Supporting Information

Lubensky et al. 10.1073/pnas.1015302108

SI Text

S1. Experimental Materials and Methods. Eye imaginal discs were dissected from late third-instar larvae, fixed, and subjected to indirect immunolabeling as described previously (1). Preparations were imaged using a Bio-Rad Radiance 2000 confocal microscope, and data processed using ImageJ v1.36b and Adobe Photoshop v8.0. Confocal images were z-projected to include labeled nuclei. In addition to wild type (Canton S), the following genetic strains were used: sca^{BP2} (2) and N^{ts1} (3). Antibody labeling procedures were performed as described (4, 5). Primary antibodies were rabbit anti-atonal serum (6), guinea pig anti-senseless serum (7), and mAb323, recognizing several bHLH proteins encoded by E(spl)-C (8). Cy2-, Cy3-, or Cy5-conjugated secondary antibodies were multilabel grade from Jackson Immunoresearch.

Normal *Drosophila* culture was at 25 °C. Immersion of culture vials in a 31.5 °C water bath for 2 h or 3 h was used to perturb Notch function in the N^{ts1} mutant. As described previously, intermediate groups maintain Atonal expression at the restrictive temperature but resolve to single R8 cells in most cases when returned to the permissive temperature (18 °C) within 4 h (9). Although N^{ts1} develops normally at 18 °C, N^{ts1}; sca^{BP2}/sca^{BP2} eye discs exhibit a slightly more disordered pattern of R8 cells than eye discs from sca^{BP2}/sca^{BP2} larvae at the same temperature.

S2. The Model. The time evolution of the variables $a_{\mathbf{j}}$, $s_{\mathbf{j}}$, $h_{\mathbf{j}}$, and $u_{\mathbf{j}}$, where $\mathbf{j}$ indexes the lattice sites, is given by Eqs. **1**–**4** of the main text, which we repeat here for completeness:

$$\frac{\partial a_{\mathbf{j}}}{\partial t} = f_{n_a}\left(\frac{a_{\mathbf{j}}}{A_a}\right) - a_{\mathbf{j}} + Ff_{m_s}\left(\frac{s_{\mathbf{j}}}{S}\right) + Gf_{m_h}\left(\frac{h_{\mathbf{j}}}{H}\right)\left[1 - f_{m_u}\left(\frac{u_{\mathbf{j}}}{U}\right)\right] \quad \textbf{[S1]}$$

$$T_s\frac{\partial s_{\mathbf{j}}}{\partial t} = f_{n_s}\left(\frac{a_{\mathbf{j}}}{A_s}\right) - s_{\mathbf{j}} \quad \textbf{[S2]}$$

$$T_h\frac{\partial h_{\mathbf{j}}}{\partial t} = f_{n_h}\left(\frac{a_{\mathbf{j}}}{A_h}\right) - h_{\mathbf{j}} + D_h\Delta(h_{\mathbf{j}}) \quad \textbf{[S3]}$$

$$T_u\frac{\partial u_{\mathbf{j}}}{\partial t} = f_{n_u}\left(\frac{a_{\mathbf{j}}}{A_u}\right) - u_{\mathbf{j}} + D_u\Delta(u_{\mathbf{j}}). \quad \textbf{[S4]}$$

A. Disordered lattice and diffusion operator. The discretized diffusion operator Δ encodes the spacing of the underlying lattice. To create a lattice that mimics the 2D packing of cells in the imaginal disc epithelium, we start with a regular triangular lattice, in which each site has six nearest neighbors. With the lattice constant set to 1, the sites of the triangular lattice have positions of the form $\mathbf{r}_{\mathbf{j}}^{(0)} = j_x\hat{\mathbf{x}} + j_y(\hat{\mathbf{x}}/2 + \sqrt{3}\hat{\mathbf{y}}/2)$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal basis vectors and $\mathbf{j} = (j_x, j_y)$ is a pair of integers. The final, disordered lattice is then obtained by adding a random displacement to each of these vectors: $\mathbf{r}_{\mathbf{j}} = \mathbf{r}_{\mathbf{j}}^{(0)} + \eta_{\mathbf{j}}\hat{\mathbf{x}} + \xi_{\mathbf{j}}\hat{\mathbf{y}}$, where each $\eta_{\mathbf{j}}$ and $\xi_{\mathbf{j}}$ is an independent random variable chosen from a Gaussian distribution with mean = 0 and standard deviation = 0.15.

The set of $\mathbf{r}_{\mathbf{j}}$'s specifies a disordered lattice on which one can define a diffusion operator Δ. Qualitatively, a particle undergoing a discretized version of diffusion should hop randomly from one lattice site to other nearby sites, with no overall bias in the hopping direction. To give this idea concrete mathematical form, we first use a Delaunay triangulation [calculated with the Qhull algorithm (10) as implemented in Matlab] to define the nearest neighbors of each lattice site. We then allow hopping from a site to its neighbors with a rate that goes like the inverse square of the distance between the sites. This inverse square scaling is dictated by the basic physics of diffusion, in which the distance traveled goes like the square root of time; a rate, as the inverse of a time, should then go like 1/(distance)^2. Our choice of hopping rates leads immediately to the diffusion operator

$$\Delta(u_{\mathbf{j}}) = \sum_{\text{n.n.}\mathbf{k}} \frac{1}{|\mathbf{r}_{\mathbf{j}} - \mathbf{r}_{\mathbf{k}}|^2}(u_{\mathbf{k}} - u_{\mathbf{j}}), \quad \textbf{[S5]}$$

where n.n. stands for "nearest neighbors," and an identical expression holds for $\Delta(h_{\mathbf{j}})$. Each lattice site may also be associated with a polygon through the Voronoi tesselation dual to the Delaunay triangulation (10); it is these polygons that appear in the simulation results shown in Figs. 2–4 in the main text. We have considered an alternative assignment of hopping rates in which the diffusive flux between adjacent polygons is proportional to the length of the interface between them; this yields qualitatively similar results.

B. Simulation domain and initial conditions. Unless otherwise noted, we solve Eqs. **S1**–**S4** on a rectangular domain 120 cells long in the direction of MF propagation by 32 cells wide, with reflecting boundary conditions. As discussed in the main text, the initial conditions must include a template of R8 cells with high a and s levels that, through the u they synthesize, position the next column of R8s. This template is chosen to be either a set of isolated, roughly evenly spaced R8s or a uniform stripe, as appropriate (Fig. 2*B*). In the imaginal disc, the primary source of the Hh protein that ultimately drives MF motion is not the *ato*-expressing cells actually in the MF but differentiating cells posterior to the MF. To reflect this fact, we generally choose the parameters T_h and D_h to be large; the h field then averages the contributions of multiple columns of h-producing cells. One consequence is that several columns of cells with high a are initially required in our simulations for there to be enough h to start front motion. We thus use eight parallel columns of template as our initial condition. (Of these, only the first, however, has any significant effect on the u distribution that templates the pattern; the others are present solely to act as initial sources of h.) Their spacing in the direction of front motion is chosen to match the expected density of the pattern that will be dynamically generated by the simulations. Occasionally, our first guess for the source density in the initial condition is clearly wrong (i.e., it produces either too little h to start the front moving or so much h that the front outruns the pattern formation, leading to a field in which every cell is an R8) (11); we then adjust the density accordingly. In agreement with the pattern densities generated by our simulations, we found that an initial density of one column of R8s for every four columns of cells for the isolated R8, "hexagonal" template (Figs. 2*B* and 4 *A* and *C*), and one column of R8s for every five columns of cells for the uniform a, "stripe" template (Fig. S4*B* and Fig. 4 *B* and *D*), worked well. Within the region bounded by our eight initial columns of R8s, we initially set u to its steady-state value (given those R8s as sources) and h to the constant value it would have in a in an infinite domain with the same source density and $D_h \to \infty$.

