

# Geometry, epistasis, and developmental patterning

Francis Corson and Eric Dean Siggia<sup>1</sup>

Center for Studies in Physics and Biology, The Rockefeller University, New York, NY 10021

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2009.

Contributed by Eric Dean Siggia, February 6, 2012 (sent for review November 28, 2011)

**Developmental signaling networks are composed of dozens of components whose interactions are very difficult to quantify in an embryo. Geometric reasoning enumerates a discrete hierarchy of phenotypic models with a few composite variables whose parameters may be defined by in vivo data. Vulval development in the nematode *Caenorhabditis elegans* is a classic model for the integration of two signaling pathways; induction by EGF and lateral signaling through Notch. Existing data for the relative probabilities of the three possible terminal cell types in diverse genetic backgrounds as well as timed ablation of the inductive signal favor one geometric model and suffice to fit most of its parameters. The model is fully dynamic and encompasses both signaling and commitment. It then predicts the correlated cell fate probabilities for a cross between any two backgrounds/conditions. The two signaling pathways are combined additively, without interactions, and epistasis only arises from the nonlinear dynamical flow in the landscape defined by the geometric model. In this way, the model quantitatively fits genetic experiments purporting to show mutual pathway repression. The model quantifies the contributions of extrinsic vs. intrinsic sources of noise in the penetrance of mutant phenotypes in signaling hypomorphs and explains available experiments with no additional parameters. Data for anchor cell ablation fix the parameters needed to define Notch autocrine signaling.**

embryonic development | modeling | signaling dynamics

**D**uring development, a single cell gives rise to organized structures comprised of specialized cell types. The molecular description of these transformations has focused on the transcriptional networks controlling gene expression (1). However we still cannot predict the time course of gene expression in a developing embryo. The quantitative description of signaling is more problematic because it is more enmeshed in cell biology and there are many genes that modulate the behavior of the common pathway components by position/aggregation/phosphorylation, depending on the cellular context and organism (2). Evolutionary bricolage (3) can alter many genetic linkages leaving the phenotype invariant. For instance, somitogenesis in all vertebrates is well described by the “clock and wavefront” model, yet there is little overlap among the genes that actually oscillate in the Wnt, Notch, and FGF pathways when comparing fish, chick, and mouse (4). Thus gene regulatory networks have to be worked out species by species, and it will be problematic to transfer numerical parameters between systems.

Classical embryology developed a suite of concepts that broke down the steps from signaling to cell differentiation and motivated many experiments, even though they are difficult to assay precisely. During development, cells or tissues become competent when they can sense signals that direct them toward a particular fate; they are specified or committed to a fate when the signal can be removed with no change in outcome. Cells are determined when they cannot be directed to alternative fates by other signals. Differentiation ensues when specialized gene batteries are induced along with characteristic morphology. None of these concepts can be tested in a completely controlled way in vivo, and the recent creation of induced pluripotent stem cells

(5) shows that even differentiation can be reversed. Yet they have provided a useful guide to experiments.

These concepts admit a natural geometric representation, which can be formalized in the language of dynamical systems, also called the geometric theory of differential equations (Fig. 1). When the molecular details are not accessible, a system’s effective behavior may be represented in terms of a small number of aggregate variables, and qualitatively different behaviors enumerated according to the geometrical structure of trajectories or topology. The fates that are accessible to a cell are associated with attractors—the valleys in Waddington’s “epigenetic landscape” (6)—to which neighboring trajectories converge. The set of points that tend to a given attractor forms its basin of attraction, and the state of commitment of a cell can be defined by its position relative to the basins of different fates. Along the boundaries between basins of attraction are saddle points, where the flow splits between two attractors, marking a “decision point” between different outcomes. Certain fates become accessible only at a particular time during development, so one should think of a landscape that changes over time. The external signals to which cells respond during competence transiently shift the boundaries between attractors, biasing trajectories toward one fate or other.

The appeal of this type of mathematics for developmental biology was recognized long ago (7) because the description is phenotypic and the mathematical concepts are formulated without reference to parameters. However, the applications never went beyond metaphor. The tools of dynamical systems theory have been used to analyze the qualitative behavior of detailed gene network models, but this assumes a correct model is available to begin with. Our aim here is to show that we may get a better grasp of developmental phenomena by stepping back from the molecular scale and reasoning directly on the phenotype expressed in geometrical terms.

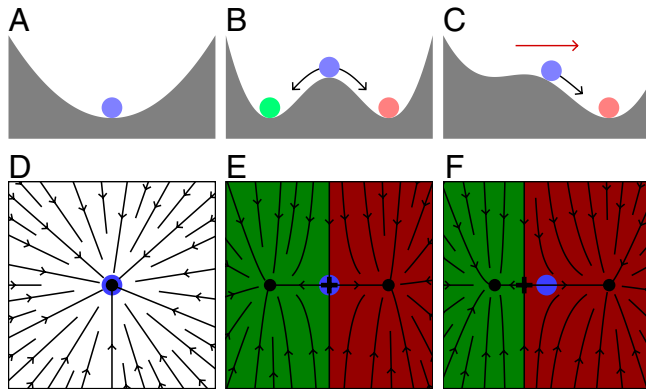
Even if the ultimate goal is a gene centric description, geometric or phenotypic models are a useful intermediate step. They organize rather clearly the minimal parameters necessary to realize the observed phenotypes and provide a context for the definition of epistasis. The conventional test for epistasis is whether the effects of two alleles are additive. But if the data pertain to protein binding, and epistasis is interpreted as cooperativity, then the relevant comparison is between binding energies. Thus the logarithm of the protein abundance or affinity should be taken before checking for additivity. It is the context that dictates that log of expression is the relevant data transformation. In the case of development, where the data are a function of space and time, the analogous transformation is not so obvious. The geometric description builds in the correct qualitative structure of the initial and final states and naturally provides for internal variables in terms of which signaling pathways combine linearly. Epistasis appears when the dynamics projects these variables onto discrete fates. Thus as our example of protein affinity shows, epistasis can be a matter of variable choice.

Author contributions: F.C. and E.D.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: siggiae@mail.rockefeller.edu.

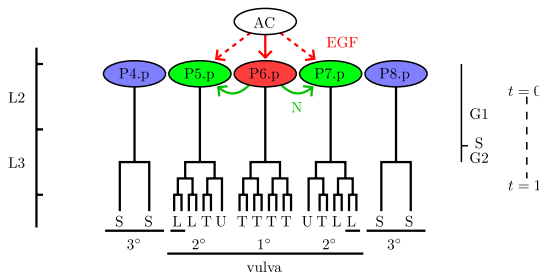
This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201505109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201505109/-DCSupplemental).



