Supplemental Materials and Methods

Analysis of the Simplified Model

Our simplified model begins as Eqs 1-4. The next step is to non-dimensionalize the equations by applying the following transformations:

$$n \equiv \frac{C_{Dl,nuc}}{C_{Dl}^0}, \qquad u \equiv \frac{C_{Dl,cyt}}{C_{Dl}^0}, \qquad w \equiv \frac{C_{DC}}{C_{Dl}^0}, \qquad v \equiv \frac{C_{cact}}{C_{cact}^0},$$

where C_{Dl}^0 is the initial concentration of Dl in the embryo, and $C_{cact}^0 \equiv V_{Cact}/k_{deg}$. After transforming the state variables into their non-dimesionalized forms in this manner, we arrive at the following equations:

$$\frac{\partial n}{\partial t} = \hat{k}_{in}u - \hat{k}_{out}n, \qquad (S1).$$

$$\frac{\partial u}{\partial t} = \hat{D}_{dl}\frac{\partial^2 u}{\partial x^2} - \hat{k}_{in}u + \hat{k}_{out}n - \hat{k}_{bind}uv + \hat{k}_{diss}(x)w, \qquad (S2).$$

$$\frac{\partial w}{\partial t} = \hat{D}_{DC}\frac{\partial^2 w}{\partial x^2} + \hat{k}_{bind}uv - \hat{k}_{diss}(x)w, \qquad (S3).$$

$$\tau_{Cact}\frac{\partial v}{\partial t} = \hat{D}_{Cact}\frac{\partial^2 v}{\partial x^2} + 1 - v - \mu(\hat{k}_{bind}uv - \hat{k}_{diss}(x)w), \qquad (S4).$$

In these equations, time is measured in minutes and space is measured in embryo DV axis lengths. Therefore, the following dimensionless constants appear in the equations:

$$\hat{k}_{in} \equiv k_{in}T, \qquad \hat{k}_{out} = k_{out}T, \qquad \hat{D}_{Dl} \equiv \frac{D_{Dl}T}{L^2}, \qquad \hat{D}_{DC} \equiv \frac{D_{DC}T}{L^2}, \qquad \hat{D}_{Cact} \equiv \frac{D_{Cact}}{k_{deg}L^2}$$

$$\hat{k}_{bind} \equiv k_{bind}C_{Cact}^0T, \qquad \hat{k}_{diss}(x) \equiv k_{diss}(x)T, \qquad \tau_{Cact} \equiv (k_{deg}T)^{-1}, \qquad \mu = \frac{C_{Dl}^0T}{C_{Cact}^0k_{deg}},$$

where $T = 1$ min and $L = 280$ μm (the length of the DV axis from ventral midline to dorsal midline).

These equations are simplified in the following manner. First, as Cact has a high turnover rate, we assume that τ_{Cact}, $\hat{D}_{cact}$, and μ are small compared to unity. This results in a spatially uniform, pseudo steady state for free Cact, or $v = 1$. Next, we assume that nuclear import and export are in pseudo equilibrium, so that $n \approx \hat{k}_{out}u/\hat{k}_{in}$, or $n \approx K_{eq}u$, where $K_{eq} \equiv k_{in}/k_{out}$ is the equilibrium constant for net nuclear import of Dl. If we then sum equations (S1) and (S2), we arrive at:

$$\frac{\partial}{\partial t}(u + K_{eq}u) = \hat{D}_{Dl}\frac{\partial^2 u}{\partial x^2} - \hat{k}_{bind}uv + \hat{k}_{diss}(x)w,$$

or,

$$\frac{\partial u}{\partial t} = \frac{1}{1 + K_{eq}}\left[\hat{D}_{Dl}\frac{\partial^2 u}{\partial x^2} - \hat{k}_{bind}u + \hat{k}_{diss}(x)w\right], \qquad (S5).$$

Note that v has dropped out of the equation. It should also be noted that the criterion for shuttling to occur in this model is now more clear. The ratio of $\hat{D}_{cact}$ must be greater than effective diffusivity in Eqn

S5, which is $\widehat{D}_{Dl}/(1 + K_{eq})$. Together, these two simplifications reduce our system to two differential equations, one for cytoplasmic Dl (u; Eq S5), and one for cytoplasmic Dl/Cact complex (w; Eq S3).

Formulation of Toll saturation term

According to Fig. 4G, the dissociation of Dl/Cact complex can be expressed in the following steps:

$$Dl/Cact + Toll \xrightleftharpoons{k_{on}/k_{off}} [Toll - Dl/Cact] \xrightarrow{k_r} Dl + Cact + Toll.$$
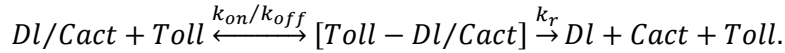
This reaction scheme is formally identical to the traditional Michaelis-Menten reaction scheme. In the case in which Toll levels are limiting (and roughly constant in total), if a pseudo steady state is assumed about the intermediate tripartite complex, the resulting expression for the rate of dissociation is:

$$r_{diss} = k_{diss}(x)\frac{w}{K_R + w},$$

where $K_R = (k_{off} + k_r)/k_{on}$ and $k_{diss} = k_r[Toll]_{tot}$, and $[Toll]_{Tot} = [Toll] + [Dl/Cact/Toll]$ (assumed roughly constant). Therefore, this Toll saturation regime replaces the final terms of Eqs S3 and S5 with

$$\widehat{k}_{diss}(x)w \rightarrow \widehat{k}_{diss}(x)\frac{w}{K_R + w}.$$

Random parameter screen

Parameters were varied randomly on a log space within reasonable constraints, and each parameter set was scored to determine whether shuttling was taking place. Both diffusivities were varied from 10^{-3} to 10^3 (see Parameter Estimation subsection), $\widehat{k}_{bind}$ was varied from 10^{-1} to 10^2, $\widehat{k}_{diss}$ was varied from 10^{-4} to 10^2, and K_R was varied from 10^{-2} to 10^2. The lower limit for $\widehat{k}_{bind}$ was higher than for other parameters because the strength of the Dl/Cact interaction needed to high enough to prevent Dl from entering the nuclei everywhere. The equilibrium constant was held fixed at $K_{eq} = 4$, which roughly reflects the ratio of nuclear to cytoplasmic Dl levels near the ventral midline, and ϕ was held fixed at 0.15, which roughly reflects the width of wildtype Dl gradients.

The metric to determine if a parameter set results in shuttling behavior is given in Eqn 5 in the Materials and Methods. Additionally, to score whether a parameter set results in widening of the gradient when diffusion is lowered, the same small fold-change perturbation was made to both diffusivities, and the width of the gradient was measured as the location where the gradient passes 60% max. If this width was higher than the width measured for the wildtype (non-diffusivity perturbed) gradient, then the parameter set was scored as positive for diffusion-based widening.

A parameter set was scored as positive for a split peak when the dosage of Dl was halved when the maximum of the gradient did not occur at $x = 0$, but did occur there for the wildtype (non-dosage perturbed) gradient. The same procedure was used for Toll-domain based peak splitting, for when ϕ was doubled.

Analysis of the Full Model

We also analyzed a more complete model of Dl/Cact interactions, which is an extension our previously-published mathematical model (O'Connell and Reeves, 2015). Here we explicitly account for both active Toll receptors and active Toll bound to Dl/Cact complex (Eqns S12 & S13, respectively) , which allows for the possibility that Toll activity is limiting. The full model of Dl/Cact/Toll interactions consists of a set of eight differential equations for each nucleus h (below). The equations describe the nuclear/cytoplasmic exchange of Dl, Cact and Dl/Cact complex, as well as the interactions between Dl, Cact, Dl/Cact complex and Toll receptors (see also Figs. 2A, 4G). The equations are simulated over nuclear cycles (NCs) 10-14 and for each nucleus $h = 1 \dots M_{NC}$, where M_{NC} is the number of nuclei in the given NC, as previously described (O'Connell & Reeves 2015). Thus, there are a total of $8M_{NC}$ equations during each nuclear cycle. By NC, $M_{10} = 13, M_{11} = 19, M_{12} = 26, M_{13} = 36, M_{14} = 51$.

$$V_n \frac{d}{dt}[Dl]_{nuc}^h = k_{in,Dl}A_n[Dl]_{cyt}^h - k_{out,Dl}A_n[Dl]_{nuc}^h - k_b V_n[Dl]_{nuc}^h[Cact]_{nuc}^h + k_d V_n[Dl/Cact]_{nuc}^h$$

(S6)

$$V_c \frac{d}{dt}[Dl]_{cyt}^h = k_{m,Dl}A_c\left([Dl]_{cyt}^{h-1} - 2[Dl]_{cyt}^h + [Dl]_{cyt}^{h+1}\right) + k_{Toll,2}A_{Toll}[Toll:Dl/Cact]^h$$
$$- k_{in,Dl}A_n[Dl]_{cyt}^h + k_{out,Dl}A_n[Dl]_{nuc}^h - k_b V_n[Dl]_{cyt}^h[Cact]_{cyt}^h + k_d V_n[Dl/Cact]_{cyt}^h$$

(S7)

$$V_n \frac{d}{dt}[dl/Cact]_{nuc}^h$$
$$= k_{in,dl/Cact}A_n[dl/Cact]_{cyt}^h - k_{out,dl/Cact}A_n[dl/Cact]_{nuc}^h + k_b V_n[dl]_{nuc}^h[Cact]_{nuc}^h$$
$$- k_d V_n[dl/Cact]_{nuc}^h$$

(S8)

$$V_c \frac{d}{dt}[Dl/Cact]_{cyt}^h$$
$$= k_{m,Dl/Cact}A_c\left([Dl/Cact]_{cyt}^{h-1} - 2[Dl/Cact]_{cyt}^h + [Dl/Cact]_{cyt}^{h+1}\right)$$
$$- k_{Toll,b}A_{Toll}[Toll]^h[Dl/Cact]_{cyt}^h + k_{Toll,1}A_{Toll}[Toll:Dl/Cact]^h$$
$$- k_{in,Dl/Cact}A_n[Dl/Cact]_{cyt}^h + k_{out,Dl/Cact}A_n[Dl/Cact]_{nuc}^h + k_b V_c[Dl]_{cyt}^h[Cact]_{cyt}^h$$
$$- k_d V_c[Dl/Cact]_{cyt}^h$$

(S9)

$$V_n \frac{d}{dt}[Cact]_{nuc}^h = k_{in,Cact}A_n[Cact]_{cyt}^h - k_{out,Cact}A_n[Cact]_{nuc}^h - k_b V_n[Dl]_{nuc}^h[Cact]_{nuc}^h$$
$$+ k_d V_n[Dl/Cact]_{nuc}^h$$