C. Decoupling slow front motion and fast pattern formation. To be able to work on much smaller lattices when we are scanning many parameter sets or many realizations of disorder in two dimensions, in *SI Text*, sections S4*B* and S5 below we work in the limit of large D_h and T_h and separate pattern formation on the scale of

a few cells from the front motion on much larger scales. In this case, Eq. **S3** is replaced by the imposed h field (11)

$$h_{\mathbf{j}} = \begin{cases} 1 - \left(\frac{vT_h + c}{2c}\right) \exp\left[\frac{-vT_h + c}{2D_h}\left(y_{\mathbf{j}} - vt\right)\right] & y_{\mathbf{j}} < vt \\ \left(\frac{-vT_h + c}{2c}\right) \exp\left[\frac{-vT_h - c}{2D_h}\left(y_{\mathbf{j}} - vt\right)\right] & y_{\mathbf{j}} \geq vt, \end{cases} \quad \textbf{[S6]}$$

where $c = \sqrt{v^2 T_h^2 + 4D_h}$, $y_{\mathbf{j}}$ is the y component of $\mathbf{r}_{\mathbf{j}}$, the MF propagates in the positive y direction, and we have introduced the dimensionless paramater v giving the front speed. This expression arises as the solution to the continuum reaction–diffusion equation $T_h \partial_t h(y,t) = \Theta(vt - y) - h + D_h \partial_y^2 h$, where Θ is the Heaviside step function, valid when $\sqrt{D_h}$ is much larger than both the lattice spacing and the pattern period. The source term $\Theta(vt - y)$ can then be viewed as averaging over the small-scale spatial variation of a pattern of a expression growing with speed v. Eq. **S6** assumes a source density of unity; H should accordingly be increased from the value used for a lower density hexagonal or stripe template (see below). For this multiple-scales approximation to be exact in the limits $D_h \to \infty$ and $T_h \to \infty$, the speed v should in principle be chosen so that h reaches the threshold value to flip a cell from the low to the high (R8) state at exactly $y = vt$ (11), but in practice we instead use an estimate of v derived from one-dimensional simulations of the full model, Eqs. **S1**–**S4** (see next section).

D. Numerical methods. We solve Eqs. **S1**–**S4** with a program written by us in Matlab that implements a first-order Euler method in which the linear diffusive and decay terms are treated implicitly and the nonlinear source terms are treated explicitly. The loss of speed from the fact that this algorithm is only first order in time is more than compensated by the semi-implicit method's ability to take larger time steps and by the simplifications that arise because the terms treated implicitly together make a constant, linear operator. This operator can be Cholesky decomposed (12) into upper and lower triangular parts once at the outset of each simulation, and the inversions necessary for an implicit method are then quite fast.

When we do numerical parameter screens (*SI Text*, section S4), we find that the equations' characteristic time scales can vary quite substantially, and it is then important to have a well-designed and robust adaptive time-stepper. For these screens, we thus use Matlab's built-in explicit Runge–Kutta solver ode45 instead of our home-written program.

S3. Parameter Values

***A.* Prefered Parameter Set.** Table S1 lists the wild-type parameters used in this paper. The nondimensionalization of Eqs. **S1**–**S4** sets the characteristic time scale of a to unity. In these units, the parameters of Table S1 create a new column of R8s roughly every 120 time units; because a new column appears roughly every 120 min in real imaginal discs (13), 1 time unit corresponds to ~1 min. To create a pattern reminiscent of the *sca* mutant (Fig. 4 *C* and *D*), we increase T_u to 10. Importantly, this single-parameter change is sufficient to give the correct *sca*-like behavior with both a hexagonal and a stripe template.

***B.* The Choice of *H* and Feedback of the Pattern on *h* Levels.** As we note in the main text, the direct regulatory connection from a to h found in our model stands for a whole cascade of events leading from *ato* expression to neural differentiation and Hh secretion and thence to MF progress. This simplification allows for a concise model and emphasizes the positive feedback loop that makes self-sustained front motion possible. It can, however, lead to an unrealistically strong feedback between changes in R8 density and h levels. One can imagine two limiting cases. In the one, h is directly proportional to the density of a-expressing cells in the immediate vicinity of the MF; in the other, h is roughly independent of this density. We believe that the latter is the more biologically plausible choice. Indeed, Hh is secreted only by cells behind the MF that have begun neural differentiation in earnest and then by photoreceptors other than R8s (14). At the disc posterior, a genetically distinct mechanism initiates Hh expression, and this Hh is sufficient to move the MF some distance even in mutants that completely lack *ato* (15). The MF is moreover driven not only by Hh, but also by the morphogen Decapentaplegic (Dpp), which is expressed by all cells in the MF and does not depend directly on Ato, and by proneural Notch signaling (16, 17). Finally, in our experiments with the conditional *Notch* allele, the majority of Hh protein present after the return to the permissive temperature presumably comes from neural cells behind the MF that are not even in a stripe pattern. A quick examination of Eq. **S1** shows that the relevant quantity is not h but h/H. Thus, to implement this second option, we want to make h/H insensitive to the density of a-expressing R8 cells in the MF; this outcome can easily be accomplished by rescaling H in proportion to this source density. Because the density of R8 cells in different templates can vary quite substantially (e.g., a factor of 5 or 6 between our hexagonal and striped templates), one cannot safely ignore this subtlety, and we thus must use different H values for different templates. Specifically, the value $H = 0.0088$ quoted in Table S1 is used with a hexagonal template for both wild-type and *sca* simulations; in both cases, with a stripe template, we set $H = 0.0283$ (a value that reflects the presence of some breaks in the perfect stripe pattern).

S4. Numerical Parameter Screens and Robustness

To investigate how robust the switch and template mechanism is to changes in its parameters, we performed a systematic numerical screen of randomly chosen parameter sets in a large window around the preferred "wild-type" parameters of Table S1. Because our simulations are much faster in one dimension and it is much easier for a simple computer program to identify a regular one-dimensional pattern than a 2D pattern (11), we began by studying each parameter set in one dimension. Those sets that were able to make regular one-dimensional patterns were then further examined in two dimensions.

***A.* Screens in One Dimension.** In one dimension, we began by simulating each parameter set on a lattice of 400 evenly spaced cells with separation 1 and absorbing boundary conditions. Because some initial source of h is required to start front motion in our model, the first (posteriormost) 100 cells in this lattice were randomly and independently assigned with probability 1/2 to initially be either low ($a = s = 0$) or high ($a = 1 + F, s = 1$). We set $a = s = 0$ in the remaining 300 cells and $h = u = 0$ everywhere. We then simulated Eqs. **S1**–**S4** up to time $t = 4{,}096$, saving the variables at 12 logarithmically spaced time points. If, at $t = 4{,}096$ and ignoring the 25 lattice sites closest to the anterior end of the lattice, the last 5 cells to be flipped to high (defined, for the purposes of this screen, as any state with $a > 1 + F/10$) had a constant spacing between them and were each isolated (i.e., their nearest neighbors were both low), then the parameter set was considered likely to be able to generate a pattern in one dimension. To check that the observed patterns were not transients driven by an initial condition that provided too much or too little h, we ran a second simulation for those parameter sets that passed this first test. To perform this second simulation, we first estimated the MF speed by looking at how far a expression had progressed between the last two time points before the front reached the anterior limits of the domain in the initial simulation. Working on the same lattice of 400 cells and starting with all variables equal to zero, we then mimicked the progression of a regular pattern by flipping isolated cells to high with the observed spacing and moving at the estimated front

speed. We continued this procedure until the front had progressed through the first 200 lattice sites. During this period, we kept track of the evolution of our four variables, with the hand-flipped cells acting as sources, but did not allow h to activate any additional cells. Once the front had passed the 200th lattice site, however, we reverted to the normal dynamics, with the cells that had been forced on by hand now serving as a source of h to drive further front progression. We then ran the simulation long enough for 10 more cells to be flipped to high at the estimated front speed. If at least 3 cells with the expected spacing were indeed flipped in this time, we declared that the parameter set had succeeded in generating a regular pattern in one dimension. For future use, we recorded an improved estimate of the front speed v from the times that the first and third cells flipped and their positions. (We required that only 3 more cells be activated because our initial estimates of the front speed could be much higher than the true asymptotic speed when the pattern density was relatively low and the initial conditions in our first round of simulations thus provided a large burst of excess h.)

