expression domain. However, the studies aimed at understanding the cell behavior and analyzing cell types in these mutant contexts are restricted to single time-point observations and mostly restricted to inflorescence meristem, which represents a terminal phenotype. The function of these gene products can be best understood by analyzing the effects of their transient dysfunction, in real time, on both cell types and cell behavior. Such studies might yield new insights into the kinetics of re-organization of cell types in SAM in relation to the altered cell behavior. Such a system can also be employed to study the effect of altered cell types on cell division patterns in relation to cell position. We have used the hormone inducible (dexamethasone) constructs, which result in mis/overexpression of WUS to perturb meristem maintenance function (constructs provided by Frank Wellmer). The effect of WUS misexpression was tested on CLV3 expression pattern. This analysis has revealed that the expansion of CLV3 expression upon WUS induction is due to the re-specification of PZ cell identity and not due to the additional cell divisions within CLV3 expression domain. The detailed analysis of PZ cell identity is being carried out by utilizing the PZ marker gene such as UFO and LFY, which is expressed in developing organ primordia within the PZ.

Modeling Phyllotaxis

Henrik Jönsson, Marcus Heisler, et al.

An important problem connected to the shoot apical meristem is the mechanism of phyllotaxis. New leaf and flower primordia are initiated from the periphery of the SAM in an amazingly symmetric fashion, which has attracted scientists dating back to antiquity. The plant hormone auxin is believed to play a major role in the initiation of the new primordia, and recent research shows that membrane proteins involved in auxin transport, such as PIN1, are expressed and polarized in a phyllotactic pattern in the epidermal layer of cells within the SAM (Figure 5A).

A large number of models have been proposed for phyllotaxis, based both on molecular and mechanical ideas. A common factor of many of these models is a regular spacing mechanism in combination with a continuous growth of the apex, but the models have had no or little connection to molecular experiments in the plant. The recent experiments described have illuminated a transport mechanism to be essential for the patterns to form. In an ongoing work, we are developing a model based on this data. In vivo imaging of a GFP-fused PIN1 protein is used to explore the expression and polarization of the PIN1 protein within cells (Figure 1A). The data indicates that PIN1 is polarized to walls closer to new primordia, which is assumed to be where the auxin concentration is highest. Models taking this into account were developed and resulted in formation of evenly spaced auxin concentration peaks, resembling phyllotactic patterns (Figure 5). Currently we are testing the model in experiments where micro-addition of auxin to the SAM is applied, followed by analysis of how the PIN1 polarization is affected.

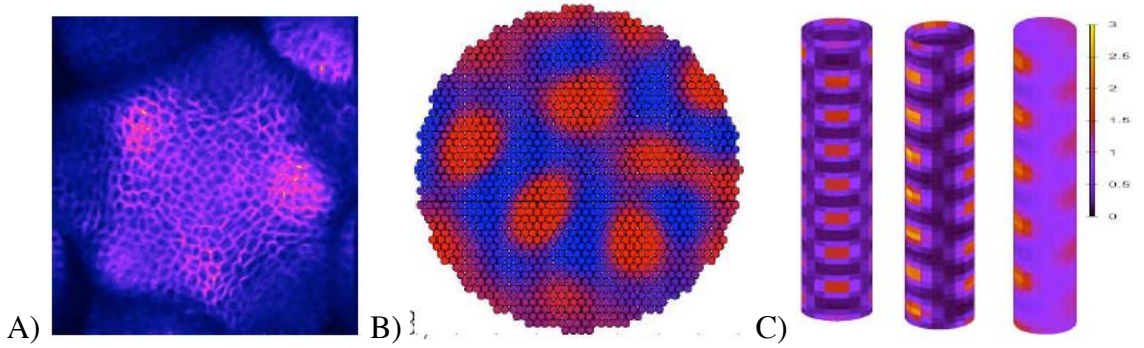

Figure 5: Plant phyllotaxis. A) Top view of a SAM, where a PIN1-GFP protein fusion is used. The membrane protein is expressed in a phyllotactic pattern. B) 2D simulation of the model described in the text. A spiral phyllotactic pattern is continued from an initial state where five of the peaks are present. New peaks appear at the "correct" positions. C) A small auxin concentration disturbance in one cell at the bottom of the cylinder is sufficient to create phyllotactic patterns in the simulations. Different parameter sets lead to different patterns such as pairwise, spiral, and opposite, all present in different plant species.