**Fig. 1.** A geometric view of cell fate specification. (A) Prior to competence, the cell occupies a precursor state symbolized by the blue ball in the bottom of the well. (B) During the competence period, there are several states accessible to the cell indicated by separate wells. (C) The transition between wells is governed by signals, represented by the red arrow, that deform the landscape and favor a particular state. The geometric viewpoint treats the signal transduction machinery as a black box, and reduces the signal to a single parameter. (D–F) The same configurations as in A–C shown as a flow field in the plane. Fixed points are shown as solid circles throughout, the saddles defining the division point between the two states are shown as crosses, and the basins of attraction for each fate are shown as colored regions.

Choosing variables that allow pathway contributions to be simply added makes the model predictive because a new interaction parameter is not needed every time a genetic cross is made. The single alleles have to be fit, because they are otherwise numerically undefined, but then all combinations of alleles are predicted.

As a classic example of cell fate patterning, the nematode vulva (8) is an instructive model for the application of geometric methods. The vulva of *Caenorhabditis elegans* forms from a row of vulval precursor cells (VPCs), which are competent to adopt three possible fates. In normal development, the VPCs adopt an invariant pattern of fates (as seen in the lineage or division pattern), which is specified by two signaling pathways: inductive EGF signaling from the neighboring anchor cell (AC) and lateral Notch signaling among VPCs (Fig. 2). When development is perturbed, other patterns of the three fates are obtained, often with variable results for genetically identical animals. Thus experimental outcomes are described as probabilities. The mechanism by which EGF and Notch specify VPC fates has been



**Fig. 2.** Vulval development in *C. elegans* (8, 27). The vulva forms from a row of six VPCs, P3.p–P8.p (P3.p fuses with the epidermis in 50% of animals, ref. 32, and is omitted here). The VPCs are competent to adopt three possible fates, inner vulval or primary (1°), outer vulval or secondary (2°), and nonvulval or tertiary (3°), which are identified by characteristic cell lineages (the orientation of terminal divisions is indicated by letters: L, longitudinal; T, transverse; U, undivided; S, fused with the epidermal syncytium; underlined cells are adhered to the cuticle). In WT, inductive EGF signaling from the AC (red arrows) and lateral Notch signaling among VPCs (green arrows) specify an invariant pattern of fates, 32123. VPC response to external signals begins in the L2 larval stage (cell cycle G1) and extends shortly after the first round of division. In the model, this competence period is defined to be the time interval  $[0, 1]$  (dashed line).

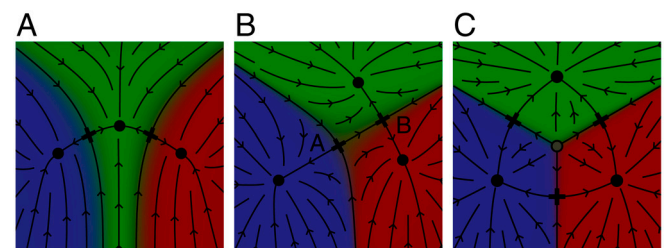
subject to debate. In the “morphogen” or “graded” model (9), different levels of EGF specify the different fates, 1°, 2°, 3°, at increasing distance from the anchor cell. In the “sequential” model (10, 11), the anchor cell induces the 1° fate in the closest VPC, which in turn induces the 2° fate in its neighbors through Notch. There is experimental evidence supporting both mechanisms and both are present in the quantitative model we develop here; they are not dichotomous.

In the following sections, we first formulate a geometric representation of vulval fate specification. Possible topologies are enumerated and discriminated on the basis of experimental observations. The favored topology is then translated into a minimal quantitative model, defining equations for the dynamics of individual VPCs along with the coupling between cells. Variability among animals is an intrinsic part of the model, and in contrast to genetic inference in most contexts, partially penetrant phenotypes are most informative, because they highlight the boundaries between fates. We show graphically how each parameter in the model is fit to a specific experiment, and can then “postdict” several experiments not used in the fit, particularly those involving correlated variability in fates (“extrinsic noise”). Genetic experiments ostensibly showing pathway interactions can be quantitatively fit with our model which lacks explicit interaction parameters. We predict nonintuitive outcomes for crosses between various hypo- and hypermorphs of the two pathways plus ablations in diverse backgrounds. Our modeling approach extends to any context where a multiway cell decision is controlled by two signaling pathways.

## Results: Model Formulation

Experiments define the competence period for vulval induction and the conditions under which the precursor cells can be switched among their three possible terminal fates by signals (*SI Appendix, Methods: Model and numerical simulation*). We can therefore generalize our binary decision model from Fig. 1 to a situation with three possible outcomes. Because the fate adopted by each VPC depends on the activities of two signaling pathways, it is natural to describe its effective dynamics in a two-dimensional space (Fig. 3). Different topologies can be enumerated by how the basins of attraction of the three fates fit together in the plane and only one of these is plausible for the vulva (Fig. 3C). With the model’s topology fixed from qualitative aspects of the experiments, we formulate a minimal quantitative model, which parameterizes the flow governing each precursor cell according to the levels of the two signaling pathways.

**Individual VPCs.** The state of each cell is described by a two-dimensional vector  $\vec{r} = x\vec{e}_x + y\vec{e}_y$ , and its dynamics is parameterized as



**Fig. 3.** Possible topologies for a decision between three fates. (A) The three basins of attraction are ordered from left to right as suggested by the response to a graded morphogen. This topology is essentially the graded model (9) and is ruled out by the equal proportions of 1° and 2° fated cells in P6.p for ablation times prior to 0.6 in *SI Appendix, Fig. S5*. (B) Two consecutive decisions (i.e., between 3° and vulva at saddle A and 1° vs. 2° at saddle B); this requires parameter tuning and is thereby excluded. (C) All transitions between two fates are permissible, there is a central source (open circle) and saddles between each pair of fates. This topology is our candidate for vulval patterning.

follows:

$$\vec{f}(\vec{r}) = 2\vec{r} + c_2[-2xy\vec{e}_x + (y^2 - x^2)\vec{e}_y] \quad [1]$$

$$\frac{d\vec{r}}{dt} = \frac{1}{\tau}[\vec{\sigma}_1(\vec{f} + \vec{m}) - \vec{r}] + \vec{\eta}(t) \quad [2]$$

$$\vec{\sigma}_1(\vec{f}) = \tanh(\|\vec{f}\|) \frac{\vec{f}}{\|\vec{f}\|} \quad [3]$$

$$\vec{m} = \vec{m}_0 + l_1\vec{m}_1 + l_2\vec{m}_2 \quad [4]$$

$$\langle \eta_i(t)\eta_j(t') \rangle = 2D\delta_{ij}\delta(t-t'). \quad [5]$$

The polynomial vector field  $\vec{f}(\vec{r})$ , which derives from a threefold-symmetric cubic potential, is the simplest flow that diverges away from the origin into three “valleys.” In the definition of the dynamics (Eq. 2), it is mapped through a saturating sigmoidal function and combined with a convergent radial flow, so that trajectories are bounded ( $\|\vec{r}\| < 1$ ) and each valley terminates in a stable fixed point (Fig. 4A). The constant term  $\vec{m}$  then biases this symmetric flow according to the external signals received by the cell, with the time constant  $\tau$  setting the maximum response rate. In the absence of external signals,  $\vec{m}_0$  biases the flow toward the default 3° fate (Fig. 4B). In keeping with our minimalist approach, the responses to EGF and Notch signaling are represented by the vectors  $\vec{m}_1$  and  $\vec{m}_2$  (Fig. 4 C–F). Signaling strength is parameterized by  $l_1$  and  $l_2$ , which vary between zero and one in WT. For instance, the reduction in EGF signal from the anchor cell in a hypomorphic allele will multiply  $l_1$  in all cells by a number specific to that allele. In these variables, there is no explicit pathway coupling, contributions from the two pathways simply add. Finally, the simplest yet natural way to introduce variability in

outcomes is a random noise term (Langevin noise),  $\vec{\eta}(t)$ , which is parameterized by a coefficient of diffusion in phase space,  $D$ .

