Morpheus Unbound: Reimagining the Morphogen Gradient

Arthur D. Lander

Department of Developmental and Cell Biology, Developmental Biology Center and Center for Complex Biological Systems, University of California, Irvine, CA 92697, USA

Correspondence: adlander@uci.edu

DOI 10.1016/j.cell.2007.01.004

The theory that the spatial organization of cell fate is orchestrated by gradients of diffusing molecules was a major contribution to 20th century developmental biology. Although the existence of morphogens is no longer in doubt, studies on the formation and function of their gradients have yielded far more puzzles than answers. On close inspection, every morphogen gradient seems to use a rich array of regulatory mechanisms, suggesting that the tasks carried out by such systems are far more extensive than previously thought.

Few long-standing problems in biology have been the focus of more curiosity, or the source of greater frustration, than embryonic pattern formation. The problem, simply put, is one of determining how the large-scale organization of cell types in space is dictated by a set of instructions—the genes—that is the same in every cell. This question has captivated biologists, physical scientists, and mathematicians who have infused the field with viewpoints from their own disciplines: physicists see the emergence of large-scale properties from small-scale elements, computer scientists see distributed processing of information; engineers see robust control systems, and mathematicians see coupled partial differential equations with interesting behaviors.

In over a century of study, the most influential concept to emerge in the field of pattern formation has been that of the morphogen. In the oldest sense of the word, a morphogen is a substance that is produced by cells and organizes pattern by spreading to other cells. Most modern biologists adopt narrower definitions in line with particular theories about how morphogens work. The view prevalent in the experimental literature comes from Wolpert (1969), who proposed that smoothly declining gradients, formed by the diffusion of morphogens from sources to sinks, assign positional values to cells, which then adopt different fates depending on the values they were assigned. In this view, a morphogen is not just an instructive molecule but one that gives qualitatively different instructions depending on its concentration. What ultimately determines pattern, therefore, is where morphogen gradients cross threshold values at which genes are turned on or off.

A different conception of morphogens comes from the theoretical work of Meinhardt and Gierer (Gierer and Meinhardt, 1972; Meinhardt and Gierer, 2000), which extended the earlier efforts of Turing (1952). This work showed how two morphogens that influence each other’s synthesis could trigger the spontaneous emergence of stable, long-range patterns of morphogen activity. In numerical simulations, Meinhardt-Gierer mechanisms produce patterns of repeated stripes and spots that bear an uncanny resemblance to some of those found in nature.

The impact of such theories on developmental biology has been great, yet the relationship of experimental biologists to them has been rocky, at best. For years, a failure to identify any animal morphogens led to widespread doubt that such substances exist. By the mid 1990s, this situation changed as a result of studies on bicoid, an intracellular morphogen (Driever and Nusslein-Volhard, 1988), and Decapentaplegic (Dpp), an extracellular morphogen (Ferguson and Anderson, 1992; Nellen et al., 1996), both of which contribute to Drosophila development. In the last decade, peptides of the fibroblast growth factor (FGF), epithelial growth factor (EGF), Wnt, Hedgehog, and transforming growth factor (TGF)-β families, as well as the vitamin A metabolite retinoic acid, have all emerged as confirmed or likely morphogens (see, for example, Green, 2002; Tabata and Takei, 2004; Schier and Talbot, 2005).

Despite such progress, a lack of complete comfort with the morphogen concept persists among many biologists. Some prefer to explain pattern formation at the level of gene regulatory networks, attaching minor importance to movements of the molecules that genes encode (e.g., Davidson et al., 2002). Others accept that morphogen gradients exist but question whether diffusion is adequate, or reliable enough, to create them (Kerszberg and Wolpert, 1998; Pfeiffer and Vincent, 1999).

Indeed, the need for reliability—or as engineers call it, robustness—in patterning has become something of an obsession among experimentalists and theorists alike, leading many to seek fresh approaches to how morphogens do their jobs (Kerszberg, 1996; Pages and Kerridge, 2000; Pfeiffer et al., 2000; Belenkaya et
al., 2004; Bollenbach et al., 2005). Although this is a healthy trend, there has been something of a disproportionate emphasis on discovering new molecular mechanisms, rather than thoroughly exploring what known mechanisms do. For example, no less than four separate mechanisms have been proposed as remedies for the (perceived) deficiencies of diffusion in transporting morphogens from one location to another (Kerszberg and Wolpert, 1998; Ramirez-Weber and Kornberg, 1999; Entchev et al., 2000; Greco et al., 2001; Belenkaya et al., 2004; Kruse et al., 2004). Of course, the impulse to seek solutions for complicated problems in novel mechanisms is nothing new in science. One is reminded of Schrödinger’s conviction that only new quantum physics could explain the permanence of genetic material in living beings (Schrödinger, 1944).

With this as a backdrop, we may ask whether it is possible to retool existing theories of morphogens and pattern formation to determine to what extent the findings and concerns of modern experimental biologists can be accommodated and to what extent new mechanisms must be sought. This task can, I believe, be aided by better integrating morphogen theory into the wider context of spatial dynamics problems in biology. It can also be aided by formulating a richer, more complete description of what the performance objectives of patterning systems are. These points are explored below.

**Space: A Final Frontier?**

Life is dynamic on many timescales. Molecules bind and react, cells come and go, organisms are born and die, species evolve. Molecular biologists have built solid frameworks for understanding the dynamic processes underlying life by applying concepts from chemistry, such as kinetics, thermodynamics, and affinity. Using such tools, areas such as metabolism, gene expression, and intracellular signaling have been explored in sophisticated, quantitative ways.

