Formulation of a model of the segment polarity network as a system of first-order ordinary differential equations using Ingeneue

Supplementary Information for "The Segment Polarity Network is a Robust Developmental Module"

G. von Dassow, E. Meir, E. M. Munro, and G. M. Odell.

Goals and Caveats

Our mathematical models consist primarily of *macroscopically observable* kinetic relationships among gene products. Consider a secreted signal known to activate some target genes. Without knowing the pathway between signal reception and target activation, one expects a dose-response relationship that is macroscopically observable in the same sense that one can measure V_{max} for an enzyme without knowing the transition-state intermediates. For a signal activating a target gene, one could measure the maximal response, the half-maximal activity of the signal, and the shape of the dose-response curve. These may or may not be easy to measure in practice, but since one could *in principle* do so, one can write qualitatively accurate equations describing such interactions with the measurable quantities as parameters. For our applications, gene networks are best described in terms of such interactions. Typically developmental biologists want to be able to say something like "wingless activates engrailed" without detailing all the (possibly unknown) steps in between, but one must be able to expand the model to account for such steps, should they be revealed. Here we present a general formulation for representing a gene network mathematically as a set of ordinary non-linear differential equations to facilitate testing hypotheses about systems-level properties of complex epigenetic processes. Our software package Ingeneue encapsulates this formulation, and we use it here to construct a model of the segment polarity gene network of *Drosophila*.

The primary limit to methods described here is that we assume cells and cell compartments are well-stirred reaction beakers in which interactions among macromolecular species follow mass-action kinetics. These methods may not be suitable when these assumptions can't be justified. First, for very small absolute numbers of interacting molecules, the continuum approximation of ordinary differential equations breaks down, and approaches that model stochastic effects afflicting individual molecules will provide more reliable descriptions of system behavior in such cases. Second, mass-action kinetics may not be appropriate for "solid-state" reactions such as conformational changes in large-scale macromolecular aggregates. Third, effects due to transient co-localization or other aggregative and compartmentalization phenomena are cumbersome in our framework. All three of these issues may be better treated by individual-based simulation methods, like those of Bray and colleagues.[1] Also, our

methods will not scale easily to hundreds or thousands of interacting species without major advances in numerical solving power. For such applications (e.g., DNA microarray data analysis) Boolean idealizations[2,3] may be more appealing.

Basic Forms

We represent each molecular species separately (mRNA, protein, protein complex, or small molecule), keeping track of which cellular compartment each resides in. Each cell in an arbitrary-sized two-dimensional field has one indexed copy of each equation for intracellular molecules; for cell-surface or extracellular molecules, each hexagonal cell has six indexed equations, one for each cell face. At present cells have equal volume, six equal faces, and do not rearrange during the simulation. This simplification, not required by the mathematical formulation, makes it simpler to implement the software. Some models may demand different notions of compartmentalization, such as individual organelles, but we have not yet needed to implement them.

Our generic formula for any molecular species is:

$$\frac{d[x]_{i,j}}{dt} = synthesis - decay \pm transformations \pm transport \qquad (1)$$

The index i refers to the cell, j to the cell face (for membrane-associated molecules). The "synthesis" term represents transcription for mRNAs, translation for proteins; for mRNAs and proteins there can be only one synthesis term unless multiple genes give rise to equivalent transcripts. This unlikely situation is irrelevant for the model developed here. Only primary transcripts and primary translation products are governed by synthesis terms; accumulation of derived forms (e.g. phosphoproteins) is governed by terms in the "transformation" category. This seeming subtlety has important consequences for the non-dimensionalization scheme discussed below. The "decay" category represents a first-order decay process; no matter how else a given species disappears, it must also exhibit first-order decay, even if very slowly. Many molecules undergo additional transformations, such as targeted cleavage, ligand-binding reactions, or phosphorylation. Finally, many molecules participate in various transport processes; exo– or endocytosis, cell-to-cell diffusion, and diffusion of membrane proteins within the plane of the cell surface, for examples.

Before illustrating the equations that constitute our segment polarity model, we discuss a simplified fragment to make clear how we translate empirical facts about kinetic interactions into mathematical equations. Consider the kinetic equations governing the mRNA and protein encoded by *hedgehog*:

$$a)\quad \frac{d[hh]_i}{dt} = T_{\max}\rho_{hh}\left(\frac{[EN]_i^{\nu_{ENhh}}}{K_{ENhh}^{\nu_{ENhh}} + [EN]_i^{\nu_{ENhh}}}\right) - \frac{[hh]_i}{H_{hh}}$$

$$b)\quad \frac{d[HH]_{i,j}}{dt} = \frac{P_{\max}\sigma_{HH}[hh]_i}{6} - \frac{[HH]_{i,j}}{H_{HH}} - k_{PTCHH}[HH]_{i,j}[PTC]_{n,j+3} \qquad (2)$$

$$c)\quad \frac{d[PH]_{i,j}}{dt} = k_{PTCHH}[HH]_{n,j+3}[PTC]_{i,j} - \frac{[PH]_{i,j}}{H_{PH}}$$

Eq. 2a governs *hedgehog* mRNA concentration in the ith cell, 2b governs Hedgehog protein concentration on the jth face of the ith cell, and 2c governs a complex formed between Hh and its receptor, Ptc (the product of the gene *patched*) on the jth face of the ith cell; the complex is denoted by $PH_{i,j}$. Each cell has one indexed copy of Eq. 2a, and 6 each of Eqs. 2b and 2c. In 2a the global parameter T_{max} is the maximum transcription rate, determined by RNA Polymerase, which can only travel so fast along DNA and can only pack so tightly along a transcription unit. A similar thing holds for ribosomes; P_{max} is the maximum possible translation rate. The parameter ρ_{hh} is a dimensionless parameter that determines how efficiently this particular gene is transcribed. The dimensionless rational function multiplying these two parameters represents activation of *hh* transcription by En; K_{ENhh} is the En concentration at which *hh* is half-maximally activated, and ν_{ENhh} is a "cooperativity" coefficient. For simplicity we postpone the regulation of *hh* transcription by Ci, which is a feature of the full model.

This kind pseudo-steady-state approximation is fundamental in our formulation. The dose-response curve it produces, illustrated in Box 1c, is the key to establishing a straightforward and empirically-accessible parameter space. For essentially every direct regulatory relationship between two molecules this form provides an approximation. No matter what the detailed structure of the *hedgehog* promoter and enhancers, there is some maximal rate at which *hh* transcription can be activated by this particular regulator; it follows that there is some concentration of that regulator (Engrailed protein here) at which half the maximum rate is achieved. Certainly gene regulatory regions may be more sophisticated, but it is generally possible to approximate the essential relationships by nesting and linear combination of dose-response curves (discussed below). An important parameter is the "cooperativity" coefficient ν. Although cooperativity is only one of several possible sources of non-linearity, we stick with this conceptual association because many regulatory relationships that involve ligand binding behave kinetically as if they are literally cooperative;[4,5] in which case ν is the Hill coefficient.

