# General Theory of Genotype to Phenotype Mapping: Derivation of Epigenetic Landscapes from N-Node Complex Gene Regulatory Networks

Carlos Villarreal,[1,4] Pablo Padilla-Longoria,[2,4] and Elena R. Alvarez-Buylla[3,4,*]

[1]*Instituto de Física, Universidad Nacional Autónoma de México, D.F. México, Mexico*

[2]*Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, D.F. México, Mexico*

[3]*Instituto de Ecología, Universidad Nacional Autónoma de México, D.F. México, Mexico*

[4]*Centro de Ciencias de la Complejidad, C3, Universidad Nacional Autónoma de México, D.F. México, Mexico*

(Received 29 November 2011; published 12 September 2012)

We propose a systematic methodology to construct a probabilistic epigenetic landscape of cell-fate attainment associated with $N$-node Boolean genetic regulatory networks. The general derivation proposed here is exemplified with an *Arabidopsis thaliana* network underlying floral organ determination grounded on qualitative experimental data.

DOI: 10.1103/PhysRevLett.109.118102 PACS numbers: 87.18.−h, 87.16.Yc

Understanding how the information contained in genes is mapped onto the organisms' traits, or the phenotypes, implies deriving a formal framework about the generic aspects of developmental constraints on biological evolution. This remains one of the main challenges of theoretical physical approaches to systems biology [1]. Crucial aspects of such mapping depend upon the topology and dynamics of gene regulatory networks (GRNs) [2]. The morphogenetic constraints derived from gene interactions in such complex GRNs are reflected in what C. H. Waddington called in 1957 the epigenetic landscape (EL) [3]. Waddington initially proposed the metaphor of an EL as an image of morphogenesis in which a developmental process is represented by a ball rolling down along a landscape with peaks and valleys. In the postgenomic era it is now possible to pursue formal and quantitative derivations of ELs that may be validated for experimentally grounded GRNs. In the EL, the steady-state multigenic profiles or attractors are found at the bottom of the basins, and these profiles characterize each cell type or cellular condition [2], with as many dimensions in the landscape as the number of genes considered in the underlying GRN. The relative size, shape, and distribution of the basins of attraction in the $N$-dimensional hyperspace of gene configurations restrict the patterns with which the system will transit from one attractor to another in time and space. It is assumed that random fluctuations move the ball along different pathways until it reaches the bottom of a valley, a final developmental or cell state. The need for mathematical models of attractors and ELs associated to cellular networks has been pointed out as one of the main outstanding issues in systems biology and medicine [4,5]. Recent efforts have provided mathematical approaches to analyze small gene circuits as dynamical systems. For example, Wang and collaborators [6] have computed a quasipotential landscape of the stochastic dynamics of a well-studied two-gene case. However, simulation or analytical approaches to derive and quantify EL for larger networks that comprise developmental modules are still lacking. In this Letter, we propose a method to construct a probabilistic EL by calculating the probability distribution of stable gene expression configurations arising from the topology of a general $N$-node GRN. The core of the method consists of two steps: first, the derivation of a stochastic continuous system from a Boolean GRN, and second, the definition of an analytically solvable Fokker-Planck (FP) equation. This method is in principle applicable to any finite-dimensional GRN. In this Letter, we first put forward the general approach to derive the EL, and then we apply it to the *Arabidopsis thaliana* floral organ determination GRN, which has been studied from alternative perspectives before [7–9].

*Derivation of a stochastic continuous system from a Boolean GRN.*—Consider a Boolean GRN defined by a group of $N$ interacting nodes (genes, proteins, or other type of molecule) and logical input functions, which formalize experimental data on gene interactions. The state of gene $i$ at a given time is described by discrete logical relations $q_i(t+1) = W_i[q_1(t), \ldots, q_n(t)]$. Here, the quantities $q_i(t)$ are dichotomic variables representing the expression level of node $i$, given as 0 or 1, while the input $W_i$ is determined by logical propositions with a Boolean structure. According to Kauffman's proposal, the equilibrium states given by the fixed-point condition $q_i(t+1) = q_i(t) \equiv q_i^s$ correspond to the phenotypes or cell types determined by a GRN.

