

Signaling in Multicellular Models of Plant Development

Henrik Jönsson,^{1,3} Bruce E. Shapiro,² Elliot Meyerowitz¹ and Eric Mjolsness⁴

¹Division of Biology and ²Jet Propulsion Laboratory, California Institute of Technology, CA, USA; ³Complex Systems Division, Dept. of Theoretical Physics, Lund University, Sweden; ⁴Department of Information and Computer Science, and Institute of Genomics and Bioinformatics, University of California, Irvine, CA, USA

Abstract

The Shoot Apical Meristem (SAM) of plants is the biological target for a mathematical modeling of developmental systems. The full model incorporates cell growth, proliferation and mechanical interaction, as well as a model for the gene regulatory network (GRN). The GRN-model includes intracellular protein interactions, and also intercellular interactions in various forms. Also transportation of proteins between cells are allowed for. The resulting multicellular model framework is used for mimicking expression domains of known genes in the SAM. Our simulations address the widely discussed regulatory network of the CLAVATA1, CLAVATA3 and WUSCHEL genes, which is important for controlling the development of the SAM and thereby the complete plant. It includes interaction between genes expressed in non-overlapping spatial regions of the SAM, requiring a model where information flows between cells. We show the applicability of the proposed model as a helpful tool to go from current insufficient assumptions about the interactions, to a genetic network which is able to produce gene expression domains *in-silico*, agreeing with experimental data.

1 Introduction

The Shoot Apical Meristem is the source of the complete aboveground part of a plant. *Arabidopsis thaliana* has become a model system for dicot plants [The-Arabidopsis-Genome-Initiative, 2000, Meinke et al., 1998], and it has a SAM constituted of about 10^3 cells. It retains this size and its almost half spherical shape throughout the post embryonic life of the plant. The SAM can be divided into cytologically defined zones where the central zone is at the very apex, the peripheral zone is on the sides, and the rib meristem is in the central parts of the meristem [Meyerowitz, 1997, Steeves and Sussex, 1989]. The central zone is the stem cell domain, the peripheral zone is where new leaf and flower primordia originate, and the rib meristem cells are differentiating into cells contributing to the stem and inner parts of the plant. An interesting property of the SAM is that it is self organizing. For example, if the shoot is divided into two parts, both parts can reorganize into new functional SAMs [Steeves and Sussex, 1989]. In recent years, analysis of expression patterns of important genes have refined the regions of different cell types in the SAM (see e.g. [Bowman and Eshed, 2000]). The mutant phenotypes and expression patterns in plants are used to explore the roles of different genes in the development of the SAM. These experiments are far from exhaustive but parts of a genetic network, controlling the development of the SAM, have been identified [Fletcher et al., 1999, Brand et al., 2000, Schoof et al., 2000].

The modest size and the amount of experimental data of the SAM, makes it an appropriate candidate for developmental modeling. The focus of this chapter will be on interactions between the CLAVATA1 (CLV1), CLAVATA3 (CLV3), and WUSCHEL (WUS) genes [Fletcher et al., 1999, Brand et al., 2000, Schoof et al., 2000], which is widely discussed in the recent literature. Since these genes are expressed in non-overlapping spatial domains of the SAM, a model incorporating intercellular interactions and signaling is required. We here extend a multicellular model framework [Shapiro and Mjolsness, 2001, Mjolsness et al., 2002] to incorporate a more thorough description of intercellular protein signaling. This allows for simulations of the WUSCHEL/CLAVATA network in the SAM which *in-silico* is able to reproduce gene expression domains from experiments on wild type plants.

We argue and show that a model framework with interaction and signaling between cells, such as the one we describe, is a powerful tool to investigate the behavior of various parts of a developmental system. The model can be used to falsify or discard hypotheses, and also to introduce and test new ones.

2 The Shoot Apical Meristem

All cells of the aboveground part of a plant originate from a group of stem cells found at the growing tip of the shoot called the shoot apical meristem. The SAM forms during plant embryogenesis, and after seed germination it remains a collection of undifferentiated cells approximately uniform in shape and size. At the same time it provides at its flanks the cells that will become lateral structures (leaves with attendant second-order meristems and flowers), and at its base the cells that will make the stem, including pith and vasculature. Thus, through the life of the plant the addition of cells to the meristem by cell division,

and the departure of cells to form differentiated structures, must be closely balanced. Furthermore, the pattern of cell divisions must be highly controlled, to maintain the uniform meristematic shape and to provide flanking structures in appropriate positions (e.g. spiral phyllotaxis).