The second term in Eq. 2a represents first-order decay, in which the only free parameter is the half-life (inverse of the decay rate) H_{hh}. The first term of Eq. 2b represents synthesis of Hedgehog protein, $HH_{i,j}$, via translation. σ_{HH} is the translational efficiency of HH. The divisor is the number of sides of the cell, so that HH is "secreted" homogeneously to each of six cell faces. The second term in

Eq. 2b is the obligate first-order decay, and the third term represents a second-order reaction in which Hh binds to Ptc. K_{PTCHH} is the second-order binding rate. Note that Hh on the jth face of the ith cell interacts with available Ptc protein on the neighboring cell (index n) face j+3 modulo six (j+3 for short). Eq. 2c has primary synthesis term; the complex between Ptc and Hh is neither transcribed nor translated, but forms by dimerization of two monomers. However, it still exhibits first-order decay. Strictly speaking Eqs. 2b and 2c could include another term representing the reverse binding reaction. However, in the model developed here we omit dissociation, assuming that Hh and Ptc bind with reasonably high affinity.

Some extracellular molecules are free to diffuse either within the plane of the cell membrane or from one cell to another. Terms for diffusive transport appear in the complete segment polarity model but we do not include them in the "tutorial" version of Eqs. 2. Ingeneue uses a coarse but serviceable approximation of diffusion, in which the finest spatial resolution is an individual cell face. Cell surface molecules can transfer from one cell face to neighboring cell faces, a process governed by symmetrical first-order kinetic terms in which the free parameter is a transfer (i.e., diffusion) rate. Molecules that transfer from cell to cell via the extracellular medium, we allow such molecules to transfer from one cell face to another as well as laterally, again governed by first-order kinetic terms. This is because in the model developed here diffusing molecules cannot act except through a cell face.

Non-dimensionalization

While not necessary, it reduces the parameter count and verifies the dimensional correctness of the model to transform algebraically the differential equations to dimensionless form. This entails replacing every occurrence of dimensioned state variables (concentrations of molecular species and time) with scaled products yielding new state variables free of units. One may rearrange the equations seeking ways to combine parameters into new, dimensionless parameters. Among the benefits are that one may choose to scale state variables such that they vary within common bounds (a help to both numerical integration and human inspection) and one generally eliminates redundant free parameters. Furthermore, the dimensionless parameters, if the relation is chosen well, may be easier to interpret and measure experimentally than the dimensional ones.

The dimensionless model is identical to the dimensional one since it is merely an algebraic transformation. If the model maker knows the scaling constants used to introduce dimensionless variables, then the state of the dimensionless model is easily converted to the state of the dimensional one. Since the parameters of the dimensionless model are products of dimensioned parameters, known values of dimensioned parameters uniquely determine values for corresponding dimensionless parameters. The opposite is *not* true: the values of dimensional parameters cannot be retrieved from a specified point in the dimensionless parameter space. Rather, each value of a dimensionless parameter

determines a manifold in the dimensional parameter space. Although this will not generally affect the usability of the model, in some circumstances a dimensional model may be preferred.

For reasons that will become clear below we choose to non-dimensionalize the equations by substituting

$$t = T_\circ \tau;\ [x]_{i,j} = [x]_\circ x_{i,j}(\tau) \tag{3}$$

τ is the dimensionless time, $T_\circ$ a characteristic time constant relating t and τ, $[x]$ is the dimensional concentration, $x(\tau)$ a dimensionless replacement, and $[x]_\circ$ a characteristic concentration of x. We choose $[x]_\circ$ such that it equals the "maximal steady state concentration"; that is, the greatest dimensional concentration possible given the differential equation governing x. This entails solving the algebraic equation with the left side equal to zero and everything that contributes to synthesis set at *its* maximum value. In solving for the maximal steady state we only consider primary synthesis and decay terms; otherwise many equations would be impossible to resolve into simple, usable forms. We set the characteristic concentration for "derived" species (complexes, cleavage products, and so forth) to the characteristic concentration for the primary gene product from which they are derived. This facilitates comparisons such as, "what fraction of x is free and what fraction bound?" For an mRNA like *hh*:

$$[hh]_\circ = T_{\max} \rho_{hh} H_{hh} \tag{4}$$

For proteins:

$$[HH]_\circ = P_{\max} \sigma_{HH} H_{HH} [hh]_\circ = P_{\max} \sigma_{HH} H_{HH} T_{\max} \rho_{hh} H_{hh} \tag{5}$$

In these example equations, we assign the characteristic constant for PH the same as for Ptc so that we can easily assess the relative amounts of these related species. When we substitute through Eqs. 4 and 5, cancel, and re-name dimensionless groups of parameters, Eqs. 2 are rendered as:

$$a)\quad \frac{d\,hh_i}{d\tau} = \frac{T_\circ}{H_{hh}}\left(\frac{EN_i^{\nu_{ENhh}}}{\kappa_{ENhh}^{\nu_{ENhh}} + EN_i^{\nu_{ENhh}}} - EN_i\right)$$

$$b)\quad \frac{d\,HH_{i,j}}{d\tau} = \frac{T_\circ}{H_{HH}}\left(\frac{hh_i}{6} - HH_{i,j}\right) - T_\circ k_{PTCHH} HH_{i,j} \cdot [PTC]_\circ PTC_{n,j+3} \tag{6}$$

$$c)\quad \frac{d\,PH_{i,j}}{d\tau} = T_\circ k_{PTCHH} [HH]_\circ HH_{n,j+3} PTC_{i,j} - \frac{T_\circ PH_{i,j}}{H_{PH}}$$

This transformation is remarkably convenient because each state variable varies between 0 and 1 (with the exception of secondary forms such as PH that are scaled according to the primary form). Thus, the model is expressed in terms of *fractions of maximum concentration* for each component rather than absolute concentration. In addition, by combining dimensional parameters into dimensionless groupings, we eliminate roughly one third of the parameters without losing any realism. The dimensionless parameters often are more intuitive than the dimensional ones. For example, in Eq. 6a there appears the parameter κ_{ENhh}, which determines how avidly En activates *hh* transcription:

$$\kappa_{Xy} = \frac{K_{Xy}}{[X]_\circ} = \frac{K_{Xy}}{P_{\max}\sigma_X H_X T_{\max}\rho_x H_x} \tag{7}$$

T_{max}, P_{max}, and all ρ's and σ's have been subsumed in dimensionless groupings renamed κ, and instead of saying how many *molecules* of En are required to maximally activate *hh*, as in the dimensional equations, we say what *fraction* of the maximal Engrailed protein concentration is required to do so. Unlike the *K* parameters they replace, all κ parameters mean the same thing: a high value (near 1.0) means the regulator in question is weak because it must be present at close to its maximal concentration to significantly activate (or repress) the target in question, whereas a low value (10^{-3}) means a potent regulator. The ν parameters, dimensionless to begin with, remain unchanged. Half-lives remain dimensional, and we leave them so because it is natural to speak of the half-life in minutes, and the best we could do with further manipulation is to eliminate one of them by setting it equal to the characteristic time constant, $T_\circ$. We thus take $T_\circ$ to equal one minute.

