Dynamics and precision in retinoic acid morphogen gradients
Thomas F Schilling\textsuperscript{1,2}, Qing Nie\textsuperscript{1,3} and Arthur D Lander\textsuperscript{1,2}

Retinoic acid (RA) regulates many cellular behaviors during embryonic development and adult homeostasis. Like other morphogens, RA forms gradients through the use of localized sources and sinks, feedback, and interactions with other signals; this has been particularly well studied in the context of hindbrain segmentation in vertebrate embryos. Yet, as a small lipophilic molecule derived from a dietary source — vitamin A — RA differs markedly from better-studied polypeptide morphogens in its mechanisms of transport, signaling, and removal. Computational models suggest that the distinctive features of RA gradients make them particularly robust to large perturbations. Such features include combined positive and negative feedback effects via intracellular fatty acid binding proteins and RA-degrading enzymes. Here, we discuss how these features, together with feedback interactions among RA target genes, help enable RA to specify multiple, accurate pattern elements in the developing hindbrain, despite operating in an environment of high cellular and biochemical uncertainty and noise.

Addresses
\textsuperscript{1} Center for Complex Biological Systems, University of California, Irvine, CA 92697-2280, United States
\textsuperscript{2} Department of Developmental and Cell Biology, University of California, Irvine, CA 92697-2300, United States
\textsuperscript{3} Department of Mathematics, University of California, Irvine, CA 92697-3875, United States

Corresponding author: Schilling, Thomas F (tshilli@uci.edu)

Introduction
A critical issue in developmental biology is how morphogen gradients are established and interpreted by cells. Most studies have focused either on cytoplasmic morphogens of syncytial embryos (e.g. Bicoid (Bcd) in \textit{Drosophila}), or secreted proteins such as members of the TGF\textbeta, FGF, EGF, Hedgehog (Hh), and Wnt families [1–5]. Non-polypeptide morphogens also exist, however, the best-studied example of which is retinoic acid (RA). RA influences the behaviors of numerous cell types and tissues during embryonic development, as well as adult stem cells (neuronal, pancreatic, cancers (leukemia) and regenerating organs (cardiomyocytes) [6–10]. RA derives from vitamin A, and its lipophilic nature and use of nuclear receptors make its movements within tissues and signal transduction properties distinctive among morphogens. Here we discuss how RA gradients, and the cellular responses to them, are established and maintained. We review literature suggesting that several unique features of RA signal regulation make it extraordinarily robust, yet precise, in defining patterning and sharp boundaries of target gene expression.

Robust RA gradient formation
Most morphogen gradients are thought to form through the action of localized sources of production (e.g. sites of synthesis or deposition) and localized or distributed sinks (e.g. uptake, degradation). Because morphogens act at a distance from their source of production, eliciting distinct cellular responses in a concentration-dependent manner [11], a traditional focus of both experimental and theoretical work on their roles in pattern formation has been on how steady-state gradients form, and how their shapes are controlled. More recently, it has also become clear that cells can respond to the temporal dynamics of morphogen signaling, that is, they can sense the rate of change in morphogen concentration, and their responses to a morphogen can change over time [5,12–15].

Whereas early work on morphogen gradients treated cells as perfect detectors of invariant gradients, attention over the last decade has increasingly focused on the robustness of morphogen-mediated patterning [16,17]. This refers to the relative insensitivity of pattern formation to variability and uncertainty in the molecular processes underlying both morphogen gradient formation and readout. Such variability arises from a multitude of sources, including environmental factors (e.g. temperature and nutrition), individual genetic differences, and the intrinsically stochastic nature of biochemical processes (such as RNA and protein synthesis). Achieving robustness to such variation is also related to the problem of making morphogen gradients scale, that is, making their steady-state shapes automatically expand or contract in response to variations in the size of the tissue field being patterned [18,19].