One may object that the interpretation of a signal surely depends on the state of the cell (i.e., the  $\vec{m}_i$  are functions of  $\vec{r}$ ), and the mutual repression between the EGF and Notch pathways (12, 13) is one instance of this. However, as already anticipated, we do not need such parameters to fit the available data, including data purporting to show mutual repression between the two pathways. The latter effect is implicitly subsumed by the bistability between fates 1 and 2.

**Interacting Cells.** We simulate five cells centered on P6.p (Fig. 2), thus our full model has 10 variables, two for each cell. The EGF signal is fixed in time and decays exponentially with distance from the AC (*SI Appendix, Inductive signal gradient*),

$$l_1 = \{\gamma^2, \gamma, 1, \gamma, \gamma^2\}. \quad [6]$$

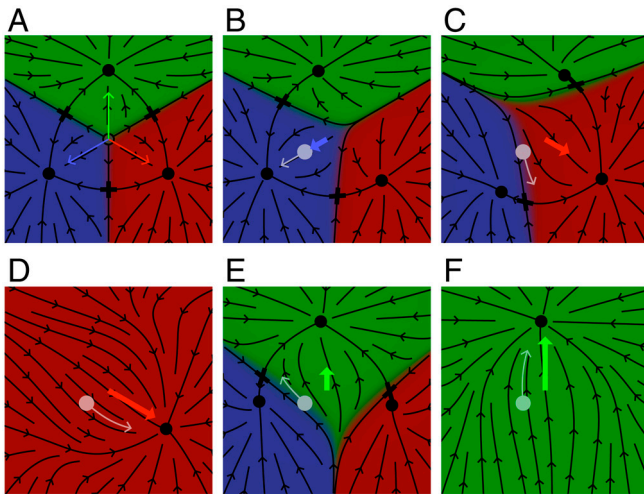
On the other hand, the lateral signal emitted by each VPC depends on its current state. The lateral signal is defined by a function  $L_2(\vec{r})$ , which varies as a sigmoid from zero to one when  $\vec{r}$  crosses a threshold line that runs diagonally through the 3°, 1° domains:

$$L_2(\vec{r}) = \sigma_2(n_0 + \vec{n}_1 \cdot \vec{r}) \quad [7]$$

$$\sigma_2(u) = \frac{1 + \tanh 2u}{2}. \quad [8]$$

To relate the signal,  $L_2$ , produced by a cell and its neighbors to Notch signaling,  $l_2$ , in that cell, we introduce a ratio  $\alpha$  to parameterize the relative importance of autocrine and paracrine signaling (14) (Notch ligands or Deltas include both transmembrane and diffusible ligands; ref. 15). Thus for P6.p,

$$l_2(\text{P6.p}) = L_2(\text{P5.p}) + \alpha L_2(\text{P6.p}) + L_2(\text{P7.p}). \quad [9]$$



**Fig. 4.** Flow fields from the model. (A) In the absence of any bias ( $\vec{m} = \vec{0}$  in Eq. 2), the flow is symmetric, diverging away from a central source (open circle). The colored arrows mark the default directions of the  $\vec{m}_i$  (Eq. 4). (B–F) Dynamics of a VPC under fixed levels of the two signals (autocrine signaling is not included here). (B) In the absence of ligands ( $l_1 = l_2 = 0$ ), the flow is biased by  $\vec{m}_0$  toward the default 3° fate and the initial condition (white disk) is in the basin of 3°. This flow is seen by VPCs after competence and defines their fates. (C) A small level of EGF ( $l_1 \approx 0.2$ ; red arrow) biases the flow toward the 1° fate. (D) At stronger levels of EGF (here,  $l_1 = 0.5$ ), the other fixed points disappear. So given sufficient time, this level of EGF will force any cell to go to the 1° fate regardless of its previous state of specification. (E) Weak levels of Notch signaling ( $l_2 \approx 0.2$ ; green arrow) bias the flow toward 2°. (F) At stronger levels of Notch (here,  $l_2 \approx 0.5$ ) the other fixed points disappear and any cell will eventually reach 2°.

**Time Course.** Model simulations proceed as follows (see *SI Appendix, Simulation procedure* for details). The VPCs are initially all equivalent and the equations are initialized from a Gaussian distribution consistent with the noise level in the model. The dynamics are then computed as described above for a period of time representing competence (defined in arbitrary units to be the interval from  $t = 0$  to  $t = 1$ ). Following competence, the fates of the cells are scored according to their locations relative to the basins of attraction in the absence of signaling (Fig. 4B).

We have found by fitting data that default parameters are adequate, specifically  $c_2 = 1$ ,  $\vec{m}_{0,1,2}$  point toward the fixed points, 3°, 1°, 2°, and  $\|\vec{n}_1\| = 3$ . A natural simplification would have  $\vec{n}_1$  in Eq. 7 aligned with  $\vec{m}_1$ , expressing that EGF induces lateral signaling, but the data impose a slightly different orientation. We are thus left with nine parameters ( $m_0 = \|\vec{m}_0\|$ ,  $m_1 = \|\vec{m}_1\|$ ,  $m_2 = \|\vec{m}_2\|$ ,  $\tau$ ,  $D$ ,  $\gamma$ ,  $\alpha$ ,  $n_0$ ,  $\angle \vec{n}_1 = \text{orientation of } \vec{n}_1$ ), plus an additional value of  $l_1$  or  $l_2$  for each pathway allele. This choice defines a minimal model where all parameters are directly tied to some property of the flow and thus a phenotype. Signaling, commitment, and determination appear in the model as in experiment, as a continuous progression though with substantial variability between animals. Further discussion of the model’s relevance and limitations is reserved for *Discussion*.

## Results: Fit to Data

Our fits minimize the mean-square difference between model and data, normalized by the expected variance based on the number of data points (*SI Appendix, Methods: Fit to data*). We also impose weak biases on parameters which are not normalized to  $[0, 1]$ , that favor values of order one in dimensionless units.