(S10)

$$V_c \frac{d}{dt}[Cact]^h_{cyt} = k_{m,Cact}A_c\big([Cact]^{h-1}_{cyt} - 2[Cact]^h_{cyt} + [Cact]^{h+1}_{cyt}\big) + k_{Toll,2}A_{Toll}[Toll:Dl/Cact]^h$$
$$- k_{in,Cact}A_n[Cact]^h_{cyt} + k_{out,Cact}A_n[Cact]^h_{nuc} - k_bV_c[Dl]^h_{cyt}[Cact]^h_{cyt}$$
$$+ k_dV_c[Dl/Cact]^h_{cyt} + P_{cact} - k_{deg}V_c[Cact]^h_{cyt}$$

$$(S11)$$

$$A_{Toll}\frac{d}{dt}[Toll]^h = (k_{Toll,1} + k_{Toll,2})A_{Toll}[Toll:Dl/Cact]^h - k_{Toll,b}\,A_{Toll}[Toll]^h[Dl/Cact]^h_{cyt}$$
$$+ k_{act}A_{Toll}\,f(x) - k_{deact}A_{Toll}[Toll]^h$$

$$(S12)$$

$$A_{Toll}\frac{d}{dt}[Toll:Dl/Cact]^h$$
$$= k_{Toll,b}A_{Toll}[Toll]^h[Dl/Cact]^h_{cyt}\ - (k_{Toll,1} + k_{Toll,2})A_{Toll}[Toll:Dl/Cact]^h$$

$$(S13)$$

Equation S6 describes the time evolution of the concentration of Dl in the nucleus. The first two terms on the right hand side (RHS) describe the import into and export from the nucleus, respectively, controlled by the rate constants k_{in} and k_{out}, respectively. The parameter A_n represents the surface area of the nucleus. The final two terms describe the binding and dissociation of Dl to/from Cact, controlled by the parameters k_b and k_d, respectively. The volume of the nucleus, which is where the binding/dissociation events occur, is represented by V_n, which is held fixed within a given nuclear cycle interphase, but can vary from nc to nc. Equation S7 describes the time evolution of the concentration of Dl in the cytoplasm, has similar terms as to those found in Equation S6, but also has a term that describes the intercompartmental transport (often referred to here as simply diffusion) of Dl, as well as a term that describes Toll-mediated dissociation of the Dl/Cact complex. The transport term consists of three factors. The first factor is the mass transfer coefficient, in μm/min, that describes the rate at which intercompartmental transport takes place. The second factor, A_c, is the area shared by neighboring compartments, which changes with each nuclear cycle. The third factor, $\big([Dl]^{h-1}_{cyt} - 2[Dl]^h_{cyt} + [Dl]^{h+1}_{cyt}\big) = \big([Dl]^{h-1}_{cyt} - [Dl]^h_{cyt}\big) + \big([Dl]^{h+1}_{cyt} - [Dl]^h_{cyt}\big)$, describes the concentration driving force for exchange between cytoplasmic compartments h and $h-1$ plus that for exchange between cytoplasmic compartments h and $h+1$. It should be noted that, if this factor is normalized by $(\Delta x)^2$, where Δx is the distance between two neighboring compartments, then this approximates the second derivative in space on a discretized mesh. The Toll term, $k_{Toll,2}A_{Toll}[Toll:Dl/Cact]$ is composed of three factors, the first being the rate constant describing the Toll-mediated dissociation of Dl/Cact complex, the second being the plasma membrane surface area available to the cytoplasmic compartment, and the last being the surface concentration of Dl/Cact complex bound to the cytoplasmic tail of the active Toll receptor signaling complex.

Equations S8 and S9 describe the time evolution of nuclear and cytoplasmic Dl/Cact complex, respectively. The second term on the RHS of Equation S9 describes the binding ($k_{Toll,b}$) of active Toll with cytoplasmic Dl/Cact complex, while the third term describes the dissociation (without reaction) of Dl/Cact complex from the active Toll receptor, which is controlled by the rate constant $k_{Toll,1}$.

Equations S10 and S11 describe the time evolution of nuclear and cytoplasmic Cact, respectively. The final two terms in Equation S11 describe the per-nucleus production rate of Cact and the Toll-signal-independent degradation of Cact, respectively.

The final two equations describe the time evolution of active Toll and active Toll bound to Dl/Cact complex, respectively. In Equation S12, the third term on the RHS represents the spatially-dependent production rate of active Toll receptors, which we model phenomenologically by the function $f(x)$ (see below). This phenomenology captures the rate at which free Spätzle (Spz), the ligand for Toll, binds to free Toll receptors. The final term represents the constitutive deactivation of active Toll receptors, and can be thought of the natural dissociation of Spz from Toll.

During interphase, all eight equations are used, and each cytoplasmic compartment contains a single nucleus; both the nucleus and the cytoplasmic compartment are assumed to be well-mixed. During mitosis, the nuclear species become undefined. At the interphase-to-mitosis transition, we assume the nuclear and cytoplasmic species become mixed, and thus, for a species with cytoplasmic concentration C and nuclear concentration N:

$$C_{mitosis} = \frac{V_c C_{interphase} + V_n N_{interphase}}{V_c + V_n}$$

…where V_c is the volume of the cytoplasm in the cytoplasmic compartment, and V_n is the volume of the nucleus.

At the mitosis-to-interphase transition, the number of cytoplasmic compartments increases. The cytoplasmic concentration is linearly interpolated in space to create the initial conditions for the next interphase in the new (greater number of) cytoplasmic compartments, as has been done previously (Kanodia et al., 2009; O'Connell and Reeves, 2015). We assume the nuclear concentration of each species is initially equal to the cytoplasmic concentration for each nucleus h.

The initial conditions for NC10 are zero for $[Dl]_{cyt}$, $[Dl]_{nuc}$, and $[Toll:Dl/Cact]$; C_{Dl}^0 for $[Dl/Cact]_{cyt}$ and $[Dl/Cact]_{nuc}$; C_{cact}^0 for $[Cact]_{cyt}$ and $[Cact]_{nuc}$; and $C_{Toll}^0 f(x)$ for $[Toll]$.

In the equations for the cytoplasmic species in cytoplasmic compartment $h = 1$, the intercompartmental exchange term becomes $k_m A_c(2C^{h+1} - 2C^h)$. Similarly, for $h = M_{NC}$, the term becomes $k_m A_c(2C^{h-1} - 2C^h)$. These changes are similar to no flux boundary conditions at both ends, representing the ventral and dorsal midlines.

Note that total Dl is conserved:

$$\frac{d}{dt}\left([Dl]_{cyt} + [Dl]_{nuc} + [Dl/Cact]_{cyt} + [Dl/Cact]_{nuc} + [Toll:Dl/Cact]\right) = 0$$

(9)

Toll receptors are activated by the ligand Spz on the ventral side of the embryo. We represent this by assuming active Toll receptors are produced with a Gaussian-like spatial dependence,

$$f(x) = \exp\left(-\frac{1}{2}\left(\frac{x}{\Phi}\right)^2\right)$$

(10)

where Φ is a measure of the width of the signaling domain (see Eqn S12). Peak production occurs at the ventral midline, located at $x = 0$. Active Toll receptors reversibly bind to Dl/Cact complexes to form Toll:Dl/Cact complex, which is consumed to produce Toll, free Dl and free Cact (rate constant $k_{Toll,2}$). The model equations were nondimensionalized, revealing a set of 20 free parameters:

$$\tilde{V}_n \frac{d}{d\tau} U_{nuc}^h = \zeta_{Dl}\tilde{A}_n U_{cyt}^h - \xi_{Dl}\tilde{A}_n U_{nuc}^h - \gamma\tilde{V}_n U_{nuc}^h Z_{nuc}^h + \beta_0\tilde{V}_n W_{nuc}^h$$

(S16)

$$\tilde{V}_c \frac{d}{d\tau} U_{cyt}^h = \lambda_U\tilde{A}_c\left(U_{cyt}^{h-1} - 2U_{cyt}^h + U_{cyt}^{h+1}\right) + \omega\epsilon\tilde{A}_{Toll}X^h - \zeta_{Dl}\tilde{A}_n U_{cyt}^h + \xi_{Dl}\tilde{A}_n U_{nuc}^h - \gamma\tilde{V}_c U_{cyt}^h Z_{cyt}^h$$
$$+ \beta_0\tilde{V}_c W_{cyt}^h$$

(S17)

$$\tilde{V}_n \frac{d}{d\tau} W_{nuc}^h = \zeta_{Dl-Cact}\tilde{A}_n W_{cyt}^h - \xi_{Dl-Cact}\tilde{A}_n W_{nuc}^h + \gamma\tilde{V}_n U_{nuc}^h Z_{nuc}^h - \beta_0\tilde{V}_n W_{nuc}^h$$

(S18)

$$\tilde{V}_c \frac{d}{d\tau} W_{cyt}^h = \lambda_W\tilde{A}_c\left(W_{cyt}^{h-1} - 2W_{cyt}^h + W_{cyt}^{h+1}\right) - \eta\epsilon\tilde{A}_{Toll}W_{cyt}^h Y^h + \nu\epsilon\tilde{A}_{Toll}X^h - \zeta_W\tilde{A}_n W_{cyt}^h$$
$$+ \xi_W\tilde{A}_n W_{nuc}^h + \gamma\tilde{V}_c U_{cyt}^h Z_{cyt}^h - \beta_0\tilde{V}_c W_{cyt}^h$$

(S19)

$$\tilde{V}_n \frac{d}{d\tau} Z_{nuc}^h = \zeta_Z\tilde{A}_n Z_{cyt}^h - \xi_Z\tilde{A}_n Z_{nuc}^h - \gamma\psi\tilde{V}_n U_{nuc}^h Z_{nuc}^h + \beta_0\psi\tilde{V}_n W_{nuc}^h$$

(S20)

$$\tilde{V}_c \frac{d}{d\tau} Z_{cyt}^h = \lambda_V\tilde{A}_c\left(Z_{cyt}^{h-1} - 2Z_{cyt}^h + Z_{cyt}^{h+1}\right) + \omega\epsilon\psi\tilde{A}_{Toll}X^h - \zeta_Z\tilde{A}_n Z_{cyt}^h + \xi_Z\tilde{A}_n Z_{nuc}^h - \gamma\psi\tilde{V}_c U_{cyt}^h Z_{cyt}^h$$
$$+ \beta_0\psi\tilde{V}_c W_{cyt}^h + 1 - \alpha\,\tilde{V}_c Z_{cyt}^h$$

(S21)

$$\frac{d}{d\tau} Y^h = (\nu + \omega)X^h - \eta\,Y^h W_{cyt}^h + \beta\,g(z) - \rho Y^h$$