The interactions between the genes CLV1, CLV3 and WUS are proposed as a possible explanation of how plants control the size of the stem cell region in the shoot [Bowman and Eshed, 2000, Weigel and Jurgens, 2002]. The expression domains and interactions of the CLV1, CLV3 and WUS genes are suggested from a number of experiments [Fletcher et al., 1999, Brand et al., 2000, Schoof et al., 2000], and are illustrated in figure 1. The WUSCHEL protein, which is a homeodomain transcription factor, upregulates the expressions of the CLV1 and CLV3 genes. On the other hand, CLV3 encodes a small, putative extracellular protein and CLV1 a receptor kinase protein which both act in a network down-regulating the WUS expression. This feedback network is proposed to regulate the size of the stem cell region of plants. An increased WUSCHEL domain will generate larger CLV1/3 regions which in turn will repress the WUS region. As can be seen in figure 1 the expression domains are not overlapping spatially. The stem cell marking CLV3 is expressed at the very apex, while CLV1 and WUS are expressed in a central region below, sometimes referred to as an “organizing center” [Bowman and Eshed, 2000].

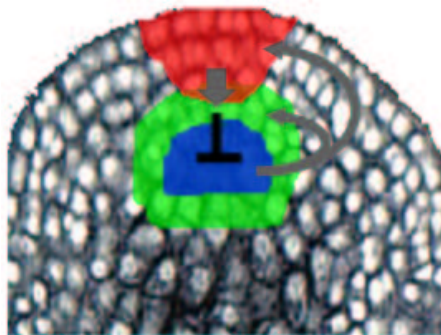


Figure 1: Illustration of suggested expression patterns of the CLAVATA1,3 and WUSCHEL genes in the SAM. Blue corresponds to WUS, green to CLV1 and red to CLV3. The arrows represent the interactions suggested by experiments [Fletcher et al., 1999, Brand et al., 2000, Schoof et al., 2000, Bowman and Eshed, 2000].

3 Modeling the SAM

Although the interaction network described in the previous section is widely discussed in the literature, some questions that arise with its assumptions are not. For example, if WUS is a positive regulator of CLV3, why is it that it only upregulates CLV3 in a region above it’s own expression domain? Why not within it, or symmetrically around it? These kinds of questions are particularly addressable using a modeling framework. In *in-silico* models it is possible to include known data in an expanded network to create expression patterns of the known genes.

Figure 2 shows a hypothesized network that includes the CLV1,3 and WUS genes, as well as yet unidentified genes. Marked in the red box is the network implied by experiments (cf. figure 1). The

added gene X is not yet identified, but experiments indicates that the WUSCHEL protein does not move between cells [Lenhard and Laux, 2001], requiring some other protein/molecule transferring the information. The X in the network could actually represent a larger subnetwork of genes that in the end upregulates both CLV1 and CLV3. The important property for the model to succeed is that a “signal” originates from the WUSCHEL region, that it decreases with distance from the region, and that it upregulates CLV1 and CLV3. In the proposed network we have implemented X as a protein which can move between cells by diffusion. For CLV1, X is the only input and CLV1 is expressed if and only if the concentration of X is above a threshold value within the cell.

The main hypothesis introduced in the network model is shown in the blue box in figure 2. It suggests how WUS can be involved in promoting the CLV3 domain. The idea is that there is a second gene, $L1$, involved which is expressed only in the surface layer (L1) of cells in the SAM. The protein Y represents a signal which originates from the $L1$ -expressing cells and diffuses into the SAM. The CLV3 gene is turned on only if the total signal (X and Y) exceeds a threshold value in the cell.

A second hypothesis, also included in the network, is marked with a green box in figure 2. This part of the network is used to create the L1-specific expression pattern for the $L1$ gene, regardless of the initial expression patterns of the included genes. Z is a gene which is expressed all over the SAM.

Since only the cells of the SAM is included in the simulations, the lowest layer of cells represents the start of the stem in the plant. The boundary is implemented as a gene expressed only in these lowest layer cells ($STEM$). Note also that the suggested repression from CLV1,3 on WUS is not yet included.

