Journal of Theoretical Biology 285 (2011) 136–146

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Distinguishing graded and ultrasensitive signalling cascade kinetics by the shape of morphogen gradients in *Drosophila*

Shev MacNamara [a,*], Ruth E. Baker [a], Philip K. Maini [a,b]

[a] Centre for Mathematical Biology, Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK
[b] Oxford Centre for Integrative Systems Biology, Department of Biochemistry, University of Oxford, South Parks Road, OX1 3QU, UK

ARTICLE INFO

Article history:
Received 27 January 2011
Received in revised form
11 May 2011
Accepted 14 June 2011
Available online 1 July 2011

Keyword:
Mitogen-activated-protein kinase

ABSTRACT

Recently, signalling gradients in cascades of two-state reaction–diffusion systems were described as a model for understanding key biochemical mechanisms that underlie development and differentiation processes in the *Drosophila* embryo. Diffusion-trapping at the exterior of the cell membrane triggers the mitogen-activated protein kinase (MAPK) cascade to relay an appropriate signal from the membrane to the inner part of the cytosol, whereupon another diffusion-trapping mechanism involving the nucleus reads out this signal to trigger appropriate changes in gene expression. Proposed mathematical models exhibit equilibrium distributions consistent with experimental measurements of key spatial gradients in these processes. A significant property of the formulation is that the signal is assumed to be relayed from one system to the next in a linear fashion. However, the MAPK cascade often exhibits nonlinear dose–response properties and the final remark of Berezhkovskii et al. (2009) is that this assumption remains an important property to be tested experimentally, perhaps via a new quantitative assay across multiple genetic backgrounds. In anticipation of the need to be able to sensibly interpret data from such experiments, here we provide a complementary analysis that recovers existing formulae as a special case but is also capable of handling nonlinear functional forms. Predictions of linear and nonlinear signal relays and, in particular, graded and ultrasensitive MAPK kinetics, are compared.

1. Introduction

The detection of chemical signals by living cells and the transduction of these signals into the interior of the cell are crucial phenomena that need to be explored as we strive to understand the complex behaviours that underlie development. Biological networks commonly use signalling cascades during these processes, but it is not clear how the individual stages of these cascades shape the input signal and how the final output differs depending on the internal workings of the transduction network.

Paradigms of signalling cascades at work can be found in development of the *Drosophila* embryo, where diffusion-trapping at the exterior of the cell membrane triggers the mitogen-activated protein kinase (MAPK) cascade to relay an appropriate signal from the membrane to the inner part of the cytosol, whereupon another diffusion-trapping mechanism involving the nucleus reads out this signal to trigger appropriate changes in gene expression. Fig. 1 is a schematic of such a cascade and represents the system studied in this paper. In this context, the term 'trapping' refers to the way an otherwise mobile and diffusing particle may become immobile or trapped. For example, Fig. 1(B) shows a mobile and diffusing extracellular molecule binding to a receptor to become immobile as part of a ligand–receptor complex. Similarly, at the next stage of the cascade, an intracellular molecule may diffuse, but it becomes trapped if it enters an immobile nuclei (Fig. 1(D)). A number of papers have investigated the role of diffusion-trapping mechanisms in the establishment of signalling gradients in *Drosophila* embryos, including Coppey et al. (2007, 2008) and Shvartsman et al. (2009). Berezhkovskii et al. (2009) study the gradients responsible for terminal patterning by the Trunk ligand binding to the Torso receptor and activating the MAPK cascade which, in turn, leads to phosphorylated MAPK shuttling in and out of the nucleus, thus triggering a transcriptional response.

In order to investigate gradients in diffusion-trapping mechanisms, Berezhkovskii et al. (2009) studied a system of partial differential equations (PDEs) in one, infinite, space dimension that describes the dynamics of a two-stage cascade of diffusion-trapping systems:

$$\frac{\partial C_m^{(1)}}{\partial t} = D_1 \frac{\partial^2 C_m^{(1)}}{\partial x^2} - (k_m^{(1)} + \alpha_1) C_m^{(1)} + \beta_1 C_{im}^{(1)} + g_1(x),$$

* Corresponding author.
E-mail addresses: shev.macnamara@maths.ox.ac.uk, Smacnamara@math.mit.edu (S. MacNamara), ruth.baker@maths.ox.ac.uk (R.E. Baker), maini@maths.ox.ac.uk (P.K. Maini).

0022-5193/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2011.06.012

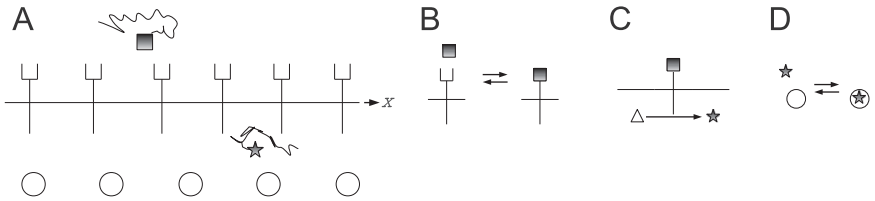

Fig. 1. Adapted from Fig. 1 of Berezhkovskii et al. (2009). (A) A diffusible ligand (solid square) is reversibly bound to cell surface receptors. A diffusible intracellular molecule (star) shuttles in and out of nuclei (circles). (B) Transitions between mobile and immobile states for a particle in the extracellular stage of the cascade. (C) Immobile particles in the first stage initiate the production of mobile particles in the second stage. In this case a ligand–receptor complex is an enzyme acting on a pool of inactive intracellular molecules (present in excess, denoted by a triangle). (D) A mobile intracellular molecule is reversibly trapped by immobile traps. In this case the traps are nuclei distributed in a shared cytoplasm of the early embryo.

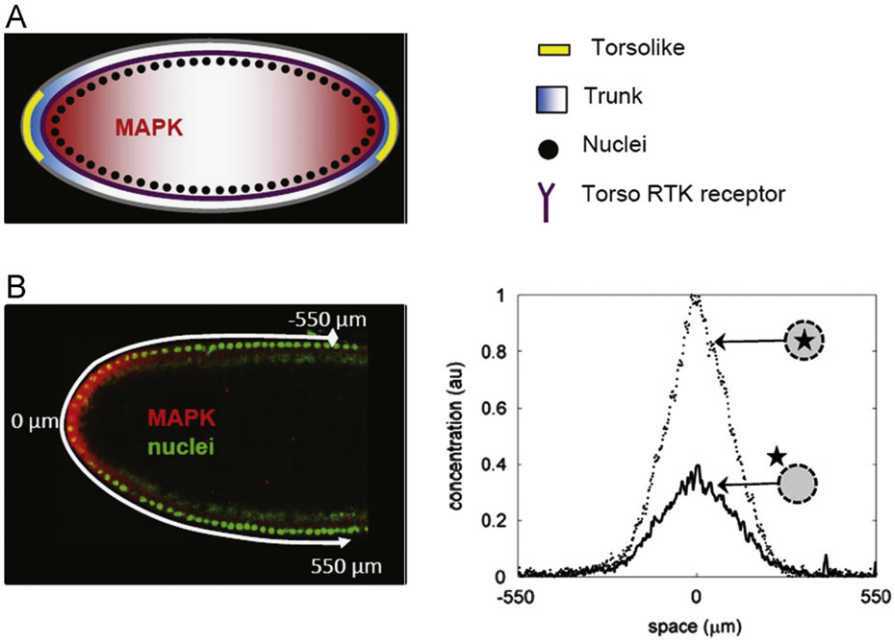

Fig. 2. Fig. 2 from Berezhkovskii et al. (2009), Copyright (2009) National Academy of Sciences, U.S.A. Terminal patterning system in the early *Drosophila* embryo. (A) Torso receptors (purple) are uniformly distributed along the plasma membrane of the embryo. Inactive ligand (Trunk) is distributed uniformly in the extracellular matrix; it is converted into an active and diffusible form (blue, light) by Torsolike (yellow, lighter), a protein localised at the poles of the embryo. The Torso–Trunk complex signals through the MAPK signalling cascade, which leads to MAPK phosphorylation and nuclear import. (B) Quantified pattern of MAPK phosphorylation. (Left) Fluorescent image of the anterior of the embryo; nuclei are stained in green (light), and phosphorylated MAPK is stained in red (dark). (Right) Gradients of nuclear and cytoplasmic phosphorylated MAPK. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\frac{\partial C_{im}^{(1)}}{\partial t} = \alpha_1 C_m^{(1)} - (k_{im}^{(1)} + \beta_1) C_{im}^{(1)},$$

$$\frac{\partial C_m^{(2)}}{\partial t} = D_2 \frac{\partial^2 C_m^{(2)}}{\partial x^2} - (k_m^{(2)} + \alpha_2) C_m^{(2)} + \beta_2 C_{im} + g_2(C_{im}^{(1)}),$$

$$\frac{\partial C_{im}^{(2)}}{\partial t} = \alpha_2 C_m^{(2)} - (k_{im}^{(2)} + \beta_2) C_{im}^{(2)}. \tag{1}$$