(S22)

$$\frac{d}{d\tau} X^h = \eta Y^h W_{cyt}^h - (\nu + \omega)X^h$$

(S23)

$$U = \frac{[Dl]}{C_{Dl}^0}, Z = \frac{[Cact]}{C_{Cact}^0}, W = \frac{[Dl/Cact]}{C_{Dl}^0}, X = \frac{[Toll:Dl/Cact]}{C_{Toll}^0}, Y = \frac{[Toll]}{C_{Toll}^0}$$
$$\tilde{A}_n = \frac{A_n}{A_n^{14}}, \tilde{A}_{Toll} = \frac{A_{Toll}}{A_n^{14}}, \tilde{A}_c = \frac{A_c}{A_n^{14}}, \tilde{V}_n = \frac{V_n}{V_n^{14}}, \tilde{V}_c = \frac{V_c}{V_n^{14}}, \tau = \frac{t}{T}, z = \frac{x}{L}$$

$$\zeta_i = \frac{A_n^{14} k_{in,i} T}{V_n^{14}}, \xi_i = \frac{A_n^{14} k_{out,i} T}{V_n^{14}}, \lambda_i = \frac{k_{m,i} A_n^{14} T}{V_n^{14}}, \beta = \frac{k_{act} T}{C_{Toll}^0}, \epsilon = \frac{C_{Toll}^0 A_n^{14}}{C_{dl}^0 V_n^{14}}, \psi = \frac{C_{dl}^0}{C_{Cact}^0}, \alpha = k_{deg} T$$

$$g(z) = \exp\left(-\frac{z^2}{2\phi^2}\right), \phi = \frac{\Phi}{L}, \gamma = k_b C_{Cact}^0 T, \beta_0 = k_d T,$$

$$\nu = k_{Toll,1} T, \omega = k_{Toll,2} T, \eta = k_{Toll,b} C_{Dl}^0 T, \rho = k_{deact} T$$

…where C_i^0 are the initial concentrations of species i at nc 10, A_n^{14} is the surface area of the nucleus during nc 14, V_n^{14} is the volume of the nucleus during nc 14, L is the half-circumference of the embryo at 50% AP position (roughly equal to 280 μm), and $T = 1\ min$. It should be noted that C_{cact}^0 is set to be $P_{cact} T / V_n^{14}$.

In the scaled equations, initial conditions for NC10 are zero for U_{cyt}, U_{nuc}, and X; unity for $W_{cyt}, W_{nuc}, Z_{cyt}$, and Z_{nuc}; and $g(z)$ for Y. When perturbing dosages of *dl* or *Toll*, these initial conditions are scaled accordingly. For example, in *dl* heterozygotes, the initial conditions for W_{cyt} and W_{nuc} are 0.5. Additionally, in Toll heterozygotes, β is also halved.

The model was then made consistent with the dynamics of the system as published by Reeves et al. (2012). We used the Improved Stochastic Ranking Evolution Strategy (ISRES) algorithm from Thomas Philip Runarsson (downloaded from https://notendur.hi.is/tpr/index.php?page=software/sres/sres) (Runarsson and Yao, 2000, 2005) for optimization, which proceeded by the least squares error combined with penalty functions designed to reject parameter sets in which the Dl gradient width decreased upon halving the dosage. Optimization runs also incorporated a penalty function designed to reject parameter sets where nuclear Dl and Dl/Cact complex displayed reversed roles (i.e. Dl/Cact forming the ventral-to-dorsal gradient), as the algorithm is otherwise blind to the differences between the two species. In other words, if we included additional information in the error calculation, such as Dl-mediated gene expression (which depends on Dl but not Dl/Cact), these parameter sets would be rejected as a matter of course, but instead we chose to simply reject results known to be unrealistic via penalty function. The 2-D version of the model used to investigate competing AP and DV Toll domains (Figure 6) uses the same equations as above, except that intercompartmental exchange can happen along two axes. The system is modeled as a rectangular array of compartments (instead of a linear array), using a 5-point stencil to discretize the Laplacian as a central difference, with reflective boundary conditions along the perimeter. For simplicity, the number of compartments along each axis is held constant, as is the distance between them (Δx & Δy).

Relationship between intercompartmental exchange, flux, and diffusion
As mentioned above, intercompartmental exchange can be viewed as a coarse-grained diffusion term. For cytoplasmic species C, the intercompartmental exchange term centered at nucleus h is

$$k_m A_c \left(C^{h-1} - 2C^h + C^{h+1} \right)$$

If we multiply and divide by $(\Delta x)^2$, and also divide by the volume of the cytoplasm, this term becomes equal to:

$$\frac{k_m A_c (\Delta x)^2}{V_c} \frac{C^{h-1} - 2C^h + C^{h+1}}{(\Delta x)^2} \approx \frac{k_m A_c (\Delta x)^2}{V_c} \frac{d^2 C}{dx^2}$$

Thus, the effective diffusivity is $D_{eff} = k_m A_c (\Delta x)^2 / V_c$.

In terms of the scaled equations, the intercompartmental exchange term is:

$$\lambda \tilde{A}_c \left(c^{h-1} - 2c^h + c^{h+1} \right)$$

…where c is the scaled version of the cytoplasmic species concentration, C. Recall that λ is defined as $k_m A_n^{14} T / V_n^{14}$, so that $\lambda \tilde{A}_c = D_{eff} \tilde{V}_c T / (\Delta x)^2$. If we define the scaled effective diffusivity as

$$\widehat{D}_{eff} = \frac{D_{eff} T}{L^2} = \frac{k_m A_c T (\Delta z)^2}{V_c} = \lambda \frac{\tilde{A}_c (\Delta z)^2}{\tilde{V}_c}$$

…then the scaled intercompartmental exchange term becomes:

$$\tilde{V}_c \widehat{D}_{eff} \frac{(c^{h-1} - 2c^h + c^{h+1})}{(\Delta z)^2} \approx \tilde{V}_c \widehat{D}_{eff} \frac{d^2 c}{dz^2}$$

Note that while k_m and λ are fixed parameters within a single simulation, D_{eff} and $\widehat{D}_{eff}$ vary with nuclear cycle.

Given these relationships, we can derive the relationship between approximate diffusive flux and intercompartmental transport:

$$\text{effective flux} = -\widehat{D}_{eff} \frac{dc}{dz} \approx -\widehat{D}_{eff} \frac{\Delta c}{\Delta z} = -\lambda \frac{\tilde{A}_c (\Delta z)^2}{\tilde{V}_c} \frac{\Delta c}{\Delta z}$$

…where $\Delta c = c^{h+1} - c^h$.

The "fluxes" plotted in Fig S2 are not actual fluxes. Instead, for species i, we plotted

$$\text{plotted flux} = -\lambda_i \frac{\Delta c_i}{\Delta z}$$

The remaining factors in the effective flux that are not included in the calculations are $\tilde{A}_c (\Delta z)^2 / \tilde{V}_c$, which are not species-specific (but they do change with nuclear-cycle). Thus, as we are comparing the plotted fluxes of two species within the same nuclear cycle, the plots are each proportional to their respective effective fluxes with the same proportionality constant.

Analysis of Kanodia and Ambrosi models
We recreated the model published by Kanodia et al. (2009) and the subsequent analysis performed by Ambrosi et al. (2014). As published, the Kanodia model assumes an equal rate of diffusion for Dl, Cact and Dl/Cact. However, to test the contribution of diffusion for each of the three species, we relax this assumption. Otherwise, the model remains unchanged for our analysis.

We used the parameter values published in Table 1, column 7 of Ambrosi et al. (2014) as the basis for a perturbation analysis. We chose this parameter set, termed "gyn 1" by the authors, because it is simulated in Figure 5C of Ambrosi et al., and in that scenario, it is shown that lowering the diffusivity widens the Dl gradient (as expected of a shuttling system). By perturbing the diffusion coefficients of this parameter set, we reproduce this widening phenomenon and show that, in the Kanodia/Ambrosi model, the widening is due to shuttling, as a decrease in the diffusion rate of Dl causes the gradient to contract ventrally, and a decrease in the diffusion of Dl/Cact causes the gradient to expand dorsally (Figure S2). In other words, shuttling is a property of this model, even when the authors did not intend it to be, and is not solely a property of our more detailed model. This is because shuttling arises naturally from the topology of a system that has a binder that, when in complex, can diffuse, protects the active species from capture, and dissociates from the active species in a spatially-dependent manner.

We also performed simulations of *dl* heterozygous embryos using the Ambrosi model (Figure S4C,D). In the paper (Ambrosi et al., 2014), the authors note that the normalized Dl gradient shape is independent

of the initial dose of Dl. The authors then suggest a possible mechanism for the widening (but not flat-top or double-peaked) phenotype of *dl/+* embryos. Namely, that in these embryos, the rate constant describing the Toll signal-independent degradation of free Cact protein by the proteasome increases four-fold in *dl/+* embryos. Therefore, their base parameter set is used to simulate wildtype embryos, and the alternative parameter set (with a four-fold increase in k_{Deg}) is used to simulate *dl/+* embryos. This results in a widening of the Dl gradient in *dl/+* embryos. However, when we compare simulations of wildtype vs. *dl/+* embryos using just the base parameter set (Table 1, Column 2 from Ambrosi et al., 2014) the normalized gradients collapse on each other (Figure S4D). Similarly, when we compare simulations of wildtype vs. *dl/+* embryos using just the alternative parameter set (Table 1, Column 3 from Ambrosi et al., 2014), the normalized gradients also collapse on each other (Figure S4D). In comparison, our model using explicit active Toll receptors reproduces the *dl/+* phenotype simply by reducing the dosage of *dl* (Figure S5EF), and does not require assuming that any other rate constants in the model change when the dosage of *dl* changes.

Fly lines
dl-paGFP was injected into the fly line yw; VK33 (landing site: third chromosome 65B2) by Duke University Model Systems, Durham, NC. *dl-dGFP* was inserted on the second chromosome at landing site attP40 by Genetic Services, Inc, Sudbury, MA. The *dl-dVenus* BAC was injected into the fly line yw; VK33 (landing site: third chromosome 65B2) by Genetic Services, Inc, Sudbury, MA.

The plasmid carrying *FRT-stop-FRT hsp83> Toll 10b: bcd 3'UTR* was injected into fly line attP2 with a landing site at (3L) 68A4 by Genetivision, Inc. (Houston, TX). To remove the FRT-stop-FRT cassette, we crossed male flies carrying this construct to virgins carrying hsFLP on both X chromosomes (BS# 8862). Females were allowed to lay embryos for 2-3 days before they were removed to a new vial. Vials containing 3-5 day old larvae (F1 generation) were heat shocked at 37°C for 2 hours. Males with red eyes (F1 generation) were crossed to virgin yw flies. Flies from the F2 generation were crossed to virgin yw flies to create the F3 generation, which was screened for female sterility, used as an indication of removal of the FRT-stop-FRT cassette.