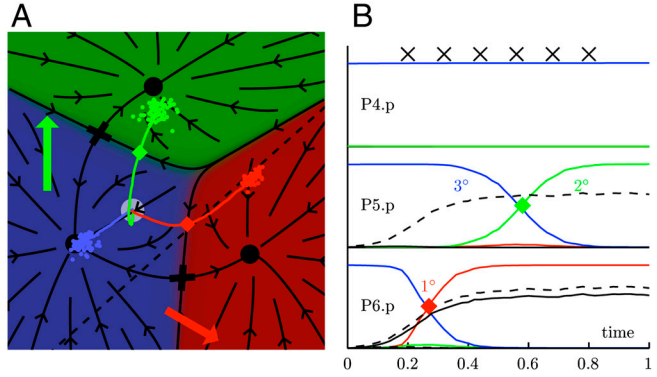


**Table 1. Summary of experiments being fit and the constraints they impose on model parameters**

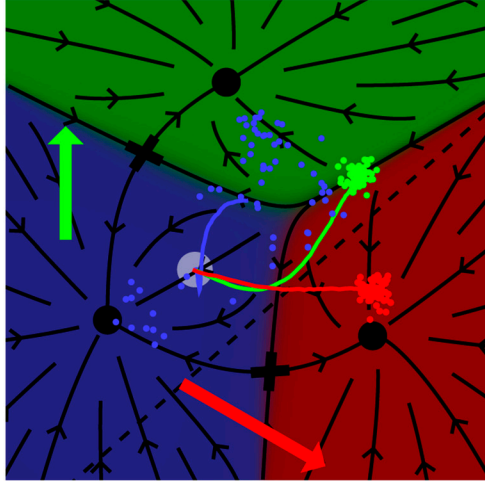
| Experiment  | Outcome         | Parameters                  |
|---|-----------------|-----------------------------|
| <i>Wild-type outcomes under reduced signaling</i>   |                 |                             |
| Wild-type   | 32123           |                             |
| EGFR in P6.p only   | WT              |                             |
| Single copy EGF/Notch   | WT              | $m_1, m_2$                  |
| <i>Phenotypes under reduced Notch/excess EGF</i>  |                 |                             |
| No Notch, WT EGF  | 33133           | $\gamma$                    |
| Notch null (2 ACs)  | 31113           | $\gamma$                    |
| EGF overexpression  | induction ca. 4 | $m_1/m_2, \angle \vec{n}_1$ |
| Ectopic EGF   | 1°/2° fates     |                             |
| <i>Phenotypes following anchor cell ablation</i>  |                 |                             |
| At successive times   | 33333 to WT     | $\tau, D, \alpha, n_0$      |
| Detailed in <a href="#">SI Appendix, Table S1</a> and <a href="#">Constraints on parameters</a> |                 |                             |

The use of priors is justified by standard Bayesian inference—i.e., a preference for typical parameters rather than more specific ones that yield the same quality fit. Thus the fits would favor a substantial value of the noise (diffusion constant) even if zero noise placed the initial conditions precisely on the boundaries between fates and thereby accounted for the partially penetrant phenotypes. The zero noise fit would very tightly constrain the deterministic part of the flow and Bayesian methods would deem this improbable. Thus Bayesian methods naturally implement a requirement of evolution that tightly constrained parameters impose a higher mutational load and thus are disfavored.

[SI Appendix, Table S1](#) states the numerical data we fit to and a summary is provided in Table 1 along with the parameters most directly determined by each experiment (see [SI Appendix, Constraints on parameters](#)). Fig. 5 shows the WT dynamics corresponding to a typical parameter set (see also [Movie S1](#)). Conditions generating unique fate assignments (e.g., WT for a half dose of EGF or Notch; refs. 16 and 17) just give inequalities on parameters. Most information comes from alleles that give variable outcomes, which serve to fix the initial conditions relative to the saddle points in the flow as in Fig. 4 C and E. Typically we do not know numerically (relative to WT) the signaling



**Fig. 5.** Our model fit displayed for WT parameters. Because the inductive signal is symmetric around P6.p, we show just P4-6.p. (A) The trajectories of P4-6.p are represented by blue, green, and red, respectively (according to their WT fates) and overlaid on the zero ligand flow field which applies at the end of the competence period and determines fates (see Fig. 4B). The cell trajectories begin from a common point and are averaged over the noise, except at the final time. The green and red arrows denote the fixed vectors  $\vec{m}_{1,2}$  in Eq. 4 and the dashed line is where Delta expression is half maximum. (B) The fractional induction of the three cell fates 1°/2°/3° is shown as red/green/blue curves as a function of time for each cell during the competence period. The upper crosses represent the AC ablation times. The solid (respectively dashed) black curves show the levels of Delta expressed (respectively received) by each cell. The red and green diamonds correspond to those in A and indicate 50% induction.

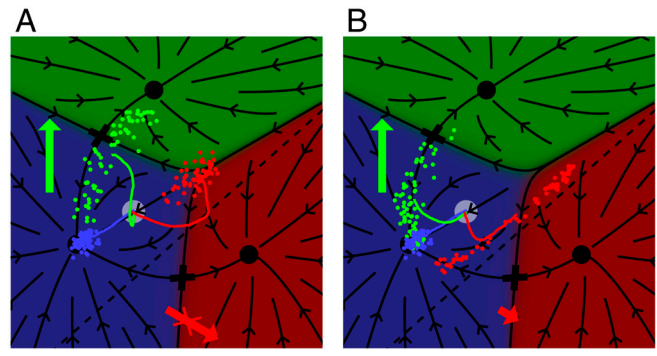


**Fig. 6.** EGF overexpression ([SI Appendix, Table S1, line 7](#)). Inductive signaling was increased by a factor of approximately seven relative to WT (arrow not to scale). Pattern frequencies for model and experiment are given in [SI Appendix, Table S5](#).

strength in the mutants, so that has to be fit too, but parameters in the base model are obtained also.

Overexpression of EGF to a level that induces on average four VPCs (18) is instructive in this regard (Fig. 6). Because P5/7.p yield a mix of 1°/2° fates, we infer that the enhanced inductive signal from the AC balances the lateral signal from P6.p. P4/8.p yield a fraction of all three fates, which implies a relation on EGF signaling at those positions as well as the orientation of the threshold for Delta expression,  $\vec{n}_1$  in Eq. 7 (from the induction of 2° by P5/7.p).

AC ablations (19) probe the time course of induction (Fig. 7A and [SI Appendix, Fig. S5](#)). Because we only know the ablation times relative to developmental stages, we model them by terminating the inductive signal at a uniformly spaced series of points in the interval [0.2, 0.8] (where the induction period is defined as [0, 1] in model units; see [SI Appendix, Table S1](#)). This choice is not a serious approximation because trends in the data as the ablation time increases match those in the model ([SI Appendix, Fig. S9](#)).