Here $C_{im}^{(i)}$ and $C_m^{(i)}$ denote concentrations of immobile and mobile molecules of species i ($i=1, 2$), respectively, D_i is the diffusion constant, $k_m^{(i)}$ and $k_{im}^{(i)}$ denote the rate constants for annihilation of immobile and mobile molecules, and α_i and β_i are the rates of binding and release, respectively (each for $i=1, 2$). In relation to Fig. 1, the concentration of solid squares is represented by $C_m^{(1)}$, *i.e.* the diffusible ligand, D_1 is its associated diffusion coefficient and $k_m^{(1)}$ is the rate of degradation. The rates of binding to and release from the receptor are α_1 and β_1, respectively. Similarly, the intracellular molecule, $C_m^{(2)}$, is represented by the star, D_2 is its associated diffusion coefficient and $k_m^{(2)}$ is the rate of degradation. The rates of shuttling into and out of the nuclei are α_2 and β_2, respectively. Essentially, the two systems are identical, with $g_1(x)$ being the input to the first system, and the output of the first system used as the input to the second via the term $g_2(C_{im}^{(1)})$. $C_{tot}^{(2)} = C_m^{(2)} + C_{im}^{(2)}$ is interpreted as the output and the system is closed by specifying suitable initial and boundary conditions: we take all species to have zero concentration initially, and impose the constraint $\lim_{|x|\to\infty} C_{m/im}^{(i)} = 0$. Note that here we restrict attention to the time-independent case, that is, we assume g_1 is only a function of x.

We will be concerned exclusively with systems at equilibrium, where we have, for $i=1, 2$,

$$C_{im}^{(i)} = K_i C_m^{(i)}, \tag{2}$$

with constant of proportionality

$$K_i = \frac{\alpha_i}{k_{im}^{(i)} + \beta_i}. \tag{3}$$

We see that the gradients of mobile and immobile species have the same shape so it is natural to distinguish three regimes:

$$K < 1, \quad K = 1, \quad K > 1.$$

The specific parameter regime can be distinguished by comparing the gradients of mobile and immobile species: experiments in Fig. 2B (right) of Berezhkovskii et al. (2009) (see Fig. 2) indicate that the second diffusion-trapping system, involving nuclear and

cytoplasmic phosphorylated MAPK, lies in the third regime. Moreover, by comparing the peaks of the two gradients we can estimate the constant of proportionality. There is some noise in the measurement so by estimating the minimum and maximum of the peak we obtain an estimate of the range of the constant: $K_2 \approx 2.5-3.3$. This gives some information about the relative strengths of the binding, release and annihilation rates: the phosphorylated MAPK species shuttles into the nucleus at more than double the rate that it shuttles out, for example. If data became available, one could similarly determine in which regime the other diffusion-trapping system, involving Torso/Trunk interactions, lies.

An important question in understanding signalling cascades, such as the MAPK cascade, is whether the associated dose–response curves are graded or ultrasensitive. Fig. 3 shows an example of a graded and of an ultrasensitive dose–response curve. This provides one of the motivations for the analysis here, which is to investigate the effect of different choices of g_2 on the gradient at equilibrium. In particular, the previous study assumed g_2 to be a simple positive multiple of its input, *i.e.*

$$g_2(x) = g_l C_{im}^{(1)}(x), \tag{4}$$

for some constant $g_l > 0$. Based on the authors' previous experiments (Coppey et al., 2007), it has been suggested that the length scale of the first system is much less than that of the second (Berezhkovskii et al., 2009). However, this study did not measure Torso occupancy directly; instead, dpERK was measured. Thus $C_{tot}^{(2)}(x)$ was measured instead of $C_{im}^{(1)}(x)$. When g_2 is linear this should not provide a problem, but when g_2 is nonlinear the length scale may be significantly affected. Fig. 4 shows that with ultrasensitive kinetics, the gradient of dpERK may have a small length scale even if the gradient of Torso occupancy has a large length scale. We are interested in understanding whether the assumption of linearity is consistent with experimental observations, and develop a mathematical framework that allows us to do this.

The outline of the rest of this paper is as follows. In the next section, an alternative formulation of the problem, which may still capture key experimental findings but which is flexible enough to explore alternatives to linear kinetics, is described. Predictions of graded and ultrasensitive models are then compared. As a general rule-of-thumb it is found that, compared to the linear case, gradients formed under ultrasensitive kinetics tend to have much shorter length scales. Next, predictions of both models are compared to data. This allows us to suggest three key features of experimental data that may be used in order to distinguish whether graded or ultrasensitive kinetics are acting in a given system. A procedure for carrying out this analysis is demonstrated on a particular

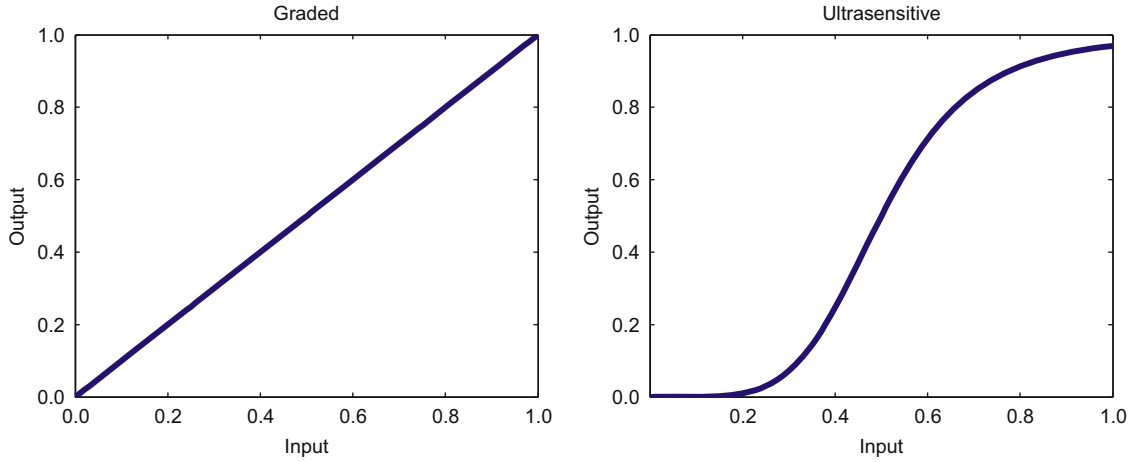

Fig. 3. Example of the relationship of output to input in a graded ($f(x) = cx$) and ultrasensitive ($f(x) = c/(1+(x_{50}/x)^n)$) function. Parameters: $c=1$, $x_{50} = 0.5$ and $n=5$.