In order to ablate the native DV Dl gradient, we generated a *gd⁷* line null for white. *gd⁷/FM3* flies (BS# 3109) were crossed to *yw* and the progeny crossed into stable lines that were screened for white eyes and females that were homozygous sterile. These flies (*gd⁷w⁻/FM3*) were then crossed to males carrying *toll¹⁰ᵇ: bcd 3'UTR/+* (weak *bcd* promoter construct on the second chromosome, strong *hsp83* promoter on the third chromosome). Males from this cross with the phenotype *gd⁷/Y; toll¹⁰ᵇ: bcd 3'UTR/+* were crossed again to *gd⁷w⁻/FM3* virgins, generating females that are homozygous null for *gd* (thus abolishing the wt Dl gradient) and provide their embryos with only the AP Dl gradient. We screened these females for the absence of bar (present on FM3) and the presence of *white* (present on *toll¹⁰ᵇ: bcd 3'UTR*).

In Fig. S3B,C, we evaluated the effect of a *dl-lacZ* transgene on the Dl gradient (Govind et al., 1992). As β-galactosidase (β-gal) tetramerizes, this fusion should slow diffusion of Dl to a greater extent than dGFP. Due to its anti-morphic nature (Govind et al., 1992), *dl-lacZ* is suspected to be expressed at low levels in surviving fly lines (Govind et al., 1996). Therefore, two copies of this transgene in a *dl/+* background were analyzed (Fig. S3B,C).

The *cact-lacZ* transgene studied in Fig. S3B,C was present in single copy and was in a heterozygous *cact* background (Fernandez et al., 2001).

Collection of Egfr^{t1}; dl^{RC} embryos

As shown in Figure S5H,I, only 3 and 1 embryos were imaged for the *Egfr^{t1}; dl^{RC}/+* and *Egfr^{t1}; dl^{RC}* fly lines, respectively. This resulted from multiple problems. First, there was a limited quantity of females found with the correct genotype. The females were collected from the *Egfr^{t1}/+; dl^{RC}/+* fly line. The relative lack of fitness of the *Egfr^{t1}; dl^{RC}/+* and *Egfr^{t1}; dl^{RC}* adults resulted in roughly ~10 such females found over the span of two weeks of collection from multiple bottles. Second, the fecundity of these females was severely compromised. Embryos were collected and fixed over 24 hour periods. The grape juice agar plates were checked for embryos every two hours. If any embryos were found, they were aged for two hours and fixed. Third, high attrition rates during fixation, especially during the devitellinization step, were observed (Kosman et al., 2004). Finally, upon imaging, most embryos were not found to be in the midst of nc 14. All of these factors contributed to the low sample size.

BAC Recombineering

Please note that residue 206 in GFP is actually residue 207 in the Venus protein as Venus has a Valine residue after the initial Methionine that is not present in the original GFP protein. For consistency, researchers refer to these residues by their location in GFP.

Fluorescent in situ *Hybridization*

Both *sna*-biotin and *sna*-fluorescein anti-sense RNA probes were used. Antibodies used were anti-dorsal 7A4 (deposited to the DSHB by Ruth Steward (DSHB Hybridoma Product anti-dorsal 7A4)) (1:10), donkey anti-mouse- 488 (Invitrogen A21202, Lot 81493) (1:500), rabbit anti-histone (abcam ab1791, Lot 940487) (1:5000), donkey anti-rabbit-546 (Invitrogen A10040, Lot 107388) (1:500), goat anti-biotin (ImmunoReagents, Raleigh, NC, GtxOt-070-D, Lot 19-19-112311) (1:50,000), donkey anti-goat-647 (Invitrogen A21447, Lot 774898) ((1:500), goat anti-fluorescein (Rockland 600-101-096, Lot 19458) (1:500), rabbit anti-fluorescein (Life Technologies A889, Lot 1458646) (1:500), goat anti-histone (Abcam, ab12079, Lots GR6952-4 and GR129411-1) (1:100)

Sequencing dl$^{1.2.5}$

Genomic DNA was extracted from males homozygous for *dl$^{1.2.5}$* according to standard protocols. We PCR amplified the entire *dl* region, then used 9 different primers to ensure complete sequencing coverage. Sequencing was performed by GENEWIZ, RTP, NC. The resulting sequence was compared to the consensus sequence available at FlyBase. Each mismatched codon was investigated as a potential source of an altered or truncated sequence. We found that a mutation at nt 3256 (entire *dl* genomic sequence) from G->T results in a premature stop codon, which we presume is the source of the amorphic allele.

Image analysis of toll^{10b}: bcd 3'UTRembryos

The Dorsal gradient and *snail* expression in embryos from mothers carrying the *toll^{10b}: bcd 3'UTR* transgene and homozygous for the *gd^{7}* mutation were analyzed using the following procedures. First, the z-stack images (taken as described in Experimental Procedures) were background subtracted assuming the mode of the image corresponded to zero fluorescence levels.   Next, a maximum intensity projection was created, and the intensities from the three color channels (Dl, histone h3, and *snail*) were summed. The resulting image was Gaussian filtered in both the x and y directions using ten pixels as a kernel. This created an image I1 with the embryo as a single, bright object to facilitate discovery of the embryo boundary.

The embryo boundary was found according to (Jermusyk et al., 2016).

Once the boundary was found, an inner boundary was constructed by moving the 60 updated boundary points inward by 30 pixels in the direction of the local normal. This defined 60 quadrilaterals that encompass the outer periphery of the embryo. These quadrilaterals were laid on top of the image slice corresponding to the mid sagittal plane of the embryo, and each of the three color channels were unrolled using a affine transformation on the 60 quadrilaterals to result in 60 rectangles (see (Liberman et al., 2009; Trisnadi et al., 2013) for more information).

The nuclear channel was then segmented using a local thresholding (Trisnadi et al., 2013). The Dorsal intensity in each nucleus was then computed as the intensity in the Dorsal channel, normalized by the intensity in the nuclear channel. The intensity of *sna* expression was found as the average intensity of the *sna* channel within this unrolled strip loosely bounded by the nuclei.

Photobleaching experiments
Embryos were dechorionated, mounted, and imaged using the same protocol as described in the "Activating paGFP in Live Embryos" section. Individual nuclei were chosen at random across the embryo. Bleaching box: ~700 pixels (26.46 microns by 26.46 microns), bleaching time (amount of time the laser bleached the nuclei): ~20 seconds, number of cycles: 30. Each bleaching session lasted about 30 minutes (Movie S3) and was followed by imaging the entire depth of the embryo. Two nuclei per embryo were imaged in a single session. Laser power: 50% for each nucleus. A 488 nm laser was used for the bleaching of GFP.

Each FRAP experiment was analyzed according to the following procedure. First, the image sequence was stabilized using a standard optical flow protocol, which was necessary because the embryo sometimes moved very slightly during imaging. The optical flow protocol is as follows. In any xy frame at time point t, let the brightness at any pixel with coordinates x, y be $f(x, y, t)$. The goal is to find the values of dx and dy that make $f(x + dx, y + dy, t + dt) \approx f(x, y, t)$, where dt is the time difference between two successive frames. In a general optical flow protocol, there will be a different set of dx, dy depending on which pixel (or local pixel region) you are examining. However, for our particular application, we assume that the entire frame is translocating with the same dx and dy; that is, all nuclei move together as a rigid body. We also assume object brightness is roughly constant from frame to frame, save possibly a small average difference between two successive frames, df.

Performing a first order Taylor series expansion of f about (x, y, t), we obtain,
$$f(x + dx, y + dy, t + dt) = f(x, y, t) + f_x dx + f_y dy + f_t dt$$
...where f_x is the partial derivative of f in the x direction evaluated at pixel (x, y) in frame at time t, f_x is the partial derivative of f in the x direction evaluated at pixel (x, y) in frame at time t, and f_y is the partial derivative of f in the y direction evaluated at pixel (x, y) in frame at time t, and f_t is the partial derivative of f in the x direction evaluated at pixel (x, y) in frame at time t. Using finite differences, for each pixel save one row/column, f_x and f_y can be approximated. According to our objective, for the correctly chosen dx, dy:
$$df = f(x + dx, y + dy, t + dt) - f(x, y, t)$$

Therefore,
$$f_x dx + f_y dy - df = -f_t dt$$
This equation can be written for every pixel in the frame, save one row and one column. Using linear least squares, the best-fit dx, dy, df can be found. Each frame from frame 1 to frame $n_t - 1$, where n_t is the total number of frames, will have different values of dx, dy, df.

The optical flow algorithm is applied to image frames that have been first morphologically opened with a disk structuring element of 5 pixels (to remove small artifactual objects) and Gaussian filtered (to blur/smooth the image) with a width of 10 pixels. Applying the optical flow protocol always resulted in an adequately stabilized image, which implies that our assumptions were adequately satisfied.

Next, the stabilized image was segmented to detect the nuclei. If the image sequence was taken with both Dl-GFP and H2A-RFP (to mark the nuclei), then the RFP channel was used to segment the nuclei. If only Dl-GFP was imaged, then the GFP channel was used, which only works if the focus is on the ventral side. To detect the nuclei, the entire image sequence was summed to find an aggregate frame (Figure S3D). This aggregate frame was morphologically eroded by the pixel equivalent of 2 microns, then morphologically dilated by the pixel equivalent of 0.5 microns. This resulted in an aggregate frame where small, non-nuclear objects were removed, but objects representing nuclei that remain were ensured to have a diameter of at least 0.5 microns. After Gaussian blurring with a width of the pixel equivalent of 0.5 microns (Figure S3E), a watershed algorithm was applied to the complement of the blurred image. The watershed image was a label image, where the pixels corresponding to each nucleus plus the cytoplasm surrounding it were given a distinct numerical label (Figure S3F). The boundaries between cytoplasmic compartments were pixels of zero intensity (white in Figure S3F,G). This watershed matrix served to delineate the boundaries between cytoplasmic compartments for the remainder of the analysis.

We then analyzed each individual frame of the image sequence. We split the frame into the pixel sets that corresponded to each cytoplasmic compartment found by watershed. Each cytoplasmic compartment was hard-thresholded at 35% intensity. The remaining object with the largest area was declared to be the nucleus, after eroding by a disk of three pixels (Figure S3G, blue). All pixels outside of the largest-area object (pre-erosion) were also eroded by a disk of three pixels then declared to be cytoplasm (Figure S3G, orange). The two erosion operations were performed to get conservative estimates of what is the nucleus and what is the cytoplasm. This process was repeated for each cytoplasmic compartment in each frame of the image sequence, which resulted in our ability to track the nuclear and cytoplasmic fluorescence in the Dl-GFP channel, over time (Figure S3H).