We screened two distinct samples of parameter sets using the above procedure. Both were based on the preferred parameter set of Table S1, modified in two ways to account for the transition from two to one dimension: First, D_h and D_u from Table S1 were each multiplied by a factor of 1.5 to ensure that the diffusion speed on long length scales (i.e., in the continuum limit) was the same on the one-dimensional lattice as on the (approximately) hexagonal lattice in two dimensions. Second, we rescaled H as described above to account for the difference in source density between our 2D hexagonal pattern, in which only one of every 32 cells is in the high state, and the stripe-like pattern the same parameters produce in one dimension, in which 1 of every 6 cells is high. In both cases, we kept Hill coefficients fixed but chose random values for all other parameters. In the first sample, the logarithm of each parameter was independently drawn from a flat distribution on an interval that stretched from a factor of 2 below the preferred parameter value to a factor of 2 above it. We chose 10,000 random parameter sets in this manner, of which 1,462 could make a regular pattern in one dimension by the criteria described above. The second sample of 40,000 random parameter sets was chosen in the same manner as the first, except that parameters were now allowed to range up to a factor of 4 above or below their reference values. Of these parameter sets, 3,119 could support one-dimensional pattern formation. In both random samples, all of the pattern-forming parameter sets, without exception, had bistable a–s subsystems at $h = 0$, and these bistable systems could all be flipped to the high state by h no larger than $2H$. Moreover, all patterning parameter sets from the 10,000-set sample and all but 9 from the 40,000-set sample had fast u, in the sense that a cell that had just been flipped to the high state could produce enough u to completely prevent its neighbor from flipping in a time $<1/v$, where v is improved estimated front speed described in the preceding paragraph. Thus, a remarkably large number of parameter sets within our high-dimensional parameter space can perform at least rudimentary one-dimensional pattern formation.

***B*. Parameter Performance in Two Dimensions.** To verify whether the ability to make a pattern on an ordered one-dimensional lattice translates into pattern formation on a disordered 2D lattice, we focused on the 1,462 patterning parameter sets from the first random sample within a factor of 2 of the reference parameter set. We simulated each of these parameter sets on 10 different random hexagonal lattices of length 12 cells in the direction of MF motion and width 16 cells perpendicular to MF motion with reflecting boundary conditions, using the model variant of *SI Text*, section S2*C*, in which an explicit spatiotemporal h profile is given and the focus is entirely on pattern formation by the a–s–u subsystem. (Because we already know that the parameter sets in question are able to form self-sustained fronts in one dimension using the full model, Eqs. **S1**–**S4**, and D_h is large enough that h in our simulations varies very little perpendicular to MF motion, we expect that any parameter set that can form a hexagonal pattern in two dimensions with the modified model can also support a moving front with the full model.)

To translate the one-dimensional parameter sets back into two dimensions, we set v in Eq. **S6** equal to the value estimated from the one-dimensional simulations and rescaled H by the ratio of the fraction of high cells in the one-dimensional pattern to that in the estimated 2D pattern. This estimated pattern was determined as follows: We first translated D_u into an exponential decay length for steady-state u on a hexagonal lattice of spacing 1. On the basis of this exponential decay, we then determined the distance from a point source of unit strength at which u has decayed to a value $U/2$, and we defined a basic interaction radius R for the parameter set as this distance + 0.5 (in our dimensionless length units). We expect that the linear separation between template R8s and newly created R8s in the next column will be about R; the added 0.5 is a rough correction for the fact that a new R8 feels inhibitor from two template cells rather than just one. To create a hexagonal pattern with a plausible aspect ratio, we separate template R8s by an average distance $4R/\sqrt{7}$ in the direction perpendicular to MF motion, which implies a predicted column separation $\sqrt{(3/7)}R$ in the direction of MF motion (Fig. S1*A*). More specifically, for each disordered 16×12 lattice, with the origin at the lower right corner and axes as shown in Fig. S1*A*, we place template R8s at the cells closest to coordinates $(8-2R/\sqrt{7}, 3.5)$ and $(8+2R/\sqrt{7}, 3.5)$. If R is small enough that $(8 \pm 4R/\sqrt{7}, 3.5)$ is still within the simulation domain, we also add template R8s at these points, but we measure only the templating efficiency of the central two R8s.

We then simulate the model, with initial conditions that $u = 0$ everywhere and that $a = s = 0$ except at the template R8 positions, starting from $t = -15/v$ (where Eq. **S6** implies that $t = 0$ when the center of the h front has reached the posterior edge of the simulation domain) and running until the h front has completely crossed the simulation domain. For each of the 10 random cell packings for each parameter set, we record the number of high cells with $a > 1 + F/10$ within a distance $R/2$ of the expected location for the new R8. If there is exactly one high cell in this region, indicating that the R8 in the new column has been templated correctly, we also record the number of cells neighboring this cell that have $a > 0.2$ at some point during the simulation run. We take the existence of such cells to indicate at least the transient presence of an IG. (An IG should ideally also include cells that are not nearest neighbors of the eventual R8, but transient activation of neighboring cells is at least a necessary condition for an IG.) In almost every case in which the R8 has not templated correctly, there are twinned R8s (or worse). Some summary statistics appear in Table S2. Clearly, a tendency to form twinned R8s is the greatest challenge to our model's robustness. If we require, for example, that twins occur in no more than 20% of ommatidia, then only 147 of the 1,462 parameter sets that work in one dimension are still acceptable. This is still >1% of the parameter sets tried in the original screen—quite a substantial fraction in a high-dimensional parameter space—but it is also a large drop-off from the one-dimensional performance.