In order to derive the stochastic continuous system, we first need to describe the Boolean GRN as a continuous system. Several approaches have been used for this purpose [7,10–12]. For that sake, in this Letter, we replace the original set $\{q_i\}$ in terms of a set of continuous variables $\tilde{q}_i$ $(0 \leq \tilde{q}_i \leq 1)$ which satisfies Zadeh's rules [13] of fuzzy propositional calculus: $q_i$ or $q_k \rightarrow \max[\tilde{q}_i, \tilde{q}_k]$, $q_i$ and $q_k \rightarrow \min[\tilde{q}_i \tilde{q}_k]$, not $q_i \rightarrow 1 - \tilde{q}_i$. It is straightforward to

 

show for $\tilde{q}_i \simeq 0$ or 1, Zadeh's algorithms define a Boolean algebra [13]. Using these replacements, for each sate variable $i$ (e.g., gene), we derive a function $w_i(\tilde{\mathbf{q}})$ which expresses with continuous variables the way its regulators interact. For simplicity, in the following we omit the tilde over $q_i$. We provide a dichotomic structure to the continuous GRN inputs by introducing steplike (differentiable) activation functions

$$\Theta[w_i] = \frac{1}{\exp[-2b(w_i(\mathbf{q}) - w_i^{\mathrm{thr}})] + 1}, \tag{1}$$

where $w_i$ is the input function for node $i$, $w_i^{\mathrm{thr}}$ is a threshold level, and $b$ the input saturation rate. In particular, for $b \gg 1$, $\Theta[w_i]$ becomes a Heaviside step function.

The stochastic continuous system can then be described by means of Langevin equations [14,15]

$$\frac{dq_i}{dt} = \Theta[w_i(\mathbf{q})] - \alpha_i q_i + \xi_i. \tag{2}$$

Here, $\Theta[w_i]$ represents activation of node $i$, $\alpha_i$ its relaxation rate, and $\xi_i$ is a random variable with ensemble average $\langle \xi_i(t) \rangle = 0$, and time-dependent correlation $\langle \xi_i(t)\xi_k(t') \rangle = Q_{ik}\delta(t - t')$. We assume that $Q_{ik} = Q\delta_{ik}$, so that fluctuations pertaining to different variables are uncorrelated.

In absence of fluctuations, the steady states of the continuous GRN are determined by the condition $dq_i/dt = 0$, leading to $q_i^s = \Theta[w_i]/\alpha_i$, so that $q_i^s = 0$ or $1/\alpha_i$, depending on the specific set of initial conditions $q_i(0)$. Because the inputs $w_i$ satisfy a Boolean algebra, the attractor sets obtained in the continuous approach result in values equivalent to those derived in the discrete logical approach (they coincide for $\alpha_i \to 1$), with the possible exception of those arising from $q_i^s = 1/2$ (see Ref. [16]). By introducing the noise term again in the dynamics, the formal solution of (2) is

$$q_i(t) = q_i(0)e^{-\alpha_i t} + \int_0^t dt' e^{-\alpha_i(t-t')}\Theta[w_i(\mathbf{q}(t'))] + \zeta_i(t) \tag{3}$$

with $\zeta_i(t) = \int_0^t dt' e^{-\alpha_i(t-t')}\xi_i(t')$. It follows that the state variables may be expressed as $q_i = \langle q_i \rangle + \zeta_i$, where the ensemble average

$$\langle q_i(t) \rangle = q_i(0)e^{-\alpha_i t} + \int_0^t dt' e^{-\alpha_i(t-t')}\langle \Theta[w_i(\mathbf{q}(t'))] \rangle. \tag{4}$$

A Taylor expansion shows that in the small-noise limit $\langle \Theta[w_i(\mathbf{q})] \rangle \simeq \Theta[w_i(\langle \mathbf{q} \rangle)]$ [up to terms $\mathcal{O}(\zeta^2)$], which we introduce in (4). We now employ Haken's adiabatic hypothesis [15], which involves considering that the state variables react instantaneously to the orders dictated by $\Theta[w_i(\langle \mathbf{q}(t) \rangle)]$ at the actual time $t$, not depending on the past history of the system. In that case the function $\Theta[w_i]$ may be extracted out from the integral in (4), which leads to

$$\langle q_i(t) \rangle \simeq q_i(0)e^{-\alpha_i t} + \frac{1}{\alpha_i}\Theta[w_i(\langle \mathbf{q}(t) \rangle)](1 - e^{-\alpha_i t}). \tag{5}$$

The adiabatic approximation has been checked numerically for Eq. (4) with reasonable accuracy. We observe that for $t \gg 1/\alpha_i$, $\langle q_i \rangle = q_i^s = \Theta[w_i(\langle \mathbf{q}^s \rangle)]/\alpha_i$. Following similar steps, it may be shown that the mean square fluctuations $\langle q_i^2(t) \rangle - \langle q_i(t) \rangle^2 \simeq Q(1 - e^{-2\alpha_i t})/\alpha_i$.