The timecourse data for the nuclear intensity were then fit to the solution of a differential equation that described the nuclear concentration of Dl:

$$\frac{d[Dl]_{nuc}}{dt} = k_{in}[Dl]_{cyt} - k_{out}[Dl]_{nuc}$$

…where $[Dl]_{nuc}$ and $[Dl]_{cyt}$ are the nuclear and cytoplasmic concentrations of Dl, respectively, and k_{in}, k_{out} are the nuclear import and export rate constants for Dl, respectively. The cytoplasmic concentration timecourse measurements served as input to this equation, and the nuclear concentration timecourse measurements were the target of the fit.

For a general function $[Dl]_{cyt}(t)$, the solution to this differential equation is:

$$[Dl]_{nuc}(t) = c_0 \exp(-k_{out}t) + k_{in}\exp(-k_{out}t) \int_0^t [Dl]_{cyt}(t') \exp(k_{out}t') \, dt'$$

…where $t = 0$ corresponds to the time point directly after bleaching, and c_0 is an adjustable parameter that corresponds to the concentration of nuclear Dl at time $t = 0$. The fits were performed with Matlab's lsqcurvefit function.

Parameter Estimation

Some biophysical parameters that appear in our model have crude estimates available. These estimates helped us set acceptable bounds for variation of these parameters in our parameter search scheme. First, photobleaching experiments (this work and Delotto et al., 2007) help to constrain k_{in} and k_{out} (from the full model; these correspond to dimensionless parameters ζ and ξ). Our photobleaching observations put our estimate of both $k'_{in} = k_{in} A_n/V_n$ and $k'_{out} = k_{out} A_n/V_n$ at roughly 0.1-0.4 min⁻¹ (Fig S3I). Analysis of photobleaching recovery curves found in Delotto et al., 2007 give an estimate of k'_{out} as 1 min⁻¹. Note that k_{in} and k_{out} in the simplified model are the same as k'_{in} and k'_{out} discussed here.

The dimensionless parameters ζ and ξ are defined as:
$\zeta = (A_n^{14}T/V_n^{14})k_{in} = k'_{in}T$ and $\xi = (A_n^{14}T/V_n^{14})k_{out} = k'_{out}T$, where in both cases, the final equality holds true during nuclear cycle 14, and where $T = 1$ min. of these Therefore, preliminary estimates of these parameters are $\zeta \sim 0.1 - 0.4$ and $\xi \sim 0.1 - 1$.

However, recall that there is a ζ for Dl and a ζ for Dl/Cact complex (and similar for ξ). Because of this, we did not have utter confidence in using the photobleaching estimates to directly constrain the $\zeta's$ and $\xi's$. Instead, we assume the estimation of k'_{in} and k'_{out} from these experiments is some weighted average between those parameters for free Dl and Dl/Cact complex with unknown weightings. Therefore, we took these estimates as justification to center our evolutionary search algorithm (for the full model) around 10^0. For the simplified model, we held $K_{eq} \equiv k_{in}/k_{out}$ fixed at 4, which roughly reflects the ratio of nuclear to cytoplasmic Dl near the ventral midline.

Next, there are several rough estimates and/or scaling arguments available for the intercompartmental exchange coefficients, λ. As defined above, $\lambda = k_m A_n^{14} T/V_n^{14}$, where k_m is the mass transfer coefficient (in dimensions of length per time), A_n^{14} and V_n^{14} are the surface area and volume of a nuclear cycle 14 nucleus, respectively, and $T = 1$ min. If the nucleus is roughly spherical and its radius is roughly 5 microns, then this definition implies $k_m = \lambda \times 1$ micron/min $\sim \lambda \times 10^{-6}$ cm/s. Engineering literature suggests that mass transfer coefficients of proteins through membranes, in protein separations processes, can span $10^{-3} - 10^{-6}$ cm/s, which implies λ could fall within the range of order 1 up through 10^3. On the other hand, theoretical scaling arguments in which λ is controlling the shape of a gradient suggest λ could fall within the range of 10^{-4} through order 1. As we have no reason to believe that the literature measurements (in separations) represent a lower bound, nor do we need to assume that a simple scaling argument represents the upper bound (considering the complexity of our model), it is not unreasonable to allow in our parameter searches several decades of variation centered on 10^0 (see Fig 2B).

As further confirmation of this, we used our photoactivation experiments to give a very crude estimate of the effective diffusivity. We observed that Dl moves over 7-10 nuclear diameters in the time span of 90 minutes. As the average distance between the centroids of two neighboring nuclei is roughly 7 μm, this translates to an effective diffusivity of ~0.4-0.9 μm²/s and a time scale to cross one cell diameter that is on the order of minutes. However, this crude estimate stems from an experiment not designed specifically to measure the diffusivity, and also does not take into account the changing distances between nucleocytoplasmic compartments due to mitosis during the 90 min period. Even so, this estimate also suggests that the λ values be centered around 10^0. Using scaling arguments, we suggest the effective diffusivity is on the order of 1 micron² per second. As detailed above, the relationship between effective diffusivity and "lambda" is as follows:

$$D_{eff} = \lambda \frac{A_c V_n^{14}}{V_c A_n^{14}} \frac{(\Delta x)^2}{T}$$

…where A_c is the surface area between nucleocytoplasmic compartments available for intercompartmental exchange, V_c is the volume of cytoplasm inside a nucleocytoplasmic compartment (so, total volume minus the volume of the nucleus), Δx is the diameter of a nucleocytoplasmic compartment (roughly 7 microns), and, as before, A_n^{14} and V_n^{14} are the surface area and volume of a nucleus, and $T = 1$ min. This relationship can be rearranged to give

$$D_{eff} \frac{T}{(\Delta x)^2} = \lambda \frac{A_c V_n^{14}}{V_c A_n^{14}}$$

…where the left hand side is approximately equal to one (dimensionless). Assuming each nucleocytoplasmic compartment is roughly cylindrical with a radius R and height h, the right hand side becomes:

$$\frac{\lambda}{0.75 \, \Delta x/r - 24/(h\Delta x)} \sim 1$$

If $\Delta x = 7$ microns, $r = 2.5$ microns, and $h = 15$ microns, then this means lambda ~ 2. However, recall that this was originally based on a very rough estimate for D_{eff} from an experiment that was not designed to estimate this parameter. However, this can be taken as an independent argument that the range of possible values for λ should be centered around 10^0.

Fitting Dl gradients and sna peaks
Each Dl nuclear gradient curve was fit to a modified Gaussian-like curve with five adjustable parameters:
$$Dl(x) = A \exp(-(x - \mu)^2/(2\sigma^2)) + B - M|x|$$
Here A is the gradient amplitude, B represents the basal levels, σ is the gradient width, and M is the (typically non-zero) slope of the tails of the gradient. The parameter μ is the location of the ventral midline within the image. See also (Liberman et al., 2009; Reeves et al., 2012; Trisnadi et al., 2013) for more information.

Plots of normalized Dl gradients were generated by subtracting B value and 70% of the M value, then dividing by A. In other words:

$$Dl_{normalized}(x) = \frac{Dl(x) - (B - 0.7M)}{A}$$

Only 70% of the slope was subtracted because it is the average between the value of the gradient tail at $x = 1$ (lowest value) and at $x = 0.4$ (where the contribution from the Gaussian terms in wildtype embryos become negligible).

The average normalized intensity curve was generated by averaging the normalized curves of all embryos in the specified genotype. After this procedure, the averaged curve was not re-normalized, which is why these averaged curves do not always fall exactly between zero and one.

The width of the *sna* domain for each embryo was computed as described in (Liberman et al., 2009; Reeves et al., 2012; Trisnadi et al., 2013). Briefly, a canonical profile of *sna*, called $sna_0(x)$, was generated by averaging the *sna* domain of many wildtype embryos together. Next, local background subtraction was performed (top hat morphological transform) with a structuring element of width of 25% DV axis length, which removes background intensity variations that are wider than the structuring element. Peaks of *sna* expression are preserved in height and width because they are narrower than the structuring element. Finally, all *sna* expression domains were then fit to this canonical profile by the following equation:

$$sna(x) \approx A\, sna_0(x/\delta) + B,$$

where A and B are the amplitude and background levels, and δ is a "stretching factor" that defines how wide or narrow the individual *sna* expression domain is with respect to the canonical form. If $\delta > 1$, then the individual profile is wider than the canonical form, and if $\delta < 1$, then it is narrower. The final width measurement of the *sna* domain was computed as the width of the canonical profile (measured at half-max) times δ. In a similar manner to Dl, when average *sna* curves are generated, the resulting curves were not re-normalized, which again is why these averaged curves do not always fall exactly between zero and one.
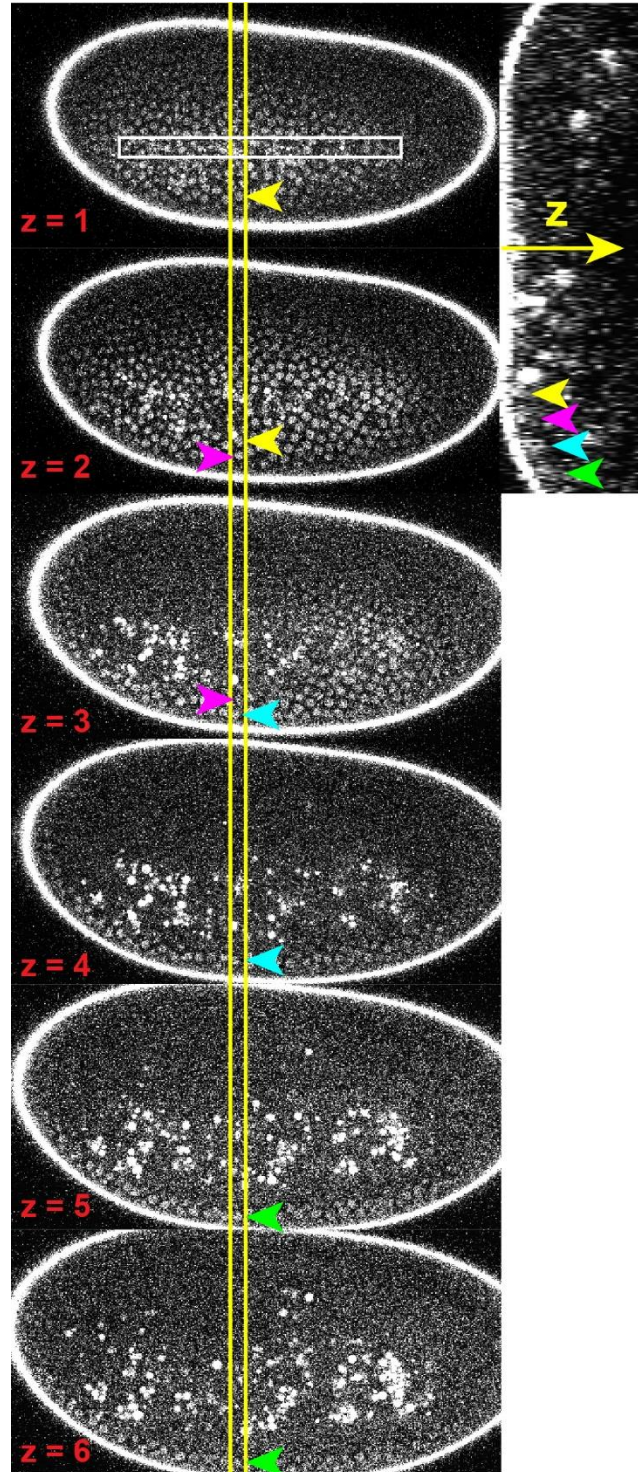
Figure S1: Photo-Activated Dl Can Be Seen at Least Seven Nuclei away from the Activation Box on the Ventral Side, Related to Figure 1. Six z-slices (xy sections) of embryo found in Figure 1H are depicted, as well as a yz-section (right), where the arrow indicates the z-direction. Yellow bars indicate the x-coordinate fixed in the yz-slice. Arrowheads point to individual nuclei. Arrowheads of the same color refer to the same nucleus in different views. Three nuclei can be counted between the activation box and the nucleus referenced by the yellow arrowhead.
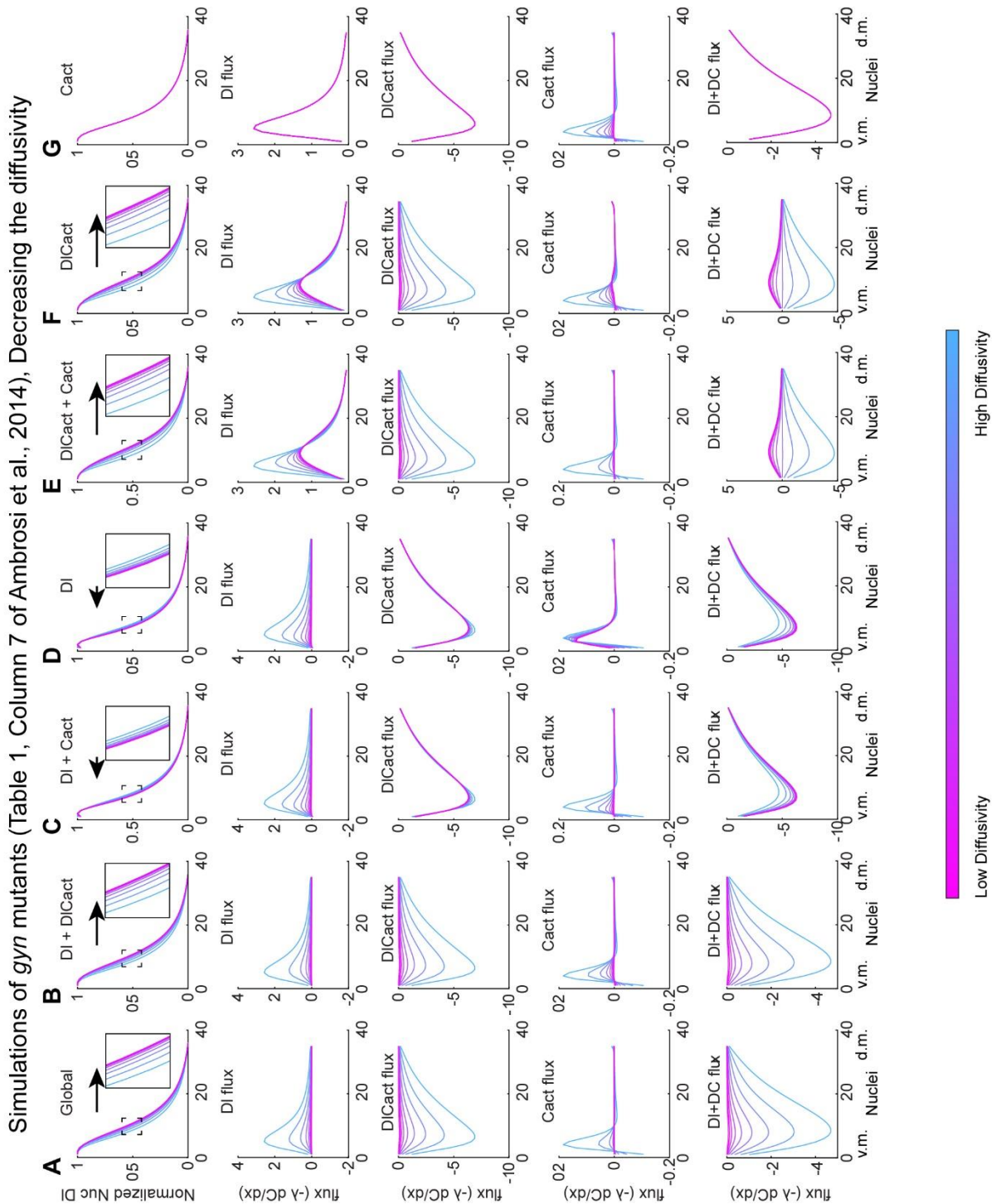
Figure S2: Previous Models of the Dl Gradient Also Exhibit Shuttling, Related to Figure 2. (A) The model from Ambrosi et al., 2014 predicts the widening of the gradient when the diffusivity of all three species (Dl, Cact, and Dl/Cact complex) is simultaneously lowered (top panel; see arrow for direction and inset for better view of the changes to the gradient). Note that in the Ambrosi model, there is only one diffusivity parameter, which dictates the mobility of all three species. This

widening is apparently the result of changes to the Dl/Cact complex flux, as the flux of Dl is towards the dorsal midline (curves are above zero in the second panel), yet the flux of Dl/Cact complex is towards the ventral midline (curves are below zero in the third panel). The flux of Cact has minimal effect (based on order of magnitude; fourth panel). (B) If the diffusivities of both Dl and Dl/Cact complex are lowered, the widening effect is also present. The fluxes are qualitatively similar to (A). (C) If the diffusivities of Dl and Cact are both lowered, the gradient becomes slightly narrower. This highlights the fact that the widening (shuttling) effect comes from Dl/Cact complex. Indeed, in this case, the Dl/Cact complex flux is only minimally effected by this change (third panel), so the narrowing effect must come from the change in the Dl flux (second panel). (D) Changing the diffusivity of Dl only has nearly the same effect as that seen in (D). (E) Changing the diffusivities of Dl/Cact complex and Cact widens the gradient, to a greater extent compared to those seen in (A) and (B). This effect again must be the product of the change to the flux in Dl/Cact complex (third panel), as the change to the Dl flux is minimal (second panel). (F) If the diffusivity of only Dl/Cact complex is changed, the result is nearly identical to that seen in (E). (G) Changing the diffusivity of Cact alone has an insignificant effect on the Dl gradient. Parameter set used can be found in Table 1, Column 7 of Ambrosi, et al., 2014, which corresponds to the simulation for the *gyn* mutant. Colorbar at bottom highlights that lower diffusivities are indicated by magenta. (Note that the first/last simulated nucleus in not exactly at DV = 0/1, respectively, because of the discretization mesh. Thus, graphs of the flux do not necessarily start and end at exactly zero for the DV coordinate.)
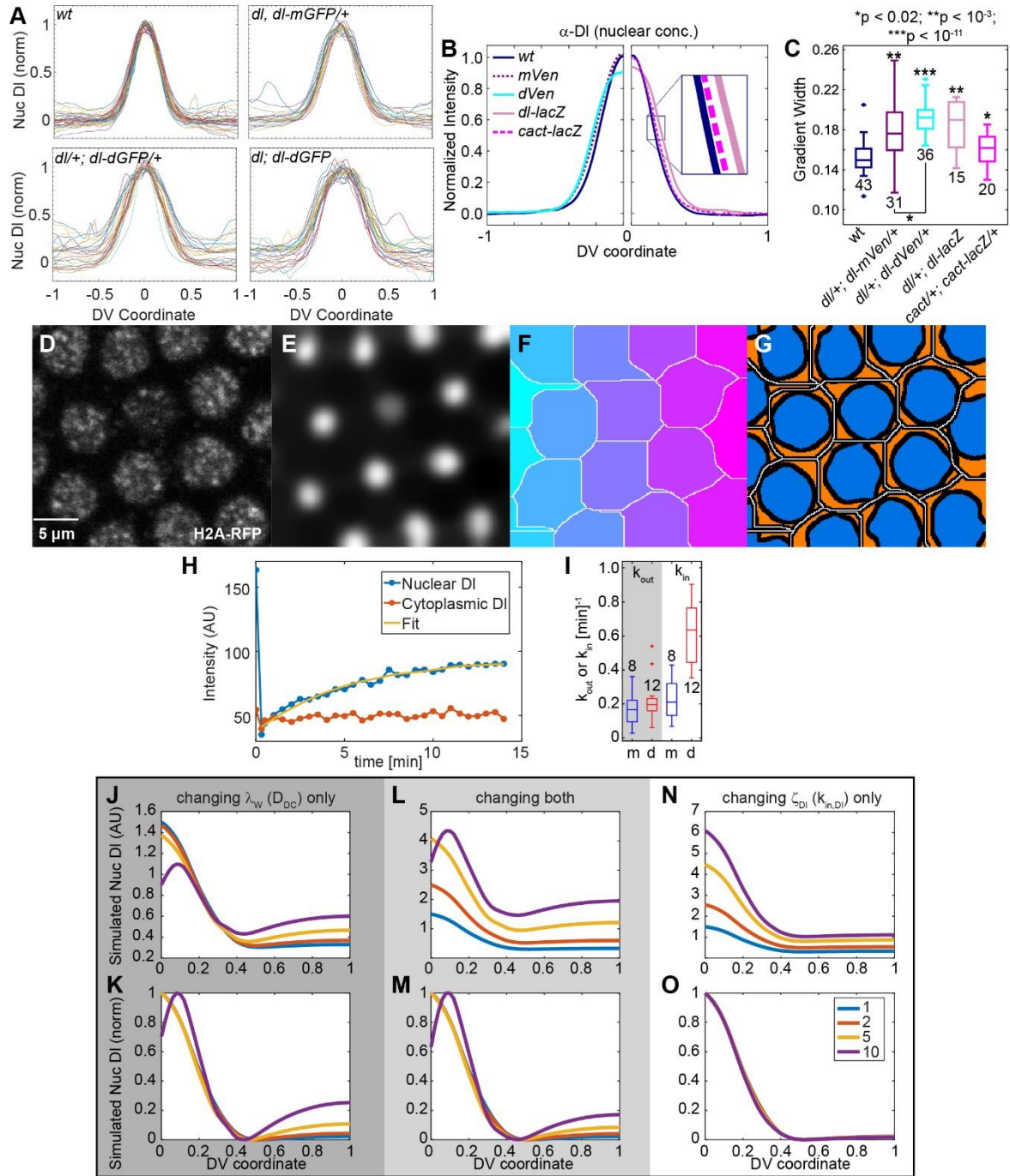
Figure S3: Decreasing Diffusion of Dl/Cact Widens the Dl Gradient, Related to Figure 3. (A) Full data sets of Dl gradients that were averaged to result in the curves found in Fig. 3A. (B) Normalized average Dl nuclear gradients in embryos from several different genetic backgrounds. Plots are only shown for half of the DV axis so that more gradients can be plotted without too many being plotting on top of each other. Genetic backgrounds include wildtype (both sides of graph), *dl/+; dl-mVen/+* (mVen; left), *dl/+; dl-dVen/+* (dVen; left), *dl/+; dl-lacZ* (dl-lacZ; right), and *cact/+; cact-lacZ/+* (cact-lacZ; right) embryos. Inset shows more clearly that *dl-lacZ* and *cact-lacZ* are slightly