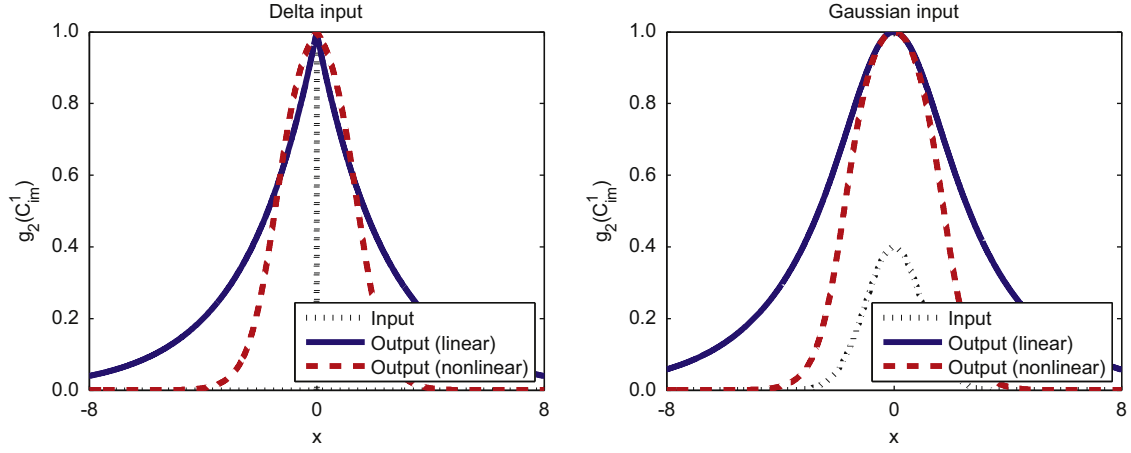

Fig. 4. Comparison of the inputs to the second stage, under graded (linear) and ultrasensitive (nonlinear) signalling kinetics. Input g_1 is a Delta distribution on $x=0$ (left), or a Gaussian distribution with mean zero and variance σ^2 (right). For the Delta distribution input the effective length scale is reduced by a factor of $n=5$ by the ultrasensitive MAPK signalling kinetics. Parameters are as follows: $D_1=0.5$, $\alpha_1 = 1.5$, $\beta_1 = 2$, $k_{im}^{(1)} = 0.1$, $k_m^{(1)} = 0.01$, and $\sigma^2 = 1.0$. For the Hill function parameters in (7) we set $C_0=1.0$ and $n=5$ and in each case we choose the scale factors g_l and g_h so the peak is unity: $g_l=0.565$ and $g_h=1.0576$ (left), and $g_l = 0.7586$ and $g_h=1.2512$ (right).

example: the MAPK cascade in the terminal patterning system of the early *Drosophila* embryo, as shown in Fig. 2. Although the focus is largely on this particular example, the framework is relatively generic and it can be applied to understand signalling cascades that arise in many other systems. Finally, the utility of the results is discussed, including speculation on how they may be extended. The major contributions and findings are summarised in the conclusion.

2. Equilibrium solution

We define

$$\lambda_i = \sqrt{\frac{D_i}{k_m^{(i)} + k_{im}^{(i)} K_i}},$$

so that equilibrium solutions of the PDE (dropping subscripts for ease of notation)

$$\frac{\partial C_m}{\partial t} = D\frac{\partial^2 C_m}{\partial x^2} - (k_m+\alpha)C_m + \beta C_{im} + g(x) \quad \text{with } C_{im} = KC_m,$$

and $x \in (-\infty,\infty)$, $t \in [0,\infty)$ may be studied via an ordinary differential equation (ODE)

$$v''(x) - \frac{1}{\lambda^2} v(x) = -\frac{1}{D} g(x), \tag{5}$$

with $x \in (-\infty,\infty)$, $\lim_{|x|\to\infty} v(x) = 0$ and g the given input function. The solution v is the equilibrium solution $C_m^{(i)}$ to (1), from which we immediately obtain the equilibrium solution $C_{im}^{(i)}$ from (2). The associated Green's function for (5) is

$$G(x,z) = -\tfrac{1}{2}\lambda e^{-|x-z|/\lambda},$$

and the solution to (5) is

$$\begin{aligned} v(x) &= -\frac{1}{D}\int_{-\infty}^{\infty} G(x,z)g(z)\,\mathrm{d}z, \\ &= \frac{1}{2D}\lambda\left(e^{-x/\lambda}\int_{-\infty}^{x} e^{z/\lambda}g(z)\,\mathrm{d}z + e^{x/\lambda}\int_{x}^{\infty} e^{-z/\lambda}g(z)\,\mathrm{d}z\right). \end{aligned} \tag{6}$$

The framework of Green's functions is well-suited to this analysis since we are interested in the ramifications of various functional forms of the input, g. Studying various inputs is straightforward because the same Green's function may be applied to different choices of g via (6). In particular, nonlinear forms of g may be handled and the method may accommodate finite domains, though this may require the Green's function to be adjusted for different sets of boundary conditions.

The main issue with the Green's function approach is evaluating the integral. In many important special cases the integration may be performed analytically, as we show below. Otherwise quadrature provides an easy numerical solution, though this is required for each x and some care must be taken with the unbounded domain. In this model the integrand is weighted by an exponential decay, so for physiologically reasonable choices of g, restricting the numerical solution on a finite domain should give very accurate results.[1]

2.1. Example

By choosing $g_1(x) = Q\delta(x)$, a delta distribution on the origin, to represent spatially restricted production, substituting into (6) yields

$$v(x) = \frac{Q}{2D}\lambda e^{-|x|/\lambda},$$

which we identify as the equilibrium solution for $C_m^{(1)}$, and recalling (2) gives the equilibrium for $C_{im}^{(1)}$. Thus we recover the equilibrium solution given in Berezhkovskii et al. (2009) Eqs. (19)–(23) therein for the spatial gradients of the first system. Next, we may substitute the equilibrium solution just found for $C_{im}^{(1)}$ into $g \equiv g_2$, with the authors' linear form of g_2, so that (6) now recovers the solution given in Berezhkovskii et al. (2009) Eq. (43) therein for the spatial gradients, $C_m^{(2)}$ and $C_{im}^{(2)}$, of the second system. Repeating the same procedure but this time choosing a Gaussian distribution for g_1, we recover the remaining equilibrium formula of Berezhkovskii et al. (2009) Eq. (47) therein. Sometimes there are limited data available on the precise details of the input distribution that is stimulating a given system. In this case, a delta distribution may be employed as a crude approximation to a source that is thought to be localised, while a Gaussian distribution may be employed as an approximation to a source that is thought to be more diffuse (see Fig. 4).

2.2. Comparing graded and ultrasensitive MAPK kinetics

In the model g_2 represents the dose–response properties of the MAPK cascade. In previous work g_2 was assumed to be linear, which, in analogy with enzyme kinetics, is termed graded. However, there is a significant body of experimental and theoretical evidence suggesting that the MAPK cascade exhibits an ultrasensitive response curve, as described in the landmark paper by Huang and Ferrell (1996). Thus it is important also to consider an ultrasensitive form of g_2. To model ultrasensitive kinetics we choose $g_2(x)$ to be

$$g_2(x) = g_h \frac{1}{1 + (C_0/C_{im}^{(1)}(x))^n}, \tag{7}$$

for given g_h, $C_0 > 0$ and Hill coefficient, n (see Fig. 3). In order to represent ultrasensitive kinetics, we require $n > 1$: in fact, Huang and Ferrell (1996) found the Hill coefficient of the MAPK cascade to be about $n = 4-5$, so we are especially interested in these cases.

Fig. 4 compares the gradient under graded and ultrasensitive signalling kinetics. We plot $g_2(C_{im}^{(1)}(x))$ with g_2 from (4) for linear signalling, and with g_2 from (7) for ultrasensitive signalling. In Fig. 4, left, the input g_1 is a Delta distribution on $x=0$, for which previous analysis shows the resulting equilibrium is

$$C_{im}^{(1)}(x) = \frac{K_1}{2D_1}\lambda_1 e^{-|x|/\lambda_1}, \tag{8}$$

whilst in Fig. 4, right, the input g_1 is a Gaussian, $1/\sqrt{2\pi\sigma^2}e^{-x^2/2\sigma^2}$, for which previous analysis shows the resulting equilibrium is

$$C_{im}^{(1)}(x) = \frac{K_1}{4D_1}\lambda_1 e^{-\sigma^2/2\lambda_1^2}\left[e^{x/\lambda_1}\operatorname{erfc}\left(\frac{\sigma^2+\lambda_1 x}{\sqrt{2}\lambda_1\sigma}\right) + e^{-x/\lambda_1}\operatorname{erfc}\left(\frac{\sigma^2-\lambda_1 x}{\sqrt{2}\lambda_1\sigma}\right)\right], \tag{9}$$

where erfc is the complementary error function. In both cases the width of the distribution with nonlinear signalling is much narrower than that with linear signalling, though this effect is more pronounced on the left. Thus, if the MAPK cascade exhibits ultrasensitive kinetics, then the length scale may be large before passing through the cascade and small afterwards. Depending on the parameter regime the result may look very different so it will be important to experimentally constrain model parameters in order to distinguish between ultrasensitive and graded kinetics, and thus to make inferences about length scales such as λ_1.