S5. Noise and Disorder

In our model, noise or disorder, whether from chemical fluctuations, random cell packings, or some other source, has effects on two scales. First, locally on the scale of a single ommatidium, it can lead to a misplaced or twinned R8. Second, the effects of these template errors then can propagate through to subsequent columns. Given the already coarse-grained nature of our model, it is difficult to attach any clear molecular interpretation to the rate

of spontaneous errors within individual ommatidia, which thus becomes simply another unknown parameter. On the other hand, the manner in which errors propagate from one column to the next is specific to the switch and template mechanism and can in principle be understood within our existing deterministic model. In this section, we take a first step toward understanding these error effects by gathering quantitative simulation data on the consequences of errors in the template column for R8 placement and twinning rates in the next column. As in the previous section, we carried out simulations using the modified model of *SI Text*, section S2*C* on a 16×12 grid of cells (Fig. S1*B*). The x and y axes were defined as in Fig. S1*A*, and we took $\nu = 0.025$ and rescaled H appropriately, but otherwise used the parameters from Table S1. For each of 1,000 randomly generated cell packings, we chose as one template R8 cell the one located closest to $(x_1, y_1) = (3.5, 3.5)$ (Fig. S1*B*). We then in succession set every cell within a distance 1.6 of $(x, y) = (11.5, 3.5)$ to be the other template R8 and simulated the emergence of the next R8 column; this procedure allowed us to study how the position of the new R8 depends on the location of the two template cells that most affect it. To examine how much R8 twinning perturbs subsequent columns, we also ran one simulation on each of the 1,000 random grids in which the template cell at ~(11.5, 3.5) was replaced by a pair of R8s, chosen to minimize the product of the distance between the two twinned R8 cells, the distance of their center of mass from (11.5, 3.5), and the sine of the angle between a line joining them and a line perpendicular to the direction of MF motion. (Thus, in essence, we chose a twinned pair of adjacent cells near the optimal position and oriented roughly perpendicular to the direction.) In each case, after the simulation had run its course, we recorded the positions of all cells that had been flipped to the high state within a circle of radius 3 centered on the position where the next R8 would appear in a perfect lattice, $(7.5,\ 3.5 + 4\sqrt{3}/2)$. In all, we ran 9,533 simulations with two single R8s as the template (or an average of 9.5 choices for the location of the second template cell for each of the 1,000 random cell packings), of which 8,611 gave rise to a single R8 in the next column, 882 led to twinned R8s in the next column, and the remained had either more than two R8s or (in two cases) no R8s at all. Similarly, of the 1,000 simulations with one single template R8 and one pair of twinned template R8s (one for each of the 1,000 random cell packings), 930 created a single R8 cell in the next column, 69 a pair of twinned R8s, and 1 more than two R8s.

Fig. S2 shows some results of these simulations. A full analysis of the data is quite challenging. For example, a single error in one column affects two ommatidia in the next column, which, in turn, can then each affect the same ommatidium located midway between them in the third column, and so on; indeed, if we view the column number as akin to a time variable, then it is easy to see that in the most general case the interaction of errors in our system must map onto a statistical–mechanical model of an interacting, driven, one-dimensional system. Moreover, even our relatively extensive simulations did not succeed in fully sampling the tails of the relevant distributions. Nonetheless, we were able to draw some rough conclusions at the level of a sort of mean field approximation. Broadly speaking, we find that the effects of errors both in R8 positioning and in twinning tend to decay from column to column. For example, even if a template R8 cell is displaced by a full cell diameter from its optimum position, the probability that a twin will result in the next column is <10% (Fig. S2*B*). Similarly, only 69 of 1,000 simulations with one template cell twinned led to a twin in the next column, so a well-placed twinned R8 in one column does not beget a twin with any higher probability than does a normal, single-template R8 cell. And when the twin leads to an isolated R8 in the next column, this R8 is on average displaced from its optimum position by only about a third of cell diameter (Fig. S2*F*). Although in any probabilistic system there is always some chance of a rare event (or confluence of events) that completely destroys the template, for small error rates our model thus usually does not experience a catastrophic accumulation of small perturbations.

S6. Mathematical Properties of the Model

***A*. Bistability and R8 Selection.** The preferred wild-type parameter set of Table S1 has several qualitative features that are important to the functioning of our model, of which the foremost is that the cell-autonomous a–s subsystem acts as a bistable switch. Specifically, we can view the inputs from the diffusible factors h and u as a bifurcation parameter

$$\mu_{\mathbf{j}} = G f_{m_h}\left(\frac{h_{\mathbf{j}}}{H}\right)\left[1 - f_{m_u}\left(\frac{u_{\mathbf{j}}}{U}\right)\right]. \qquad \textbf{[S7]}$$

At fixed μ, the a–s subsystem in a given cell evolves according to

$$\frac{\partial a}{\partial t} = f_{n_a}\left(\frac{a}{A_a}\right) - a + F f_{m_s}\left(\frac{s}{S}\right) + \mu \qquad \textbf{[S8]}$$

$$T_s \frac{\partial s}{\partial t} = f_{n_s}\left(\frac{a}{A_s}\right) - s, \qquad \textbf{[S9]}$$

where we have dropped the subscript **j** because we are looking at only one cell. Fig. S3 *A*–*C* plots the nullclines of this system for the parameters of Table S1 and three different values of μ. One can easily see that there is a saddle node bifurcation as a function of μ. For $\mu = 0$, the system has three fixed points, but as μ is increased to a critical value $\mu_c = 0.265$, the two lower fixed points merge, so that for $\mu > \mu_c$ only the high fixed point remains. This behavior is further illustrated in the bifurcation diagram of Fig. S3*D*, which shows allowed steady-state values of a as a function of μ.

***B*. Pattern Multistability.** The cartoon mechanism of Fig. 2*D* in the main text suggests that, depending on the spacing of the R8 cells in the initial template, the same model parameters should sustain multiple different patterns: If one moves the R8s in a given column closer together while maintaining the radius of the inhibitory discs, a new, but still regular, pattern should result, with a large distance between successive columns to compensate for the closer packing within columns (Fig. S5). Simulations of our model bear out this prediction: Fig. S5*B* shows two simulations with the wild-type parameters of Table S1, but with an initial spacing between R8s along a column of 5.2 in the *Left* panel and of 8 in the *Right* panel. These spacings are maintained as the pattern grows.

S7. Model Variants

***A*. Inhibiting *a*'s Self-Activation.** The switch and template pattern formation mechanism requires that the inhibitor u be able to block cells from flipping from the low to the high state, but that the same u levels not be enough to force high cells to flip back down to the low state. This is most easily accomplished when u directly inhibits activation by h while leaving the core a–s positive feedback loop unaffected, but is also possible when u inhibits a's autoactivation. To demonstrate this, we considered a variant model in which Eq. **S1** is replaced by

$$\frac{\partial a_{\mathbf{j}}}{\partial t} = f_{n_a}\left(\frac{a_{\mathbf{j}}}{A_a}\right)\left[1 - f_{m_u}\left(\frac{u_{\mathbf{j}}}{U}\right)\right] - a_{\mathbf{j}} + F f_{m_s}\left(\frac{s_{\mathbf{j}}}{S}\right) + G f_{m_h}\left(\frac{h_{\mathbf{j}}}{H}\right). \qquad \textbf{[S10]}$$

We found that this model can generate a hexagonal pattern of a expression behind a moving front. For example, Fig. S4*C* shows the results of simulation of this modified model with most of the parameters the same as in Table S1, but with the changes $A_a = 0.4$, $A_u = 2.4$, $G = 0.28$, $F = 1.8$, and $T_u = 1$.

***B*. Alternative *sca*-Like Parameters.** In the main text, we show results from simulations with T_u increased from 2 for the wild-type

parameter set of Table S1 to 10 and argue that the results are very similar to those seen for the *sca* phenotype. Although seeing *sca*-like behavior when only T_u is changed makes it very clear that the cause is a change in the response time of u, not in the steady-state strength of inhibition, it is not so likely that a real mutation would have precisely the effect of changing T_u and nothing else. In fact, any parameter change whose major effect is to slow inhibition will have qualitatively similar effects. For example, Fig. S4*E* shows the results of simulations with the parameters of Table S1 changed so that $U = 7 \times 10^{-6}$, $D_u = 0.48$, and $m_u = 2$. As is evident from Fig. S4*E*, these parameter sets also give a striped pattern when simulations are initiated with a stripe template.