Using these techniques, we are developing a set of alternative few-player models to explain the known expression patterns and to guide discriminating experiments. These models include auxin gradient and flux models, an internal auxin gradient model, a model based on an intermediate signaling molecule "X", and a model based on an unknown

hysteresis loop with a memory factor "M". These alternative models can be expressed or reformulated in terms of reaction notation (see Model Generation section, Figure 10).

Inferring Model Parameters

Inference of models from data
Another important aspect of a modeling approach is to study the inverse problem where a model is inferred directly from data, with little or no restrictions on the model. In this reverse engineering scenario an optimization algorithm is used to find the best possible model that describes data, and the tool can be used for computationally induced hypotheses. In previous work, we have used a simulated annealing approach for a model describing development in *Drosophila*, and a first application to a growing shoot apical meristem has been explored in (Jönsson *et.al*. 2003).

Our main goal is to find gene network interactions by simulation of different networks and comparing with template data. To be able to measure similarity between a simulation and a template, a quality measure is defined as

$$E = \sum_{t_a}^{N_{time}} \sum_{i}^{N_{cell}} \sum_{k}^{N_{protein}} \sqrt{\left((v_{i,k}(t_a) - v_{i,k}^{(template)}(t_a))^2\right)}$$

E measures the difference of target protein concentrations between a simulation and the template for each cell at each template time point.

Simulations vs. L1-Template Comparison
A 2D template of L1 expression in a growing plant is produced by simulation of the L1-part of the SAM model network.

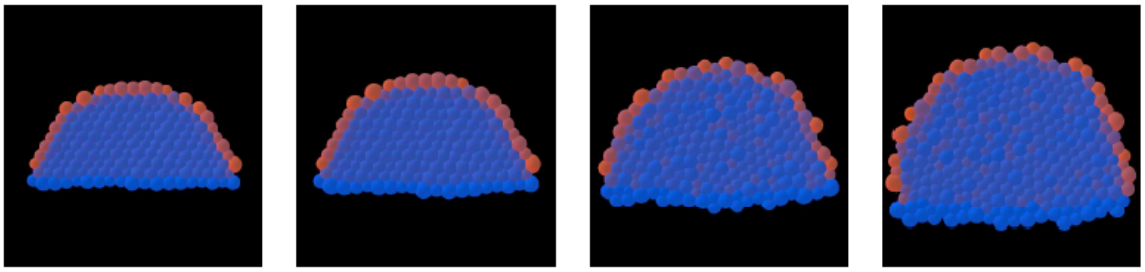

Figure 6. Template data for an L1 expressed gene at four time points.

Simulations with parameter values similar to those used when creating the template are performed.

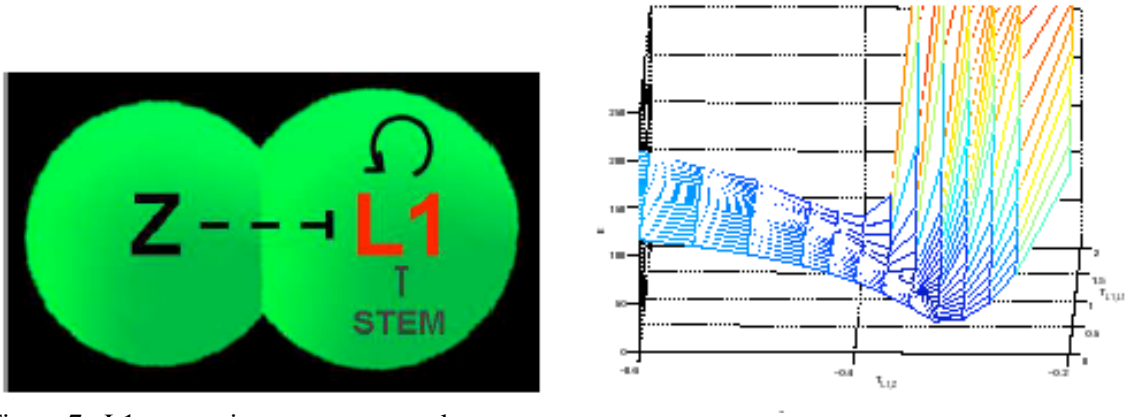

Figure 7. L1 expression pattern network.