Consider Fig. 4, left, in more detail. When the input is a Delta distribution, the output is given by Eq. (8) so under linear signalling, $g_2(C_{im}^{(1)}(x))$ is proportional to $e^{-|x|/\lambda_1}$ and thus has length

[1] Even though physiologically realistic models must be finite the original model was formulated on an unbounded domain in order to accommodate Fourier-Laplace transforms. The identification of exponential decay behaviour in the Green's function provides reassurance that this original approximation was not unreasonable because in many cases the behaviour of the model will not be very sensitive to truncation to a large, finite domain. With this in mind, Figs. 4 and 6 are the results of computations on suitably truncated domains.

scale λ_1. In contrast, for fixed λ_1, under nonlinear signalling,

$$g_2(C_{im}^{(1)}(x)) \propto e^{-|x|/(\lambda_1/n)}\left(\frac{1}{e^{-|x|/(\lambda_1/n)}+\text{constant}}\right). \quad (10)$$

Away from $x=0$, the term in parenthesis is very nearly constant so, to a good approximation, $g_2(C_{im}^{(1)}(x)) \propto e^{-|x|/(\lambda_1/n)}$, and thus the effective length scale under nonlinear kinetics is reduced by a factor of n compared to the linear case. Thus if the MAPK cascade is ultrasensitive, as predicted by Huang and Ferrell (1996), then it effectively reduces the length scale of the second input gradient to be about $n=4-5$ times smaller than that of the primary input gradient.

We note that analytic expressions for the output gradient under ultrasensitive kinetics may also be found for Delta distribution initial conditions: more details may be found in Appendix A.

3. Examining gradient dynamics in nuclear cycles 10–14

We compare how well graded and ultrasensitive signalling kinetics fit experimental measurements of the spatial gradient formed during terminal patterning in *Drosophila*. We focus on the data given in Fig. 3B of Berezhkovskii et al. (2009) (shown in Fig. 5) in order to distinguish between graded and ultrasensitive kinetics. The dots are experimentally observed intensities of dpERK levels at different points along the embryo. A description of the protocol can be found in Coppey et al. (2008). The *Drosophila* embryo develops in stages of rapid mitotic divisions in which the number of nuclei doubles with each cycle. For more details see, for example Coppey et al. (2008), Shvartsman et al. (2008) or Wolpert et al. (2002). The data we analyse quantify the gradient at cycles 10 and 14.

3.1. Tails of the distribution

In Fig. 3B of Berezhkovskii et al. (2009) (shown in Fig. 5), at the late stage, in the tails of the distribution, the experimental data (dots) lie above the line predicted by the graded mathematical model (solid line). This is especially noticeable at the $+550$ μm end. While this is not a perfect fit, it is not unreasonable to imagine that ultrasensitive kinetics would only exaggerate this effect and make the fit worse because we have seen that they tend to attenuate the signal even more sharply than linear kinetics. For example, ultrasensitive kinetics reduce the effective length scale of $g_2(C^{(1)}(x))$, as shown in Fig. 4.

3.2. Three striking features of the gradient

Three striking qualitative features of the experimentally recorded gradients at the early and late stages of nuclear division shown in Fig. 3B of Berezhkovskii et al. (2009) are

(i) the peak of the distribution rises by about 2/3;
(ii) the distribution narrows by about 1/2;
(iii) the curvature at the peak sharpens.

This experimentally observed gradient should be identified with

$$C_{tot}^{(2)} \equiv C_{im}^{(2)} + C_m^{(2)} = (1+K_2)C_m^{(2)},$$

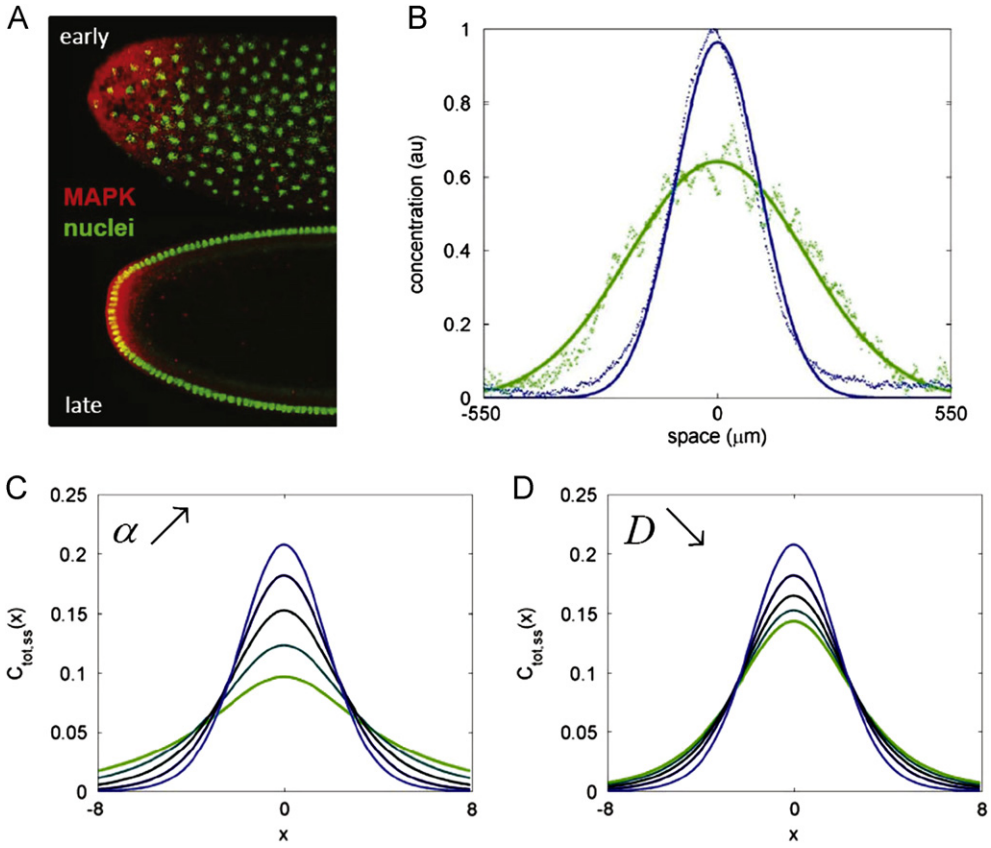

Fig. 5. Fig. 3 from Berezhkovskii et al. (2009), Copyright (2009) National Academy of Sciences, U.S.A. Dynamics of MAPK phosphorylation profiles in the terminal system. (A) Fluorescent images of nuclei (green, light) and phosphorylated MAPK (red, dark) at two different nuclear densities. (B) Quantified gradients of total (nuclear and cytoplasmic) phosphorylated MAPK. The green (light, short and fat) and blue (dark, tall and thin) curves show the patterns of MAPK phosphorylation at nuclear cycles 10 and 14, respectively. Increase in the nuclear density amplifies the total level of MAPK phosphorylation near the poles of the embryo and attenuates it in the rest of the system. (C) Increasing the nuclear trapping rate sharpens the profile of $C_{tot,ss}^{(2)}(x)$, computed with Eqs. (53) and (40). (D) Decreasing in the diffusivity sharpens $C_{tot,ss}^{(2)}(x)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and the analytic expressions derived earlier will help in comparing how well this mathematical model fits the data under graded and under ultrasensitive kinetics.[2]

The three features of the mathematical model corresponding to the conditions (i)–(iii) are as follows:

(i) The height of the peak is $C^{(2)}_{tot}(0)$.
(ii) One measure of the width is an effective characteristic length scale, λ. In order to estimate the change in width we crudely approximate the gradient by $ke^{-x/\lambda}$ for $k>0$ so that the effective characteristic length scale is estimated by reading off from the data the x-value that gives a concentration-value of $ke^{-1}\approx k\times 0.37$. From Fig. 3B, at the early stage $k\approx 0.65$ while at the late stage $k\approx 1$. This procedure estimates that λ at the early stage is about twice as big as at the late stage: crude estimates of λ^E (early) and λ^L (late) are 0.5×550 and 0.24×550 μm, respectively.
(iii) The 'curvature' is measured by the derivative near the peak at $x=0$.