Yet life is also dynamic on spatial scales. By this I mean that the stuff of life is not uniform in space—“well-stirred,” to use the chemistry expression—but is arranged in complex forms, from macromolecular assemblies to migratory herds. Questions involving spatial dynamics—how structure and pattern arise and are used—are as important in biology as those involving temporal dynamics. Spatial dynamics can be placed on a firm quantitative footing, too (the tools more often come from physics than chemistry), but it is a vastly more difficult endeavor. This is partly because space has three independent dimensions, but also because objects in it move forward and backward (things only go forward in time) and because interesting spatial questions often involve both space and time.

The mathematical, computational, and bookkeeping hurdles associated with describing, analyzing, and simulating spatial phenomena can be formidable, especially when spatial organization cannot be approximated by a few well-stirred compartments. Problems of this type tend to be ones in which there is continuous, rather than stepwise, variation of items in space, e.g., those in which molecular gradients matter. Embryonic pattern formation obviously falls into this category, but so do other areas, such as the control of cell movement and shape by chemoattractants, the exploitation of paracrine and autocrine signaling, and the interaction of intracellular signaling with cell shape and structure. Progress toward quantitative understanding has been accelerating in each of these areas (e.g., Iglesias and Levchenko, 2002; Wiley et al., 2003; Reas and Ballaro, 2004; Meyers et al., 2006), in part because of improvements in computing speed and power. Examination of the literature suggests that general themes and strategies, common to many spatially dynamic biological systems, are beginning to emerge. Accordingly, work being done on morphogens and morphogen gradients is increasingly relevant to the interests of biologists of all kinds.

**Walking the (Random) Walk**

A common thread among many spatial dynamics problems in biology is that interesting behaviors arise out of the randomly directed movement of molecules. In the macroscopic world, the kind of motion we typically encounter is what physicists call ballistic: objects stay at rest until acted upon by force and then move in the direction of the force for some time until, usually through friction, they come to rest. In contrast, the molecules in and around cells are in constant motion at extremely high velocities (the result of thermal energy) but travel only minuscule distances before colliding with other objects and randomly changing direction (Berg, 1993). Because the trajectory of any individual molecule approximates a random walk, a set of statistical rules can describe their collective behavior. For example, if a group of molecules is placed at one location, it will spread out in a predictable way to fill the remaining space. That such molecules are driven from high to low concentration prompts a view of a concentration difference as a sort of force (“driving force”). This seems to suggest that, in thinking about aggregate molecular motion, everyday ballistic intuitions can be applied. Nothing could be less true.

Consider that objects moving ballistically have a speed, whereas sets of objects spreading randomly do not. If it takes 10 min for a set of randomly moving molecules to travel an average of 10 μm, it will take 40 min to go 20 μm and 90 min to go 30 μm (a quadratic relationship between time and distance is a cardinal feature of random walks). What captures how molecules spread out is not a velocity but a diffusion coefficient (or diffusivity). Extracting this number from experimental data can be tricky, especially when molecules are not just moving but also undergoing binding, degradation, or chemical modification.

In the case of morphogen gradients, the problem is not just that transport can be difficult to measure but that the application of ballistic thinking to experimental data so easily leads to misapprehensions. A striking
and time (from 0 to 200 s). Diffusion coefficients for the two cases were $5 \times 10^{-7}$ cm$^2$ s$^{-1}$ and $10^{-9}$ cm$^2$ s$^{-1}$. A very similar effect of diffusivity on gradient shape may also be observed for steady state gradients, depending on the conditions that produce the steady state (unpublished observations).

(B and C) Simulations of ballistic and diffusive transport, and their responses to barriers. In both cases, 10$^3$ moving objects are released—10 per second for 10,000 s—from a single point into a field possessing an impenetrable barrier at one location (arrow). Object positions at the end of the time period are shown by individual dots (upper images), and via histograms of object density along the horizontal axes at the level of the barrier (lower images). In (B) ballistic motion is shown: the objects are tennis balls hit in random directions. Note the large number of balls that accumulate in front of the boundary. In (C) diffusive motion is shown: the objects are morphogen molecules undergoing random walks. Although the slope of the morphogen gradient flattens near the boundary, which may give a visual impression of slight accumulation, the histogram shows that there is no significant buildup of molecules at that location.

An example is shown in Figure 1A. The solid curve is the spatial gradient that would be formed by a freely diffusing morphogen with the diffusivity of a typical protein that is continuously produced for 200 s in a 20-µm-wide domain. The dashed curve shows what would happen under the same conditions if morphogen diffusion were decreased by a factor of five. If this morphogen induces a particular gene at a concentration threshold of 8, then slowing its diffusion will reduce the width of the domain of gene induction from about 100 µm to 65 µm. But if the threshold for gene expression is 20, the domain of gene induction will increase from 15 µm to 35 µm. How can making a morphogen move slower cause it to act farther away? The explanation is that lowered diffusivity allows the morphogen to accumulate to much higher levels near its source. In a recent model of the sonic hedgehog (Shh) gradient that patterns the ventral neural tube of the chick embryo (Saha and Schaffer, 2006), a version of just this situation arose, in which it was calculated that reducing Shh diffusion should increase the range of Shh action. This is a finding of immediate practical importance: current wisdom is that heparan sulfate proteoglycans promote the transport of Drosophila hedgehog (Hh), because removing them markedly decreases the range of Hh action (Bellaiche et al., 1998; Lin, 2004). The findings of Saha and Schaffer (2006) tell us that these very observations could just as easily mean that proteoglycans inhibit Hh transport!