The $T_{L1,L1}$ (self inducing intracellular) and $TL_{1,Z}$ (intercellular repression from Z) parameters are varied. In the plot there is a minimum for E at the template parameter values (1.0,-0.35), but also a "valley" of low values showing that an increased Z repression can be compensated by an increase of self activation of $L1$.

A Receptor-Ligand Network

Another network that produces an L1 expression similar to the template is shown in the figure below. It is using a receptor(B) - ligand(A) signal for repressing the L1 expression.

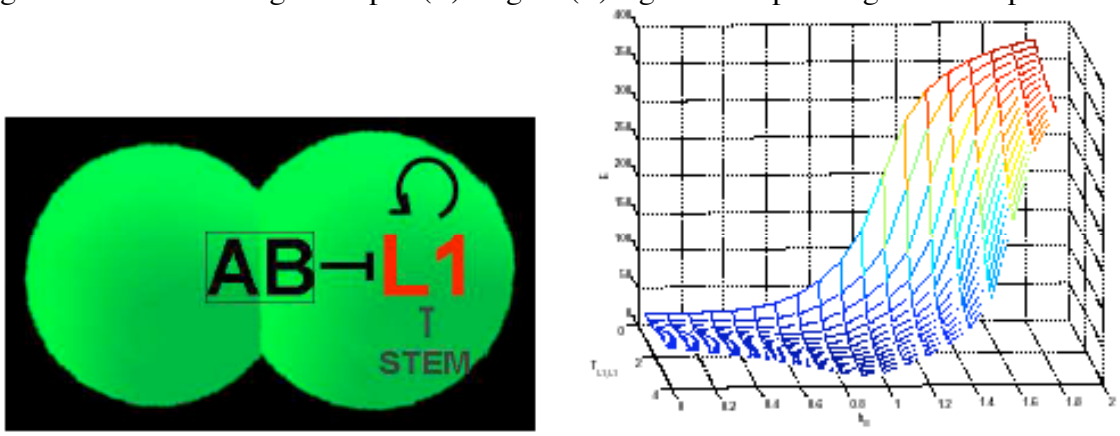

Figure 8. L1 expression pattern network with receptor-ligand interactions.

The $T_{L1,L1}$ (self inducing intracellular) and hB (the default B expression) parameters are varied. In this case there is a "plateau" of values that produce a good result. The parameter space can be searched, and good networks can be extracted from low E simulations. Instead of an exhaustive search, we plan to use an optimizing schema such as Lam-Delosme simulated annealing to find good parameter sets.

Model generation software development

Bruce Shapiro and Victoria Gor

We have undertaken the design and prototyping of new simulation software that will in the future allow exchange of developmental models between research groups, by augmenting the Systems Biology Markup Language (SBML). Our software architecture will also will allow the large and expanding collection of biological mechanism models available in Cellerator (using computer-algebra manipulation) to be applied as needed to developmental modeling. It is illustrated in Figure 9. A prototype of this architecture now runs, and a near-term goal is to create and disseminate a simple reference application involving our previous models of CLV3 expression domain constancy.