3.3. Gradient at the end of the first stage

We now check whether $g_2(C^{(1)}_{im})$ can exhibit properties (i)–(iii). Of course we should be comparing the experimental results of Fig. 3B of Berezhkovskii et al. (2009) with $C^{(2)}_{tot}$, not $g_2(C^{(1)}_{im})$. However, it is beneficial to start with a simplified example in order to gain some insight from which to then build up to analysing the full model. Furthermore, our goal is to provide a framework for distinguishing graded and ultrasensitive kinetics and this task will usually be more sensibly accomplished by analysis of $g_2(C^{(1)}_{im})$ than of $C^{(2)}_{tot}$. One reason for this is that diffusion in the second stage will tend to smooth out any differences between the distributions due to graded and ultrasensitive forms of g_2. This point is emphasised in the Discussion. In the next two subsections, as a demonstration of the procedure we propose, we consider a Delta distribution as an input g_1, and compare graded and ultrasensitive kinetics.

3.3.1. Graded kinetics are consistent with nuclear dynamics

First consider graded kinetics, for which

$$g_2(C^{(1)}_{im}(x))=\frac{g_l K_1}{2D_1}\lambda_1 e^{-|x|/\lambda_1}.$$

The width of the peak is λ_1 and the height of the peak is $g_2(C^{(1)}_{im}(0))=g_l K_1\lambda_1/2D_1$. Thus in order to match (i) and (ii) we require that λ_1 decreases and that K_1/D_1 increases. These two criteria are simultaneously possible, for example, by decreasing the diffusivity D_1 and/or increasing the capture rate α_1. To match the relative magnitude of the changes estimated in (i) and (ii), we can be more precise and require that λ_1 decreases to about half its original value, while K_1/D_1 must increase to about triple its original value in order to compensate.

Berezhkovskii et al. (2009) observe the same principle (shown in Fig. 5), though this is with reference to the second stage. In the second stage, in the cytosol, the increase in the number of nuclei from early to late stages naturally motivates the increase in α_2 and decrease in D_2. However, the first stage takes place at the membrane where potential mechanisms for these changes in α_1 and D_1 are perhaps more complicated, as we note in the Discussion.

Finally, to investigate (iii), note that for $x>0$, say,

$$\frac{\mathrm{d}}{\mathrm{d}x}(g_2(C^{(1)}_{im}(x)))\propto -\frac{K_1}{D_1}e^{-x/\lambda_1},$$

and thus by the late stage, at $x=0$, $|\mathrm{d}/\mathrm{d}x(g_2(C^{(1)}_{im}(x)))|$ increases to about triple its value at the early stage, and by continuity also increases in a neighbourhood of the peak.[3] Thus all three criteria (i)–(iii) are met by the graded model.

3.3.2. Ultrasensitive kinetics are difficult to match with nuclear dynamics

Next, consider the ultrasensitive model. Recall (10), which describes the gradient for a fixed λ_1. The behaviour away from $x=0$ is approximately proportional to $e^{-|x|/(\lambda_1/n)}$, so even in the nonlinear model the length scale is still proportional to λ_1, though it may be reduced. Thus in order to match (ii) we again require that λ_1 halves between early and late stages.

Let $r_p=p_L/p_E$ be the ratio of the early stage peak magnitude to late stage peak magnitude, and let $E=K_1^E/D_1^E$ and $L=K_1^L/D_1^L$ be the ratios of the key parameters that change from early to late stages, respectively. We obtain

$$L=2\sqrt[n]{r_p}\left(2\frac{C_0}{\lambda_1}\right)\sqrt[n]{\frac{E^n}{\left(2\frac{C_0}{\lambda_1}\right)^n+(1-r_p)E^n}}.$$

We see that in the simple case where $r_p=1$ the ratio L/E must double to compensate for the halving of λ_1, and that for $r_p>1$ the ratio must more than double. From the estimate in (i), $r_p\approx 1.67>1$ so we see that the ratio must indeed more than double. However, due to the nonlinearity in the Hill function the question of how much extra is required over and above doubling depends on the relative magnitudes of C_0, λ_1 and E. We introduce the constant $C_1=(K_1\lambda_1)/(2D_1)$. For C_1 small (large) relative to C_0 the extra amount required will be less (more) than the corresponding extra amount in the linear setting. Notice that there is a bound on the maximum increase:

$$r_p<r_{max}=1+\left(\frac{2C_0}{E\lambda_1}\right)^n.$$

This saturation of the response is a property of the Hill function and it is a point of difference between the linear and nonlinear mathematical models. If the parameter regime can be constrained experimentally, this bound may help to distinguish the two kinetic settings. In our example, a necessary condition for ultrasensitive kinetics to be consistent with the data is

$$0.45\approx\frac{1}{2}\sqrt[n]{r_p}<\frac{C_0D_1}{K_1\lambda_1}. \tag{11}$$

Thus, if one can experimentally observe an increase in the peak of the gradient between early and late stages that is too large, with reference to this bound, then ultrasensitive kinetics (7) can be ruled out.

In the absence of parameter information, we suppose that the observed increase satisfies the bound, $r_p<r_{max}$, so that it is possible for ultrasensitive kinetics to simultaneously satisfy (i) and (ii). Finally we check (iii): recall (10) but now consider different values of λ_1 as well as behaviour near $x=0$. We write

[2] Observing that $1+K_2$ is merely a scaling factor it does no harm to compare these three qualitative features with, say, the immobile fraction $C^{(2)}_{im}$ instead of the $C^{(2)}_{tot}$, if this is more convenient.

[3] We take derivatives for $x>0$ but, by symmetry, it is trivial to address the case that $x<0$ and also by symmetry, at $x=0$ the signs of the left- and right-hand sided limits may not agree but the magnitude will. The derivative at $x=0$ may not exist. Thus, although we abuse notation to indicate the evaluation of a derivative at $x=0$, in order to be well-defined, we must agree to identify this as the right-sided limit. For the purpose of this investigation we are only interested in the magnitude, so one could just as easily choose the left-sided limit.

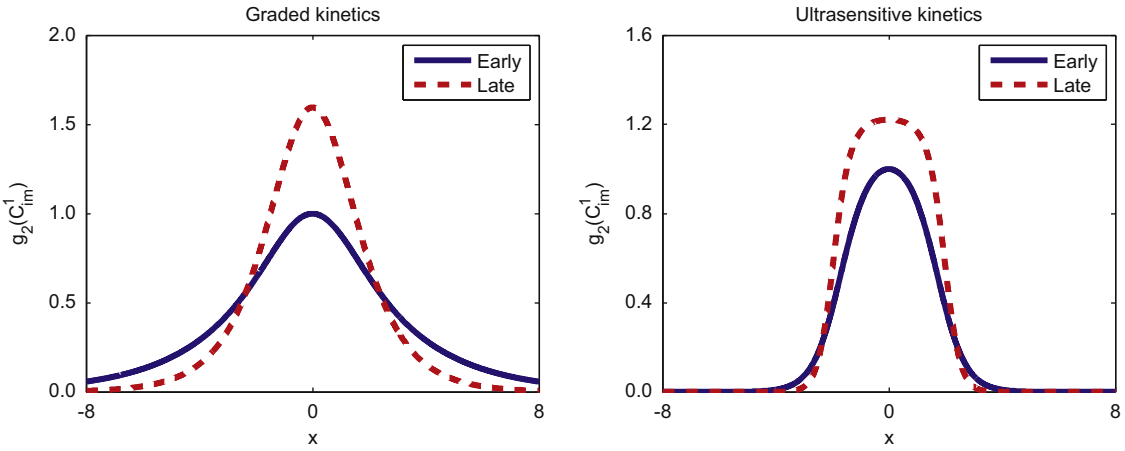

Fig. 6. Comparison of the gradient at the end of the first stage, under graded and ultrasensitive signalling kinetics. Input g_1 is a Gaussian with mean zero and variance σ^2. For these parameters, with ultrasensitive kinetics, the peak broadens at a late stage, which is not consistent with the experimental data, (ii). Parameters for the early stage are the same as Fig. 4. Late stage: $D_1^L = D_1^E/3$, and $\alpha_1^L = 4\alpha_1^E/3$.