**Fig. 7.** Anchor cell ablation vs. EGF hypomorph: dynamic vs. static reduction in EGF signaling compared for the same vulval induction of P6.p, 58% [corresponding to the hypomorphic allele *lin-3(e1417)*, [SI Appendix, Table S4](#)]. (A) The anchor cell was ablated at time 0.4 (intermediate between the early L3 and DU divided stages, [SI Appendix, Table S1](#)). The trajectories are WT until the kink, most evident in the red line for P6.p, after which they continue to move upward under the influence of the autocrine Notch signal and thus become partly 2°. (B) Our fit to the *lin-3(e1417)* allele multiplies  $I_1$  by 0.24 in all cells. In contrast with the ablation experiment in A, P6.p is never exposed to autocrine Notch signaling and does not become 2°. Whenever P6.p is induced, P5/7.p can adopt the 2° fate, but because of the later onset of lateral signaling, this is less frequent than in the ablation experiment. Pattern frequencies for model and experiments are given in [SI Appendix, Table S5](#).

To attempt greater precision here would be illusory because the fate of EGF protein after AC ablation is unknown. Whatever decay rates influence the EGF time course after ablation will be the same in other genetic backgrounds we consider, and thus we can predict the consequences of ablation in any defined background with no additional parameters. Induction levels of P5/7.p following ablation fix the threshold for lateral signaling ( $n_0$  in Eq. 7). The mixed 1°/2° fates adopted by P6.p at intermediate times rely on the diffusing Delta fraction and autocrine signaling to drive cells into the 2° domain, fixing the autocrine-to-paracrine ratio  $\alpha$ . The variability of outcomes determines the numerical noise  $D$ .

We have confirmed the association between parameters and experiments by Monte Carlo sampling of parameter space while imposing various subsets of the data. Because we have relatively few parameters and they are mostly tied individually to phenotypes, we are able to converge the fits using standard Levenberg–Marquardt gradient descent. The number of data points for partial phenotypes—excluding experiments with WT or trivial outcomes—is 40, so the typical minimum Chi-squared (see *SI Appendix, Likelihood of experimental data*) of approximately 20 is well within statistical errors. Repeating the fit with different initial parameter sets confirms the existence of a unique global optimum.

Our standard parameter set is given in *SI Appendix, Table S3* with most uncertainties under 20%. Monte Carlo parameter sampling shows that most residual variation resides in  $m_1$  and  $m_2$  and could be resolved if we knew the numerical degree of under/over-expression in several of our signaling pathway mutants.

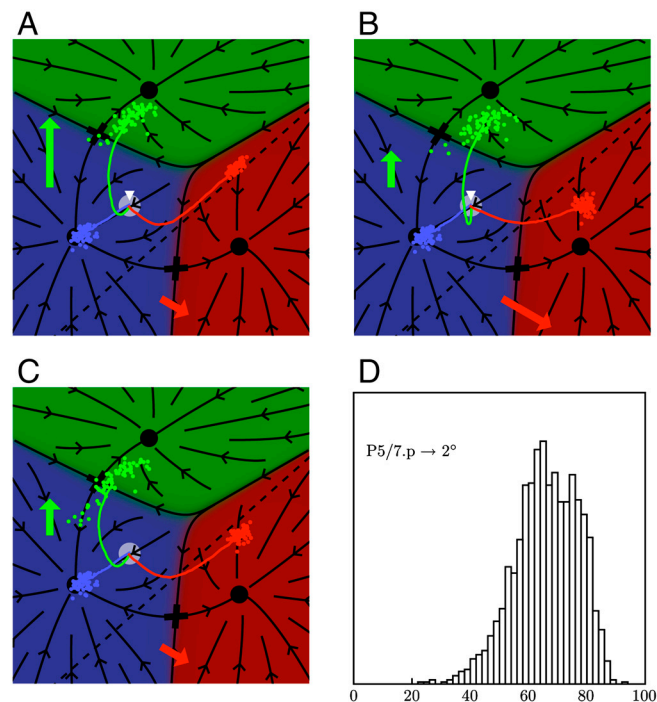
### Results: Predictions

We have been parsimonious in selecting the fitting data so as to delineate what is strictly necessary to determine parameters. As a consequence, several existing datasets are available for prediction. We first consider conditions that reduce inductive signaling, then the results of crosses. Epistasis is a natural test of the model, which has no explicit pathway coupling. There are many experiments on vulva whose interpretation involved gene interactions, because weak genetic effects are often revealed by expressing them in a sensitized background.

**Signaling Hypomorphs.** EGF signaling is required for vulval induction and a trivial 33333 pattern is obtained in its absence—e.g., when the AC is ablated at an early stage (20). The phenotypes that result from partially reduced signaling are more informative, particularly when compared with AC ablation (Fig. 7 and *SI Appendix, Figs. S5 vs. S6*). If we chose the ablation time to match the 58% induction level of P6.p for a particular *lin-3* hypomorph (*SI Appendix, Table S4*), we find 2° fates in P6.p only for ablation and never in the hypomorph. Furthermore P5/7.p are less induced in the hypomorph than in ablation, because P6.p signals earlier in ablation experiments, when P5/7.p are closer to the 2° fate and more susceptible to induction. Our predicted range of P5/7.p induction ( $6 \pm 6\%$ ) agrees with the weak but nonzero induction (15%; *SI Appendix, Fig. S7*) observed.

**Extrinsic vs. Intrinsic Noise.** Under conditions of partial penetrance we can compute the probabilities of the full VPC patterns and ascertain whether cell fates are spatially correlated. We expect that extrinsic noise, where variation in P6.p affects both P5/7.p, will generate correlation, whereas intrinsic noise such as in EGF overexpression experiments, where P6.p is robustly induced to a 1° fate, will lead to uncorrelated fates. This prediction is indeed realized in *SI Appendix, Table S5* and shows that the model correctly predicts data for high and low levels of extrinsic noise.

**Epistasis from Fate Projection.** One apparent source of epistasis in our model is the threshold effect that results from the assignment to a discrete fate after the simple addition of two signals.



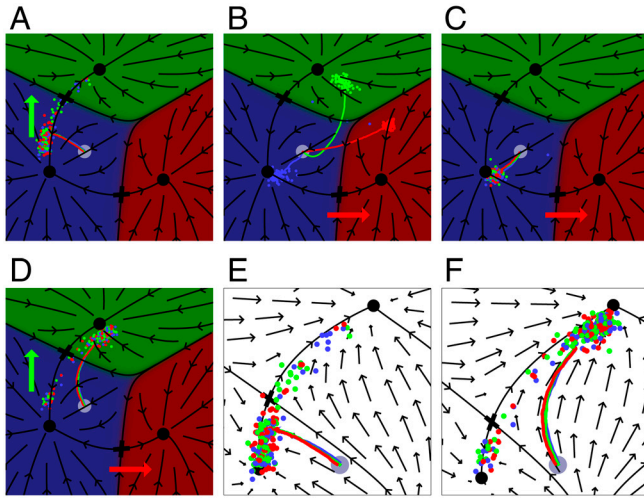
**Fig. 8.** The cross of two marginally WT hypomorphs in EGF ( $l_1 \approx 0.5 \times \text{WT}$ ) and Notch ( $l_2 \approx 0.5 \times \text{WT}$ ) yields a substantial reduction in P5/7.p induction. (A and B) are the EGF and Notch hypomorphs with the red/green arrows showing the changed strength in the two pathways. The white arrow in A and B highlights that EGF signaling normally keeps P5/7.p dynamics poised between the 1° and 3° fates. This effect enhances the efficacy of Notch signaling in WT, as contrasted with the longer path to the 2° fate in the cross (C). D shows the fractional induction of P5/7.p is robust to Monte Carlo sampling of parameter space.