wider than wt. Legend: full genotypes depicted at bottom of boxplot in (C); color scheme the same between (B) and (C). (C) Gradient widths in the embryos from (B). Note that the same "allelic series" approach is also valid for Dl-Venus constructs. Embryos from *dl/+; dl-mVenus/+* mothers have slightly wider gradients than wildtype, (σ = 0.179 ± 0.005; mean ± s.e.m.), while embryos from *dl/+; dl-dVenus/+* mothers had further widened Dl gradients (σ = 0.192 ± 0.003). Furthermore, both *dl-lacZ* and *cact-lacZ* expand the gradient. In particular, the fact that *cact-lacZ* expands the Dl gradient indicates that that this fusion affects the accumulation of Dl. (D) Example aggregate frame (sum of all frames in the nuclear channel) for bleaching experiment (*dl/+; dl-mGFP/+* embryo shown). (E) Aggregate frame after erosion, dilation, and Gaussian blurring. The complement of this image was used for a watershed operation. (F) Output of watershed protocol. Each cytoplasmic compartment is individually labeled with a different numerical value (shown here as a different color). White pixels delineate boundaries between the compartments. (G) Segmented image of frame 1. Each compartment is analyzed individually and is subdivided into nucleus (blue) and cytoplasm (orange). The black pixels belong to neither, due to labeling nucleus and cytoplasm conservatively. (H) Timecourse of nuclear (blue) and cytoplasmic (red) Dl for one photobleaching experiment. The least-squares fit is shown in yellow. (I) Boxplot of measured values of k_{out} (left side, gray) and k_{in} (right side, white) for *dl/+; dl-mGFP/+* embryos (m; blue) or *dl/+; dl-dGFP/+* embryos (d; red). Values in inverse minutes. The difference in k_{out} between mGFP and dGFP is not significant, whereas the difference in k_{in} clearly is: k_{in} for dGFP is roughly twice as large as that for mGFP. (H-O) Model analysis of perturbing the import rate, the diffusion rate, or both. This analysis was performed to determine whether the difference in import rate could explain the difference in gradient width. (J) Effect of lowering the diffusivity of Dl/Cact complex on the nuclear Dl gradient (simulation using the "Full model"). The hallmark shuttling phenotype is observed. (K) Curves from (J), but normalized. (L) Same as (J), but k_{in} is increased by the same fold-change as the D_{DC} is decreased. The hallmark shuttling phenotype is still observed. (M) Curves from (L), but normalized. (N) Same as (L), but only k_{in} is perturbed (D_{DC} is held fixed). The shuttling phenotype is not observed. Thus, a shuttling phenotype is observed when diffusion is altered, but not when only capture is altered. (O) Curves from (N), but normalized.
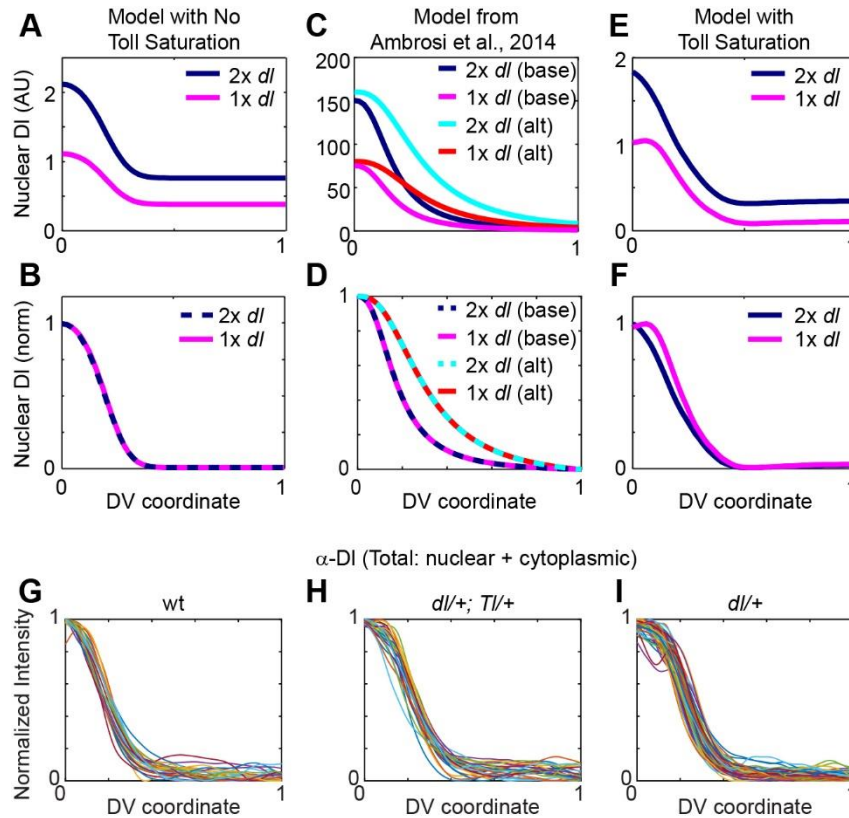
Figure S4: Updated Model Formulation Recreates the *dl* Heterozygous Phenotype, Related to Figure 4. (A) Plots of simulated Dl gradients for wildtype (2x *dl*) and *dl/+* (1x *dl*) embryos in a model with no toll saturation taken into account (O'Connell and Reeves, 2015). (B) Same plots from (A), but normalized. The two plots collapse onto each other. (C) Same as (A), but for a different model formulation found in (Ambrosi et al., 2014). Two separate parameter sets are represented: base and alternative. See Supplemental Experimental Methods for explanation. (D) Same plots from (C), but normalized. As with (A,B), the plots from corresponding parameter sets collapse onto each other. (E) Same as (A), but for a model that considers the possibility of active Toll saturation ("Full model" formulated here). Even without normalization, the difference in peak shape can be seen. (F) Same plots from (E), but normalized. The difference in peak shape is more clearly seen. Similar results can be obtained using the simplified model with Toll saturation. (G-I) Full data sets of Dl gradients that were averaged to result in the curves found in Fig. 4H.
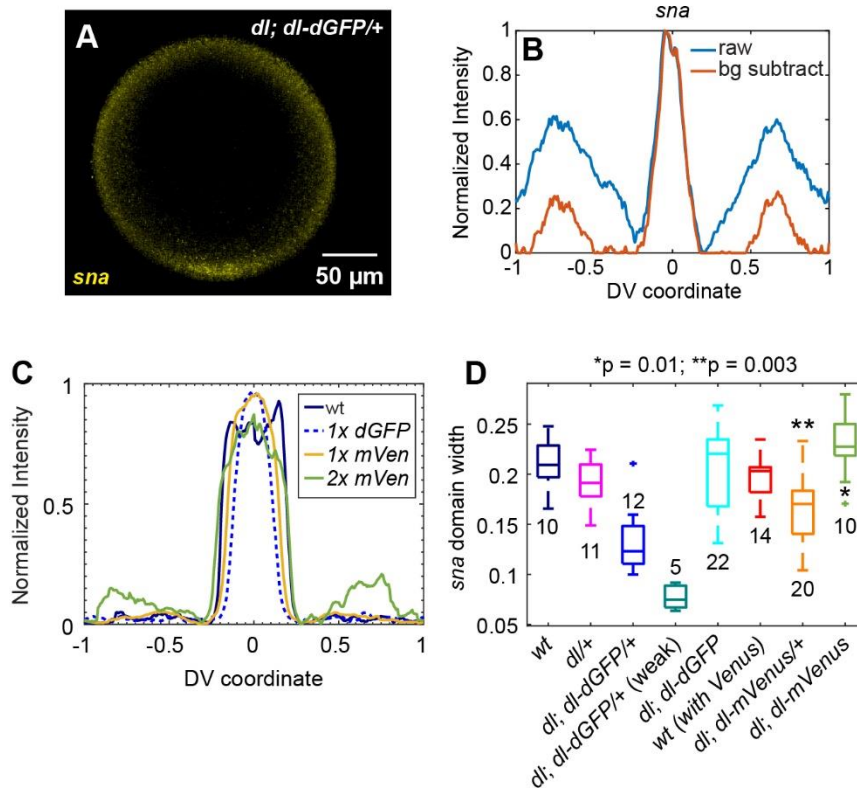
Figure S5: Severe Defects in Dl Gradient Formation from Simultaneous Shuttling Perturbations Result in Compromised *sna* Domains, Related to Fig. 5. (A) Representative embryo from among the five *dl; dl-dGFP/+* embryos that expressed sna very weakly. The *sna* domains in these embryos could not be detected by unsupervised image analysis (Trisnadi et al., 2013). Manual analysis was required to extract the width of the *sna* domain. (B) Quantification of the sna domain from embryo in (A). The width of the *sna* domain in this embryo was measured as the half-width of the peak at half-max. The red curve is background subtracted with a structuring element of the size of 25% of the DV axis arclength. This procedure preserves the peak (which is narrower than 25% DV axis arclength) while removing most of the noise outside of the peak. (C) The average *sna* domain from wildtype, *dl; dl-dGFP/+* (1x *dl-dGFP*), *dl; dl-mVen/+* (1x *dl-mVen*), and *dl; dl-mVen* (2x *dl-mVen*) embryos. Of these, only *dl; dl-dGFP* are non-viable, and these also have a visibly narrower *sna* domain. Wildtype and *dl; dl-dGFP* curves the same as found in Fig. 5C, and do not include the "weak" sna domain embryos found in (A,B). (D) Boxplot of *sna* domain widths for embryos of various genotypes depicted at bottom. The first, second, third, and fifth box-and-whisker sets are also found in Fig. 5E. This boxplot is for further comparison of the "weak" *sna* domains (embryos in A,B) and embryos with *dl-mVenus* constructs. Most relevant pair-wise statistical tests can be found in Fig. 5E. The "wt (with Venus)" corresponds to measurements of wildtype embryos that were stained with the mVenus embryos in the same experiment. These are statistically indistinguishable from the wildtype population in the first column. Asterisks indicate pair-wise statistical difference from "wt (with Venus)" population.
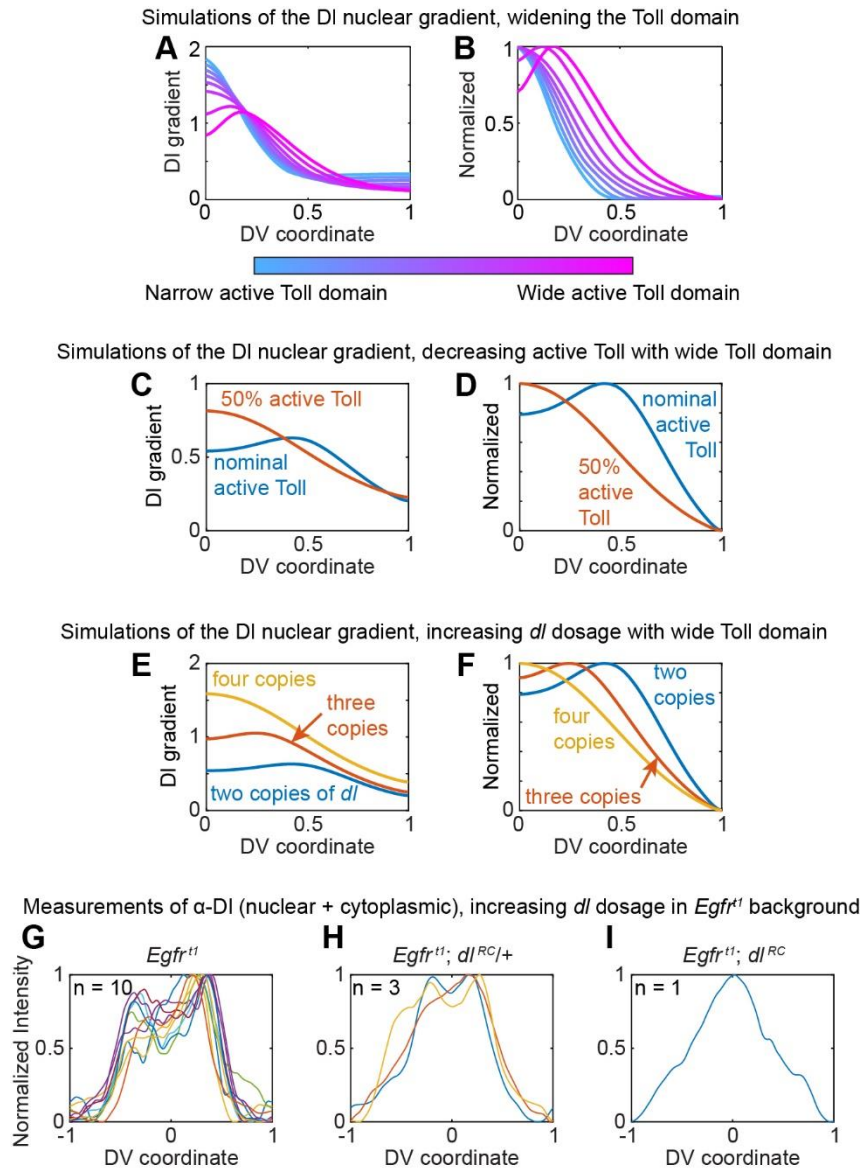
Figure S6: Rescuing The Single Dl Gradient Peak in Embryos with Wide Toll Domains, Related to Fig. 6. (A,B) Simulations of the Dl nuclear gradient, using the "Full model", with different widths of the Toll domain. Note that wider Toll domains result in wider gradients, and split peaks when the Toll domain is wide enough. The non-normalized plots are in (A), and normalized in (B). Colorbar indicates spectrum of Toll domain width. Note this colormap runs opposite to those for the diffusivity in Figs 2 and S2, as in all plots with this colormap, moving towards magenta refers to larger shuttling perturbations. Similar results can be obtained using the simplified model with Toll saturation. (C,D) Simulations of the Dl gradient with a wide Toll domain and decreased Toll receptors. The shuttling hypothesis with Toll saturation predicts that the split peak of the Dl gradient found when the Toll domain is widened is abolished if active Toll levels are reduced by 50%. This can be seen in both the non-normalized (C) and normalized (D) plots. (E,F) Simulations of the Dl gradient with a wide Toll domain and with three different *dl* dosages: wt (2x), 3x, and 4x. Increasing *dl* dosage results in a tall gradient (E), but also one with a less severe split peak (3x) or a