1. Firth LC, Li W, Zhang H, Baker NE (2006) Analyses of RAS regulation of eye development in *Drosophila melanogaster*. *Methods Enzymol* 407:711–721.
2. Mlodzik M, Baker NE, Rubin GM (1990) Isolation and expression of *scabrous*, a gene regulating neurogenesis in *Drosophila*. *Genes Dev* 4:1848–1861.
3. Foster GG, Suzuki DT (1970) Temperature-sensitive mutations in Drosophila melanogaster. IV. A mutation affecting eye facet arrangement in a polarized manner. *Proc Natl Acad Sci USA* 67:738–745.
4. Tomlinson A, Ready DF (1987) Neuronal differentiation in *Drosophila* ommatidium. *Dev Biol* 120:366–376.
5. Lee EC, Hu X, Yu SY, Baker NE (1996) The *scabrous* gene encodes a secreted glycoprotein dimer and regulates proneural development in *Drosophila* eyes. *Mol Cell Biol* 16:1179–1188.
6. Jarman AP, Grell EH, Ackerman L, Jan LY, Jan YN (1994) *Atonal* is the proneural gene for *Drosophila* photoreceptors. *Nature* 369:398–400.
7. Nolo R, Abbott LA, Bellen HJ (2000) Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in *Drosophila*. *Cell* 102:349–362.
8. Jennings B, Preiss A, Delidakis C, Bray S (1994) The Notch signalling pathway is required for *Enhancer of split* bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development* 120:3537–3548.
9. Baker NE, Yu S, Han D (1996) Evolution of proneural *atonal* expression during distinct regulatory phases in the developing *Drosophila* eye. *Curr Biol* 6:1290–1301.
10. Barber CB, Dobkin DP, Huhdanpaa HT (1996) The quickhull algorithm for convex hulls. *ACM Trans Math Softw* 22:469–483.
11. Pennington MW, Lubensky DK (2010) Switch and template pattern formation in a discrete reaction-diffusion system inspired by the *Drosophila* eye. *Eur Phys J E Soft Matter* 33:129–148.
12. Press WH, Teukolsky SA, Vettering WT, Flannery BP (1992) *Numerical Recipes in C* (Cambridge Univ Press, Cambridge, UK).
13. Wolff T, Ready DF (1993) Pattern formation in the *Drosophila* retina. *The Development of Drosophila melanogaster*, eds Bate M, Arias AM (Cold Spring Harbor Lab Press, Plainview, NY), pp 1277–1325.
14. Rogers EM, et al. (2005) Pointed regulates an eye-specific transcriptional enhancer in the *Drosophila hedgehog* gene, which is required for the movement of the morphogenetic furrow. *Development* 132:4833–4843.
15. Domínguez M, Hafen E (1997) Hedgehog directly controls initiation and propagation of retinal differentiation in the *Drosophila* eye. *Genes Dev* 11:3254–3264.
16. Baonza A, Freeman M (2001) Notch signalling and the initiation of neural development in the *Drosophila* eye. *Development* 128:3889–3898.
17. Fu W, Baker NE (2003) Deciphering synergistic and redundant roles of Hedgehog, Decapentaplegic and Delta that drive the wave of differentiation in *Drosophila* eye development. *Development* 130:5229–5239.

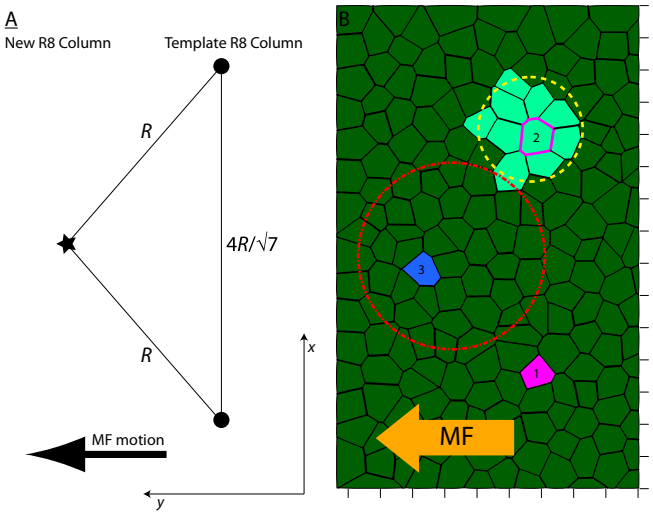


Fig. S1. (*A*) Template setup for 2D simulations as part of the numerical parameter screen (*SI Text*, section S4*B*). The preferred distance R between successive R8s is estimated from one-dimensional simulations as described in *SI Text*. Once R is known, two template cells (circles) are spaced a distance $4R/\sqrt{7}$ from each other along a line perpendicular to the direction of MF motion. This template then is expected to produce a new R8 in the next column at roughly the position indicated by the star. Axis orientation: x axis perpendicular to the direction of MF propagation, y axis parallel to the direction of MF propagation. (*B*) Procedure for simulations testing the effect of noise in template position on the position of the next R8 (*SI Text*, section S5). 1: On a 16×12 array of randomly packed cells, the cell closest to (3.5, 3.5) (purple) is chosen as one of the two template R8s. 2: Every cell within a circle of radius 1.6 (yellow dotted line) centered on (3.5, 11.5) is considered as a possible choice for the other template R8 (light green). 3: For each choice of a second template cell (e.g., purple outline), the system is simulated, and the position of any new R8s (e.g., blue cell) within a circle of radius 3 (red dashed-dotted line) centered on the ideal position of an R8 in the new column is recorded.