For maximum effect, consider two hypomorphs of EGF and Notch each marginally WT (i.e., the single mutant is WT, but if slightly enhanced would give a phenotype). When combined, the cross exhibits a substantially reduced induction of P5/7.p to 2° (Fig. 8). We consider this prediction relatively robust because this “projection” effect depends mainly on the geometry of the domains and is insensitive to the orientations of the  $\vec{m}_i$ . This figure also illustrates that the weak EGF signal received by P5/7.p under WT conditions enhances the patterning robustness by placing the presumptive 2° cells poised between the 3° and 1° domains from where the distance to the 2° basin is as short as possible and the induction of P5/7.p as large as possible (*SI Appendix, Fig. S8*).

**Epistasis from Saddle Flows.** A sensitized background was used by Zand et al. (21) to examine the effect of low EGF on the induction of 2° fates. A mild gain-of-function Notch mutation results in weak constitutive activity of the pathway, potentiating the induction of the 2° fate. This mutant lacks an AC. Weak EGF signaling is induced uniformly across the VPCs by a temperature sensitive *lin-15* mutant, that in the absence of an AC would yield no induction. In the double mutant, induction of the 2° fate is greatly enhanced, from which the authors conclude that low EGF acts as a “pro-2°” signal (whereas high EGF is as always pro-1°).

Our model has no such property: EGF pushes cells in a fixed direction (specified by the vector  $\vec{m}_1$ ). Strikingly, it can nevertheless reproduce the observed effect (Fig. 9). For this  $\vec{m}_1$  was slightly rotated from its default direction (while still fitting our training data), so we cannot claim this result as a true prediction, but we can challenge its published interpretation. As can be seen in Fig. 9E, the flow in the sensitized background sends the initial conditions to just below the saddle between the 3° and 2° fates, with a small fraction of cells driven across the boundary by noise.



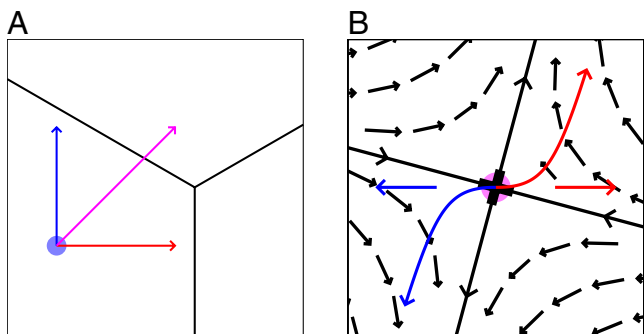


**Fig. 9.** Epistasis between *lin-12(d)* and *lin-15(ts)*. (A) A *lin-12(d)* mutant has no AC and its lateral signal level is described by an additive term in  $I_2$  (approximately 0.2), fit to yield an average induction index of 0.4 as in experiments (21) (the rare outcomes in the green region). Without an AC, all VPCs behave similarly. (B) We fit the *lin-15(ts)* state in ref. 21 by a uniform  $I_1$  (approximately 0.1). When added to the EGF signal from the AC there is no ectopic induction. (C) In the absence of an AC (as in the double mutant), the weak EGF signal from *lin-15(ts)* shows no induction. (D) In the double mutant, most cells are induced to a 2° fate. (E and F) Close-ups of the flow seen by the cells in A and D, showing its changing orientation.

A small change in the flow field induced by low EGF (pointing toward the 1° fate) then suffices to direct most of the initial conditions to the other side of the saddle and thus to the 2° fate.

Contrary to intuition, we find that a signal that pushes cells in one direction can displace outcomes in a different direction. We can understand this form of epistasis as a generic effect of the flow in the vicinity of saddle points (Fig. 10B), and it also occurs, with the roles of Notch and EGF interchanged, when an EGF hypomorph is crossed with a weak Notch hypomorph (*SI Appendix, Epistasis from saddle flows*). This phenomenon will confound the interpretation of experiments in sensitized backgrounds, which by definition are tuned to marginal outcomes, and correspond geometrically to flows into saddle points.

**Putative Epistasis Between Notch and EGF Pathways.** Sensitized backgrounds have also been used to study the mutual inhibition between the EGF and Notch pathways. For instance, the Notch pathway up-regulates the transcription of the LIP-1 MAPK phos-



**Fig. 10.** Two forms of epistasis. (A) Epistasis from the projection of continuous dynamics onto discrete fates. The plane is divided into three sectors corresponding to three fates (see Fig. 5A), and two signals push the cells to the right (red arrow) or up (blue arrow). Levels of signal that are, separately, insufficient to drive the cells to a different domain (as shown here) can together induce a different fate (magenta arrow). (B) Epistasis from the flow. In the vicinity of a saddle, a small push in one direction (horizontal arrows) can deflect trajectories in another direction (vertical).

phatase that inhibits the EGF pathway in presumptive 2° cells (13, 22). However *lip-1(lf)* has no phenotype, so the genetic strategy in ref. 22 was to cross this allele into a background with ectopic EGF [e.g., *lin-15(rf)*] and observe epistasis. We can quantitatively replicate their results by an additive contribution to  $I_1$  [for *lin-15(rf)*] followed by a multiplicative change corresponding to *lip-1(lf)*, reflecting the order of the two genes in the pathway (*SI Appendix, Table S6*).

**Mapping Genetic Effects onto Model Parameters.** Another form of coupling between EGF and Notch is the down-regulation of Notch receptors in presumptive 1° cells, which was shown to be required for lateral signaling: P5/7.p otherwise adopt the default 3° fate (12). Unlike the Notch-induced down-regulation of EGF signaling, our model transparently incorporates this effect, with no changes, if the emitted signal  $L_2$  is interpreted as the “effective signal” that reaches neighboring cells, rather than the ligand level expressed by P6.p and then adsorbed by that cell. Manipulations that decrease Notch internalization would then be parameterized by moving the threshold for lateral signaling (dashed line in Fig. 5A) further into the 1° basin. Thus multiple genetic effects can map onto the same model parameter.

Likewise, it has been suggested that the EGF receptor sequesters ligand, preventing ectopic induction, so the shape of the gradient may change in EGF receptor (EGFR) mutants (23). This effect can be expressed through our parameter  $\gamma$  (Eq. 6) which can be related to a kinetic model for the diffusion and sequestration of ligand (*SI Appendix, Shape of the EGF gradient*). To explain the observed induction of 2° fate in isolated cells also requires assumptions about how Notch ligand diffuses and sticks (*SI Appendix, Isolated cells*).