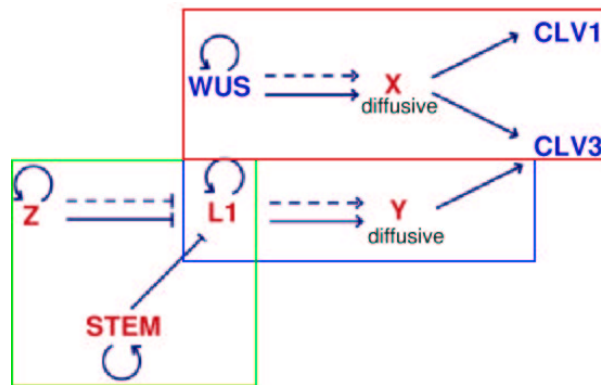


Figure 2: Hypothesised network for developmental control in the plant shoot. Solid line arrows represents intracellular interactions and dashed represent intercellular.

In figure 3, the results of a non-growth simulation of the SAM is presented. The final (stable) expression pattern for some of the genes in the network are shown.

The simulation starts with an initial concentration of WUSCHEL protein in the cells shown in figure 3a, analogous to the known WUS domain. There is no external input to the WUS gene, and the expression stays on in these cells throughout the simulation. The stem indicating protein is initiated with a nonzero protein concentration in the bottom cell layer. This gene also stays on in initiated cells

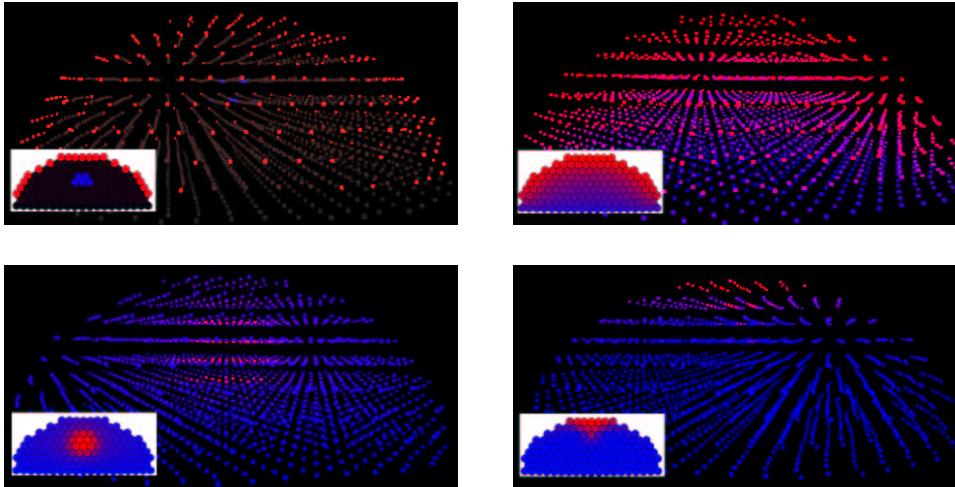


Figure 3: Results from an static SAM simulation including 1765 cells. The cells are colorcoded by protein concentration values. The WUSCHEL region is initiated and the system is simulated until the expression patterns are stabilized in all cells. (a) WUS (blue) and a L1-specific gene (red). (b) The L1 diffusive signal Y (red is high protein concentration, blue is low). (c) CLV1 (colors as in b). (d) CLV3 (colors as in b).

only throughout the simulation. All other proteins are initiated with zero initial concentration in all cells, and the expression patterns shown in figure 3 are the stable self-organizing expression domains for the model network.

Figure 3a shows the final expression of the *L1* gene in red. The expression is restricted to the cells of the surface layer. From the *L1* expressing cells there is a diffusive signal sent into the inner regions of the SAM which is shown in the concentration value of the *Y* protein in figure 3b.

Figure 3c and d shows the resulting expression patterns for the CLV1 and CLV3 genes. Both show close resemblance with the domains found in real SAMs.

4 Discussion

It is indeed encouraging that the proposed network actually produces a stable configuration where the CLV1 and CLV3 expression domains agree with experimental data. The model shows a self organization of the expression domains with the only requirement that the WUSCHEL gene is expressed in its known region.

A large portion of the network consists of hypothesized components. Some of these may correspond to real genes. There are for example an L1-sharp expressing gene (ATML1, [Lu et al., 1996]) as well as a diffusely L1-peaked gene (ACR4, [Tanaka et al., 2002]), which might be analogous to the *L1* and *Y* genes respectively. The homeodomain transcription factor ATML1 is also thought to bind to its own promoter. This is an essential component in the submodel corresponding to our second hypothesis; the network that automatically generates a L1-specific gene. ATML1 binds to the L1-box motif [Abe et al., 2001], which appears in a number of L1-specific gene promoters, including its own. Hence, it is a good target for

mutation experiments, testing our hypotheses, and we have initiated a discussion to create experiments for this.