Another counterintuitive feature of random transport is the way it responds to obstacles. Consider an experimental manipulation that causes morphogens to accumulate at a discrete location, such as at the edge of a mutant cell clone facing a morphogen source. The usual interpretation is that the morphogen accumulates because it cannot pass through the clone (e.g., Entchev et al., 2000). This is certainly how ballistic motion behaves: if I stand in a field and hit tennis balls in all directions, balls end up scattered about with a density that declines with distance; if a wall is placed at one point in the field, balls accumulate in front of it (Figure 1B). But it is not how molecules behave: if randomly moving molecules are released from a point source and encounter a barrier, they simply move away from, and around, the obstacle (Figure 1C).

The ability to sidestep obstacles also explains why freely diffusing molecules will traverse any random maze in not much more time than it takes them to cross the same distance in free space (Rusakov and Kulmann, 1998). This result helps us accept what otherwise may seem counterintuitive: that the labyrinth of tortuous intercellular spaces in most tissues poses little impediment to the free diffusion of morphogens. Limited awareness of this fact probably explains why morphogens are so often depicted moving exclusively within relatively unobstructed spaces on the apical surfaces of epithelia (Pfeiffer and Vincent, 1999; Christian, 2000; Belenkaya et al., 2004; Lin, 2004). In reality, apical transport incurs far more difficulties than basolateral, due to the potential for massive morphogen loss to the overlying medium.

**Steady or Not, Here They Come**

The strange world of randomly moving molecules gets even stranger when we add in effects of production and destruction. In well-stirred systems, if the capacity for destruction of a substance exceeds its rate of production, a steady state can be approached in which the amount of the substance tends toward a constant value. The same is true for spatially dynamic systems, except that if production and destruction occur at
different locations, stable gradients form. There are useful rules of thumb about such gradients. For example, between a discrete source and a discrete sink, a linear gradient will form. In contrast, when diffusion from a localized source is balanced by destruction that occurs with constant probability everywhere, the steady-state gradient will have an exponential shape. The length over which such a gradient decays to 1/e of its highest value—a number sometimes referred to as the length scale of the gradient—will be equal to the square root of the ratio of diffusivity to the degradation rate constant (Eldar et al., 2003; Gregor et al., 2005; Reeves et al., 2006). The length of any region of interest divided by the length scale of a molecule diffusing in it is its Thiele modulus, a term recently borrowed from engineering (Goentoro et al., 2006; Meyers et al., 2006).

Historically, most models of morphogen gradients and pattern formation—whether drawn from the “Wolpertian” or “Meinhardtian” perspectives—make the assumption that pattern is driven by steady-state gradients. Indeed, it seems logical that patterning information ought to be something stably maintained, especially because downstream responses of cells (e.g., gene expression) are relatively slow compared with the times required for diffusing molecules to move from one cell to another. Indeed, for some morphogen gradients, such as the Dpp and Wg gradients of the Drosophila wing imaginal disc, experiments show that the time over which gradients develop is short compared with the several days over which patterning occurs (Strigini and Cohen, 1999).

But other cases are less clear: in the zebrafish embryo, only 4–5 hr elapse between late blastula—when localized expression domains of bone morphogenetic proteins (BMPs), Wnts, Nodal, FGFs, and Retinoic acid emerge—and late gastrula stages, by which time an enormous amount of patterning orchestrated by these molecules has taken place (Schier and Talbot, 2005). In the Drosophila embryo, things move even faster: at 25°C, the bicoid gradient forms and does its job in antero-posterior patterning in less than 2 hr (Gregor et al., 2005), and the BMP morphogen gradient at the dorsal midline forms and specifies dorso-ventral pattern in under 1 hr (Dorfman and Shilo, 2001).

Can such gradients achieve a steady state rapidly enough? For most simple scenarios, the dominant factor determining the rate of approach to steady state is the average lifetime of morphogen molecules (i.e., the inverse of their degradation rate constant [Gregor et al., 2005; Lander et al., 2005]). For bicoid, an intracellular protein, a lifetime on the order of minutes is plausible. Of course, as Gregor et al. (2005) argue, because degradation rate also affects steady-state length scale (see above), the need to reach a steady state within 1–2 hr should constrain bicoid gradients to a maximum size of 1–2 mm (this, interestingly, seems to be about as big as insect embryos get).

For secreted polypeptide morphogens, degradation usually proceeds through sequential steps of binding, receptor-mediated endocytosis, and lysosomal proteolysis (Scholpp and Brand, 2004; Marois et al., 2006). It is questionable whether such events are fast enough for a steady-state BMP gradient to develop in the early fly embryo in less than 1 hr. Indeed, the idea that this gradient does not pattern under steady-state conditions is supported by experimental data, as there are combinatorial phenotypes of certain mutations that are difficult to explain otherwise (Mizutani et al., 2005). Other cases in which analysis suggests that steady states are not achieved by morphogen gradients include the previously mentioned model of neural tube patterning by Shh (Saha and Schaffer, 2006) and some models of the Drosophila embryo segmentation network (Gursky et al., 2004).

Why is it so important to know whether patterning by morphogen gradients occurs under transient or steady-state conditions? For one thing, the underlying mathematics tells us that responses of gradients to external manipulations can be very different in the two regimes. Consider the predicted effects of “nonspecific” binding sites: if a diffusing morphogen binds reversibly to immobile sites in its environment, then to the extent it does so, it spreads more slowly. At any time during the approach to steady state, the observed gradient will be narrower (as though diffusivity had decreased). But at steady state, the gradient profile will be the same as if there were no nonspecific binding sites (once a local equilibrium of capture and release has been achieved everywhere, it can have no net effect on morphogen flux [Eldar et al., 2003]). On the other hand, a change in the net rate of degradation of a morphogen will always affect steady-state gradient shape (length scale is inversely proportional to the square root of the degradation rate constant) but may have only small effects during the approach to steady state, especially at early times and near the morphogen source.