**Model Suggests Informative Experiments.** Although we have concentrated here on a few significant predictions, the model does supply the phenotype under any combination of mutations affecting EGF/Notch signaling, AC ablation, temperature-induced changes in expression, etc. We summarize here a few interesting experiments suggested by the model with further details in the *SI Appendix*. AC ablations are predicted to have qualitatively different consequences in EGF and Notch mutants. EGF mutations affect mostly the timing of induction, whereas Notch mutations displace the intermediate phenotypes toward different proportions of 1°/2° fates (*SI Appendix, Anchor cell ablation in mutants*). In some mutant backgrounds, cells are predicted to transition through the basin of attraction of an intermediate fate on their way to their eventual fate. We find that these intermediate fates could be uncovered by suitably timed temperature-induced changes in expression (*SI Appendix, Temperature steps*).

In some experiments, a fraction of cells adopt hybrid fates, with the two daughters giving rise to lineages characteristic of different fates. Geometrically, hybrid lineages can naturally be interpreted as precursor cells ending near a boundary, such that their daughters can take divergent paths because of noise. Thus we expect that more hybrid fates will occur in experiments where cells are clustered near a boundary, rather than spread out across it. This expectation is realized in an extension of our model that treats separately the daughters of VPCs, and we find reasonable agreement with experimental data for cases with clustered outcomes, particularly for isolated cells induced to mixed 1°/2° fates (*SI Appendix, Hybrid fates*). We predict too few hybrid fates when outcomes are spread out, however, suggesting other factors contribute to their occurrence.

When cells land near the boundary between two fates, our model also makes quantitative predictions for the increased time necessary for cells to express a given level of fate marker in comparison to WT. The effect can be easily parameterized in terms of the curvature around the ridge separating two basins in the fate plane.

Note that the connection between reporters and cell fate may be lost in mutants; e.g., in an EGF hypomorph when P6.p is induced to 1°, the cells remain close to the 1°/2° boundary and do not express Delta—a plausible 1° fate marker (Fig. 7B). Consistent with this expectation, expression of another 1° fate reporter in L3 is very strongly reduced in a mutant with reduced EGFR levels but WT fates (24).

## Discussion

Development is a dynamical process with competing fates, and it is often hard to visualize the consequences of genetic perturbations. Our phenotypic model effectively codifies the genotype to phenotype map because the influence of each gene is generally limited to one parameter. Geometric reasoning provides a compact, nonredundant parameterization of signaling and fate specification.

Experience with other problems suggests that, to effectively parameterize dynamics, the critical points (saddles, sources, and sinks) have to be correct (25). Qualitative features of the data determine the flow topology (Fig. 3), which is then parameterized in a minimal way. Our procedure is no different in principle than parameterizing gene expression with a Michaelis–Menten–Hill function for activation or repression. Phenotypic models seem well suited for problems of signaling where protein trafficking, clustering, endocytosis, and other aspects of cell biology can have a substantive effect, but are hard to quantify.

Within a geometric framework, it is very natural that epistasis is obtained by projecting a linear model (contributions from the signaling pathways are added as vectors) onto discrete fates. Pathways “interact” because they operate on a strongly nonlinear mapping described by the geometry of fate domains (Fig. 10A). But there are no additional parameters required to predict the cross between any pair of alleles (or AC ablation for an allele) once the single alleles are parameterized, giving the model considerable predictive power.

Our fits in Fig. 9 demonstrated that a geneticist’s “marginal allele” that poises the cell between two fates is mathematically a flow that passes near a saddle point. This geometry is another source of epistasis because any small change that displaces the cells normal to flow into the saddle will have a large effect on cell fate (Fig. 10B). Our model provides a quantitative way to treat sensitized backgrounds in the context of a robust model for the WT. Genetic experiments purporting to show EGF/Notch interactions are explained by the multistability of the flow.

Nevertheless, there is biochemical evidence for mutual pathway inhibition, and one might ask how this would affect model fits. Down-regulation of Notch in presumptive 1° cells (12) can be implemented (with no extra parameters) by letting their sensitivity to lateral signaling decrease when they start to signal (cross the dashed line in Fig. 5A). This modified model fits experiments equally well, with largely unchanged parameter values. Other layers of regulation include the temporal gating of lateral signaling by heterochronic genes (26) and of 2° fate commitment by the cell cycle (27). These are not included in our model, but lateral signaling and the specification of 2° fates in the model begin some time into the induction period (Fig. 5B), so that temporal gating is likely to be redundant with the time course of specification in most conditions.

Our philosophy of starting with a simplified but global model for developmental patterning meshes well with basic biology. Signaling, commitment, and determination are discrete terms applied to a continuous process. It is unnatural to imagine one model for signaling and another for commitment and determination—the same genes are involved—and the regulation of competence by external signals would fit naturally within our description. A unified model also makes it very simple to incorporate noise, which is a prominent feature of any interesting mutant and of course affects the average trajectories as well as the probabilities of the final

states. The existence of a predictive dynamic model should encourage experiments to measure crosses among the available alleles as well as temporal signaling perturbations in mutant backgrounds. Variability in the timing of fate determination is a strong consequence of the flow geometry and should also be measured.

The AC ablations (19) and the EGF overexpression experiment where P4/8.p were induced to all three fates (18) were crucial to our fits because the data were sensitive to the transition regions in the flow field. Thus we were able to predict the observed fates in an EGF hypomorph. The correlations between fates (extrinsic noise) were a previously neglected aspect of available datasets that our model correctly predicted (SI Appendix, Table S5; note that we used only single cell fate probabilities in our fits). As previously anticipated (18), the ablation data strongly constrain the diffusible Delta fraction and thus determine the induction of isolated VPCs. However our model also makes clear what additional assumptions about the transport and adsorption of diffusible Delta are needed to explain isolated VPC induction and suggests that hybrid 1°/2° fates will be prevalent in such cells, as observed. Further experiments, such as temperature shifts and ablations in mutants, are predicted to yield unexpected outcomes, making for stringent tests of our description of fate specification.

Using existing data, our model has correctly reproduced experiments on *lip-1(lf)* (22) and a *lin-12(d);lin-15(ts)* cross (21) that were interpreted as evidence for pathway interaction. Our model does not support a “battle” (28) metaphor to describe EGF–Notch interactions. There are no dueling parameters in our model and we can explain experiments showing pathway interaction by the addition of vectors and the dynamical projection in the fate plane. In fact, our results suggest that the inductive signaling seen by presumptive 2° cells makes them less sensitive to fluctuations in lateral signaling—contrary to the intuition that it must be countered for robust patterning (13, 29). Thus the mutual inhibition between EGF/Notch enhances the stability of specified fates more than the decision between fates itself.