single peak (4x), as seen in the normalized plot (F). (G,H,I) Measurements of total dl (nuclear + cytoplasmic) in single *Egfr^{t1}* embryos (maternal genotype) with two copies of *dl* (G), three (H), and four (I). The averaged gradients from these sets of embryos can be seen plotted in Fig. 6E. Note that one outlier for *Egfr^{t1}* was not included in part (G).
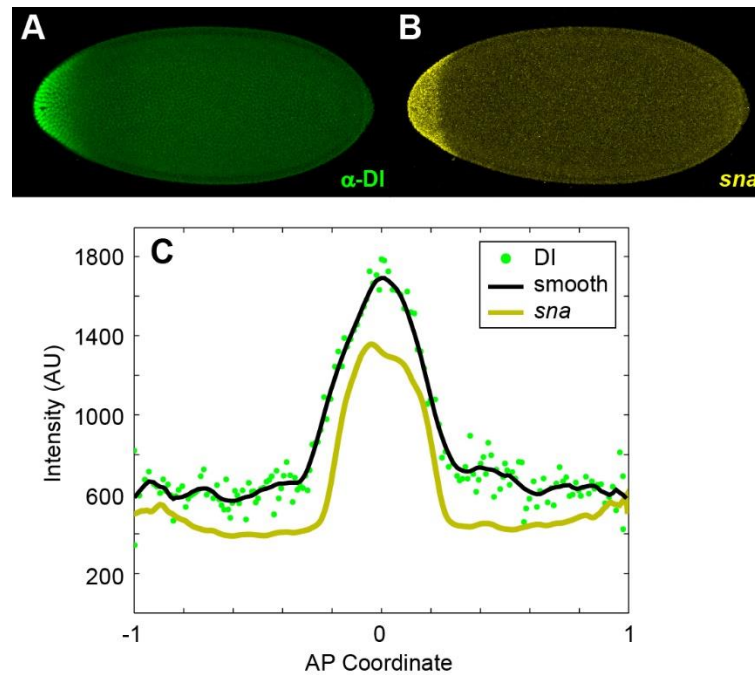

Figure S7: Embryos with AP Dl Gradients Under the Control of a Weaker (bcd) Promoter Do Not Exhibit the Double-Peak Phenomenon, Related to Figure 7. For this experiment, five embryos were analyzed, four of which had a detectable peak at the anterior pole. Note that the x-axis in (C) is the anteroposterior axis, not the dorsal-ventral axis as has been commonly used in the rest of this document.
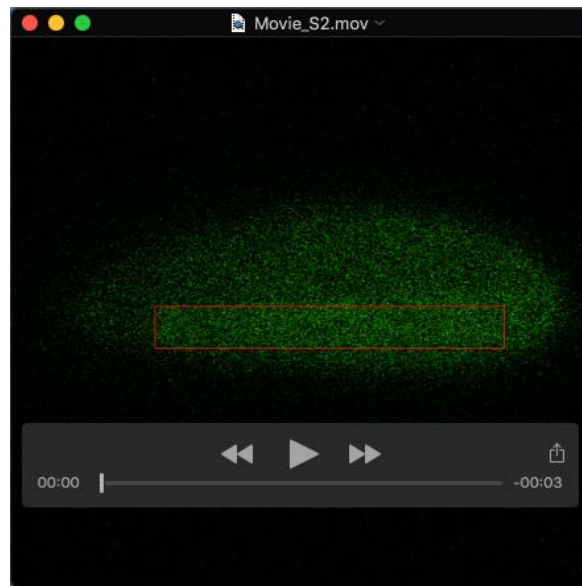
Supplemental Tables

Table S1. List of Primers Used for BAC Recombineering and Sequencing.
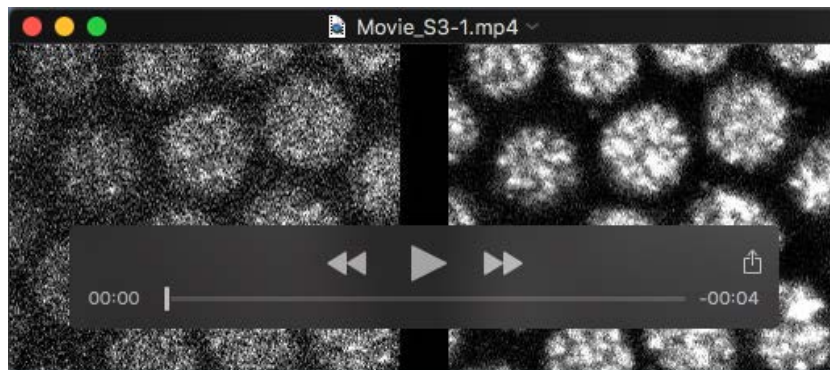
Click here to Download Table S1

Supplemental Movies
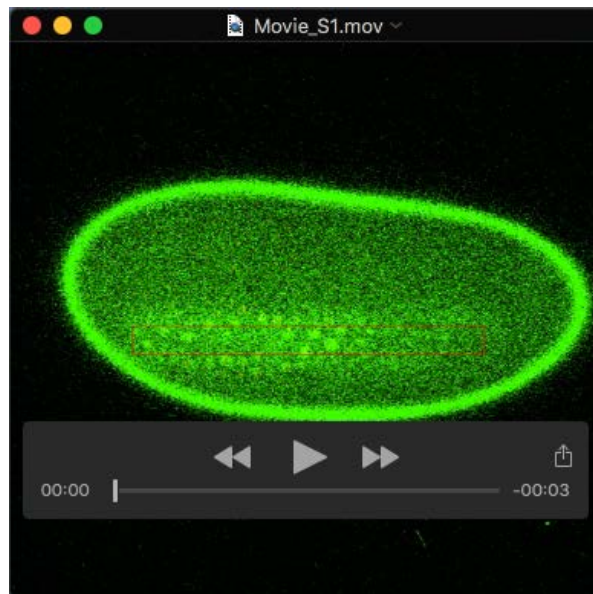
Movie S1: A Z-stack of an Embryo with Dl-paGFP That Has Been Activated Near the Ventral Midline, Related to Figure 1.

Movie S2: A Z-stack of an Embryo with Dl-paGFP That Has Been Activated on the Dorsal Side, Related to Figure 1.

Movie S3: Time course of a photobleaching experiment with a *dl/+;dl-mGFP/+* embryo. Left side: Dl-mGFP fluorescence. Right side: H2A-RFP fluorescence. Frames are 30 s apart. Movie played at 7 frames per second. Related to Figure 3.