(10) more carefully as

$$f(x) = e^{-|x|/(\lambda_1/n)}\left(\frac{1}{e^{-|x|/(\lambda_1/n)} + \hat{C}_{01}\lambda_1^{-n}}\right),$$

where $C_{01} \equiv C_0/C_1$ and $\hat{C}_{01} \equiv C_{01}/\lambda_1$ is a constant not involving λ_1. Notice $g_2(C^{(1)}_{im}(x)) \propto f(x)$. Analysis presented in Appendix B shows that in order to satisfy (iii) we have another necessary condition:

$$\lambda_1 > \lambda_1^* = \sqrt[n]{\hat{C}_{01}/\gamma^*}. \tag{12}$$

Notice that $\lambda_1^* < \sqrt[n]{\hat{C}_{01}}$ and that for $n \gg 1$ a good approximation is $\lambda_1^* \approx \sqrt[n]{\hat{C}_{01}}$, which is easy to compute. Even for moderate n, such as $n=5$, this approximation will often be within 10%, which may be useful in practice. This may help when checking the necessary condition (12) although we must bear in mind that Eq. (12) has been derived from simplifying assumptions so, while it provides some intuition about important parameter regimes, corrections would be required in more careful applications.

Thus we have shown that the peak may actually broaden under ultrasensitive kinetics if λ_1 is too small. This result is for an input of g_1 as a Delta distribution but, making use of formulae (9), (4) and (7), Fig. 6 shows this effect of peak broadening is also possible when the input is a Gaussian (note that in Figs. 4, 6–8 parameters are in arbitrary units and chosen merely to exaggerate effects).

In summary, with an ultrasensitive model, it is difficult to match the dynamics of the gradient at the end of the first stage with the three qualitative features (i)–(iii). Two necessary conditions have been derived, in Eqs. (11) and (12). Guided by these conditions, perhaps experiments can constrain parameters in such a way that ultrasensitive kinetics may be ruled out as not consistent with qualitative features of the gradient. This process of checking whether (i)–(iii) are satisfied could be repeated for $C^{(2)}_{tot}$, and we begin to address this in Section 4.2, and in the Appendix. However, as noted earlier, in many cases it will be better to do the comparison at the level of the first stage, as demonstrated here, because this gives the best chance of distinguishing between linear and nonlinear forms.

4. Discussion

For simplicity, the biochemical process that relays the signal from one stage to the next, involving the MAPK cascade, has been represented here by a single function g_2. When drawing conclusions about graded or ultrasensitive kinetics, note that the kinetics of the MAPK cascade *per se* may be different to the kinetics of the process that we represent here by g_2. Moreover, even in the simpler setting where the MAPK cascade is studied in isolation, it is not trivial to definitively classify its behaviour as always being of one type, such as always being graded. Compare, for example, notable results in stochastic and other settings, such as those described in Huang and Ferrell (1996), Mackeigan et al. (2005), Takahashi et al. (2010) and Qiao et al. (2007). Given the typically high degree of uncertainty surrounding parameters and other aspects of models in systems biology, it has been argued that it is better to focus on qualitative and systems-level properties rather than on the fine details of solutions (Gutenkunst et al., 2007; Qiao et al., 2007) and these sentiments are worth bearing in mind when we interpret the behaviour of g_2 as a proxy for the MAPK cascade. In this regard it is also worth noting that the analysis here is not restricted to the MAPK cascade in particular and provides a general framework for analysing linear and nonlinear signal relays in cascades of diffusion-trapping systems. Also note that physiologically reasonable models of signalling kinetics should eventually saturate. Ultrasensitive kinetics satisfy this property but with graded kinetics, at least in the present mathematical model, the signal may grow without bound.

In order to distinguish ultrasensitive and graded kinetics, we focused on three key features of the gradient: (i) the magnitude of the peak; (ii) the effective characteristic length scale; and (iii) the curvature near the top of the peak. These were chosen because previously published experimental results show how these features change between nuclear cycles 10 and 14, and thus show that it is at least possible to experimentally measure them. However, other good choices could also be made for the key features with which to interrogate the model. Very similar strategies – in which the sensitivity of the chosen features to key parameters is compared under graded and ultrasensitive kinetics – to those proposed here could be followed.

4.1. Sensitivity analysis

Genetic perturbation studies may be employed to interrogate the kinetics of the signalling cascade. With the analytic expressions derived earlier we may assess the sensitivity of the equilibrium distributions to perturbations in key parameters by, for example, differentiating with respect to those parameters.

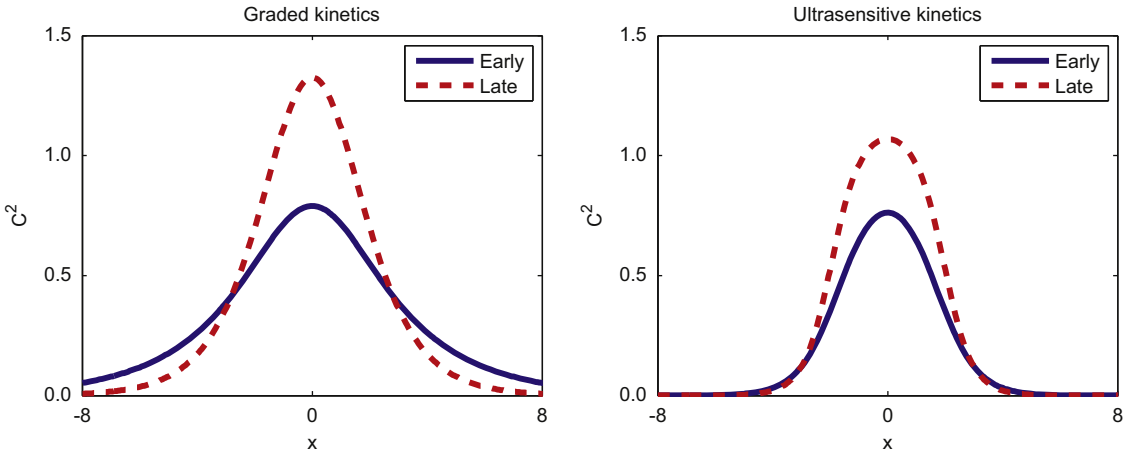

Fig. 7. Comparison of the gradient at the end of the second stage, under graded and ultrasensitive signalling kinetics. The four gradients shown in Fig. 6 are used here as four input gradients to the second stage. With D_1 and λ_1 as in Fig. 6, the parameters for the second stage here are $D_2^E = 0.9D_1^E$, $D_2^L = 0.75D_2^E$, $\lambda_2^E = 0.25\lambda_1^E$, and $\lambda_2^L = 0.9\lambda_2^E$. The distinction between graded and ultrasensitive kinetics at this second stage is not as pronounced as at the end of the first stage (compare with Fig. 6).

Suppose that the length scale is an independent variable that may be perturbed.[4] Differentiate $g_2(C^{(1)}(0))$ as a function of λ_1, the length scale of the first diffusion-trapping system, to obtain one crude measure of sensitivity. When the input, g_1, is a Delta distribution the result shows that in the linear setting a, say, 10% change in λ_1 will be accompanied by a 10% change in the height of the peak. In contrast, in the nonlinear setting, the same 10% change in λ_1 may be accompanied by a change in the height of the peak that is out of proportion. Moreover there are different ranges of λ_1 in which the sensitivity will be vastly different, whereas in the linear setting the sensitivity is constant across different ranges. This last observation suggests that small perturbations in just two well-chosen ranges of λ_1 are sufficient to distinguish ultrasensitive from graded kinetics.