Heterodimerizations, ligand-binding, and other second-order reactions require more attention. It looks in Eqs. 6 as if we could introduce a dimensionless parameter – why not:

$$\chi_{PTCHH} = T_\circ k_{PTCHH}[PTC]_\circ \quad \text{(bad idea)} \tag{8}$$

In those not at all uncommon circumstances in which second-order reactions are linked by the participation of binding partners in more than one of those reactions, introduction of parameters like χ_{PTCHH} would allow physically meaningless, contradictory combinations of parameters. Instead, to get rid of dimensions we simply group the units into one or the other of k_{PTCHH} or $PTC_\circ$ and they evaporate, leaving both parameters dimensionless but with a scalar value and range identical to their dimensioned counterparts. $T_\circ$ can be folded into k_{PTCHH} and $PTC_\circ$ to make a more convenient dimensionless parameter: the product of $T_\circ$, k_{PTCHH} and $PTC_\circ$ means, "what percent of this reaction *could* occur in a single unit of time?"

Several additional formulas haven't yet been treated explicitly. Phosphorylation and regulated cleavage are two representatives of a large class of transformations that can befall a molecule, which we typically represent like so:

$$\frac{dX}{dt} = -V_{\max}\left(\frac{Y^{\nu_{YX}}}{K^{\nu_{YX}} + Y^{\nu_{YX}}}\right)X + \cdots \tag{9}$$

Eq. 8 models *Y* as an allosteric regulator of an enzymatic process which transforms *X*. Of course more complex formulas may be required in some instances; for example V_{max} might itself vary with enzyme concentration, if we know the enzyme involved and that its availability varies. However, often we do *not* know the enzyme and must simply assume it is present and parameterize it appropriately. This is the situation in our simpler segment polarity models for Patched-regulated cleavage of Cubitus interruptus. These forms non-dimensionalize easily.

Another important class of terms governs reactions like inter-compartmental transfer (exocytosis, "diffusion", etc.). These are all modeled as first-order reactions, and the dimensionless parameters mean, again, "what fraction of the reaction takes place per unit time?"

More Complex Forms

While many interactions, such as binding reactions, cleavage, and enzymatic modifications are additive processes, this is not the case for regulated transcription. A typical gene might have a single promoter and transcription unit, but multiple regulatory sites, both activating and inhibiting, distributed throughout a large enhancer region. Since the single promoter and transcription unit imposes the upper bound on transcription rate, all regulators of transcription must be combined within a single term which saturates at the global limit imposed by the physical transcription process. Our general strategy is to approximate a kinetic model of an enhancer region using nested functions composed of two fundamental forms:

$$\Phi(X,\kappa_X,\nu_X) = \left(\frac{X^{\nu_X}}{\kappa^{\nu_X} + X^{\nu_X}}\right); \quad \Psi(X,\kappa_X,\nu_X) = 1 - \Phi(X,\kappa_X,\nu_X) \tag{10}$$

One standard combination, representing several independently-functioning activators acting, perhaps, at discrete enhancers but integrated by a single promoter, is:

$$\left(\frac{\sum_{i=1}^{n} \alpha_i \left(\frac{X_i^{\nu_{X_i}}}{\kappa_{X_i}^{\nu_{X_i}} + X_i^{\nu_{X_i}}} \right)}{1 + \sum_{i=1}^{n} \alpha_i \left(\frac{X_i^{\nu_{X_i}}}{\kappa_{X_i}^{\nu_{X_i}} + X_i^{\nu_{X_i}}} \right)} \right) \quad \text{or} \quad \left(\frac{\sum_{i=1}^{n} \alpha_i \Phi(X_i)}{1 + \sum_{i=1}^{n} \alpha_i \Phi(X_i)} \right) \tag{11}$$

Here the X_i represent activators, κ_x the activation coefficients for X_i, and α_i determines the degree to which each individual enhancer element can maximally activate the gene in question. Eq. 11 varies between 0 and 1 for all non-negative values of the X_i. To introduce inhibitors, we have several options, with $\Psi(Y)$ representing an inhibiting effect as in Eq. 10:

$$a) \quad \Psi(Y) \left(\frac{\alpha_1 \Phi(X_1) + \alpha_2 \Phi(X_2)}{1 + \alpha_1 \Phi(X_1) + \alpha_2 \Phi(X_2)} \right)$$

$$b) \quad \left(\frac{\alpha_1 \Phi(X_1 \Psi(Y)) + \alpha_2 \Phi(X_2)}{1 + \alpha_1 \Phi(X_1 \Psi(Y)) + \alpha_2 \Phi(X_2)} \right) \tag{12}$$

$$c) \quad \left(\frac{\Psi(Y)(\alpha_1 \Phi(X_1) + \alpha_2 \Phi(X_2))}{1 + \Psi(Y)(\alpha_1 \Phi(X_1) + \alpha_2 \Phi(X_2))} \right)$$

In Eq. 12a the activation quotient (the term in Eq. 11) multiplies an inhibitory quotient. This represents a global repressor that squelches transcription by lowering the promoter's efficiency. Another inhibitor might interact locally with a single activator, perhaps by competing for binding to a consensus site on the DNA; as in Eq. 12b, we can represent this situation by nesting $\Psi(Y)$ within a single $\Phi(X)$ from the sum in Eq. 11. Furthermore, a different inhibitor might prevent the enhancer from interacting with the promoter, in which case we would multiply the entire sum in Eq. 11 both on top and bottom by $\Psi(Y)$. While these formulas and their relatives don't capture all possibilities, they provide a versatile basis with which to approximate transcriptional regulation. We have developed alternate representations, though they are not used in the model developed here. For example, one alternative which is mathematically tidier than Eq. 12b but which accomplishes the same general effect, is:

$$\left(\frac{\alpha_1 \left(\frac{X_1 \Psi(Y)}{\kappa_{X_1}} \right)^{\nu_{X_1}} + \alpha_2 \left(\frac{X_2}{\kappa_{X_2}} \right)^{\nu_{X_2}}}{1 + \alpha_1 \left(\frac{X_1 \Psi(Y)}{\kappa_{X_1}} \right)^{\nu_{X_1}} + \alpha_2 \left(\frac{X_2}{\kappa_{X_2}} \right)^{\nu_{X_2}}} \right) \tag{13}$$

The virtue of Eq. 13 is that if either X_2 or X_1 are absent, Eq. 13 reduces to a simple case as in Eq. 6a. The disadvantage is that Eq. 12b better approximates a more general two-step kinetic model in which regulators bind to enhancers, then to the promoter complex. To date our experience has shown that these choices are usually matters of aesthetics.