Accordingly, whether an investigator should interpret an experimentally induced change in the shape of a morphogen gradient as evidence for more or less nonspecific binding, more or less degradation, faster or slower diffusion, or any of variety of other alternatives, depends critically on whether or not observations are being made at steady state and, in some cases, on where in the gradient one is looking. To make matters even more challenging, simple experimental manipulations commonly affect more than one important aspect of gradient formation at a time. For example, adding a soluble, morphogen binding protein into a steady-state morphogen gradient system will tend to diminish morphogen signaling globally (by lowering occupancy of morphogen receptors) but can also greatly increase the length scale of the gradient (by decreasing the rate of receptor-mediated destruction of the morphogen). The net effect can be a net decrease in morphogen signaling at some locations coupled with a net increase at others (Mizutani et al., 2005).
Complexity and Performance

The fact that some morphogen gradients may do their jobs under non-steady-state conditions is just one indication that patterning by morphogens is more complicated than we used to think. Indeed, over the past decade, an astonishing array of what seem to be regulatory mechanisms has been uncovered in a wide range of morphogen gradient systems.

Some of these seem to tie directly into the peculiar properties of random transport. For example, morphogens of the hedgehog and Wnt families are frequently lipid-modified, which provides a way to control their diffusivity (Eaton, 2006). Secreted inhibitors are widely deployed in BMP, activin, nodal, Wnt, and EGF gradients (Kawano and Kypa, 2003; De Robertis and Kuroda, 2004; Klein et al., 2004; Yamamoto et al., 2004). Non-specific (or, more accurately, nonreceptor) binding sites for BMPs, Wnts, hedgehogs, and FGFs are abundantly found in the form of cell surface heparan sulfate proteoglycans, and for retinoic acid in the form of cellular retinoid binding proteins (Ross, 1993; Lin, 2004).

Evidence for complex regulation does not stop there: almost every morphogen interacts with one or more type of coreceptor, loosely defined as a cell surface morphogen binding protein that boosts the formation of morphogen-receptor complexes and/or enhances their signaling (Kirkbride et al., 2005). In several gradient systems, pairs of morphogens act in collaboration, sometimes synergistically (e.g., Nguyen et al., 1998). For some morphogens, restricted transport paths may exist along cellular structures such as cytonemes (Ramirez-Weber and Kornberg, 1999) or through intracellular shuttles (Greco et al., 2001).

On top of this, experimental work has revealed a rich array of feedback mechanisms driven by morphogen signaling. Examples in which morphogen signaling either up- or downregulates further morphogen signaling, morphogen production, morphogen destruction, synthesis of morphogen receptors, or the synthesis of morphogen coreceptors is found in studies on patterning by TGF-βs, hedgehogs, Wnts, EGFs, FGFs, and retinoids (e.g., Cadigan, 2002; Green, 2002; Fujise et al., 2003; Dobbs-McAuliffe et al., 2004; Lai et al., 2004; Ben-Haim et al., 2006). Indeed, it is rare to find morphogen systems that lack feedback.

When classical models of morphogen gradients, such as those of Wolpert and Meinhardt, first appeared, a scarcity of mechanistic information made it necessary to formulate minimal descriptions of pattern-forming systems. Given how well such models work in theory, it seems puzzling that so much additional complexity is found in studies on patterning by TGF-βs, hedgehogs, Wnts, EGFs, FGFs, and retinoids (e.g., Cadigan, 2002; Green, 2002; Fujise et al., 2003; Dobbs-McAuliffe et al., 2004; Lai et al., 2004; Ben-Haim et al., 2006). Indeed, it is rare to find morphogen systems that lack feedback.

Complex regulation comes into focus. This is an issue faced in dealing with such tradeoffs that the need for complex regulation becomes apparent. This is why, as Doyle argues forcefully (Csete and Doyle, 2002), engineering concepts and theory, especially those derived from the study of control systems, ought to have much to offer biology. Most of the rest consists of myriad control systems, the majority are surely imposed by evolution. Because of natural selection, the morphogen gradients we observe in the world today are those that do their jobs particularly well. This will be especially true for morphogen gradient systems that have been conserved across large evolutionary distances (as so many have).

If morphogen gradient systems had to become complex in order to get the job done, then the key to unlocking their complexity is to know what the job really is. At first glance this seems trivial: morphogens create pattern. Yet in breaking this into definable units—let’s call them performance objectives—we encounter a variety of nontrivial individual tasks: meeting timing specifications, making sharp borders, assigning multiple fates at once, producing periodic patterns, linking patterns to each other, and so forth. In addition, performing such tasks reliably in the face of myriad environmental and genetic perturbations must surely be something that evolution selects for. Thus, we need to include robustness—resistance to perturbation—as a performance objective. However, because environmental changes can influence both the rate at which overall development proceeds and the sizes of patterned structures, morphogen gradients must sometimes adapt to, rather than resist, changes (and there is ample experimental evidence that they do [e.g., Teleman and Cohen, 2000]). Adaptability as a performance objective arises in another context: because the organisms we study have descended from ancestors of very different shapes, sizes, and developmental timetables, only those morphogen gradients capable of being adjusted to fit such differences will be observed; this performance objective is sometimes referred to as evolvability (Meir et al., 2002).

