Inferring Gene Regulatory Network Models for Plant Stem Cell Regulation

Henrik Jönsson[1], Marcus Heisler[2], G. Venugopala Reddy[2], Elliot M. Meyerowitz[2] and Eric Mjolsness[4]

http://www.computableplant.org

[1]Complex Systems Division, Department of Theoretical Physics, Lund University, Sweden. **henrik@thep.lu.se**
[2]Division of Biology , California Institute of Technology, Pasadena CA, USA. **{mheisler,venu,meyerow}@caltech.edu**
[4]Institute for Genomics and Bioinformatics, and School of Information of Computer Science, University of California, Irvine CA, USA. **emj@uci.edu**

http://www.computableplant.org

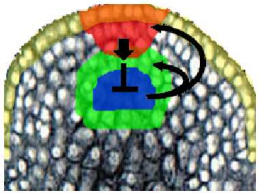
The **Shoot Apical Meristem (SAM)** of plants is the biological target for a mathematical model of multicellular organisms. The model is implemented using ordinary differential equations, and simulations of the dynamical time development are performed. We are here presenting a framework for inferring gene regulatory networks from GFP data in living plants.

The Shoot Apical Meristem (SAM) of *Arabidopsis*

- Source of the aboveground part of a plant
- Small (about 10^3 cells)
- Genes important for the development identified

Genes and Known Interactions

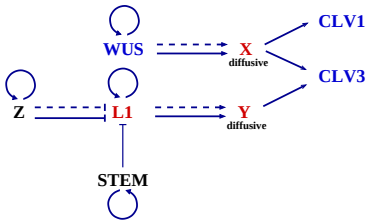
- CLAVATA3 (**CLV3**): stem cell marker
- CLAVATA1 (**CLV1**): receptor kinase
- WUSCHEL (**WUS**): homeodomain, transcription factor
- ATML1 (**ATML1**): homeodomain, transcription factor



The interactions between **CLV3**, **CLV1** and **WUS** partly regulate the development of the SAM, and thereby the complete plant. **WUS** induces both **CLV1** and **CLV3**. On the other hand **CLV3**(ligand) and **CLV1**(receptor) act in a network repressing **WUS** creating a feedback loop for the regulation. Genes exist which are expressed only in the L1 layer (such as **ATML1**).

The SAM Model Network

How can **WUS** regulate **CLV3** when the expression domains do not overlap?



A partly hypothesized network, where the inducing signal from WUS is combined with an L1-originating signal for CLV3 activation. **X** is suggested by experiments, but unknown. **L1** and **Y** have genes with analogous expression patterns (ATML1 and ACR4).

The Generic Model

Essential parts of a developmental system are introduced in a mathematical model, and continous differential equations are simulated for cell parameters and molecule concentrations.

- Cell Growth
- Cell Cycle/Proliferation
- Mechanical Cell Interactions
- Gene Regulatory Network (GRN)
- Molecular Transport

The GRN-Equations

$$\tau_a \frac{dv_a^{(i)}}{dt} = g(u_a^{(i)} + h_a) - \lambda_a v_a^{(i)}$$

where

$$u_a^{(i)} = \sum_b T_{ab} v_b^{(i)} + \sum_j \Lambda_{ij} \Big(\sum_b \hat{T}_{ab} v_b^{(j)} + \sum_{bc} \tilde{T}_{ac}^{(1)} \tilde{T}_{cb}^{(2)} v_b^{(j)} v_c^{(i)} \Big),$$

v - set of protein concentrations

T - intracellular interactions

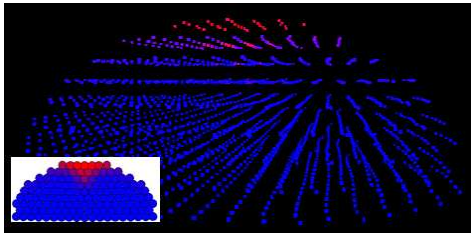
$\hat{T}$,$\tilde{T}^{(1)}\tilde{T}^{(2)}$ - intercellular interactions

$g(x)$ - a sigmoidal function

$\lambda, \tau, \Lambda, h$ - parameters

Simulation of the SAM Model Network

In previous work we have demonstrated that the SAM model network is able to produce expressions mimicking the stem cell region in the SAM.