Biologically Realistic Bounds on Parameters

A major advantage of the dimensionless model is that it is more straightforward to set biologically realistic bounds on the dimensionless parameters. This is not generally the case for the corresponding dimensional model unless many or most of the parameters can be constrained by measurement. The ease of assigning ranges to dimensionless parameters is a consequence of the choice to scale most state variables so that the maximum value they can take on is 1.0; scaling constants could be chosen arbitrarily, but our particular choice makes the parameter space come out conveniently.

In assigning ranges to parameters, one must keep several things in mind. First, while a given parameter could in principle be allowed an infinite range, this is obviously absurd because the vast majority of choices are physically unrealistic, and even more are biologically ludicrous. For example, a half-life in the millisecond range has no meaning whatsoever in the context of cellular macromolecules. Second, a choice of range must suit the context of interest. For instance, in the context of a pattern-formation process that takes place over several hours, a half-life much greater than that time is not so much unrealistic as it is irrelevent. That is, if the time-scale of the process of interest is about an hour, than a molecule with a half-life of a few hours is for most purposes equivalent to one with a half-life of a month. Thus one needn't waste computational effort looking around in the upper end of such a range. Third, mapping out ranges for each parameter defines a hyper-rectangle in the parameter space. Many of the corners of this box may be highly unrealistic. The general expectation is that one will be able to recognize such unrealistic combinations if they dominate the behavior of the model. If they do not, they are harmless except so much as they waste computer time. Ingeneue provides some crude facilities to avoid defined regions of parameter space, and we continue to investigate better strategies. Finally, the proof of the pudding is in the eating: if a model fails, it may be due to an unsuitable choice of bounding rectangle. Below we discuss explicitly several classes of dimensionless parameter, outlined in Table S1.

Table S1.

Parameter	Meaning	Realistic (General) Range	Range used for SP Model
κ	half-maximal activation coefficient	10^{-3} – 10	10^{-3} – 1
H	half-life (inverse of degradation rate)	1 – 10^4 min. (for mRNA or protein)	5 – 100 min
ν	Hill coefficient	1 – 50 (highest measured is 35)	1 – 10
α	saturability coefficient for an enhancer	0.1 – 10	1 – 10
transfer rates	how much reaction occurs per unit time	10^{-3} – 10	10^{-3} – 1.0
transform rates	ditto; but for cleavage, phosphorylation, etc.	10^{-3} – 10	10^{-3} – 10

Half-maximal activation coefficients

Small κ indicates a potent regulator. If κ is much greater than one, then there is usually no way the regulator will achieve a concentration necessary to half-maximally activate the target. Thus, values between 1 and 10 mean a very weak to an essentially absent connection, unless the regulator is scaled (e.g. relative to a precursor form) such that it can achieve higher concentrations. If all the connections are known to be effective *in vivo* then this part of the range may be omitted, as we do in the segment polarity model. The lower bound comes from considering the maximum cellular concentration of a protein produced from a single gene; it is given by the formula in Eq. 5. Some estimates for the rates of limiting processes are available in the literature.[6,7] For eukaryotic cells the maximum transcription rate is on the order of 10–100 nucleotides per second. The closest possible spacing between RNA Polymerase II complexes is about 50 nucleotides, so roughly 10–100 molecules of mRNA can be synthesized per minute from a single template. A similar rate holds for translation. If the half-lives in this formula are between 10–10^3 min. each, and the efficiencies (ρ and σ) are each between 0.01–1.0 (1–100%), this means that we expect the maximum cellular concentrations of typical eukaryotic proteins to be somewhere between 1–10^{10} molecules per cell. The low end is clearly a real limit. The high end represents a very efficiently transcribed and translated gene whose mRNA and protein are each quite stable (on the order of a day), that experiences sustained full-strength activation over a period substantially longer than the half-life, and inhabits a cell whose synthetic machinery is non-limiting and operates at peak levels.

Obviously few genes in real cells match either extreme. More realistic conditions would be that the typical protein should have a maximum cellular concentration of 10^3–10^7 molecules per cell. For a blastoderm cell in a *Drosophila* embryo, which might have a total volume on the order of a picoliter, 10^7 molecules per cell is nearly millimolar concentration. The lower bound, 10^3 molecules

per cell, is sub-micromolar concentration. For a protein at the high end a κ value near 1.0 would mean that millions of molecules per cell are required to have much impact on the target. A value near 10^{-3} means that tens of thousands are required per cell, but the more important interpretation is that a tiny trickle of expression of this hypothetical protein is quite effective. Lower values of κ can't be qualitatively much different, so we would not bother exploring further in that direction. On the other hand a protein at the low end of the possible concentration spectrum a κ value of 10^{-3} would mean that it activates or represses its target half-maximally at a only one molecule per cell. In such a case lower values of κ would be physically implausible.

Half-lives

A variety of generic processes contribute to the differential stability of biological macromolecules in living cells. No cellular macromolecule accumulates without bounds, so we model each with a first-order decay term with a single intrinsic and constant degradation rate (the inverse of a half-life). This term may indeed mask many more or less specific degradation processes, but often these will be unknown. Specific degradation processes that *are* known are represented as separate additive terms. For a few cases the half-lives of proteins or mRNAs have been measured in living cells. Some molecules are very stable; we assume that structural proteins of connective tissue might have half-lives measured in months. Others are very *un*stable; in late blastoderm *Drosophila* embryos *ftz* mRNA has an apparent half-life on the order of a few minutes.[8] Small molecules (cAMP, for instance) may have far shorter half-lives due to highly efficient enzymatic mechanisms which may need to be modeled as a first-order decay process.

Another contextual consideration goes into assigning ranges on half-lives: the time-scale of the process being modeled. First, if the process of interest takes place within an hour, then a half-life for a participating molecule of 10^3 min. is equivalent to 10^4 min: both values mean that absent new synthesis more or less the same amount of this molecule is around at the end of the simulation as at the beginning. One can typically identify a time-scale on which changes take place, and set the upper bound on half-lives only high enough to allow a comfortable margin. In our segment polarity model we know that in the living fly embryo gene expression patterns change on a time scale of 30 min. or less during the stages we're interested in. We therefore allow half-lives to range up to 100 min.; higher values might constrain the system from changing on the time-scale upon which it is known to do so.