The problem posed by performance objectives is not just that there are so many of them but that strategies that support one frequently compromise another. It is in dealing with such tradeoffs that the need for complex regulation becomes apparent. This is an issue faced constantly by engineers: to build a car that accelerates, turns, and stops is easy. To make one that sells means meeting many other performance objectives—handling, safety, appearance, cost, longevity, etc. This in turn means adding features like power steering, airbags, hybrid engines, onboard computers, etc. Today only a fraction of the machinery in the average car is actually directly involved in accelerating, stopping, and turning. Most of the rest consists of myriad control systems, the importance of which may only be revealed under specialized circumstances (e.g., airbags in a collision).

There is no reason to expect organisms to be different from cars in this respect, with control systems at least as abundant as the mechanisms that execute basic functions. This is why, as Doyle argues forcefully (Csete and Doyle, 2002; Stelling et al., 2004), engineering concepts and theory, especially those derived from the study of control systems, ought to have much to offer biology. Unfortunately for those of us working on morphogen gradients, the bulk of such theory has been developed with temporally, not spatially, dynamic systems in mind. On the positive side, this means that the study of morphogen gradient systems can be considered “cutting edge” by both biologists and engineers!
Trouble at the Border

There are many ways to start delving into issues of performance objectives and tradeoffs in morphogen gradient systems. Let us first consider a frequent goal of developmental patterning: to organize cells into discrete domains of gene expression with crisp borders (reviewed by Irvine and Rauskolb, 2001). To get a sense of what this entails, imagine that you have been asked to design, from scratch, a morphogen gradient system that achieves this objective.

To begin with, you might select a classical, Wolpert-type gradient, in which a morphogen is produced at one location and simply transported away by diffusion. Assuming the morphogen is a secreted polypeptide, binding to high-affinity cell surface receptors can both mediate signaling and take care of morphogen removal. As long as morphogen is produced slowly enough that receptors do not become saturated, an exponential gradient of both morphogen and occupied receptors will be obtained at steady state. As long as the size of the field being patterned is not too large, the approach to steady-state receptor occupancy will take about the same time as it takes for receptor bound morphogen molecules to be destroyed.

Figure 2. Strategies for Making Sharp Borders
(A–D) Each panel shows calculated shapes of signaling gradients produced at three rates of morphogen production, over a field of 20 rows of cells, using different sharpening strategies (units of signaling are arbitrary). The shaded box gives the distance from the tenth row to the location where the middle curve falls by half. The width of the box thus reflects the steepness of the curves, whereas the distance between curves reflects robustness to variations in morphogen synthesis rate. The shaded box of panel (A) is extended as a light stripe into panels (B–D) to facilitate comparison. In (A), no sharpening strategy is shown. The morphogen gradient is a declining exponential, and intracellular signaling is proportional to the amount of extracellular morphogen. In (B), the morphogen gradient is the same as in panel (A), but there is positive feedback in intracellular signaling (a key signaling molecule inhibits its own destruction). In (C), morphogen synthesis rates are high enough that receptors are saturated near the start of the gradient. In (D), morphogen signaling upregulates expression of morphogen receptors. In (A), (C), and (D), the black, blue, and red curves represent successive 2-fold increases in morphogen synthesis rate. In (B), these curves represent only 9% increases. Notice that in (B), steep declines in signaling over one cell row are easily obtained, but tolerance for variation in morphogen synthesis rate is dramatically reduced. In (C) and (D), robustness to morphogen synthesis rate is improved, but other performance objectives are compromised (not shown).

Figure 2A. Strategies for Making Sharp Borders
(A–D) Each panel shows calculated shapes of signaling gradients produced at three rates of morphogen production, over a field of 20 rows of cells, using different sharpening strategies (units of signaling are arbitrary). The shaded box gives the distance from the tenth row to the location where the middle curve falls by half. The width of the box thus reflects the steepness of the curves, whereas the distance between curves reflects robustness to variations in morphogen synthesis rate. The shaded box of panel (A) is extended as a light stripe into panels (B–D) to facilitate comparison. In (A), no sharpening strategy is shown. The morphogen gradient is a declining exponential, and intracellular signaling is proportional to the amount of extracellular morphogen. In (B), the morphogen gradient is the same as in panel (A), but there is positive feedback in intracellular signaling (a key signaling molecule inhibits its own destruction). In (C), morphogen synthesis rates are high enough that receptors are saturated near the start of the gradient. In (D), morphogen signaling upregulates expression of morphogen receptors. In (A), (C), and (D), the black, blue, and red curves represent successive 2-fold increases in morphogen synthesis rate. In (B), these curves represent only 9% increases. Notice that in (B), steep declines in signaling over one cell row are easily obtained, but tolerance for variation in morphogen synthesis rate is dramatically reduced. In (C) and (D), robustness to morphogen synthesis rate is improved, but other performance objectives are compromised (not shown).
shallow morphogen gradient into a sharp boundary of gene expression has not been lost on developmental biologists (e.g., Small et al., 1992; Dyson and Gurdon, 1998; Jaeger et al., 2004; Ashe and Briscoe, 2006). As it happens, there are many ways to handle this.