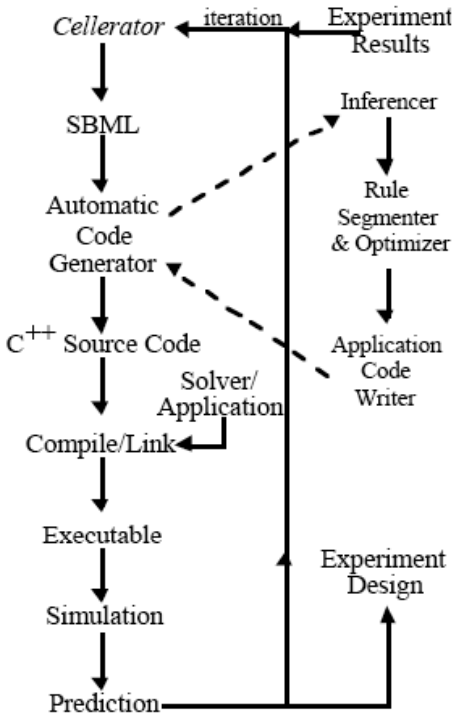

Figure 9. Software architecture and data flow for sharable developmental simulations. The *automatic code generator* is central to this architecture. It consists of an *inferencer*, a *rule segmenter and optimizer*, and *application code writer* modules (Fig. 1). It queries the parser for SBML structures and produces efficient C++ application code. The resulting C++ code is then compiled into object code optimized for the desired application. The first two modules of the automatic code generator – the *inferencer* and *rule segmenter* – are pre-processors. They are called once for each SBML model, independent of the application software to be generated. The *inferencer* receives parsed SBML structures from the parser and infers element attributes given the element name. This reflects the inverse relationships between SBML elements and their attributes. For example, the extended SBML has a parameter attribute "foreach" that indicates the compartment; the *inferencer* creates a list of inferred elements, such as the list of parameters in each compartment. Details in (Gor et al. 2004).

The newest component of our architecture is the Automatic Code Generator. It queries the parser for various SBML structures and produces efficient application code in C++. Required modifications to the SBML parser were made in routines that handle "Rule", "ASTNode", "Model", "Compartment", "Species", "Parameter", "Connection", "Dimension" and "Index" elements. The various rule elements of SBML-V2 were combined into one "Rule" element with five SBML-specified attributes: "variable", "formula", "math", "type" and "ListOfIndices". The resulting C++ code, compiled into an executable program, is the efficient adaptation of the indicated solver to the given SBML model.

Thus, the following have been achieved:

- Creation and maintenance of web site www.computableplant.org
- Extension of current SBML language to handle array formats
- Automatic simulator and optimizer code generation

In support of this architecture, numerous enhancements to our Cellerator computer-algebra based biological model generator have also been made:

- Updates to Cellerator: GMWC, SSystem, Michaelis-Menten, Numerous bug fixes
- Updates to MathSBML to support Arrays in SBML
- Cellerator plugin to write SBML file with GRN network using MathSBML
- Cellerator plugin to perform simulations on a tissue (cell) array
- Various simulations on internal web page at Caltech

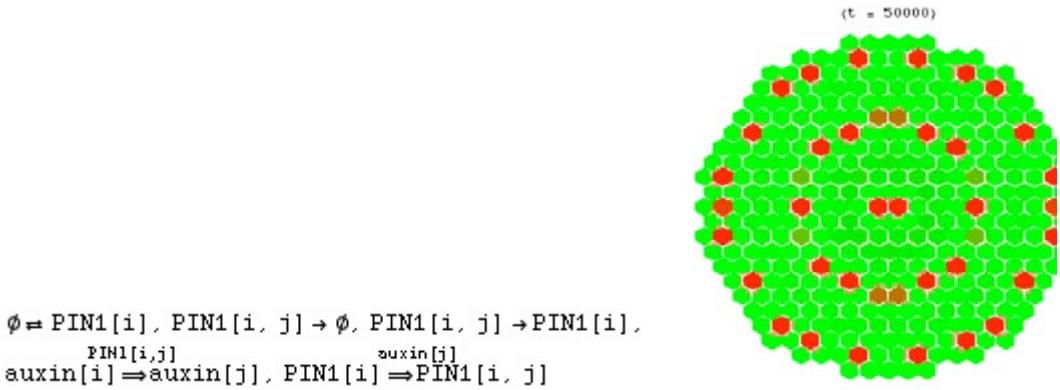

Figure 10. Cellerator input and output for highly simplified auxin/PIN1 model on a hexagonal grid. Note the use of wall/membrane compartments [i,j]: wall/membrane of cell i adjacent to cell j. Observed behaviors are diverse (including solitons) and not yet biologically satisfactory.