Cooperativity (Hill) coefficients

True allosteric cooperativity is one source of non-linearity in dose-response curves; other processes (e.g. titrations) may also result in non-linear response curves.[4,5] The parameter ν allows us

to approximate all of these, and in the case of cooperative binding is equivalent to the Hill coefficient. A value of 1.0 for ν gives a linear response that saturates at high levels of regulator. Higher values introduce a sigmoid shape. Values above about 10 behave like a switch: at low levels of regulator there is essentially no response, but at regulator levels equaling κ the target switches on rapidly. We do not know of many values below 1.0 in a real mechanism, so in the segment polarity model we provisionally set this as the lower bound. For direct interactions between molecules or between DNA binding proteins and enhancers, the realistic upper bound on ν may be around 10. However, indirect pathways, especially enzymatic ones, may have very high effective cooperativity: this value was measured at 35 for the MAP kinase cascade in *Xenopus* oocytes.[9]

Saturability coefficients

We provide for different regulators to behave additively at low levels but saturate at high levels. Not all enhancer elements are equally able to stimulate transcription. The parameter α determines for each cluster of regulators within a term such as Eq. 11 the extent to which that cluster can saturate the whole term. α values higher than 10.0 make little difference, whereas a value of 1.0 means the term in question can only half-way saturate the promoter. For most applications, including this one, we bound α on the lower end at 1.0.

Transfer rates

The ranges suitable for intercompartmental transfer rates depend, like half-lives, on the time-scale of the process of interest. Transfer rates are first-order rate constants. The dimensionless rate constants no longer mean, "how *many* molecules are transferred across a given boundary per unit time?" but rather, "what *fraction*…?" A dimensionless exocytosis rate of 10, in a model in which most processes take place on the scale of minutes or longer, therefore means that whatever item is being exocytosed is transferred from cytoplasm to cell exterior virtually instantaneously relative to other kinetic processes. We seriously doubt that most cellular processes really transfer molecules around so quickly (exceptions might include rapid exocytosis of docked vesicles at a synapse), so for most purposes the high end can be ignored. At the low end, we make sure there is sufficient room to slow transfer processes to a crawl relative to other processes in the model. This is because one often simply does not know for sure whether a particular molecule actually undergoes a possible transfer process (such as intercellular diffusion or endocytic re-uptake). For instance, many secreted molecules (like the Wingless and Hedgehog proteins in the segment polarity network[10]) may stick fairly tightly to components of the cell surface and/or extracellular matrix, thus restricting the effectiveness of diffusion to short distances or long time-scales.

Transformation rates

Rates for transformations like cleavage, phosphorylation, and ligand binding are treated like standard kinetic rate constants, with the following considerations. Unlike transfer processes, usually one knows whether or not a significant amount of cleavage or phosphorylation takes place in the biological context, so it is not usually necessary to allow parameter values so low that the reaction would be effectively eliminated. Some transformations are inherently faster than others: phosphorylation is often fast, whereas regulated proteolysis may be slower. Some transformations (e.g., phosphorylation) should be reversible whereas other (proteolysis) should not. For those that are reversible, ranges must allow the possibility that the reverse reaction rate (de-phosphorylation) may be significantly faster than the forward rate.

The Segment Polarity Network Model

We applied the formula described above to derive a mathematical description of interactions between a subset of the segment polarity genes. The map of the interactions we modeled is shown in Box 1 of the paper to which this document is a supplement. The complete system of equations is shown as Eqs. 17 at the end of this document. For the benefit of readers who are not as facile with reading equations, we will summarize in words what the equations say in maths. The primary work establishing the network topology is cited in the main paper (see Methods); additional citations below support choices made in formulating the model.

- *engrailed* (*en*) transcription is activated by extracellular Wingless (Wg) protein (EWG), but the N-terminal repressor fragment of Cubitus interruptus (full-length protein=Ci, repressor=CN) squelches this activation. Wingless stimulates a response via a signal transduction pathway involving the products of *Dfrizzled2* (the Wg receptor), *disheveled*, *zeste-white 3* (glycogen synthase kinase 3), *armadillo* (beta-catenin), and *pangolin* (a transcription factor).[11] We have completely collapsed this signal transduction pathway in an attempt to build the simplest plausible model first. Work to be published elsewhere explores the consequences of incorporating the entire signal transduction pathway. We model CN as a global repressor of *engrailed* transcription. No one has *directly* demonstrated this interaction, and thus its existence is a prediction of our model. Ci antibodies stain weakly the polytene chromosome band in which *en* resides.[12] Also, in imaginal discs *en* is activated by the highest levels of Hedgehog (Hh) signalling in a manner dependent upon the known Hh signal transduction pathway (terminating with the regulation of cleavage of Ci).[13] CN represses other targets of the full-length, activator form of Ci.[12] These facts may support our

hypothesis that CN represses *en*. Nevertheless, this link has the weakest support of any in our model, but should be readily testable. Continuing to Eq. 17b, En protein is translated in proportion to the availability of *en* mRNA.

- *wg* expression is activated (Ci) or repressed (CN) by the two different forms of Ci represented in this model. *wg* expression is also auto-activated[14-16] as a function of the level of the pre-secretion form of Wg, lWG. lWG, in addition to translation and decay, exchanges via exocytosis and endocytosis with EWG. EWG, in addition to the reciprocal effects of exo- and endocytosis, experiences flux both between neighboring faces of an individual cell and between apposite faces of neighboring cells. Note that this suite of transport fluxes encompasses the possibility for cell-to-cell Wingless traffic through transcytosis.[17]
- *ptc* is transcribed under the control of Ci and CN in the same way as *wg*. The Ptc protein is present on the cell surface, where in addition to diffusing around the membrane of the cell, it binds Hh on neighboring cells according to a second-order heterodimerization reaction to form PH, the Patched-Hedgehog complex.
- *ci* is basally expressed (at a level determined by the presence of a "dummy" variable B) but is repressed by En. The Ci protein is cleaved at a rate determined by the abundance of free Ptc on the surface of the cell. The cleavage reaction results in the accumulation of CN. In reality the product of *smoothened*, another multi-pass transmembrane protein, is thought to be the active signaling transducer; Hh binding to Ptc relieves repression of Smo signaling activity, presumably by titrating away Ptc and keeping it from binding to Smo.[18,19]
- *hh* is activated by En and repressed by CN. The protein is present on the surface of cells, where it can bind Ptc. In this simplest model this binding reaction is irreversible. There is compelling evidence in imaginal discs to believe that Hh protein diffuses some distance from the site of its production.[13] Hh is synthesized with a cholesterol moeity which causes it to be tethered to the cell membrane upon secretion unless the protein undergoes an auto-cleavage reaction whose regulation, as far as we are aware, remains poorly understood.[20] Therefore, for the sake of speed of simulations and simplicity, we do not allow Hh to diffuse; in work to be published elsewhere we will show that this choice does not appreciably affect the outcome of simulations described here.