The collective magnitude of all of this variability is expected to be in the range where the need for robustness (or scaling) places severe constraints on the design of morphogen patterning systems. One of the most generic strategies for achieving robustness to variability in morphogen production is to exploit negative feedback, which indeed seems to be used in most morphogen systems [11]. For example, in gradients formed by TGF\textbeta, FGF, EGF,
Hh and Wnt family members, there can be negative feedback through: firstly, regulation of morphogen synthesis, secretion, or transport; secondly, regulation of morphogen–receptor interaction (e.g. through induction of secreted antagonists or sequestering components of the extracellular matrix [ECM]); thirdly, regulation of receptor expression; or fourthly, regulation of signal transduction events downstream of morphogens (Figure 1a). An especially wide variety of such strategies has been described in Decapentaplegic (Dpp) and Sonic hedgehog (Shh) signaling pathways [20–27].

In the case of RA morphogen gradients, the need for robustness to variation in morphogen production is exacerbated by the fact that RA is synthesized from a precursor, vitamin A or retinol, the levels of which are highly sensitive to dietary conditions. A first line of defense against this problem seems to be negative

Figure 1

Comparing feedback mechanisms in morphogen signaling. (a) A cell responding to a typical polypeptide morphogen (purple ovals). Green dashed arrows point to extracellular inhibitory factors induced by the signal. Red lines denote intracellular mechanisms of repression. (b) An RA-producing cell showing the biosynthetic pathway for RA. Retinyl esters and beta-carotene are important stores of RA precursor. Green dashed arrows point to components of the pathway induced by RA signaling. Red lines denote repression of adhs and aldh1as by RA. (c) Polypeptide morphogens typically induce their own receptors, which are subsequently degraded through receptor-mediated endocytosis. (d) An RA-responding cell showing signaling and catabolism. RA (red circles) enters cells and binds Crabps (blue crescents), which transport it either to RARs (blue and green ovals) in the nucleus (orange) for signaling or to Cyp26s (black hexagons) for degradation. Adhs, alcohol dehydrogenases; aldh1as, aldehyde dehydrogenase 1as.
feedback control of the synthesis of RA from its precursors (Figure 1b) [6]. Indeed, it has been shown that RA downregulates, in a dose-dependent manner, the expression of Aldh1a2 (arrow 1 in Figure 1b), the major aldehyde dehydrogenase required for conversion of retinaldehyde to RA in embryos [28], and recent evidence demonstrates that it also downregulates retinol dehydrogenase (RDH10, arrow 2), the enzyme that converts retinol to retinal [29]. Microarray screens in zebrafish for RA-inducible genes recently revealed that additional negative feedback occurs through upregulation of Dhrs3 (arrow 3), a dihyroreductase that catalyzes the conversion of alltrans RA to vitamin A [30*]. Curiously, there also appears to be some positive feedback of RA on its own production, through upregulation of retinyl ester hydrolases, such as lecithin:retinol acyl transferase (Lrat, arrow 4), that produce retinol [31]. Although the roles of such mixed positive and negative feedback are unknown, the existence of so many levels of feedback on RA synthesis suggests that tight regulation is very important.

Degradation and robustness in RA gradients
Quite independent of their ability to regulate RA synthesis, RA morphogen gradients also achieve extraordinary robustness through mechanisms that act downstream of synthesis, a fact we know thanks to a fortuitous situation in the zebrafish hindbrain. During early embryogenesis, the anterior–posterior (A–P) axis of the hindbrain is subdivided into seven rhombomeres (r1–7) by a posterior-to-anterior RA gradient. The shape and orientation of that gradient is so greatly determined by the location of sites of RA degradation, that the normal, endogenous, posteriorly localized RA source can be eliminated and completely replaced with a uniform application of exogenous RA to the entire embryo. Remarkably, such embryos not only pattern the early hindbrain normally, but they also produce nearly the same pattern over a 20-fold range in applied exogenous RA concentration [32,33].