The proposed network can explain experimental data, but it remains speculative. The more important contribution is to show how the introduced model framework is valuable as a tool when evaluating a developmental system. We have used it to suggest possible explanations of experimental data in the SAM, including hypotheses open for verification. In the model framework, a gene is only an extra variable, and an interaction is a parameter in the differential equations. The simulations presented here need only a couple of minutes to finish on a typical personal computer (1GHz AMD processor). Hence, the framework provides a tractable method to refine new hypotheses *in-silico* for future verification *in-vivo*.

5 Methods

The Generic Model

The central part of the model is the proteins and the regulation of their production. The proteins are implemented by concentration values in each cell, and the dynamics are described by a neural-network inspired genetic regulatory network (GRN) model [Mjolsness et al., 1991, Marnellos and Mjolsness, 1998]

$$\tau_a \dot{v}_a^{(i)} = g(u_a^{(i)} + h_a) - \lambda_a v_a^{(i)} + \gamma_a \dot{v}_{a,ext}^{(i)}, \quad (1)$$

where

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

$$g(x) = \frac{1}{2} \left(1 + \frac{x}{\sqrt{1+x^2}} \right). \quad (3)$$

In equations (1-3), the v 's are the protein concentrations, a, b are indices for the proteins and i, j are indices for the cells. The model describes intracellular interactions encoded by the T matrix, as well as intercellular interactions between proteins of neighboring cells. The intercellular part contains a direct interaction (\hat{T}), and also a ligand-receptor type of interaction ($\tilde{T}^{(1)}\tilde{T}^{(2)}$). In our simulation, only the direct interaction term is used, but the ligand-receptor interaction would be the suitable for the CLV1/3 signal, omitted here. The λ term is a degradation term, τ is a time parameter, and h is a parameter regulating the basal expression level for a gene. Λ describes the connection between cells. The addition of the γ -term represents contributions from other kinds of processes. The simulation presented in this paper uses diffusion as a model for transferring a molecular signal between cells, which defines the γ -term in this case.

The cells are approximated to have spherical shapes. A mechanical interaction between cells is described by a spring potential between each pair of neighboring cells with a relaxing distance equal to the sum of the radii. The interaction is softly truncated for larger distances, such that there is no interaction between cells that are not neighbors.

Although the simulation presented here only represents a non-growing model, the framework does include cell growth and proliferation [Shapiro and Mjolsness, 2001]. Growth is implemented as being

radially symmetric and the rate may be a function of both the mass (size) of the cell, and the protein concentrations within the cell. Different models of the cell-cycle are implemented, where the simplest implementation uses only the cell size, and the cells divide when the mass reaches a threshold value. A more biologically interesting implementation uses the Goldbeter model [Goldbeter, 1991], and the period can be tuned by a binding of proteins to the cyclin protein as proposed in [Gardner et al., 1998]. Using this model, growth is decoupled from the cell proliferation. At a cell division the total mass is conserved. In the first step of the division, two smaller spheres are created and placed partly on top of each other and then the action of the spring force moves the new cells apart, in a direction dependent on interactions with neighboring cells.

Simulation

The model framework is implemented both in C/C++ and in Cellerator [Shapiro and Mjolsness, 2001, Shapiro et al., 2002], which is a Mathematica package. The simulation presented is from the C/C++ implementation of a nongrowing SAM. The static solution is found by integrating the time differential equations until a stable configuration is found. The differential equation solver is a 5th order Runge-Kutta using adaptive step sizes, based on the *odeint* function from Numerical Recipes [Press et al., 1992]. The simulations have been performed on a PC running Linux.

In the simulation there are many parameters, the most important of which are the parameters in the GRN-equations (1-2). Experiments only suggest whether interactions are up- or down-regulating, and no rates are given. Hence, the parameters are tuned by hand. More than one set of parameters result in the behavior shown in this paper.

6 Acknowledgments

Thanks to Barabara Wold, Marcus Heisler and Venugopala Reddy, for inspiring discussions. The research described in this chapter was carried out, in part, by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the national Aeronautics and Space Administration. Further support came from the California Institute of Technology President's Fund. HJ was in part supported by Knut and Alice Wallenberg Foundation through the Swegene consortium.