In each case where there is ambiguity about the nature of a particular link (an ambiguity in which formula we choose, not whether a link exists or a reaction takes place), we have tested to see whether or not our choices in the model make a qualitative difference in the behaviors reported here. The only choice which affects the behavior significantly is that CN must be a global, rather than a local,

repressor of *en*: CN must be able to overcome any level of Wg activation of *en* transcription or *en* will be activated on all sides of Wg-producing cells.

Ingeneue – a Java toolkit for gene network simulations

Our initial efforts employed the symbolic mathematics package Mathematica. While Mathematica is an excellent tool, for us it had several drawbacks. It was slow compared to what we believed could be achieved with custom-written code. It was difficult to modify the model or to concoct a new model without re-writing dozens of equations, an error-prone chore that we were eager to avoid. Finally, we paled at the prospect of trying to repeat this process with another gene network. We therefore began work on a custom gene network simulation program (Ingeneue) designed to eliminate these drawbacks. We chose the object-oriented language Java to develop this package. Java remains somewhat slower than traditional compiled languages, but recent advances have narrowed the performance gap between Java and C++. Java provides several advantages. First, Java code compiled on one kind of computer can run without modification on another kind, saving programmer effort. Second, Java lacks most of the common pitfalls that make C++ a difficult language for novices, an essential feature if we expect our software to be used *and extended* by working biologists. Third, Java allows run-time modification of running code and dynamic loading of code modules (objects). This means the simulation can be arbitrarily changed while the program is running, rather than having to stop and re-link the program to accomplish each change.

Ingeneue is presently a working prototype that we use as a research tool. The program reads pseudo-English text files describing a network and various simulation tasks. The core package is a set of objects that work together. "Cells" are compartmental containers for "Nodes" (mRNAs, proteins, etc. – molecular species). Each Node has its own group of Affectors; an Affector object usually constitutes a single additive term in the Node's differential equation. For example, the most commonly used Affector encapsulates the term for first-order decay. We cultivate an expanding garden of Affectors and it is a trivial programming task to write new types. The user specifies initial concentrations for every Node in every Cell, using a library of InitialCondition objects that make describing a pattern straightforward. The script also specifies other parameters and conditions. A ModelRunner object conducts the simulation by marching a numerical integrator along timesteps; each Node uses its list of Affectors to compute its time derivative as required by any integration scheme in use. A complex StoppingCondition object, assembled from a library of simpler StoppingConditions, monitors the course of the integration. StoppingConditions encapsulate recipes both for assigning a score for how well the present state matches the desired pattern and for stopping the integration run if it is either hopelessly misbehaving or behaving so well that further integration would be superfluous. By default Nodes use a standard-issue adaptive step-sizing embedded Runga-Kutta integrator. We

have recently developed a customized integration strategy that exploits stereotyped features of the equations used by Ingeneue. This strategy is based on classical predictor-corrector methods but can sometimes achieve much higher speeds, especially for complex models, and will be described elsewhere.

The Ingenue core is dressed with various utility objects that handle tasks like parsing script files, displaying model state on screen, and so on. This is sufficient if the user knows all the parameters (rate constants, half-lives, etc.) necessary to constrain the model, as might be the case for certain well-studied biochemical processes such as the life cycle of lambda phage. However, this is generally *not* the case. More often, *none* of the parameters governing epigenetic interactions has been measured, as in the segment polarity model described here. Because of this problem the Ingeneue core is wrapped in a layer of Iterator objects. Iterators encapsulate algorithms for searching the parameter space for sets of values that lead to desired behavior. The simplest possible scheme, of course, is random choice, and indeed the Iterator that implements a random search is the one we use the most. However, the parameter space is often of so very high dimension that searching it presents a combinatorical problem that mere brute force and more computers cannot hope to overcome. Thus, most Iterator objects encapsulate directed search strategies, including non-linear optimizers, population-selection schemes, and so on. This area of development remains immature at present.

Pattern recognition function for the segment polarity model

We needed to assess how well the segment polarity model could mimic expression patterns of key segment polarity genes in the *Drosophila* embryo. Since *wg*, *hh*, and *en* are the best-characterized "outputs" of the segment polarity network, we constructed a goodness-of-fit function that scored the time-evolution of numerical integrations according to how well the dynamic pattern of these three genes *in simulo* matched the known pattern *in vivo*. For convenience using optimization methods (although such methods were not used in this paper) the function reports low scores (near 0.0) for good matches, and scores near 1.0 for bad matches. For all results described here we used a threshold of 0.2 below which the pattern is deemed acceptable.

The score combines several independent components falling into two categories. The first is a set of mild threshold functions. In each simulated cell a battery of sigmoids assesses the expression of *wg*, *en*, and *hh*. Crudely, these genes are considered to be "on" if they are above 10% of their maximal expression level, and "off" otherwise. However, each threshold function returns a scalar score with the half-maximal value (in this application) at the 10% threshold according to the formulas

$$T_{off} = \alpha_{\max}\tau(x_{i,j}) = \alpha_{\max}\frac{\left(x_{i,j}/x_t\right)^3}{1+\left(x_{i,j}/x_t\right)^3} \qquad (14)$$
$$T_{on} = \alpha_{\max}\left(1-\tau(x_{i,j})\right)$$

where x_t is the threshold for x (10% here), α_{max} is a worst-possible score (0.5 here) and indices i and j refer to the species and the cell scored by this particular threshold. The rationale for using mild sigmoids is that these functions respond linearly around the threshold but provide both diminishing benefits and diminishing punishments further and further from the threshold. This feature appeals to the intuition that if a particular gene is well above some threshold it should not matter exactly how far above. Experience has also shown that this choice helps certain non-linear optimization methods navigate the parameter space using these scoring functions.

All the individual threshold scores are combined by a meta-function according to another formula that saturates at high (poor) scores:

$$\left(\frac{\sum_{i=1}^{\#x}\sum_{j=1}^{\#cells}T(x_{i,j})}{1+\sum_{i=1}^{\#x}\sum_{j=1}^{\#cells}T(x_{i,j})}\right) \qquad (15)$$

While the software objects that implement these threshold functions can use a fading memory or a time average of the value of the species they assess, we did not use those features here. Instead, the entire pattern was scored at two timepoints: typically 200 min. and 1000 min. This means that the system is required to achieve the proper on/off pattern within 200 simulated minutes, and that it must express qualitatively the same pattern at the later time. For some simulations (Table 1 of the main paper) the initial timepoint was 600 (instead of 200) min.