Mathematical models have provided some insight into this phenomenon, identifying feedback regulation of morphogen degradation as one means for compensating for unreliable levels (or locations) of morphogen production. Such regulation exemplifies ‘self-enhanced degradation’ [34] (or, more generally, self-enhanced decay), a type of negative feedback that has also been argued to make both Wnt and Hh gradients more robust to fluctuations in morphogen production [1–3,35]. For polypeptide morphogens, self-enhanced degradation typically involves activity-dependent regulation of receptor expression or receptor-mediated endocytosis (Figure 1c). For example, both fly and vertebrate Hh proteins induce their receptor, Patched, and Shh in vertebrates also induces the matrix proteoglycan Glypican3 (GPC3), and these in turn regulate Hh endocytosis and degradation [36,37]. Effective feedback control may also occur through induction of proteases or other ECM components — as long as morphogen activity increases morphogen degradation, it will reduce sensitivity to perturbations in morphogen production. However, because this type of feedback necessarily makes gradients shallower with distance from the morphogen source, it tends to exacerbate the difficulty of forming sharp boundaries of target-gene expression [16], a problem we return to later.

How is self-enhanced degradation of RA achieved in the developing hindbrain? First, RA induces the expression of a cytochrome p450 enzyme (Cyp26a1) that specifically degrades RA [32] (Figure 1d, arrow 1). Second, RA induces expression of certain intracellular RA binding proteins (Crabps), one role of which is to promote the delivery of RA to Cyp26s [38*] (arrow 2 in Figure 1d). Models that incorporate these two molecular features achieve considerably more robust gradients than models with either one alone [38**]. Interestingly, Crabp2 can also mediate positive feedback in RA signaling, by promoting the delivery of RA to RA-receptors (Figure 2a).

Robustness and interactions between RA and other morphogens
Even when the dynamics of morphogen gradient formation are sufficiently fast that they may be treated as steady-state systems, in many developing systems the sizes of the regions in which morphogens are produced and act undergo marked changes over time, due to tissue growth or cell movements. In such cases, in order for positional cues to maintain stable relative positions, it may be necessary for morphogen gradients to continually readjust themselves. In Drosophila imaginal discs, for example, there is evidence that, at least for a substantial fraction of larval life, the Dpp morphogen gradient both increases in amplitude and length-scale so as to compensate for disc growth [40*]. Length-scale changes may reflect the presence of feedback mechanisms that involve diffusible ‘expander’ molecules, such as Pentagone, whose concentration is thought to be a function of overall disc size [41].

In the hindbrain RA system there is also a gradual increase in size during patterning, due to cell movement during gastrulation, but the presence of an ‘expander’ molecule has not been described. What has been described, however, is an interaction between the RA system and other morphogen systems. In particular, A–P patterning of the hindbrain is influenced by at least three ‘posteriorizing’ signals, RA, Wnt and Fgf [42] (Figure 2b). All three are produced in the posterior mesoderm during
gastrulation, and induce posterior and suppress anterior expression of genes involved in rhombomere specification. Yet what would seem to be three independent morphogen gradient systems become intertwined at the level of RA gradient formation. This is because Fgf and Wnt inhibit the upregulation of Cyp26a1 by RA [32,42]. In this way, steadily increasing Fgf and Wnt expression may both drive the gastrulation movements that make the hindbrain field grow in size, and drive a compensatory expansion of the RA gradient, due to the resultant downregulation of Cyp26a1.

Although we can imagine how such compensatory mechanisms might broadly couple RA gradients to steady hindbrain elongation, the reality is that, as gastrulation proceeds, RA sources and sinks are much more dynamic and intricate. For example, not only is RA produced by the mesoderm that lies immediately posterior to the presumptive hindbrain (i.e. that will form somites), from which it diffuses anteriorly to pattern gene expression in the neural ectoderm (Figure 3a, upper panel), some RA is also synthesized and degraded in cranial mesoderm that lies on either side of the hindbrain. RA plays an essential role in A–P patterning of the mesoderm, which could, in turn secondarily influence hindbrain segmentation [43]. Moreover, within the cranial mesoderm, some of the targets of RA, such as Hoxa1 and Pbx1/2, are required for maintenance of expression of the RA biosynthetic enzyme Aldh1a2, that is, mesodermal RA expression appears to be autocatalytic (Figure 3a, lower panel) [44]. Whether or not this dynamically changing nearby source of RA plays a role in the robustness of hindbrain patterning remains to be investigated.