For example, you could replace the linear relationship between receptor occupancy and signaling with a steep one. One mechanism for doing this is to insert positive feedback into the signaling pathway—for example by allowing morphogen to inhibit the degradation of a component of its signaling pathway. Under the right circumstances, the cells in row 10 could easily be made to receive five times as much signal as cells in row 11 (Figure 2B). Or you could make signaling highly cooperative—e.g., by requiring dimerization of signaling molecules or cooperative action of transcription factors at their promoters. Or, you could combine both positive and negative feedback to generate a phenomenon known as bistability, with the result that morphogen signaling will jump abruptly from one potential steady-state value to another at a specific point in the gradient (Von Dassow and Odell, 2002; Ingolia, 2004; Lai et al., 2004). Or, instead of just using morphogen signaling to drive gene activation, you could use it to convert a repressor into an activator, getting two effects for the price of one. You could also take advantage of a phenomenon known as zero-order ultrasensitivity (Goldbeter and Koshland, 1981), in which two antagonistic enzymatic processes operate under conditions of high enzyme saturation, causing the output of the system to become extremely sensitive to the ratio of the levels of the two enzymes. You could even exploit a phenomenon known as stochastic focusing (Paulsson et al., 2000), in which sensitive behavior emerges out of statistical effects that become significant when the numbers of signaling molecules per cell are small.

The above mechanisms, despite their diversity, take advantage only of temporal, and not spatial, dynamics. Taking a more expansive view would allow you to uncover a host of additional strategies. Some are spatial analogs of temporal phenomena. For example, you could have an initially fuzzy boundary sharpen itself by allowing cells to rearrange their positions, or readjust their gene expression, based upon comparisons between their responses and those of their neighbors (e.g., Rogulja and Irvine, 2005). If such processes move cells with similar responses nearer to each other, the effect is essentially one of cooperativity; i.e., stable cell positions depend on there being a multiplicity of nearby, like-responding cells.

A phenomenon reminiscent of zero-order ultrasensitivity but implemented in space rather than time could be achieved by setting morphogen production rates high enough to saturate receptors over a portion of the patterned field. This strategy can produce a steep decline in morphogen signaling at any arbitrary location (Figure 2C). Unlike any of the other strategies discussed so far, this one (as well as those that follow, below) works by altering the shape of the morphogen gradient itself, not just the characteristics of the cellular response.

When it comes to implementing feedback effects in space, however, you will likely find that things get complicated quickly. Consider a morphogen that upregulates synthesis of its own receptor (Figure 2D). Because signaling increases when there are more receptors, this is positive feedback. But if receptors carry out morphogen degradation, then more receptors means the morphogen spreads less far; thus, receptor upregulation is also negative feedback. One might expect these opposing effects to cancel, but they don’t, because each has a different spatial range. An increase in receptor number elevates signaling only in the cells that get the extra receptors but depresses signaling at a distance (less morphogen makes it to other cells). This combination of short-range autoactivation and long-range inhibition may strike a chord with anyone familiar with the classical morphogen models of Meinhardt and Gierer, which use the same generative principle to form sharp stripes and spots. Traditionally, the set-up for such models involves pairs of antagonistic morphogens with different diffusivities, but the above discussion suggests that similar spatially dynamic behaviors can arise through other mechanisms, even ones involving a single morphogen.

One final spatial mechanism for producing border-like profiles of morphogen signaling deserves mention here because of its sheer ingenuity. In this strategy, which came to light as the result of experiments involving the BMP gradient that patterns the dorsoventral axis of the early Drosophila embryo, the diffusible BMP inhibitor Sog is used to carry the morphogen up its own concentration gradient, where Sog is cleaved and the morphogen is released (Holley et al., 1996; Shimmi and O’Connor, 2003). Formally, this is an example of facilitated transport, where the potential energy in one molecule’s concentration gradient is used to drive the energetically unfavorable transport of another.

**Sharp, Prompt, or Reliable?**

Of the various theoretical mechanisms described above, every one has been observed in one morphogen gradient system or another. In Hedgehog gradients we find positive feedback, conversion of transcriptional inhibition to activation, bistability, and upregulation of receptor synthesis (Saha and Schaffer, 2006). In retinoic acid signaling, we also see positive feedback and conversion of transcriptional inhibition to activation, as well as cooperativity (Kerszberg, 1996). Dpp converts transcriptional inhibition to activation indirectly, through downregulation of the repressor brinker, which can, in some cases, also drive positive feedback (Jazwinska et al., 1999). Bicoid is one of many transcription factors that acts cooperatively at its transcriptional targets (Burz and Hanes, 2001). In an EGF gradient that sharply patterns the ventral ectoderm of the Drosophila embryo, recent evidence points to the use of zero-order ultrasensitivity (Melen et al., 2005). Evidence for spatial cooperativ-
ity through cell rearrangement exists in several systems (Irvine and Rauskolb, 2001). And the BMP facilitated transport mechanism mentioned above appears to be used not only in the fly embryo but again in forming wing veins (O’Connor et al., 2006), and in the dorsoventral patterning of vertebrate embryos (De Robertis and Kuroda, 2004).

Why are so many mechanisms used to meet the same performance objective? For one thing, each mechanism has its own impact on other performance objectives. For example, some strategies, such as those based on positive feedback and zero-order sensitivity, have a tendency to increase the time required to reach steady state, which can be problematic if rapid patterning is important. Other strategies, such as facilitated BMP transport, can be very fast, but there seem to be limitations to the sharpness of the boundaries they form (Umulis et al., 2006).

The performance objective most easily compromised by all boundary-forming strategies is robustness. This is because, at root, forming a sharp boundary is about increasing sensitivity: sensitivity of signaling to morphogen level or sensitivity of morphogen level to position. In contrast, robustness is about decreasing sensitivity: sensitivity to variations in system architecture, environmental perturbations, or signaling noise. The trick, it would seem, is to create high sensitivity when desirable and low sensitivity otherwise.