The second component of the pattern recognition scheme focuses on specific pattern elements: stable stripes of *wg*, *en*, and *hh* in specified columns. This function is substantially more complicated, and not strictly necessary for crude searches. However, we used it in all the work described here because it helps the goodness-of-fit function catch almost all patterns that we humans would also consider good matches, while helping to reject the few poor matches that the thresholds alone would otherwise allow. Like the thresholds, stripe scoring also uses mild sigmoids to assess 1) the level of expression along the stripe, 2) the ratio differentiating the stripe from surrounding cells, and 3) the variance along the stripe. Several parameters control the stripe scoring function itself; these were hand-trained by a human monitoring thousands of simulations, and remain fixed for this application. In

addition, the stripe scoring function punishes temporal oscillations by assessing (after allowing some time for initial transients) peaks and valleys in the timecourse in each cell of the stripe. The oscillation measurement and the stripe score are then combined:

$$S(x_{i,j}) = \left(\alpha + (1-\alpha)\left(\frac{Amplitude(x_{i,j})}{Peak(x_{i,j})} \right) \right)^{StripeScore(x_{i,j})} \quad (16)$$

Here indices i and j indicate the species and the column in which it is expected to form a stripe, and α limits the impact of oscillations on the score (α=0.5 here). Herein we have exclusively used this function in a mode where the stripe component is time-averaged from whenever the amplitude of oscillations starts to decline and either the specified stopping point is reached for the integration or the amplitude dies away.

The final score is the maximum of the threshold component (Eq. 15) and the oscillating stripe component (Eq. 16). Final scores below 0.2 dub the pattern acceptable and the current parameter set is saved. The coefficients in the scoring formula, including the 0.2 threshold, were hand-tuned by a human watching tens of thousands of simulation runs. The result is that the goodness-of-fit function misses fewer than one in ten cases that the human would have (possibly erroneously) considered acceptable, whereas it catches (surely erroneously) fewer than 1 in 50 patterns that the human would have rejected. Thus we can use it to screen literally tens of millions of simulation runs in the course of the work reported here, reporting a conservative but close estimate of the true success rate. Although simpler functions might seem more convenient, the one described here allowed us to minimize the error inherent in automatic iterative searching of a vast parameter space.

Literature Cited for Supplementary Information

1. Morton-Firth, C. J., Shimizu, T. S. & Bray, D. A free-energy-based stochastic simulation of the Tar receptor complex. *J Mol Biol* **286**, 1059-1074 (1999).

2. Kauffman, S. A. *The origins of order: self organization and selection in evolution* (Oxford University Press, New York, 1993).

3. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. *Bioessays* **20**, 433-440 (1998).

4. Gutfreund, H. *Kinetics for the life sciences : receptors, transmitters, and catalysts* (Cambridge University Press, Cambridge ; New York, 1995).

5. Wyman, J. & Gill, S. J. *Binding and linkage : functional chemistry of biological macromolecules* (University Science Books, Mill Valley, Calif., 1990).

6. Kornberg, A. & Baker, T. A. *DNA replication* (W.H. Freeman, New York, 1992).

7. Alberts, B. *Molecular biology of the cell* (Garland Pub., New York, 1994).

8. Edgar, B. A., Weir, M. P., Schubiger, G. & Kornberg, T. Repression and turnover pattern fushi tarazu RNA in the early Drosophila embryo. *Cell* **47**, 747-754 (1986).

9. Ferrell, J. E., Jr. & Machleder, E. M. The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science* **280**, 895-898 (1998).

10. Vincent, J. P. & Lawrence, P. A. Drosophila wingless sustains engrailed expression only in adjoining cells: evidence from mosaic embryos. *Cell* **77**, 909-915 (1994).

11. Cadigan, K. M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev* **11**, 3286-3305 (1997).

12. Aza-Blanc, P., Ramirez-Weber, F. A., Laget, M. P., Schwartz, C. & Kornberg, T. B. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell* **89**, 1043-1053 (1997).

13. Strigini, M. & Cohen, S. M. A Hedgehog activity gradient contributes to AP axial patterning of the Drosophila wing. *Development* **124**, 4697-4705 (1997).

14. Yoffe, K. B., Manoukian, A. S., Wilder, E. L., Brand, A. H. & Perrimon, N. Evidence for engrailed-independent wingless autoregulation in Drosophila. *Dev Biol* **170**, 636-650 (1995).

15. Manoukian, A. S., Yoffe, K. B., Wilder, E. L. & Perrimon, N. The porcupine gene is required for wingless autoregulation in Drosophila. *Development* **121**, 4037-4044 (1995).

16. Hooper, J. E. Distinct pathways for autocrine and paracrine Wingless signalling in Drosophila embryos. *Nature* **372**, 461-464 (1994).

17. Pfeiffer, S. & Vincent, J. P. Signalling at a distance: transport of wingless in the embryonic epidermis of drosophila. *Semin Cell Dev Biol* **10**, 303-309 (1999).

18. Chen, Y. & Struhl, G. Dual roles for patched in sequestering and transducing Hedgehog. *Cell* **87**, 553-563 (1996).

19. Chen, Y. & Struhl, G. In vivo evidence that Patched and Smoothened constitute distinct binding and transducing components of a Hedgehog receptor complex. *Development* **125**, 4943-4948 (1998).

20. Porter, J. A. *et al.* Hedgehog patterning activity: role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. *Cell* **86**, 21-34 (1996).