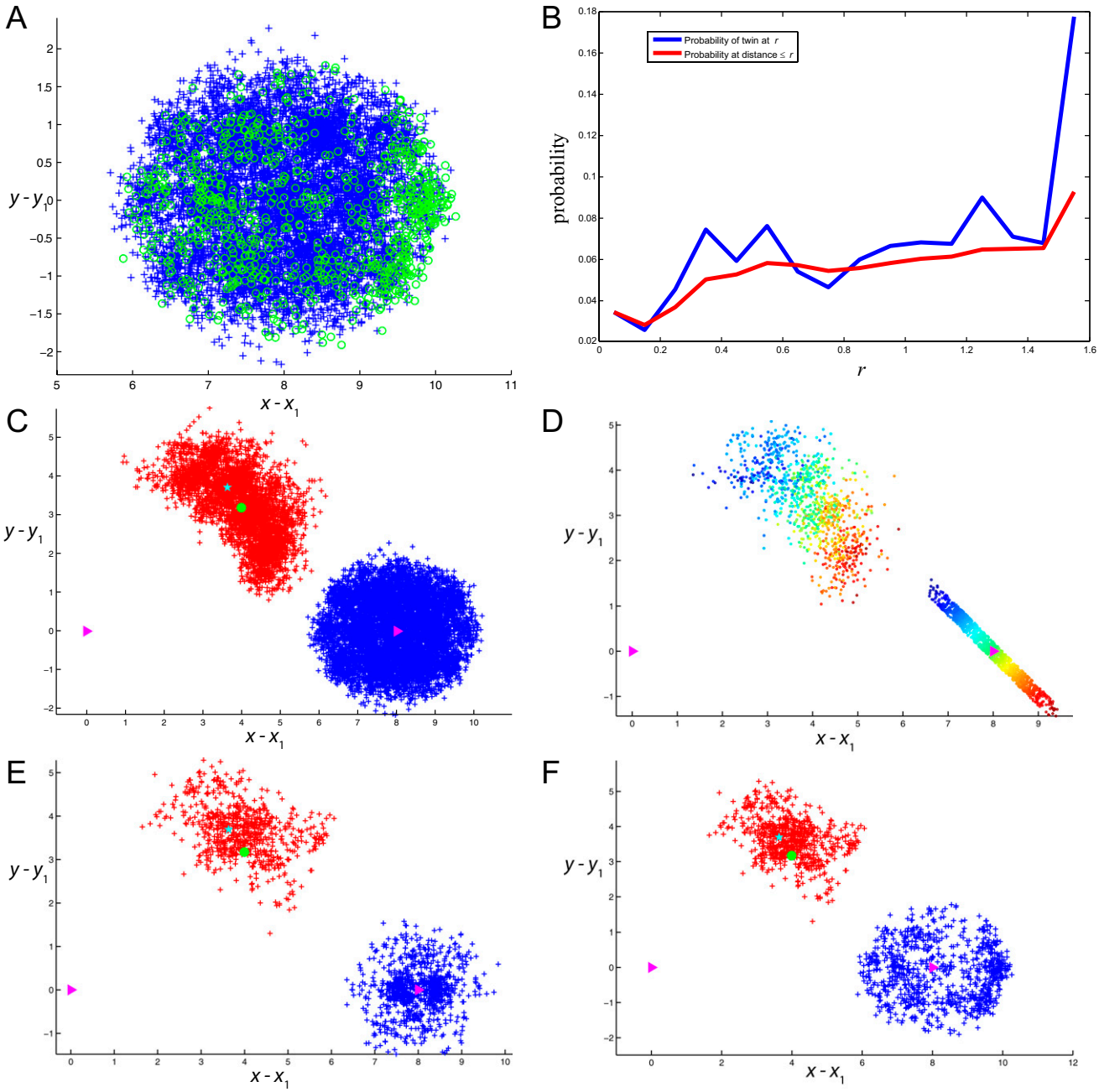


Fig. S2. Propagation of template errors (*SI Text*, section S5). (*A*) Positions (*x*, *y*) of all 9,533 choices for the second template cell (purple outline, Fig. S1*B*), plotted as a distance relative to the position (x_1, y_1) of the first template cell (purple cell, Fig. S1*B*). Blue crosses mark the 8,611 template cells that yielded a single R8 in the next column, and green circles show the 882 template cells that yielded twinned R8s in the next column. Not surprisingly, twins are most likely when the second template cell is farther from the optimum position, in particular when the spacing between the two template cells is too large. (*B*) Probability that a template cell will cause a twinned R8 as a function of its distance *r* from its ideal position (*x*, *y*) = (3.5, 11.5). Blue, fraction of template cells at a given distance (bins of size 0.1) that produce twins; red, fraction of all cells anywhere within a disc of radius *r* that produce twins. The rate of twinning remains <10% until *r* approaches 1.5. (*C*) Positions of single template R8 cells and the their images. The axes give coordinates relative to the position (x_1, y_1) of the first template R8, indicated by the left purple triangle. Blue crosses, positions of each of the 8,611 choices for the template R8s that lead to a single R8 in the next column. Red crosses, positions of each of these 8,611 new R8s. Right purple triangle, ideal position (3.5, 11.5) for the second template cell. Green circle, position the newly created R8 would have in a perfectly regular lattice. Light blue star, average position of new R8s when the second template R8 is replaced by a pair of twinned R8s. Because the twin produces more inhibitor than the first template R8, this average position is displaced away from the twin. (*D*) Subset of the data from *C*, showing how the position of the R8 in the new column correlates with the position of the second template R8 in the previous column. Only template R8s within a diagonal band are shown (lower right), and they are given a graded color to indicate their position along this band. The points marking the positions of the newly created R8s (more diffuse cluster, upper half of the plot) are labeled with the color corresponding to their template R8. (*E*) Same as *C*, except the blue crosses give the average position of the two cells in each of the 930 pairs of twinned R8s that gave rise to a single R8 in the next column, and the red crosses give the positions of these single R8s. (*F*) Same as *C*, except the blue crosses are the positions of the 882 single R8s that caused a twinned R8 to appear in the next column, and the red crosses give the mean position of the two cells in each pair of twins so created.

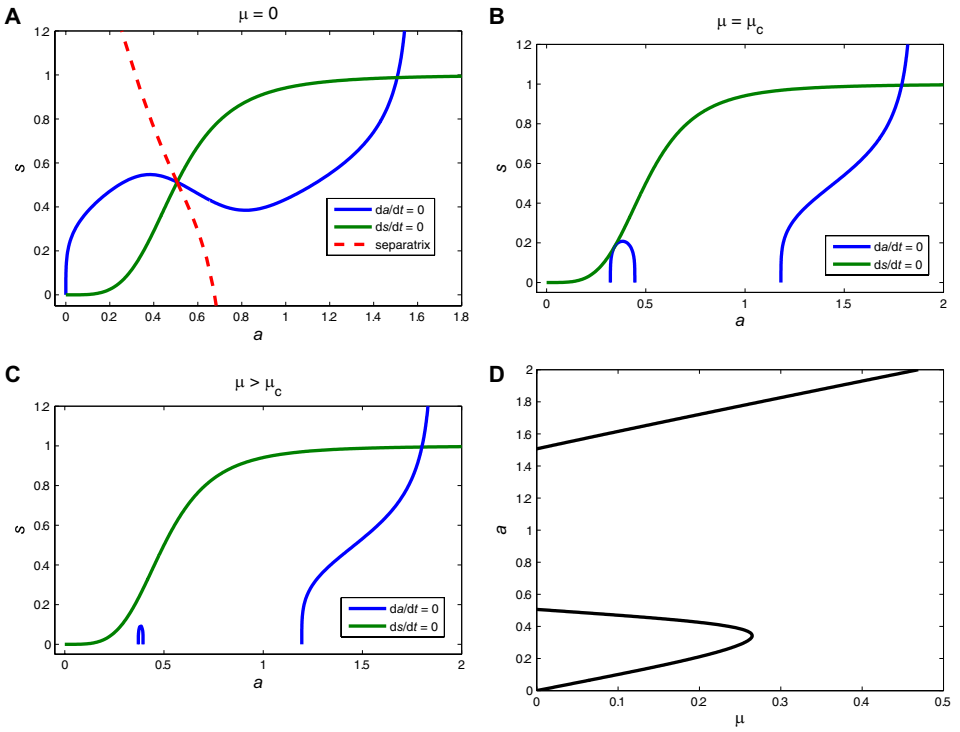


Fig. S3. Nullclines $\partial a/\partial t = 0$ (blue) and $\partial s/\partial t = 0$ (green) in the *a*–*s* plane of the cell-autonomous system of Eqs. **S8** and **S9** for $\mu = 0$ (*A*), $\mu = \mu_c = 0.265$ (*B*), and $\mu = 0.275 > \mu_c$ (*C*). The red dashed line in *A* gives the separatrix that divides the basins of attraction of the high and low stable fixed points. (*D*) Bifurcation diagram for the *a*–*s* cell-autonomous subsystem. Fixed point values of *a* are plotted as a function of the bifurcation parameter μ that gives the net input from the diffusible signals *h* and *u*. Note that the high *a* fixed point persists for any μ. Parameters are as in Table S1.