The role of downstream gene regulatory networks

Typically, the ultimate functions of morphogen gradients are to specify distinct, spatial domains of gene expression and cell differentiation. From a mechanistic standpoint, generating even a single, sharp domain boundary from a shallow morphogen gradient is tricky enough, as it requires large changes in gene expression in response to small differences in morphogen signal; such strong amplification is also likely to amplify cell-to-cell variation (i.e. noise) in the morphogen response, thereby degrading boundary sharpness [39]. Remarkably, many morphogen gradient systems solve this problem not just at one threshold morphogen concentration, but at several, so that multiple sharp boundaries are specified. In this section, we consider the roles played by gene regulatory networks in providing solutions to the multiple sharp boundary problem, including those downstream of RA in the hindbrain. In the next section we will consider how cell-to-cell noise affects the sharpness of such boundaries, and how noise itself can be used to counteract these effects.

So far, the most detailed studies of how gene regulatory networks shape the outputs of morphogen gradients have come from studies of Bed in Drosophila and the Shh morphogen gradient system of the vertebrate neural tube. In both cases, morphogen signals feed into networks of cross-regulatory gene expression that are rich in mutual
Robustness, boundary sharpening and noise attenuation in RA signaling. (a) Modified morphogen model incorporating roles of adjacent cranial mesoderm on patterning of hindbrain rhombomeres, anterior to the left. The RA gradient (dark red), RA in the cranial mesoderm (light red), and putative influences on the hindbrain neuroepithelium (blue arrows). (b) Models showing the RA gradient (red), cross-inhibition and auto-activation of Krox20 and Hoxa1/b1 at boundaries of r3-5 and influences of Fgf signaling (green) — initially from a posterior source and later from r4 — together with Spry4-mediated negative feedback, on Krox20 induction. (c) Schematic illustrating the concept of noise-induced switching at a rhombomere boundary. Fluctuations in RA levels, together with the gene regulatory network (upper panel) lead to fluctuations in target gene expression cell-by-cell (red and green squares) near a boundary, and noise in gene expression helps push cells into one stable state in this bistable region (e.g. from green to red).

Inhibition [13,23,45,46]. Such feedback is critical in creating ‘attractor states’, that is, states of gene expression to which cells are driven from any of a large number of possible starting configurations. When either of two stable states is possible depending solely upon initial conditions, we refer to the situation as bistability (two stable attractors), but for complex networks the attractors can be more numerous, and can include dynamic behaviors, such as oscillations or defined trajectories (i.e. where gene expression changes continuously in a reliable manner). In the network downstream of Bcd, where the number of mutually interacting genes is fairly large, and where additional positional information is provided by maternally-derived gradients, there are many potential attractors, both stable and dynamic [47,48]. In the Shh system that patterns the dorsal–ventral (D–V) axis of the neural tube, a more modest (so far) network of inhibitory gene regulatory interactions involving the genes Nkx2.2, Olig2 and Pax6, creates three stable states [13,46], which correspond to three distinct domains of gene expression that emerge at different distances from the ventral Shh source. The existence of attractor states downstream of morphogens has at least three important consequences. First, due to strong positive feedback created by the loops of mutual inhibition typically found in such systems, switching from one attractor to another often occurs with only very small changes in input (morphogen). Thus, such systems provide an explanation for how shallow gradient information is converted into spatially abrupt changes in cell fate (Figure 3b). Second, because a cell’s choice of attractor depends not only on morphogen concentration but on the dynamics over which that concentration changes, positioning of boundaries can be more robust than could be achieved given measurements of morphogen concentration alone [47,49]. Third, as discussed below, the effect of dynamics on attractor selection can make it possible to specify a greater number of spatial domains than there are attractor states.