Remaining uncertainties in model parameters concern the numerical magnitude of signal response strengths. To examine crosses between EGF/Notch mutants (e.g., Fig. 8), we have therefore defined hyper/hypomorphs as marginally WT, rather than by numerical values of  $l_1, l_2$  which are not yet available experimentally—i.e., we cross phenotypes. If it becomes possible to numerically define signaling alleles relative to WT, alleles yielding approximately 50% induction will give the tightest constraints on parameters because one is fitting to the steepest region of the sigmoidal curve relating fate to parameters.

The vulva has motivated numerous other modeling studies. Among these are rule-based models (30), which describe specification as sequence of events and can incorporate some variability in event ordering, but by their discrete nature cannot account for the quantitative effects underlying partially penetrant phenotypes. The quantitative models presented in refs. 18 and 29 are more suitable points of comparison. Both model only signaling, they are not multistable and define territories in the EGF/Notch activity space to assign fates. They would therefore display epistasis from the projection onto fates. They use deterministic differential equations based on Michaelis–Menten kinetics, and they do not quantitatively fit or predict data involving partial penetrance. Though a model that required a parameter specified to better than factor of two might be rejected on fitness grounds, it does not follow that exploring a range of 100 or 10<sup>5</sup> (18, 31) in parameters is biologically meaningful and should be a basis for model selection as is commonly done.

Our model describes the experiments and time window we are interested in and not much more. There are no parameters that up-regulate receptors at the beginning of competence or shut down pathways after the first VPC division. We can say nothing about the differentiation that ensues near the fixed points, which clearly requires additional dimensions. However the fixed points

are merely compact ways of creating the simplest parameterizations for the boundaries between them and enumerating the various flow topologies. We are inclined to believe predictions based on these features of the model.

- Davidson EH (2006) *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic, New York).
- Nusse R (2010) *The Wnt Homepage*, Available at <http://wnt.stanford.edu>.
- Jacob F (1977) Evolution and tinkering. *Science* 196:1161–6.
- Krol AJ, et al. (2011) Evolutionary plasticity of segmentation clock networks. *Development* 138:2783–2792.
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676.
- Waddington CH (1957) *The Strategy of the Genes* (Allen & Unwin, London).
- Slack JMW (1991) *From Egg to Embryo: Regional Specification in Early Development* (Cambridge Univ Press, Cambridge, UK), pp 34–66.
- Sternberg PW (2005) Vulval development. *WormBook: The online review of C. elegans biology*, ed 1–28, 10.1895/wormbook.1.6.1, <http://www.wormbook.org>.
- Katz WS, Hill RJ, Clandinin TR, Sternberg PW (1995) Different levels of the *C. elegans* growth factor LIN-3 promote distinct vulval precursor fates. *Cell* 82:297–307.
- Simske JS, Kim SK (1995) Sequential signalling during *Caenorhabditis elegans* vulval induction. *Nature* 375:142–146.
- Koga M, Ohshima Y (1995) Mosaic analysis of the *let-23* gene function in vulval induction of *Caenorhabditis elegans*. *Development* 121:2655–2666.
- Shaye DD, Greenwald I (2002) Endocytosis-mediated downregulation of LIN-12/Notch upon Ras activation in *Caenorhabditis elegans*. *Nature* 420:686–690.
- Yoo AS, Bais C, Greenwald I (2004) Crosstalk between the EGFR and LIN-12/Notch pathways in *C. elegans* vulval development. *Science* 303:663–666.
- Sternberg PW, Horvitz HR (1989) The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. *Cell* 58:679–693.
- Chen N, Greenwald I (2004) The lateral signal for LIN-12/Notch in *C. elegans* vulval development comprises redundant secreted and transmembrane DSL proteins. *Dev Cell* 6:183–192.
- Ferguson EL, Horvitz HR (1985) Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. *Genetics* 110:17–72.
- Greenwald I, Sternberg PW, Horvitz HR (1983) The *lin-12* locus specifies cell fates in *Caenorhabditis elegans*. *Cell* 34:435–444.
- Hoyos E, et al. (2011) Quantitative variation in autocrine signaling and pathway crosstalk in the *Caenorhabditis elegans* vulval network. *Curr Biol* 21:527–538.
- Milloz J, Duveau F, Nuez I, Félix MA (2008) Intraspecific evolution of the intercellular signaling network underlying a robust developmental system. *Genes Dev* 22:3064–3075.
- Kimble J (1981) Alterations in cell lineage following laser ablation of cells in the somatic gonad of *Caenorhabditis elegans*. *Dev Biol* 87:286–300.
- Zand TP, Reiner DJ, Der CJ (2011) Ras effector switching promotes divergent cell fates in *C. elegans* vulval patterning. *Dev Cell* 20:84–96.
- Berset T, Hoier EF, Battu G, Canevascini S, Hajnal A (2001) Notch inhibition of RAS signaling through MAP kinase phosphatase LIP-1 during *C. elegans* vulval development. *Science* 291:1055–1058.
- Hajnal A, Whitfield CW, Kim SK (1997) Inhibition of *Caenorhabditis elegans* vulval induction by *gap-1* and by *let-23* receptor tyrosine kinase. *Genes Dev* 11:2715–2728.
- Stetak A, et al. (2006) Cell fate-specific regulation of EGF receptor trafficking during *Caenorhabditis elegans* vulval development. *EMBO J* 25:2347–2357.
- Strogatz SH (2001) *Nonlinear Dynamics and Chaos* (Westview, Boulder, CO).
- Li J, Greenwald I (2010) LIN-14 Inhibition of LIN-12 contributes to precision and timing of *C. elegans* vulval fate patterning. *Curr Biol* 20:1875–1879.
- Ambros V (1999) Cell cycle-dependent sequencing of cell fate decisions in *Caenorhabditis elegans* vulva precursor cells. *Development* 126:1947–1956.
- Sundaram MV (2004) Vulval development: The battle between Ras and Notch. *Curr Biol* 14:R311–R313.
- Giurumescu CA, Sternberg PW, Asthagiri AR (2006) Intercellular coupling amplifies fate segregation during *Caenorhabditis elegans* vulval development. *Proc Natl Acad Sci USA* 103:1331–1336.
- Fisher J, Piterman N, Hajnal A, Henzinger TA (2007) Predictive modeling of signaling crosstalk during *C. elegans* vulval development. *PLoS Comput Biol* 3:e92.
- Giurumescu CA, Sternberg PW, Asthagiri AR (2009) Predicting phenotypic diversity and the underlying quantitative molecular transitions. *PLoS Comput Biol* 5:e1000354.
- Sternberg PW, Horvitz HR (1986) Pattern formation during vulval development in *C. elegans*. *Cell* 44:761–772.

**ACKNOWLEDGMENTS.** We thank John Guckenheimer, Paul Francois, Marie-Anne Félix, and Michalis Barkoulas for extensive conversations, and Marie-Anne Félix for experimental data. We were supported by National Science Foundation Grant PHY-0954398.