*The FP equation.*—To study the temporal evolution of the probability distribution of the stochastic continuous system, we introduce the FP equation $\partial p(\mathbf{q}, t)/\partial t = -\nabla_q \cdot \mathbf{J}$, where $p(\mathbf{q}, t)$ denotes the probability distribution for genetic expression, and

$$J_i = \{\Theta[w_i(\mathbf{q})] - \alpha_i q_i\}p(\mathbf{q}, t) - \frac{1}{2}\sum_{k=1}^{N} Q_{ik}\frac{\partial p(\mathbf{q}, t)}{\partial q_k}, \tag{6}$$

the probability flux within the $N$-dimensional configuration space. The continuous representation of the input functions (1) enables the use of a logistic activation function that is constant at equilibrium, thus yielding an effectively linear drift. An analytical dynamical solution of the FP equation can be found in the present case with a linear drift vector and constant diffusion tensor; the solution hence acquires a Gaussian structure:

$$p(\mathbf{q}, t) = \mathcal{N}(t)\exp\Big[-\sum_{i=1}^{N}\frac{(q_i - \langle q_i(t) \rangle)^2}{a_i(t)}\Big], \tag{7}$$

where $\langle q_i(t) \rangle$ is given by (4), the distribution width

$$a_i(t) = \frac{Q}{\alpha_i}(1 - \exp[-2\alpha_i t]) + a_i(0)\exp[-2\alpha_i t] \tag{8}$$

with initial value $a_i(0)$, and the normalization function $\mathcal{N}(t) = \pi^{-N/2}[\Pi_{i=1}^{N} a_i(t)]^{-1/2}$. Notice that for $a_i(0) \to 0$, $a_i(t)$ coincides with the mean square fluctuations arising from the Langevin approach, whereas for $t \gg 1/\alpha_i$ $a_i(t) \to a_i^s = Q/\alpha_i$.

*Attractor transitions in the EL.*—In the spirit of Waddington's original metaphor we characterize the probabilistic transit through the EL by constructing the transition probability distribution between neighboring basins of attraction: $n \to m$. We assume that the system is initially described by the probability distribution $\Pi_{i=1}^{N}\delta[q_i - q_i(0)]$, with negligible dispersion $a_i(0) \to 0$, and the point $q_i(0)$ located at a basin $n$ centered at $q_i^{s(n)}$. Because of the stochastic dynamics, the system spreads over the basin $n$ and subsequently transits to a basin centered at $q_i^{s(m)}$. With the distribution (7) we can derive the transition probability distribution $p^{nm} = \mathcal{N}^{nm}\exp[\sum_i (q_i - \langle q_i \rangle^{nm})^2/a_i^{nm}]$, where $\langle q_i \rangle^{(nm)} = q_i^{s(n)}e^{-\alpha_i t} + q_i^{s(m)}(1 - e^{-\alpha_i t})$, the width $a_i^{nm} = Q(1 - \exp[-2\alpha_i t])/\alpha_i + a_i^{s(n)}\exp[-2\alpha_i t]$, and the normalization $\mathcal{N}^{nm} = \pi^{-N/2}[\Pi_i a_i^{nm}]^{-1/2}$. We finally express the EL probability distribution as a superposition

$p_n(\mathbf{q}, t) = \sum_m g_m p^{nm}(\mathbf{q}, t)$, where $n$ denotes the starting EL basin, and $g_m$ the probability to occupy the basin $m$. This last probability is estimated as the relative size of the basins $g_n = \Omega_n/\Omega$. where $\Omega_n$ correspond to the number of alternative sets of initial conditions leading to a given attractor $q_i^{s(r)}$, and $\Omega = 2^N$ to the total number of initial conditions of the original Boolean GRN.