Notation : $X_{n,j+3} = amount\ of\ X\ on\ apposite\ cell\ face;\ X_{i,T} = \sum_{j=1}^{6} X_{i,j};\ X_{n,T} = \sum_{j=1}^{6} X_{n,j+3};\ X_{i,lr} = X_{i,j-1} + X_{i,j+1}$

$$a)\ \frac{d\,en_i}{d\tau} = \frac{T_\circ}{H_{en}}\left(\frac{EWG_{n,T}\left(1-\frac{CN_i{}^{v_{CNen}}}{\kappa_{CNen}{}^{v_{CNen}}+CN_i{}^{v_{CNen}}}\right)^{v_{WGen}}}{\kappa_{WGen}{}^{v_{WGen}}+EWG_{n,T}\left(1-\frac{CN_i{}^{v_{CNen}}}{\kappa_{CNen}{}^{v_{CNen}}+CN_i{}^{v_{CNen}}}\right)^{v_{WGen}}}-en_i\right)$$

$$b)\ \frac{d\,EN_i}{d\tau} = \frac{T_\circ}{H_{EN}}\left(en_i - EN_i\right)$$

$$c)\ \frac{d\,wg_i}{d\tau} = \frac{T_\circ}{H_{wg}}\left(\frac{\alpha_{CIwg}\cdot\left(\frac{CI_i\left(1-\frac{CN_i{}^{v_{CNwg}}}{\kappa_{CNwg}{}^{v_{CNwg}}+CN_i{}^{v_{CNwg}}}\right)^{v_{CIwg}}}{\kappa_{CIwg}{}^{v_{CIwg}}+CI_i\left(1-\frac{CN_i{}^{v_{CNwg}}}{\kappa_{CNwg}{}^{v_{CNwg}}+CN_i{}^{v_{CNwg}}}\right)^{v_{CIwg}}}\right)+\alpha_{WGwg}\cdot\left(\frac{IWG_i{}^{v_{WGwg}}}{\kappa_{WGwg}{}^{v_{WGwg}}+IWG_i{}^{v_{WGwg}}}\right)}{1+\alpha_{CIwg}\cdot\left(\frac{CI_i\left(1-\frac{CN_i{}^{v_{CNwg}}}{\kappa_{CNwg}{}^{v_{CNwg}}+CN_i{}^{v_{CNwg}}}\right)^{v_{CIwg}}}{\kappa_{CIwg}{}^{v_{CIwg}}+CI_i\left(1-\frac{CN_i{}^{v_{CNwg}}}{\kappa_{CNwg}{}^{v_{CNwg}}+CN_i{}^{v_{CNwg}}}\right)^{v_{CIwg}}}\right)+\alpha_{WGwg}\cdot\left(\frac{IWG_i{}^{v_{WGwg}}}{\kappa_{WGwg}{}^{v_{WGwg}}+IWG_i{}^{v_{WGwg}}}\right)}-wg_i\right)$$

$$d)\ \frac{d\,IWG_i}{d\tau} = \frac{T_\circ}{H_{IWG}}\left(wg_i - IWG_i\right) + T_\circ\left(r_{EndoWG}EWG_{i,T} - r_{ExoWG}IWG_i\right)$$

$$e)\ \frac{d\,EWG_{i,j}}{d\tau} = T_\circ\left(\frac{r_{ExoWG}IWG_i}{6} - r_{EndoWG}EWG_{i,j} - r_{MxferWG}EWG_{i,j} + r_{MxferWG}EWG_{n,j+3} - 2r_{LMxferWG}EWG_{i,j} + r_{LMxferWG}EWG_{i,lr}\right) - \frac{T_\circ EWG_{i,j}}{H_{IWG}} \quad (17)$$

$$f)\ \frac{d\,ptc_i}{d\tau} = \frac{T_\circ}{H_{ptc}}\left(\frac{CI_i\left(1-\frac{CN_i{}^{v_{CNptc}}}{\kappa_{CNptc}{}^{v_{CNptc}}+CN_i{}^{v_{CNptc}}}\right)^{v_{CIptc}}}{\kappa_{CIptc}{}^{v_{CIptc}}+CI_i\left(1-\frac{CN_i{}^{v_{CNptc}}}{\kappa_{CNptc}{}^{v_{CNptc}}+CN_i{}^{v_{CNptc}}}\right)^{v_{CIptc}}}-ptc_i\right)$$

$$g)\ \frac{d\,PTC_{i,j}}{d\tau} = \frac{T_\circ}{H_{PTC}}\left(\frac{ptc_i}{6} - PTC_{i,j}\right) - T_\circ k_{PTCHH}[HH]_\circ HH_{n,j+3}\cdot PTC_{i,j} + T_\circ\left(r_{LMxferPTC}PTC_{i,lr} - 2r_{LMxferPTC}PTC_{i,j}\right)$$

$$h)\ \frac{d\,ci_i}{d\tau} = \frac{T_\circ}{H_{ci}}\left(\frac{B_i\left(1-\frac{EN_i{}^{v_{ENci}}}{\kappa_{ENci}{}^{v_{ENci}}+EN_i{}^{v_{ENci}}}\right)^{v_{Bci}}}{\kappa_{Bci}{}^{v_{Bci}}+B_i\left(1-\frac{EN_i{}^{v_{ENci}}}{\kappa_{ENci}{}^{v_{ENci}}+EN_i{}^{v_{ENci}}}\right)^{v_{Bci}}}-ci_i\right)$$

$$i)\ \frac{d\,CI_i}{d\tau} = \frac{T_\circ}{H_{CI}}\left(ci_i - CI_i\right) - T_\circ C_{CI}CI_i\left(\frac{PTC_{i,T}{}^{v_{PTC\cdot CI}}}{\kappa_{PTC\cdot CI}{}^{v_{PTC\cdot CI}}+\boldsymbol{PTC}_{i,T}{}^{v_{PTC\cdot CI}}}\right)$$

$$j)\ \frac{d\,CN_i}{d\tau} = T_\circ C_{CI}CI_i\left(\frac{PTC_{i,T}{}^{v_{PTC\cdot CI}}}{\kappa_{PTC\cdot CI}{}^{v_{PTC\cdot CI}}+PTC_{i,T}{}^{v_{PTC\cdot CI}}}\right) - \frac{T_\circ CN_i}{H_{CI}}$$

$$k)\ \frac{d\,hh_i}{d\tau} = \frac{T_\circ}{H_{hh}}\left(\frac{EN_i\left(1-\frac{CN_i{}^{v_{CNhh}}}{\kappa_{CNhh}{}^{v_{CNhh}}+CN_i{}^{v_{CNhh}}}\right)^{v_{ENhh}}}{\kappa_{ENhh}{}^{v_{ENhh}}+EN_i\left(1-\frac{CN_i{}^{v_{CNhh}}}{\kappa_{CNhh}{}^{v_{CNhh}}+CN_i{}^{v_{CNhh}}}\right)^{v_{ENhh}}}-hh_i\right)$$

$$l)\ \frac{d\,HH_{i,j}}{d\tau} = \frac{T_\circ}{H_{HH}}\left(\frac{hh_i}{6} - HH_{i,j}\right) - T_\circ k_{PTCHH}[PTC]_\circ PTC_{n,j+3}\cdot HH_{i,j} + T_\circ\left(r_{LMxferHH}HH_{i,lr} - 2r_{LMxferHH}HH_{i,j}\right)$$

$$m)\ \frac{d\,PH_{i,j}}{d\tau} = T_\circ k_{PTCHH}[HH]_\circ HH_{n,j+3}\cdot PTC_{i,j} - \frac{T_\circ PH_{i,j}}{H_{PH}}$$