Publications and presentations on the software architecture are listed in the reference section of this report as (Mjolsness *et al*. 2004), (Gor *et al*. 2004), (Shapiro et al. 2004), and (Shapiro 2004). Although they weren't funded by this project, our papers from BGRS-2002 and ICSB-2002 have now been published in two collections: (Jönsson *et al*. 2003, Mjolsness *et al*. 2004).

Biomechanics

Eric Mjolsness and Alexey Vorobyov

We have begun to formulate a new biomechanical model, to refine the previous mass-spring model. Both models can be integrated with regulatory network models. The assumptions of the new model are that plant cell geometry is approximately that of "beveled polyhedra" with strong thin walls made of anisotropically oriented cellulose, a relatively weak cytoplasm, and strong spherical nuclei inside. The wall/membrane compartments are also the loci for PIN1 in sub-cellular compartmental regulatory models. We are investigating finite element discretizations for such anisotropic models that would allow for growth and cell division. A likely starting place is a 3D triangulation of the polyhedral cell compartments, with an affine deformation in each tetrahedron, and nonlinear anistropic incompressible stress/strain relationships similar to those used in (Teran et al. 2003) for muscle tissue.

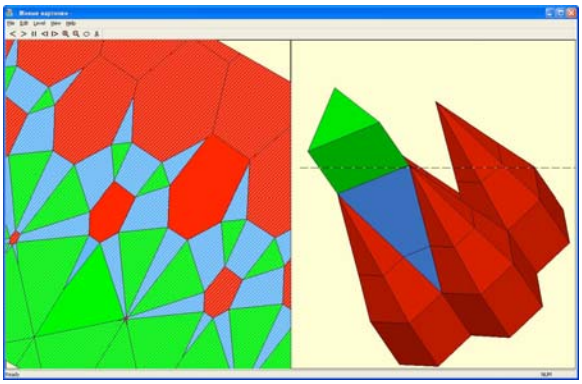

Figure 11. Visualization screenshot is from a program that allows visualization of 3D collections of cells that have a shape of polyhedra. It is possible to look at every cell or every combination of cells from any point of view in 3D, allowing rotation, shift, zoom operations. One can draw sections by arbitrary plane (left pane).

ICG Visit and Workshop

E. Mjolsness

Five visitors from the Institute for Cytology and Genetics came to UCI in April for a productive two-week visit. They included Nikolay Kolchanov (Deputy Director ICG), Nadejda Omelianchuk, Nikolay Podkolodny, Sergei Nikolaevich, and Vitali Likoshvai. Their expertise includes computing, mathematical modeling, biological databases, and bioinformatics software. On April 14 there was an all-day workshop for all project scientists at Caltech's Keck Marine Laboratory at Corona del Mar, 3 miles from the UCI campus. Highlights from this visit included presentation of ICG's relevant work on an *Arabidopsis* phenotype database (see Figure), convergent evolution to a pathway database and pathway modeling environment similar to those associated with Cellerator, and the intersection of mathematical modeling efforts relating stochastic grammars, Colored Petri Nets, network connectivity statistics, and reaction mechanism libraries.

Plans are underway to make the ICG plant phenotype database available to the Computable Plant project.

MA wild type [Fletcher J.C. et al., 1999] RT mRNA, AR RD flowering RO SAM, the central zone, L1 RO SAM, the central zone, L2 RO SAM, the central zone, L3 RL present RC expressed in the L1 and L2 cell layers, and in a few underlying L3 cells [Fletcher J.C. et al., 1999] XX	**MA** clv1-4 [Fletcher J.C. et al., 1999] **MA** clv2-3 [Fletcher J.C. et al., 1999] **MA** clv3-2 [Fletcher J.C. et al., 1999] **MA** shd [Ishiguro S. et al., 2002] **MA** CLV3::CLV3-GFP; clv3-1 [Lenhard M. and Laux T., 2003] RT mRNA, AR RD flowering RO SAM, the central zone, L1 RO SAM, the central zone, L2 RO SAM, the central zone, L3 FL enlarged domain RC domain is enlarged in accordance with the SAM central zone enlargement [Fletcher J.C. et al., 1999] [Lenhard M. and Laux T., 2003] RC CLV3::CLV3-GFP; clv3-1, mGFP4 was used as AR probe, expression was found solely in a three- to four-cells high, wedge-shaped domain in the SAM centre, this expression domain was larger than that of the endogenous *CLV3* gene in wild type, yet much smaller than in non-transgenic *clv* mutants [Lenhard M. and Laux T., 2003] XX MA CLV3::CLV3-GFP; clv3-1 [Lenhard M. and Laux T., 2003] RT protein, GFP RD flowering RO SAM, the central zone, L1 RO SAM, the central zone, L2 RO SAM, the central zone, L3 RL present FL enlarged domain RC GFP fluorescence was detectable not only in the region corresponding to the mRNA also in cells towards the periphery of the meristem in the outer layers, extending farthest in the L1 [Lenhard M. and Laux T., 2003] XX