This derivation provides a novel analytical dynamical solution of the GRN stochastic dynamics that yields explicit expressions for the temporal evolution of the probability distribution of gene states, and for the transition probability distribution between neighboring attraction basins. We refer to such derivations as the probabilistic EL, as they allow us to explore the temporal evolution of the probabilities of the cell being in each cell state (e.g., gene configuration), and the probability of transition to another cell state, given an initial state. One of the virtues of this expression is that its Gaussian structure enables us to derive a reduced $N - D$-dimensional Gaussian distributions by integrating (7) over $D$-state variables. As a consequence, we also have a method which allows us to focus the probabilistic description on a subset of the more relevant state variables. The latter may be chosen based on dynamical or biological considerations, for example, the GRN nodes are known to have a strong influence on the phenotype when mutated, or either slow modes, i.e., variables with the lowest relaxation rates which have been shown to drive the emergent dynamics of self-organized systems [15].

*Example: Arabidopsis thaliana flower organ identity GRN.*— These tools may be applied to well-defined GRN consisting of a set of genes and a defined Boolean input function for each genes' regulators. In particular, we study the GRN dynamics leading to floral organ determination in *Arabidopsis thaliana*. The main features of floral development in *Arabidopsis* and its underlying 15-node GRN are summarized in Fig. 1 (adapted from Ref. [7]); they are thoroughly described in Ref. [17]. An updated table for logical expressions for GRN interactions is presented in the Supplemental Material [18]; this network definition is the starting point. From the logical expressions for the network, it is straightforward to replace each state variable using Zadeh's rules, and derive the continuous function $w_i(\tilde{\mathbf{q}})$ as described above. Given this transformation and the specification of parameters $Q$ and $\alpha_i$, we can use the analytical expressions derived above to follow the temporal evolution of the EL probability distribution from an initial configuration. For this case, we focused the probabilistic description on a small enough subset of the state variables so that we could provide a graphical representation of the temporal evolution. Following our method, this is done by integrating (7) over those $D$-state variables which will not be the focus of the description. For this specific case, we chose the variables of interest based on a dynamical consideration. Specifically, we first derived the

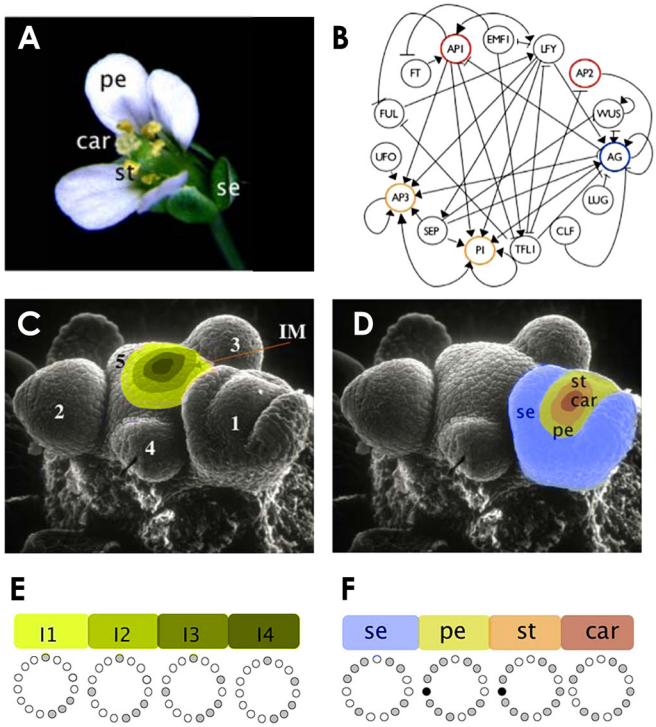


FIG. 1 (color online). Flower development and gene network underlying primordial floral organ cell-fate determination in *Arabidopsis thaliana* [7,17]. (a) Mature flower of wild-type *Arabidopsis*. (b) GRN topology. Circles correspond to genes, while arrows and bars correspond to positive and negative regulatory interactions, respectively. (c) Inflorescence meristem (IM). Four regions denoted as I1, I2, I3, and I4 from outer to inner circles can be distinguished. Flower meristems arise from the flanks of the IM (1, oldest; 5, youngest). (d) Young flower meristems are subdivided into four concentric regions that will eventually develop into mature flower organs: sepals (se), petals (pe), stamens (st), and carpels (car). (e) and (f) GRN attractors for IM and FM meristems, respectively. Expressed and nonexpressed genes are represented in by gray circles and white circles, while black circles represent a gene (UFO) that can be either expressed or not expressed.