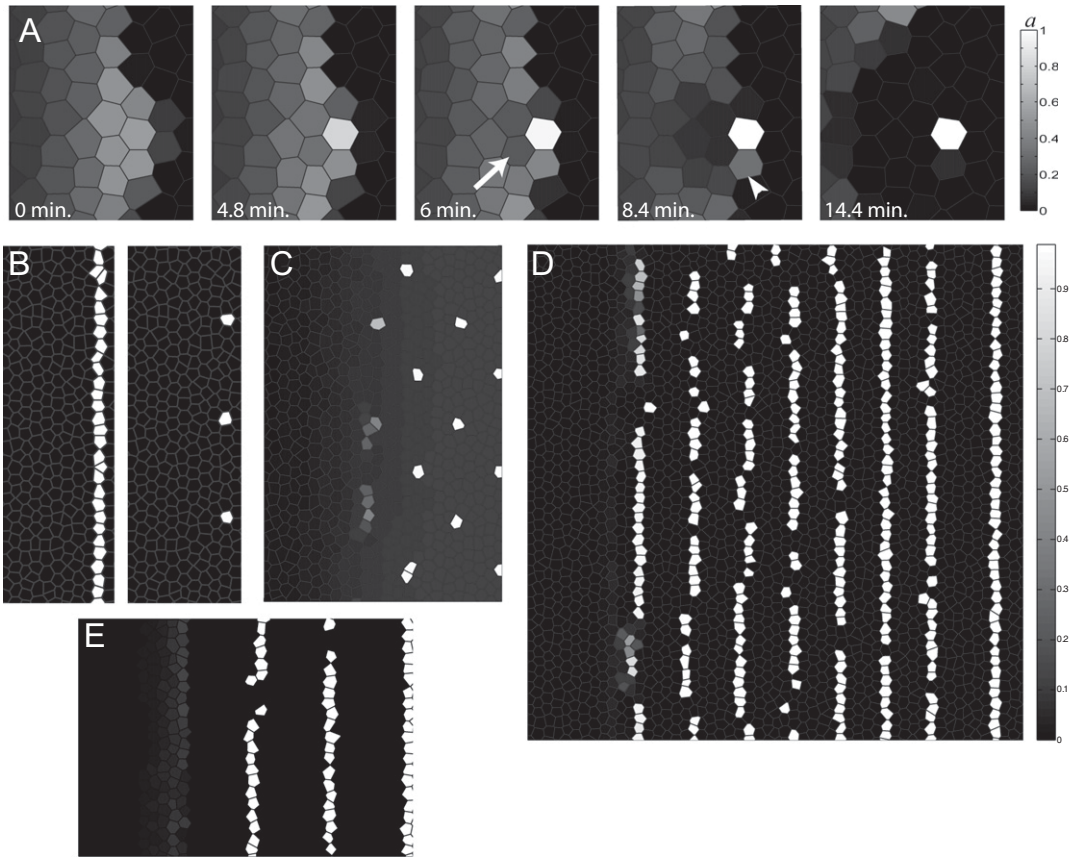


Fig. S4. (*A*) Evolution of *a* expression pattern during intermediate group (IG) resolution. To make differences within the IG clearly visible, grayscale does not cover the entire dynamic range of *a*. (*a* = 1.57 in the R8 in the rightmost panel.) Time estimated assuming 1 dimensionless simulation time unit corresponds roughly to 1 min; wild-type simulation parameters are as in Table S1. From *Left* to *Right*: initial IG, immediately before resolution begins, with no clearly visible R8; then three stages of IG resolution; and then the final state, with a single R8 selected. During resolution, cells (arrow) immediately to the anterior of the presumptive R8 turn off before those to either side, leading to characteristic arm-like intermediates (arrowhead). (Compare, e.g., figure 1d of ref. 9.) (*B*) Representative initial conditions for *a* and *s* used for simulations with a stripe-like template (*Left*) or a "hexagonal" template of isolated R8 cells (*Right*). In either case, at the start of the simulations, the white cells have $a = s = 1$ whereas the black cells have $a = s = 0$. To produce enough *h* to initiate front motion, multiple parallel columns of this form are normally required (see text), although only the anteriormost serves a genuine templating function. (*C*–*E*) *s* in three additional situations (grayscale in *D* applies to *C*–*E*). (*C*) Inhibition of *a* self-activation instead of activation by *h* (*SI Text*, section S7*A*). (*D*) Simulation of stripe parameters on a larger domain. Some breaks in stripes do appear, but these breaks can also be healed in subsequent columns. The same parameters are used as in Fig. 4 *C* and *D*. (*E*) Alternative stripe-forming parameters (*SI Text*, section S7*B*).

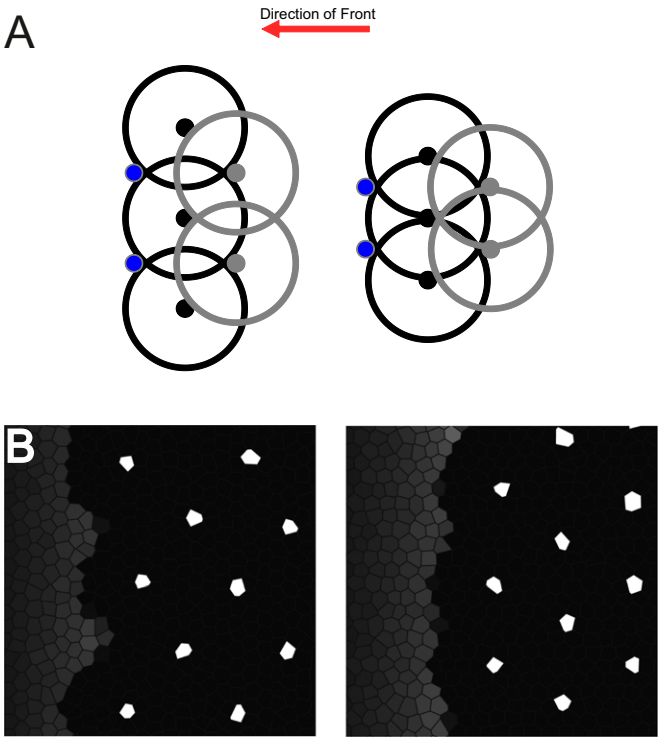


Fig. S5. (*A*) Schematic following Fig. 2*D*: Different templates can yield different patterns for the same parameter values. With the radius of the inhibitory disc held constant, moving the R8s within the template column closer together (*Right*) causes the spacing between columns to increase, yielding a regular pattern with a different aspect ratio. (*B*) (*Left* and *Right*) *a* expression in patterns of two different aspect ratios generated with the same parameters. The two simulations were done with the wild-type parameters of Table S1, but with different initial templates. *B*, *Right* has more R8s per vertical column, but more space between these columns; conversely, the R8s are more separated within each column in *B*, *Left*, but the columns are closer together.