This is no small trick. Whatever boundary-forming mechanism one chooses, one can usually find some aspect of robustness that is severely impaired. Positive feedback, for example, easily produces fragility (the opposite of robustness) to parameter variation (in Figure 2B, a mere 8% change in morphogen synthesis will cause a boundary at row 11 to shift beyond row 19). The combination of negative and positive feedback that produces bistable behavior can diminish such fragility, but not entirely (Saha and Schaffer, 2006). Strategies like zero-order ultrasensitivity can exhibit impressive robustness to certain variations (e.g., level of a saturated kinase or its phosphorylated substrate [Melen et al., 2005]), but only because fragility lies elsewhere. For many of the spatial strategies discussed above, comprehensive sensitivity analysis has not been done, yet it is still usually easy to uncover examples of substantial fragility.

Fortunately, biology doesn’t need to be robust to everything (a good thing, as engineering arguments suggest such a goal is unattainable). Natural selection can be expected to drive biological systems to be robust to the perturbations they encounter in nature. The evolutionary pressure associated with any given fragility will be a balance between the frequency of occurrence of the perturbation and the fitness disadvantage associated with failing to compensate for it. By this argument, we should expect that morphogen gradient systems that use different morphogens, affect events with different timescales, pattern territories of different sizes, or operate in species that inhabit different environments will have been selected to be robust in different ways.

Consider, for example, retinoic acid (RA), a morphogen derived by enzymatic modification of vitamin A. As vitamin A is obtained from the diet, and diets can be highly variable, we might expect RA synthesis rates to be rather unreliable. Accordingly, we might expect there to be a strong selection for mechanisms that make RA gradients robust to levels of RA synthesis. One such mechanism, an implementation of negative-feedback control, would be for the cells that make RA to downregulate expression of RA biosynthetic enzymes in response to RA signaling. An alternative strategy, which uses feedback but exploits spatial dynamics as well, would be for RA to induce, in RA-responsive cells, an enzyme that degrades RA (that self-enhanced degradation promotes robustness of morphogen gradients to variations in morphogen production has already been demonstrated for polypeptide morphogens [Eldar et al., 2003]). Interestingly, the evidence from RA patterning systems suggests that both feedback strategies are routinely implemented (Dobbs-McAuliffe et al., 2004).

Fueling New Thoughts

We may thus hypothesize that, in every morphogen gradient system, the dynamic and regulatory mechanisms we observe are directly related to the balancing of performance objectives, among them robustness to some, but not all, perturbations. Like all hypotheses that generalize from evolutionary theory, this one will occasionally be wrong—some complex states of affairs in biology may just be historical accidents. Still, such a hypothesis can guide our thinking in directions we might never have gone.

Consider, for example, neural tube patterning by Shh. The model presented by Saha and Schaffer (2006) generates a sharp signaling boundary about 70 μm from the morphogen source, on either side of which cell fates of the V3 interneuron and motoneuron type are specified. Boundary formation relies on both the intrinsic bistability of the Hedgehog signaling pathway, plus the spatial effect of short-range activation and long-range inhibition associated with receptor upregulation. Because of positive feedback, the system takes a long time to reach steady state, longer in fact than the window during which cell fates are specified. Fortunately, bistability kicks in well before steady state is achieved, with the result that a fixed cell-fate boundary is established long before the signaling difference on either side of it stabilizes (Saha and Schaffer, 2006).

Although the robustness of this mechanism has not yet been explored in depth, earlier work by the same group suggests that boundary location will likely be sensitive to many parameters (Lai et al., 2004). We may hypothesize, therefore, that the system possesses additional control loops that couple unwanted movement of the boundary to a corrective change in some boundary-affecting parameter. Such loops have not been suggested in the
literature, but the hints are there if we look for them. For example, modeling and experiment both suggest that changes in the levels of accessory molecules such as heparan sulfate proteoglycans and hedgehog-interacting-protein (Hip) can be used to shift the boundary position in either direction, and both types of molecules are known to be transcriptional targets of Hh/Shh (Chuang and McMahon, 1999; Fujise et al., 2003; Saha and Schaffer, 2006).

A second example of a morphogen gradient system in which new hypotheses can emerge from an exploration of performance objectives is the BMP gradient that pattern the dorsoventral axis of Drosophila. As mentioned earlier, this gradient uses facilitated transport to create a sharp peak of BMP signaling in the dorsal midline of the embryo. Modeling has suggested that robustness of the width of this peak is a performance objective that justifies this elaborate transport scheme (Eldar et al., 2002), but in vivo observations don’t support this view. In particular, the width of the signaling peak is actually quite fragile to Sog dosage (Mizutani et al., 2005). Moreover, experimental conditions that substantially alter peak width generally produce normal-looking flies, suggesting that selection for a highly robust peak is not likely to be strong.

What then are the performance objectives of facilitated transport? Possible insight comes from the recent observation that a positive-feedback process, initiated late in patterning through the gene-regulatory effects of BMPs, markedly sharpens the dorsal midline signaling peak (Wang and Ferguson, 2005). The mechanism underlying this process, which has been likened to bistability in Hh gradients, is unknown, although plausible models have been developed (Umulis et al., 2006). We may thus rationalize a reduced need for robustness in facilitated transport on the grounds that a subsequent, feedback-driven event independently dictates pattern. This, however, only begs the question of why facilitated transport is needed in the first place. Indeed, in pupal wing vein formation—another patterning system in which BMP facilitated transport seems to be coupled to activation of a positive-feedback loop to form narrow stripes of gene expression (O’Connor et al., 2006)—mutations that presumably abolish facilitated transport (Yu et al., 1996; Vîlmos et al., 2005) have little effect on vein width (implying that thin stripes of BMP activity can be made without facilitated transport).