Evidence and modeling suggest that all of these phenomena are important in A–P patterning of the vertebrate hindbrain by RA. For example, the RA target genes Hoxb1 and Krox20 are self-activating and mutually repressing, creating three stable states: Hoxb1-on/Krox20-off, which is adopted by the territory that will become rhombomere 4 (r4); Krox20-on/Hoxb1-off, which occurs both in r3 and r5; and a both-off state. Activation of Hoxb1 expression occurs first, in a diffuse pattern that simply reflects RA levels. Later, induction of Krox20 becomes
possible, probably reflecting the fact that it is a more indirect target of RA [50]. Recent modeling [51**] indicates that, given appropriate assumptions about relative strengths of mutual inhibition and self-activation of Hoxb1 and Krox20, and their relative sensitivity to RA, three spatial outcomes can result. Posterior to r4, where RA concentration is highest and the input to both genes approaches maximal values, Krox20 — the stronger inhibitor — dominates, and Krox20-on becomes the only available steady state. Anterior to r4, where RA concentration is lowest and therefore cells have not previously expressed Hoxb1, the greater sensitivity of Krox20 to RA, as well as its stronger autoactivation, leads to the adoption of a stable Krox20 state as well. In r4 itself, however, Hoxb1 is induced more rapidly than Krox20, so that when Krox20 is induced Hoxb1 levels are high enough to block its expression and cells fall into the hoxb1 attractor, that is, the Hoxb1-on/Krox20-off state. In this manner, three sharply defined domains of gene expression are generated (Figure 3b).

In this model, the attractor landscape, and thus the locations where cell fates switch from Krox20 to Hoxb1 and back again, emerge out of the strength and history of the mutually inhibitory interactions between these genes. One would thus expect that any perturbation that gives an ‘advantage’ to either gene — by hastening or strengthening its expression or autoactivation — would result in predictable movements of gene expression boundaries. Recent work on the effects of Fgf signaling during hindbrain patterning supports this view. Early during the patterning of r4, its cells begin to express Fgf [52–54]. LaBalette et al. [55**] observed that knocking down Sprouty4, which normally inhibits Fgf signaling, leads to a marked expansion of r3 and r5 at the expense of r4. This is accompanied by an earlier than normal onset of Krox20 expression, which they showed is a direct Fgf target. This result agrees with model predictions, which state that a greater constitutive input to Krox20 should enable it to compete more effectively with Hoxb1 and thus be expressed in a wider territory. Interestingly, because Fgf is produced by r4, yet promotes the ability of Krox20 to drive cells toward an r3 or r5 fate (Figure 3B), the Fgf effect may be seen as a form of stabilizing negative feedback on the width of r4, that is, it should help make that width more robust to the parameters of Hoxb1/Krox20 mutual inhibition.

**Overcoming noise and boundary sharpening**

When we refer to morphogen gradient robustness, we typically mean the relative insensitivity of the boundary positions established by morphogen gradients to perturbations that affect all cells equally. Yet morphogen gradients must also deal with cell-to-cell variability, which creates problems of a different type: to the extent that such variability makes cells misread the morphogen concentration in their immediate vicinity, gene expression boundaries should take on a ragged, or salt-and-pepper, appearance.

Indeed, many morphogen gradient systems display such rough boundaries, at least early on during patterning, but often the situation corrects itself. In the zebrafish hindbrain, for example, rhombomere boundaries start out rough and sharpen over a few hours — 9–11 hours post-fertilization [56]. This is accomplished only in part by cell rearrangement. The rest is due to cells switching patterns of gene expression to better reflect their true positional environment. Indeed, individual cells of forming rhombomeres have been shown to remain distinctly plastic — able to upregulate or downregulate Hox expression after initiation — for some time [57,58].

How does such plasticity help cells figure out their true locations, if the signals they receive remain corrupted by a constant level of noise? It turns out that the same gene regulatory networks that create multiple attractor states create an opportunity to implement a strategy known as ‘noise-induced switching’, which can help cells improve their positional choices. Noise-induced switching depends upon hysteresis, the property that switching from one attractor state to another will occur at one morphogen threshold, while switching back in the other direction will occur at a distinctly different threshold. If the thresholds are well-separated enough, and the noise is large enough to drive most cells across one or the other threshold at least a few times, cells will tend to settle into steady states that much better reflect the average morphogen concentration than if they had to rely upon a single noisy measurement. Stochastic modeling has shown how the interaction of noise with the hysteresis produced by the mutual inhibition between Hoxb1 and Krox20 is sufficient to explain much of the spontaneous sharpening of rhombomere boundaries during hindbrain patterning (Figure 3c) [51**].