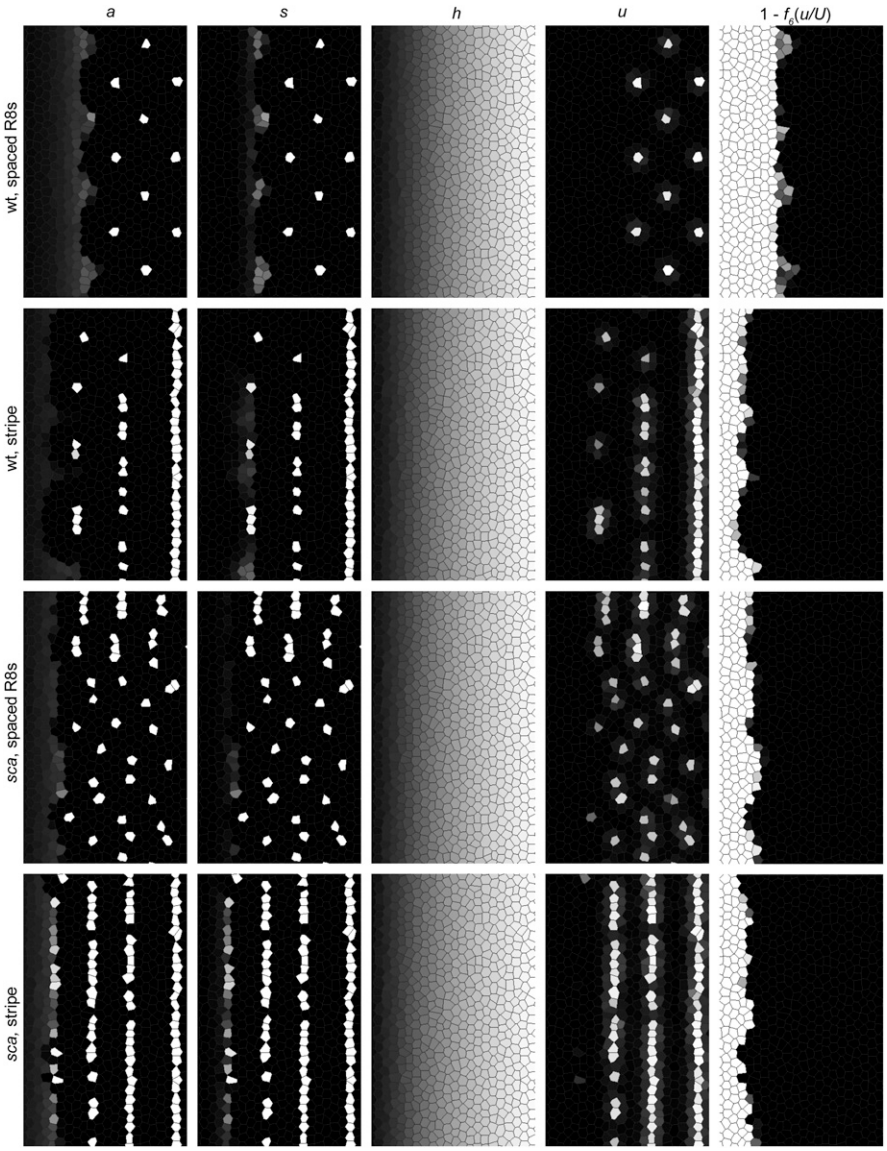


Fig. S6. Expression levels for all four fields in our model corresponding to the snapshots in Fig. 4 of the main text. The first four columns, from left to right, show the levels of *a*, *s*, *h*, and *u*. (The four panels of the *s* column match the four panels of Fig. 4.) In our model, cells are sensitive to relatively small amounts of *u* that are not easy to detect directly with the sort of grayscale plots used here. In the fifth column, we thus plot the *u*-dependent inhibitory term $1 - f_6(u/U)$ (Eq. **S1**), which gives a more direct measure of how inhibition varies across the eye field. Note that in this column dark colors correspond to inhibition and light colors to the absence of inhibition. The four rows correspond, in order, to simulations done with the wild-type parameter set (Table S1) and a template of isolated R8s (as in Fig. S4*B*), to the wild-type parameter set with a stripe-like template (Fig. S4*B*), to the *sca* parameter set with a template of isolated R8s, and to the *sca* parameter set with a stripe-like template. Throughout, the grayscale is normalized so that white cells have the maximum value of whatever variable is being plotted.

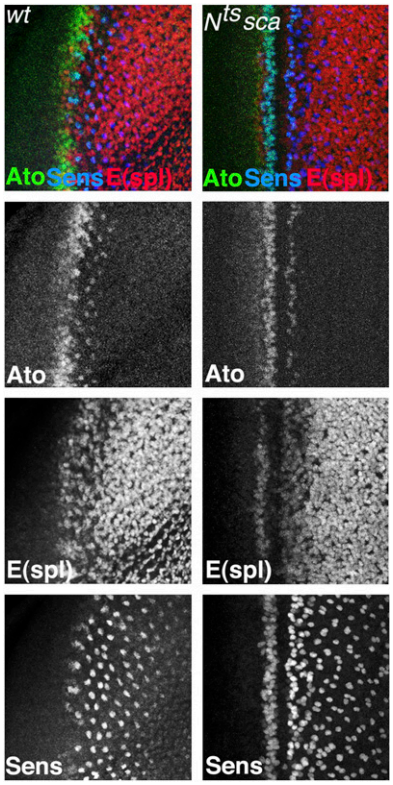


Fig. S7. (*Left*) Normal patterning in wild type (exposed to 31.5 °C for 2 h and 18 °C for 14 h). Merged channels show the overlapping expression of the transcription factors Atonal (green), Senseless (blue), and E(spl) (a readout of Notch signaling; red). Individual channels are shown separately below. This is the same preparation shown in Fig. 4*H*. (*Right*) N^{ts}; *sca* exposed to 31.5 °C for 2 h and 18 °C for 14 h. Merged channels show Atonal (green), Senseless (blue), and E (spl) (red) and separate channels are shown below. The disordered *sca* phenotype is replaced by a stripe of R8s. This is the same preparation shown in Fig. 4*G*.

Table S1. Parameter values used for simulations of wild-type imaginal discs

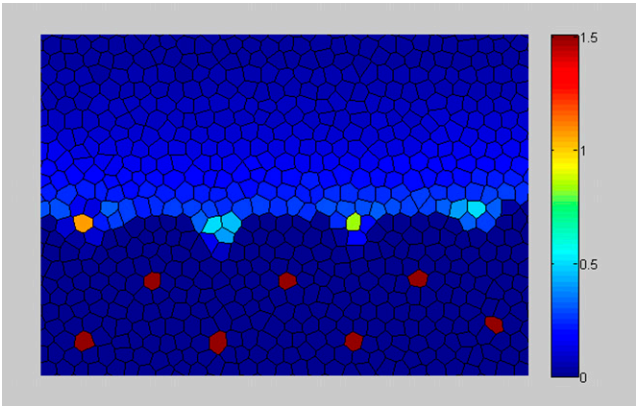
A_a	0.65	n_a	4	U	4×10^{-6}	m_u	6
A_s	0.5	n_s	4	H	0.0088	m_h	4
A_h	1.5	n_h	8	G	0.8	F	0.6
A_u	2.2	n_u	8	D_h	200	D_u	0.16
S	0.57	m_s	4	T_h	101	T_u	2
T_s	4						

With these parameters, the MF makes a new column roughly every 120 time units, implying that 1 dimensionless time unit corresponds to ~1 min of real time. Our simulations of *sca* mutant discs leave all parameters the same except for T_u, which is set to 10. As explained in *SI Text*, the parameter H is also adjusted so that variations in source density between different patterns do not lead to spurious results.

Table S2. Accuracy of 2D patterning by parameter sets able to produce a regular one-dimensional pattern

n	R8	1 IG cell	1 IG cell per R8
1	1,084	846	846
2	917	729	614
3	783	612	437
4	629	487	309
5	473	361	203
6	348	263	122
7	248	184	71
8	147	105	34
9	68	42	13
10	24	17	7

The columns list the number of patterning parameter sets (of a total of 1,462; *SI Text*, section S4*B*) for which at least *n* of the 10 random cell packings produced a single R8 in the correct position, produced IG-like activation in at least one cell adjacent to at least one R8, and produced IG-like activation in at least one cell adjacent to all successful single R8s.



Movie S1. Time-dependent *a* levels during the establishment of a new column of R8 cells. The color in each cell gives the instantaneous *a* level in that cell, as indicated by the scale bar at right.

Movie S1