The implication that the true performance objectives of facilitated BMP transport remain to be found challenges us to think in new ways. For example, in the very fast patterning environment of the fly embryo, it is plausible that positive-feedback-driven bistability is too slow to do the job on its own. Perhaps the embryo solves this problem by first using a very fast facilitated transport process to create a rough approximation of the final pattern, after which the positive-feedback system can begin its job with initial conditions that are already close to the desired goal. Achieving both speed and accuracy through the sequential application of coarse then fine control is a strategy commonly used in engineering, so there is every reason to expect to encounter it in the world of cells.

Yet even this explanation is likely to be incomplete. As discussed previously, facilitated movement of BMP is driven by the repeated binding and cleavage of Sog. In Drosophila, Sog is cleaved much faster when bound to BMPs than when not (Shimmi and O’Connor, 2003), which introduces positive feedback into Sog destruction (Sog cleavage releases BMPs, which then bind Sog, causing it to be cleaved). Such a feedback loop is not needed for facilitated transport per se, but it is predicted to have the interesting effect of making the dorsal midline BMP peak appear very abruptly (Mizutani et al., 2005), which agrees with experimental observations (Ross et al., 2001). In effect, the system seems to have a built-in time-delay switch. Might this further some performance objective? Perhaps there is a need to coordinate dorsoventral patterning with other events in the early embryo, such as anteroposterior patterning or nuclear division. The pace of these other fast-moving processes is certainly sensitive to environmental conditions (e.g., temperature), so perhaps the time delay is used to adjust the pace of dorsoventral patterning to match. If so, we ought to be looking out for mechanisms by which anteroposterior patterning events, or the nuclear division cycle, feed into just those parameters that are predicted to control the duration of the time delay (such as the ratio of the rates of Dpp and Sog production [Mizutani et al., 2005]).

**Toward a Systems Biology of Pattern Formation**

In the preceding pages, I have used selected examples to discuss the many regulatory mechanisms found in morphogen gradient systems, the relationship of those mechanisms to phenomena in spatial dynamics, and the utility of the concept of performance objective in explaining why morphogen gradient systems are constructed as they are. I have not tried to present a comprehensive, or even a balanced, view of all that is known about morphogen gradients. Instead, I have tried to use examples from the literature to illustrate how some of the complexity of morphogen gradient systems may be dissected and, to some degree, understood.

Many of the same points could have been made using other morphogen gradient systems. For example, the performance objectives of precise timing and coordination could have been introduced with a discussion of the vertebrate somite “clock” (Dubrulle and Pourquie, 2004). Bistability could have been discussed with reference to dorsal appendage formation in the insect egg (Shvartsman et al., 2002), or the insect segment polarity network (Von Dassow and Odell, 2002; Ingolia, 2004). Additional performance objectives could have been introduced, such as robust amplification of subtle differences (a key step in left-right patterning [Nonaka et al., 2005]), suppression of spatial noise (Ashe and Briscoe, 2006), or coordination of pattern with growth (Lawrence, 2001).
Some readers may be troubled by the appearance of terms with teleological overtones, such as job, task, and performance objective, throughout this article. In modern biology (especially molecular biology), suggestions that objects are “designed” to do jobs, or have “purposes,” tend to arouse suspicions of an antievolutionist agenda. In fact, and to the contrary, it is because of evolution that teleological language succeeds in biology (Ayala, 1999; Lander, 2004). Saying something is designed to do a particular job is more natural and less roundabout than saying its structure is maintained by natural selection through its ability to contribute to the fitness of the species in a particular way.

Yet my reliance on teleological language in this article furthers not just brevity, but another agenda: in recent years, biologists have increasingly sought to pose problems at a “systems” level. The trend toward “systems biology” is commonly linked to explosive increases in the sizes of biological data sets and data-gathering capacity. In science, new data are always useful, but an avalanche of observations only hinders understanding unless what’s important can be identified and sorted from the rest. What distinguishes systems biology from earlier traditions is the tendency to define importance less in operational terms (e.g., necessary or sufficient to produce a behavior) than in terms of relevance to the goals of a system. In making this leap, systems biology inextricably binds itself to teleology (Lander, 2004). Indeed, without the presupposition of goals or purposes, the very notion of “system” itself is hollow.

In focusing much of this review around notions of performance objectives, it should thus be clear that I favor greater integration of systems biology approaches into developmental biology. To some extent, this is already happening: One need only look at how many of the recent papers cited here combine experimental approaches with mathematics or computation to appreciate how much the field of developmental biology is changing along these lines. Yet I also believe that developmental biology has much to offer systems biology. Philosophically, developmental biologists have a long tradition of viewing embryos in terms of interacting, goal-oriented systems, and they know a great deal about the advantages and disadvantages of doing so. Technically, developmental biologists are among those most responsible these days for pushing the envelope of mathematical and computational analysis of complex biological systems. In part this is because spatially dynamic problems, which are mathematically so challenging, figure so heavily in developmental biology. In part it is because quantitative models of developmental events are often demanded long before accurate values of parameters are known, forcing developmental modelers to become leaders in conducting large-scale in silico explorations of parameter spaces (e.g., von Dassow et al., 2000; Eldar et al., 2003). Such explorations are more than just technical achievements. They teach a lesson about biology that is as important as it is surprising: sometimes, answering the most qualitative of questions—“Why does the organism do it that way?”—succeeds only through the most quantitative of approaches.

ACKNOWLEDGMENTS

The author is grateful to K. Arora, Q. Nie, T-M. Yi, T. Schilling, and F. Wan for helpful discussions and comments on the manuscript. This work was supported by the NIH (GM66051, GM67247, and HD38781) and the Human Frontiers Science Program.

REFERENCES


Driever, W., and Nusslein-Volhard, C. (1988). The bicoid protein deter-