What makes noise-induced switching such a counter-intuitive strategy is that it requires noise to work, that is, it takes uncertainty to overcome uncertainty. Although several instances have been described in which noise-induced switching influences the behaviors of isolated cells undergoing differentiation [59*], its role in improving the precision of patterning boundaries has not generally been appreciated. Given that the gene regulatory networks downstream of many morphogen systems are distinctly hysteretic [46,47], it seems likely that this strategy is quite broadly exploited in vivo.

**Conclusions and perspectives**

In this review we have highlighted how RA signaling is regulated — in ways that are sometimes distinct from, and sometimes similar to, those employed by polypeptide morphogens — so that it can form gradients that are...
surprisingly robust, precise, and capable of inducing multiple sharp boundaries of target gene expression. Mechanisms that enhance robustness include: firstly, tight feedback regulation of RA synthesis, secondly, multiple paths of self-enhanced degradation, and thirdly, interactions between RA and other morphogens. Mechanisms that enhance the precision of boundary formation include: target gene regulatory networks that drive cells toward distinct attractor states and a surprising beneficial role for noise in facilitating the switching of target gene expression between bistable states, enabling individual cells to choose their fates more accurately.

Acknowledgements
This work was supported by the Center for Complex Biological Systems and the National Institutes of Health P50GM076516 (AL, QN, TS), R01DE13828 (TS), R01GM067247 (AL, QN), and the National Science Foundation DMS-0917492 (QN) and DMS-1161621 (QN).

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest
•• of outstanding interest


This paper describes a dihydroxysteracetase that converts all-trans RA to vitamin A, which is RA-inducible and provides additional negative feedback which may influence signal robustness.


Dynamics and precision in retinoic acid morphogen gradients

Schilling, Nie and Lander 569


38. Cai AQ, Radtke K, Linville A, Lander AD, Nie Q, Schilling TF:


This work combines mathematical modeling and genetic studies in zebrafish to examine robustness of the RA morphogen gradient. A model incorporating Crabp1 was generated, and its main predictions were validated experimentally.


This paper showed that Signed Activation Time (SAT) is a key feature that can enable noise reduction in a regulatory system. Since SAT is straightforward to obtain, it is an efficient computational tool for studies of noise propagation in gene regulatory networks.

40. Wartlick O, Mucmcu P, Kicheva A, Bittig T, Seum C, Jülicher F,


This paper demonstrates that the morphogen gradient of Dpp in the Drosophila wing disc increases in amplitude and length-scale, which can help compensate for disc growth during larval development.


44. Vitobello A, Ferretti E, Lampe X, Vilain N, Ducet S, Ori M, Spetz JF,


This paper demonstrates that some RA targets within the cranial mesoderm (i.e. Hoxa1 and Pbx1/2), maintain Alx1a2 expression, which may then influence hindbrain patterning and signal robustness.


This work combines mathematical modeling and gene expression studies in zebrafish to examine how sharp boundaries between rhombomeres can form in the face of noise in the RA morphogen gradient and in its downstream targets, Hoxb1 and Krox20. In part, because these two genes cross-inhibit one another and auto-activate themselves, noisy expression can actually facilitate boundary sharpening.


55. Labalette C, Bouchouca YX, Wassif MA, Gongal PA, Le Men J,


This paper demonstrates that initiation and maintenance of krox20 expression occur through separate enhancer elements, and that Fgf not only promotes initiation of expression but also modulates it through negative feedback. This work reveals robustness mechanisms downstream of RA and Fgf in the hindbrain.


This paper provides an example in which noise-induced switching helps facilitate the differentiation of isolated cells.