In fact, genetic perturbation studies have been performed in which the number of copies of the Torso gene was increased (Coppey et al., 2008, Fig. 2). No effect on the distribution or amplitude of the signalling output (dpERK) was observed in these experiments, which suggests receptor numbers are not a limiting factor. One way to represent increasing receptor numbers in the present mathematical model is to increase α_1 but this does change the distribution (as shown, for example in previous work by Berezhkovskii et al., 2009). A possible explanation is that the experimental system corresponds to a parameter regime in the mathematical model in which the output gradient at the end of the second stage is insensitive to α_1. However, reconciling this experimental observation with the mathematical model remains an issue for future research. Precisely identifying all of the parameters in the model is a related issue and in this regard we note that estimates of effective diffusion constants have been obtained by Sample and Shvartsman (2010).

4.2. Gradient at the end of the second stage

As noted earlier, it will often be challenging to distinguish differing forms of the relay function, g_2, by examination of the gradient at the end of the second stage because, in general, diffusion tends to smear out any differences in an initial distribution. The larger the diffusivity in the second stage, the less sensitive $C_{im}^{(2)}$ will be to differing forms of the relay, g_2. Analytic expressions for $C_{tot}^{(2)}$ may aid in assessing the extent to which such differences will be smeared out and in this way it may be possible to extend the framework here to consider whether it is practical to distinguish different g_2 kinetics solely from measurements of equilibrium solutions, $C_{tot}^{(2)}$. We begin to address this in the Appendix. However, it is clear that as the diffusivity D_2 is increased, eventually, it will no longer be practical to distinguish between differing forms of g_2.

In order to provide a comparison with Fig. 6, which focused on the end of the first stage, Fig. 7 shows the gradient at the end of the second stage. The gradients shown are computed via quadrature on (6). Recall from (5) that the result of one such computation is proportional to $C_{im}^{(2)}$. The constant of proportionality (3) is not specified here because this would require more assumptions about the parameter values of the second diffusion-trapping system and because this particular result is only intended to illustrate qualitative properties. The main qualitative property to note is that, for the purposes of distinguishing between graded and ultrasensitive regimes, differences in g_2 are not as easily recognised at this stage as they are at the end of the first stage.

4.3. Gradients with constant shape and changing amplitude

Recently, Kanodia et al. (2009) showed that the Dorsal gradient in *Drosophila* is dynamic, maintaining a constant shape with increasing amplitude. Although this result applies to a different system to the one being studied here, some of the underlying principles, notably diffusion-trapping involving nuclear dynamics, are common to both. Therefore the general framework of the diffusion-trapping model outlined in this paper may also be applicable to the study of other signalling mechanisms and our purpose here is merely to suggest that this may be the case.

As an indication of this generality and motivated by the example of the gradient that maintains constant shape with changing amplitude, we investigate the potential of our model to exhibit similar amplification properties. We ask what input gradients, $g(x)$, result in purely signal amplification, giving an output of the form $\rho g(x)$? In this respect observe that in (6), solutions v are obtained by applying a linear operator to an input function g, so it is enlightening to consider associated eigenfunctions. For example, for special cases in which $v(\lambda,D,\rho) = \sqrt{1/\lambda^2 - 1/\rho D}$, an input gradient of the form $g(x) = e^{\pm vx}$ results in an output gradient ρg. Clearly this is not

[4] Of course this is not strictly correct because the length scale is a function of physically meaningful parameters such as diffusion so this example is for the purpose of illustration only.

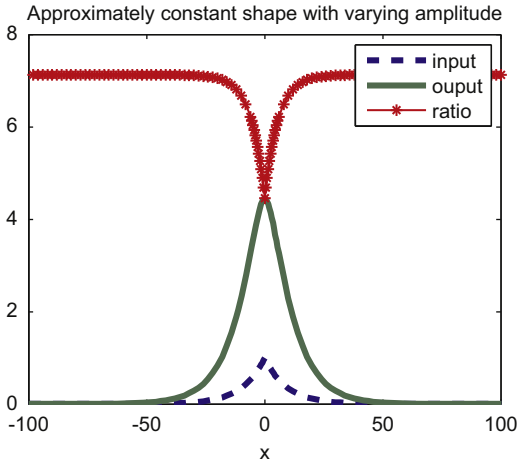

Fig. 8. Quadrature on (6) reveals an example of an input gradient, $g=e^{-\nu|x|}$, that results in an output gradient, v, which maintains an approximately constant shape with increasing amplitude. The ratio of output to input, v/g, is also plotted. Parameters are as follows: $\lambda=3.5$, $D=2$, $\rho=\lambda^2/D+1=7.125$, and $\nu=\sqrt{1/\lambda^2-1/(\rho D)}\approx 0.1$.

a physically realisable solution. However, it does suggest that gradients that are close, such as $g(x)=e^{-\nu|x|}$, might also be close to exhibiting the desired amplification property.

For example, as shown in Fig. 8, after truncation to a finite domain, quadrature on Eq. (6) reveals that it is possible that the input $g(x)=e^{-\nu|x|}$ results in an output $v(x)\approx\rho e^{-\nu|x|}$, which has approximately the same shape as the input. If the output had exactly the same shape then the ratio $v(x)/g(x)$ would be exactly ρ, at all points x. In the example of Fig. 8, away from zero, the ratio $v/g\approx\rho$. Near the peak of the gradient, the ratio $v/g<\rho$. The ratio is smallest at the peak and increasing away from zero. This observation suggests that for experimentally observed gradients that appear to maintain a constant shape with growing amplitude, it may be worthwhile measuring the ratio of the output to the input gradient more carefully because similar behaviour may be present. This could be accomplished, for example, by measuring the ratio in two places, near the peak, and away from the peak.

5. Conclusions

Morphogen gradients play a key role in differentiation and development. We have analysed the role of the MAPK cascade in the mechanisms that establish such gradients and allow the developing embryo to detect them and respond appropriately. Cascades of two-state diffusion-trapping systems are one of the mechanisms that establish gradients in *Drosophila*. Recently, the question of how MAPK kinetics influence this spatial gradient was raised in the literature, with particular reference to whether the MAPK cascade exhibits graded or ultrasensitive kinetics in this context. Previously, analytic expressions have been obtained in the graded, linear setting. One of the major contributions of this work is to provide a complementary framework for analysing the same system in both ultrasensitive and nonlinear settings. One result is an analytic expression for the spatial gradient under ultrasensitive MAPK kinetics. The techniques developed are reasonably general and the same principles can be applied to handle various functional forms of the input gradient. Suggestions have been made for extending the analysis to various qualitative cases, such as those in which the length scales of consecutive diffusion-trapping systems are approximately equal, or in which they differ by about an order of magnitude.

As a demonstration of the utility of the approach, a case of special biological relevance has been treated. In particular, we show that it is difficult to match the change in shape of the gradient between nuclear cycles 10 and 14, shown in Fig. 3B of Berezhkovskii et al. (2009), with ultrasensitive kinetics. Thus, compared to ultrasensitive kinetics, our analysis supports a graded model of signal relay between the two stages as a better fit to the data.

Finally, previous studies have concluded that the characteristic length scale of the Torso/Trunk receptor/ligand diffusion-trapping system is much less than that of the cytosolic/nuclear diffusion-trapping system. While this conclusion is very reasonable, increased confidence would come with more direct measurements. This is because the conclusion is based on measurements of the output of the signalling cascade (dpERK) but not on direct measurements of Torso occupancy. However, under ultrasensitive MAPK kinetics, dpERK may exhibit a small characteristic length scale even if the gradient of Torso occupancy exhibits a large length scale. For example, a useful approximation presented here is that, as a rule-of-thumb, ultrasensitive signalling kinetics reduce the effective length scale by about a factor of n where n is the Hill coefficient characterizing the signalling kinetics. For the MAPK cascade, this has been suggested to be a factor of about $n=4$.

Acknowledgements

S.M. acknowledges financial support from St John's College, Oxford. P.K.M. was partially supported by a Royal Society Wolfson Research Merit Award. This research was partially supported by the Oxford-Princeton Collaborative Research Grants Scheme.