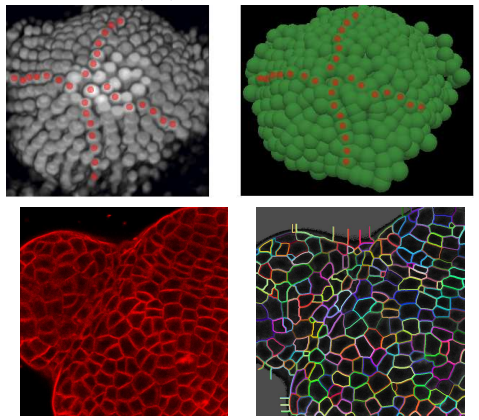
Simulation of a 3D nongrowing SAM of 1765 cells. The final (stable) expression level of the stem cell marking **CLV3** is shown.

Acknowledgements

HJ was in part supported by Knut and Alice Wallenberg Foundation through Swegene. Experimental meristem work was supported by a U.S. NSF grant to EMM. Additional support provided by National Science Foundation Frontiers in Integrative Biological Research program.

Extracting Data from Living Plants

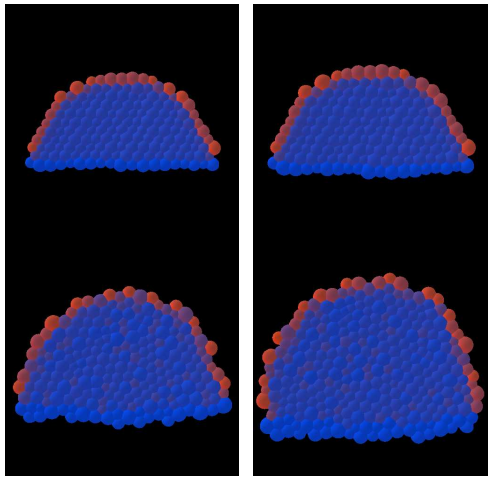
We are developing software for automatic extraction of cell data from a stack of GFP images of living plants. This can be used to create templates of cell positions, cell sizes, and protein/mRNA concentrations at a (sub)cellular level.



Experimental (left) and extracted (right) data.

A Template for L1 Expression

A 2D template of L1 expression in a growing plant is produced by simulation of the L1-part of the SAM model network (Fig. down left, also compare Fig. top right).



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

Inferring Model Parameters

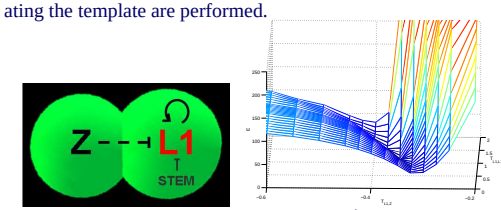
Our main goal is to find gene network interactions by simulation of different networks and comparing with template data. To be able to measure similarity between a simulation and a template, a quality measure is defined as

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

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

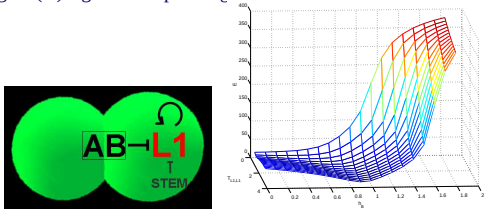
Simulations with parameter values similar to those used when creating the template are performed.



The $T_{L1,L1}$ (self inducing intracellular) and $\hat{T}_{L1,Z}$ (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.



The $T_{L1,L1}$ (self inducing intracellular) and h_B (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-anneal to find good parameter sets.