attractor sets by applying Boolean algebra rules, and then, based on these sets, we eliminated those variables which do not play a central role in the GRN. The analytical derivation of the attractors is presented in the Supplemental Material [18]. The attractors and basins could also be derived with numerical simulations. The final expressions for the GRN attractors are tabulated in Table I, which indicate that the gene expression patterns for the inflorescence meristem (IM) and the floral meristem (FM) in *Arabidopsis* are uniquely determined by only five transcription factors: LFY, UFO, WUS, AP3, and AG. According to Table I, the gene configuration space is split into two quasi-independent subspaces depending on the values LFY = 0 or LFY = 1, so that LFY becomes the central regulator of the GRN. The two subsets are characterized by the expression values of UFO and WUS, and

TABLE I. Steady-state values of the gene regulatory network (GRN) as given by the Boolean logical rules. The explicit value of the steady states are given by substituting at the right-hand side the possible input values (0 or 1) of the GRN-state variables. Boolean expressions are recovered from the substitution $\max[q_i, q_k] \to q_i$ or $q_k$, $\min[q_i, q_k] \to q_i$ and $q_k$, $1 - q_i \to$ not $q_i$.

| | |
|---|---|
| LFY = LFY | |
| FUL = min[AG, LFY] | FT = LFY |
| AP1 = min[1 − AG, LFY] | AP2 = LFY |
| WUS = min[WUS, 1 − AG] | TFL1 = 1 − LFY |
| AG = min[LFY, max[AG, WUS]] | EMF1 = 1 − LFY |
| PI = min[LFY, max[AG, AP3]] | SEP = LFY |
| AP3 = min[LFY, max[AP3, UFO]] | UFO = UFO |
| CLF = 1 | LUG = 1 |

AP3 and AG, respectively. For LFY = 0, UFO and WUS induce four IM attractors depicted in Fig. 1: I1, I2, I3, and I4. Similarly, for LFY = 1, AP3 and AG induce four FM attractors: se, pe2, st2, and car, for UFO = 0, and two additional attractors, pe1 and st1, for UFO = 1. Notice that for a fixed expression of UFO, the four floral organs are completely specified by the $2^2$ different combinations of the expression values of AG and AP3. It may be verified that these and other different predictions inferred from Table I show an excellent correspondence with experimental observations on *Arabidopsis* compiled in Ref. [17].

The former results allow us to consider the reduced probability distributions on the 2D subspaces expanded by UFO and WUS, and AP3 and AG, and eliminate the other variables by integration. We assigned a small level of noise for $Q \approx 10^{-2}$, while for the decay rates of UFO and WUS we assumed (in absence of experimental information) that $\alpha_{\mathrm{UFO}} = \alpha_{\mathrm{WUS}} = 1$; concerning AP3 and AG, experimental measurements indicate that $\alpha_{\mathrm{AP3}} > \alpha_{\mathrm{AG}}$ [19]. This latter consideration induces asynchronous transitions between the EL attraction basins and thus defines a temporal ordering for floral organ attainment.

We present in Fig. 2 the most probable temporal dynamics associated to floral expression under the assumption that the distribution function is originally set at the sepal