References

- [Abe et al., 2001] Abe, M., Takahashi, T., and Komeda, Y. (2001). Identification of a cis-regulatory element for l1 layer-specific gene expression, which is targeted by an l1-specific homeodomain protein. *The Plant Journal*, 26(5):487–494.
- [Bowman and Eshed, 2000] Bowman, J. L. and Eshed, Y. (2000). Formation and maintenance of the shoot apical meristem. *Trends in Plant Science*, 5(3):110–115.

- [Brand et al., 2000] Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M., and Simon, R. (2000). Dependence of stem cell fate in arabidopsis on a feedback loop regulated by *clv3* activity. *Science*, 289:617–619.
- [Fletcher et al., 1999] Fletcher, J. C., Brand, U., Running, M. P., Simon, R., and Meyerowitz, E. M. (1999). Signaling of cell fate decisions by *clavata3* in *arabidopsis* shoot meristems. *Science*, 283:1911–1914.
- [Gardner et al., 1998] Gardner, T. S., Dolnik, M., and Collins, J. J. (1998). A theory for controlling cell cycle dynamics using a reversibly binding inhibitor. *Proc. Natl. Acad. Sci. USA*, 95:14190–14195.
- [Goldbeter, 1991] Goldbeter, A. (1991). A minimal cascade model for the mitotic oscillator involving cyclin and *cdc2* kinase. *Proc. Natl. Acad. Sci. USA*, 88:9107–9111.
- [Lenhard and Laux, 2001] Lenhard, M. and Laux, T. (2001). Role of *wuschel* in specifying stem cell identity in the arabidopsis shoot apical meristem. In *Abstractbook of the 12th International Conference on Arabidopsis Research*.
- [Lu et al., 1996] Lu, P., Porat, R., Nadeau, J. A., and O’Neill, S. D. (1996). Identification of a meristem 11 layer-specific gene in arabidopsis that is expressed durin embryonic pattern formation and defines a new class of homeobox genes. *The Plant Cell*, 8:2155–2168.
- [Marnellos and Mjolsness, 1998] Marnellos, G. and Mjolsness, E. D. (1998). A gene network approach to modeling early neurogenesis in drosophila. In *Pacific Symposium on Biocomputing ’98*, pages 30–41. World Scientific.
- [Meinke et al., 1998] Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D., and Koornneef, M. (1998). *Arabidopsis thaliana*: a model plant for genome analysis. *Science*, 282:678–682.
- [Meyerowitz, 1997] Meyerowitz, E. M. (1997). Genetic control in cell division patterns in developing plants. *Cell*, 88:299–308.
- [Mjolsness et al., 1991] Mjolsness, E., Sharp, D. H., and Reinitz, J. (1991). A connectionist model of development. *Journal of Theoretical Biology*, 152:429–454.
- [Mjolsness et al., 2002] Mjolsness, E. D., Jönsson, H., and Shapiro, B. E. (2002). Modeling plant development with gene regulation networks including signaling and cell division. In *Proceedings of the Third International Conference on Bioinformatics of Genome Regulation and Structure (BGRS’2002)*.
- [Press et al., 1992] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C The Art of Scientific Computing*. Cambridge University Press, New York.
- [Schoof et al., 2000] Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Jurgens, G., and Laux, T. (2000). The stem cell population of arabidopsis shoot meristems is maintained by a regulatory loop between the *clavata* and *wuschel* genes. *Cell*, 100:635–644.

- [Shapiro et al., 2002] Shapiro, B. E., Levchenko, A., Meyerowitz, E., Wold, B., and Mjolsness, E. D. (2002). Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. *Bioinformatics*, (accepted for publication).
- [Shapiro and Mjolsness, 2001] Shapiro, B. E. and Mjolsness, E. D. (2001). Developmental simulation with cellerator. In *Proceedings of the Second International Conference on Systems Biology*, pages 342–351.
- [Steeves and Sussex, 1989] Steeves, T. and Sussex, I. (1989). *Patterns in Plant Development*. Cambridge University Press, New York.
- [Tanaka et al., 2002] Tanaka, H., Watanabe, M., Watanabe, D., Tanaka, T., Machiada, C., and Machida, Y. (2002). Acr4, a putative receptor kinase gene of *arabidopsis thaliana*, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. *Plant Cell Physiology*, 43(4):419–428.
- [The-Arabidopsis-Genome-Initiative, 2000] The-Arabidopsis-Genome-Initiative (2000). Analysis of the genome sequence of the flowering plant *arabidopsis thaliana*. *Nature*, 408:796–815.
- [Weigel and Jurgens, 2002] Weigel, D. and Jurgens, G. (2002). Stem cells that make stems. *Nature*, 415:751–754.