Figure 12. Example entry in ICG *Arabidopsis* phenotype database.

Outreach

Martha Kirouac and Jill Andrews

Dr. Martha Kirouac is coordinating outreach and instructing our first week-long workshop, "*The ABCs of Developmental Botany: Integrating plants into the classroom*". Fifteen high school teachers will participate in the class from August 23 – 27, at The Huntington Library, Art Collections, and Botanical Gardens. Topics will include: fundamentals of plant development, plant cell biology, classification, mutation, variation, adaptation, cell differentiation, genetically modified crops, and use of dissecting and compound microscopes. A follow-up three-day workshop will be scheduled for winter. The Los Angeles Unified School District has approved the course for three salary point credits, and is advertising it on their site.

Additionally, Caltech's Educational Outreach program has established a course web site (see http://www.outreach.caltech.edu/computableplant/, where a tentative agenda can be found), and is advertising the course through its Local Educator's Network and San Francisco's Exploratorium. We are currently in discussion with Dr. David Drew, Joseph B. Platt Chair in Management of Technology at the Claremont Graduate University, over assessment of the baseline knowledge of participating teachers.

Summary

The accomplishments enumerated in the Overview section have been substantiated in detail. Plans for the coming year include the commencement of serious work in image processing, as well as further development in all areas covered above.

References

Gor V., Shapiro B.E., Jönsson H., Heisler M., Reddy G.V., Meyerowitz EM, and Mjolsness E, A Software Architecture for Developmental Modeling in Plants: The Computable Plant Project. Abstract accepted for BGRS (Bioinformatics of Genome Regulation and Structure, Novosibirsk Russia 2004):

Jönsson H, Shapiro BE, Meyerowitz EM, Mjolsness E (2003). Signaling in Multicellular Models of Plant Development, in On Growth, Form, and Computers, ed. Bentley P and Kumar S, Academic Press, pages 156-161.

Mjolsness E, Jönnson H, Shapiro BE, Meyerowitz EM (2004). Modeling plant development with gene regulation networks including signaling and cell division, in Bioinformatics of Genome Regulation and Structure, ed. N. A. Kolchanov, Kluwer Publications.

Mjolsness E, Shapiro B, Gor V, Jönsson H, Reddy GV, Heisler M, Meyerowitz EM (2004) The Computable Plant: A Software Architecture for Developmental Modeling in Plants. Poster presentation at the Pacific Symposium on Biocomputing (PSB-2004) Big Island, Hawai'i, USA.

Reddy, G.V., Heisler, M.G., Ehrhardt, D.W., and Meyerowitz, E.M. (2004). Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana* (Manuscript accepted, DEVELOPMENT).

Bruce E Shapiro, Eric Mjolsness and Elliot Meyerowitz, The Computable Plant: A Mathematical Framework for Developmental Modeling in Plants. Abstract accepted for SIAM Meeting on Life Sciences 2004.

Bruce E Shapiro, 2004: "Using Computer Algebra for Developmental Modeling: Introduction to Signal Transduction, Cellerator, and the Computable Plant". Invited tutorial at the International School on Complex Systems in Computer Science and Biology, Rovereto, Italy, in coordination with the University of Trento. (No expense to the Computable Plant project was required for this event.)

Teran J., Blemker S., Ng V., Hing T., and Fedkiw R. (2003). Finite Volume Methods for the Simulation of Skeletal Muscle. Eurographics/SIGGRAPH Symposium on Computer Animation D. Breen, M. Lin (Editors).