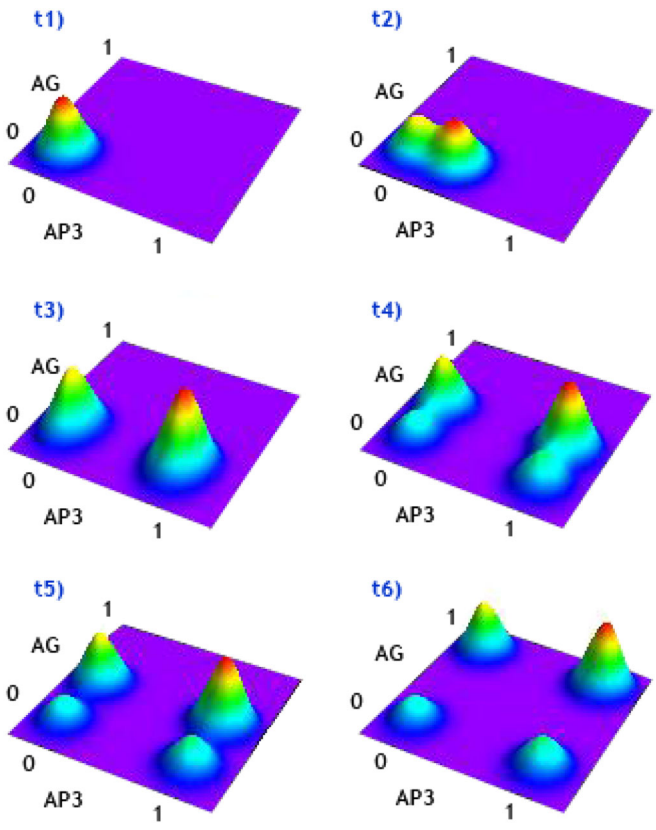


FIG. 2 (color online). Temporal evolution ($t_1 \to t_6$) of the EL probability distribution on the floral organ domain expanded by AP3 and AG. The system is originally set at the sepal basin (AP3 = 0, AG = 0). Afterwards, the system transits, first, to the petal basin (AP3 = 1, AG = 0), and subsequently to the stamen (AP3 = 1, AG = 1) and carpel basins (AP3 = 0, AG = 1), as empirically observed in *Arabidopsis thaliana* flower development (see Ref. [7]).

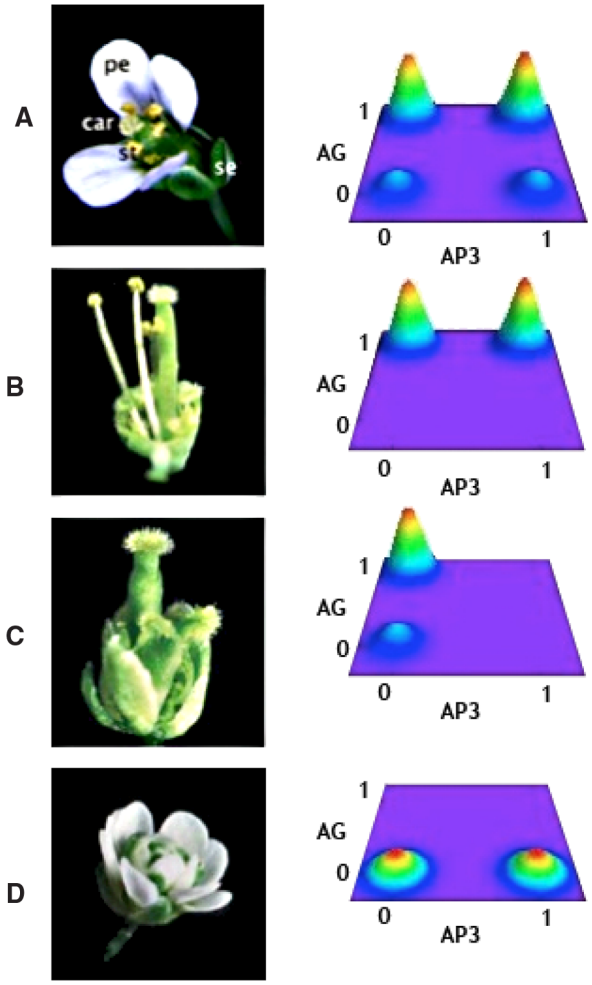


FIG. 3 (color online). Probabilistic ELs for wild-type (a) and loss-of-function mutants: AP1 = 0 (b), AP3 = 0 (c), and AG = 0 (d).

basin. We observe that, starting at the sepal basin, the system first transits to the petal basin (even though it is smaller than those of stamens and carpels), and subsequently, to the stamen and carpel basins, in agreement with the temporal development pattern of floral organs attainment in *A. thaliana* and most flowering species. This temporal ordering is insensitive to the specific values of $\alpha_i$ as far as the inequality $\alpha_{\mathrm{AP3}} > \alpha_{\mathrm{AG}}$ holds. The temporal order in flowering had been previously addressed in Ref. [7] using an alternative scheme based on discrete-time Markov chains.

We have further validated the model by exploring the EL associated to loss-of-function mutants of the so-called ABC model (see Ref. [7] for details). Mutations are simulated by setting AP1 = 0 (A mutant), AP3 = 0 (B mutant), and AG = 0 (C mutant) in the logical rules appearing in Table I. We present in Fig. 3 the probabilistic ELs with corresponding images of the wild-type and loss-of-function mutants of *Arabidopsis*. In addition to validating our model, such mutant analysis provides an explanation for a morphogenetic constraint observed in ABC mutants, namely, those pairs of contiguous verticils that are transformed in homeotic ABC mutants.

In conclusion, we have put forward an analytical derivation of the probabilistic EL for an arbitrary $N$-dimensional GRN grounded on experimental data. We have exemplified our method with *Arabidopsis thaliana* floral organ GRN. The methods proposed here successfully recovered the steady-state gene configurations characteristic of primordial cells of each floral organ type in wild-type and ABC mutants. In addition, we recovered the observed temporal pattern with which these configurations are attained in real flowers.

The authors acknowledge the valuable discussions and comments of M. Benítez, C. Espinosa-Soto, and J. Dávila. E. R. A.-B. was supported by Conacyt 81542 and Papiit IN229009-3 grants. E. R. A.-B. is also supported by Miller Institute for Basic Research in Science, University of California, Berkeley.

*Corresponding author.
eabuylla@gmail.com

[1] G. von Dassow, E. Meir, E. M. Munro, and G. M. Odell, Nature (London) **406**, 188 (2000); R. Albert and R. G. Othmer, J. Theor. Biol. **223**, 1 (2003); P. M. Brakefield and J. C. Roskam, Am. Nat. **168**, S4 (2006); R. Lewontin, *The Triple Helix: Gene, Organism, and Environment*, (Harvard University, Cambridge, MA, 2001); K. Kaneko, BioEssays **33**, 403 (2011).

[2] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, Oxford, 1969).

[3] C. H. Waddington, *The Strategy of the Genes*, (George Allen & Unwin, Bristol, 1957).

[4] S. Huang and D. E. Ingber, Breast Dis. **26**, 27 (2007); J. X. Zhou, L. Brusch, and S. Huang, PLoS ONE **6**, e14752 (2011).

[5] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber,Phys. Rev. Lett. **94**, 128701 (2005); Z. Wang, Z. Hou, H. Xin, and Z. Zhang, Biophys. Chem. **125**, 281 (2007); B. D. MacArthur, A. Ma'ayan, and I. R. Lemischka, *Toward Stem Cell Systems Biology: From Molecules to Networks and Landscapes*, Cold Spring Harbor Symposia on Quantitative Biology Vol. 73 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2008), p. 211. A. Sakata, K. Hukushima, and K. Kaneko, Phys. Rev. Lett. **102**, 148101 (2009); K. Kaneko, J. Biosci. **34**, 529 (2009); M. Hoffmann, H. H. Chang, S. Huang, D. E. Ingber, M. Loeffler, J. Galle, PLoS ONE **3**, e2922 (2008); C. Furusawa and K. Kaneko, J. Theor. Biol. **209**, 395 (2001); F. H. Willeboordse and K. Kaneko, Phys. Rev. E **72**, 026207 (2005).

[6] J. Wang, K. Zhang, L. Xu, and E. Wang, Biophys. J. **99**, 29 (2010); , Proc. Nat. Acad. Sci. U.S.A. **108**, 8257 (2011).

[7] E. R. Alvarez-Buylla *et al.*, PLoS ONE **3**, e3626 (2008).

[8] E. R. Alvarez-Buylla, E. Balleza, M. Benítez, C. Espinosa-Soto, and P. Padilla-Longoria, in *Practical Systems Biology*, edited by C. Grierson and A. Heterington, SEB Experimental Biology Series Vol. 61 (Taylor & Francis, London, 2008).

[9] R. A. Barrio, A. Hernández-Machado, C. Varea, J. R. Romero-Arias, and E. Álvarez-Buylla, PLoS ONE **5**, e13523 (2010).

[10] L. Mendoza and I. Xenarios, Theor. Biol. Med. Model. **3**, 13 (2006).

[11] E. Azpeitia, M. Benítez, I. Vega, C. Villarreal, and E. Alvarez-Buylla, BMC Syst. Biol. **4**, 134 (2010).

[12] L. Glass, J. Theor. Biol. **54**, 85 (1975).

[13] L. A. Zadeh, Inf. Control **8**, 338 (1965).

[14] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[15] H. Haken, *Synergetics, An Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology* (Springer-Verlag, New York, 1983), 3rd ed.

[16] E. H. Snoussi and R. Thomas, Bull. Math. Biol. **55**, 973 (1993).

[17] E. R. Alvarez-Buylla *et al.*, *The Arabidopsis Book* (American Society of Plant Biologists, 2010), Vol. 8, p. e0127, http://my.aspb.org/.

[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.109.118102 for a table of Boolean expressions of GRN interactions in *A. thaliana*, including basic steps for analytical derivation of GRN attractors.

[19] *Arabidopsis: An Atlas of Morphology and Development*, edited by J. Bowman (Springer, New York, 1994).