Appendix A. Analytic expressions for the gradient under ultrasensitive kinetics

Suppose the input, g_1, is a Delta distribution, as in the first example of Berezhkovskii et al. (2009). Recall that in this case the equilibrium solution is

$$C_{im}^{(1)}(x)=\frac{K_1}{2D_1}\lambda_1 e^{-|x|/\lambda_1}.$$

Applying (6), (2), and (7) the equilibrium solution for $C_{im}^{(2)}(x)$ is

$$C_{im}^{(2)}(x)=\frac{g_h K_2}{2D_2}\lambda_2\left(e^{-x/\lambda_2}\int_{-\infty}^{x}\frac{e^{z/\lambda_2}}{1+\left(\frac{C_0}{C_1 e^{-|z|/\lambda_1}}\right)^n}\,dz + e^{x/\lambda_2}\int_{x}^{\infty}\frac{e^{-z/\lambda_2}}{1+\left(\frac{C_0}{C_1 e^{-|z|/\lambda_1}}\right)^n}\,dz\right), \tag{A.1}$$

where we introduce the constant $C_1=(K_1\lambda_1)/(2D_1)$ to tidy up the integral. Observing the symmetry of the problem, $C_{im}^{(2)}(x)=C_{im}^{(2)}(-x)$, it suffices to evaluate the integrals for the case that $x>0$. In this case, the first integral is the sum of the integrals on either side of $z=0$, with $|z|$ replaced by $\pm z$ as appropriate. Define

$$m=\frac{\lambda_1}{\lambda_2}. \tag{A.2}$$

With the substitutions $u=e^{\pm z/\lambda_1}$ Eq. (A.1) becomes

$$C_{im}^{(2)}(x)=\frac{g_h K_2}{2D_2}\lambda_1\lambda_2 e^{-|x|/\lambda_2}\left(\int_0^1\frac{u^{m-1}}{1+\left(\frac{C_0}{C_1 u}\right)^n}\,du+\int_1^{e^{|x|/\lambda_1}}\frac{u^{m-1}}{1+\left(\frac{C_0 u}{C_1}\right)^n}\,du\right)$$

$$+\frac{g_h K_2}{2D_2}\lambda_1\lambda_2 e^{|x|/\lambda_2}\int_0^{e^{-|x|/\lambda_1}}\frac{u^{m-1}}{1+\left(\frac{C_0}{C_1 u}\right)^n}\,du. \tag{A.3}$$

The key integral we require is

$$\int u^{m-1}[1+(C_{01}u^{\pm 1})^n]^{-1}\,du=\frac{u^m}{m}\,{}_2H_1\left(\left[1,\pm\frac{m}{n}\right],\left[1\pm\frac{m}{n}\right],-(C_{01}u^{\pm 1})^n\right), \tag{A.4}$$

where ${}_pH_q(\{a_i\}_{i=1}^p,\{b_i\}_{i=1}^q,z)$ is the hypergeometric function and we introduce the constant $C_{01}=C_0/C_1$. We use Maple for this calculation and for (A.5), (A.9) and (A.10). Note the '-' part of the solution does not apply in the special case that m/n is an integer, so we treat this case separately. Otherwise, when $p=q+1$, the hypergeometric function is defined by a power series for $|z|<1$ and extended to the rest of the complex plane by analytic continuation. When $C_{01}u^{\pm 1}<1$ we can compute (A.4) via the convergent series

$$\frac{u^m}{m}\left(1\pm\frac{m}{n}\sum_{k=1}^{\infty}\frac{(-C_{01}^n u^{\pm n})^k}{k\pm\frac{m}{n}}\right).$$

For $u=1$, two of the limits in (A.3) simplify to ${}_2H_1(\cdot,\cdot,-C_{01}^n)$ and two more of the limits are ${}_2H_1(\cdot,\cdot,-(C_{01}e^{\pm|x|/\lambda_1})^n)$. Some care must be taken with the limit at zero for the '-' part of the solution.

A.1. Integer values of m/n

Experimentally, it has been observed that the length scale of the first system is less than that of the second, so $0<m<1$, and that the MAPK cascade operates in a regime of $n=4$–5. Thus, typical values of the fraction m/n are in the range $0<m/n<1/5$, so perhaps for biologically relevant values the case of integer values of m/n is not critical. Nevertheless, this case can also be accommodated in the current framework. The issue is that ${}_pH_q(\{a_i\}_{i=1}^p,\{b_i\}_{i=1}^q,z)$ may not be defined if b_i is a negative integer for some i. In the case of integer m/n, introduce the substitution $w\equiv u^n$, and use the following in place of (A.4)

$$\int u^{m-1}(1+(C_{01}u^{-1})^n)^{-1}\,du$$
$$=\frac{1}{n}\int\frac{w^{m/n-1}}{1+C_{01}^n w^{-1}}\,dw=\frac{1}{m}w^{m/n}\left[1-{}_2H_1\left(\left[1,\frac{m}{n}\right],\left[1+\frac{m}{n}\right],-\frac{w}{C_{01}^n}\right)\right]. \tag{A.5}$$

As usual, the limits must be adjusted for the substitution.

A.2. Systems with widely differing length scales

The case $\lambda_2\gg\lambda_1$ is of special importance because it has been observed experimentally that the length scale of the second system is much larger than that of the first. A convenient approach to this case is to change the definition of (A.2) to $m=\lambda_2/\lambda_1$ and use the substitutions $u=\exp(\pm z/\lambda_2)$. In this case, up to sign changes, the integrals in (A.1) transform to

$$\lambda_2\int(1+C_{01}^n u^{\pm mn})^{-1}\,du, \tag{A.6}$$

which may also be evaluated with (A.4) by replacing n by nm, C_{01} by $\sqrt[m]{C_{01}}$, and m by 1.

A.3. Comparing graded and ultrasensitive kinetics

In order to compare the spatial gradient under graded and under ultrasensitive signalling kinetics of the MAPK cascade, we may compare (A.1) with the solution obtained in the linear setting. Following the procedure outlined in Example 1, we obtain the solution in the linear setting as, for $\lambda_1\neq\lambda_2$,

$$C_{im}^{(2)}(x)=\frac{g_l K_1 K_2\lambda_1^2\lambda_2^2}{2D_1D_2(\lambda_1^2-\lambda_2^2)}(\lambda_1 e^{-|x|/\lambda_1}-\lambda_2 e^{-|x|/\lambda_2}), \tag{A.7}$$

or, for $\lambda=\lambda_1=\lambda_2$, as

$$C_{im}^{(2)}(x)=\frac{g_l K_1 K_2\lambda^2}{2D_1D_2}e^{-|x|/\lambda}(\lambda+|x|). \tag{A.8}$$

Thus, for a particular value of the Hill coefficient, n, we evaluate (A.1) via (A.4) and compare the result with (A.7), if the length scales are different, or (A.8) if the length scales are the same. As noted, for the MAPK cascade, important values of n are 4–5. To illustrate this process we now fix $n=4$ and consider three special cases in which the length scales of the two stages are approximately the same, the length scale of the first is much larger than that of the second, and *vice versa*. The following examples are merely to illustrate the sorts of issues that may arise in this process.

A.3.1. The case of $\lambda_1=10\lambda_2$

This corresponds to the choice $m=10$ in (A.2). Following the procedure outlined above we obtain an expression for the equilibrium solution under ultrasensitive kinetics:

$$\begin{aligned}C_{im}^{(2)}(x)=\frac{g_h K_2}{2D_2}\lambda_1\lambda_2\Big[e^{-x/\lambda_2}\Big\{&-\frac{1}{2}C_3C_{01}^{10}+\frac{1}{2}C_{01}^8-\frac{1}{6}C_{01}^4+\frac{1}{10}\\&-\frac{1}{6}\frac{C_1^4}{C_0^{10}}(-3C_1^4c_0^2+C_0^6+3C_4C_1^6+3C_1^4C_0^2e^{2x/\lambda_1}\\&-C_0^6e^{6x/\lambda_1}-3C_1^6\arctan(C_{01}^2e^{2x/\lambda_1}))\Big\}\\&+\frac{1}{30C_1^{10}}e^{x/\lambda_2}\{-5C_0^4C_1^6e^{-6x/\lambda_1}+3C_1^{10}e^{-10x/\lambda_1}+15C_0^8C_1^2e^{-2x/\lambda_1}\\&-15C_0^{10}\arctan(C_{01}^{-2}e^{-2x/\lambda_1})\}\Big],\end{aligned} \tag{A.9}$$

where $C_3=\arctan(C_{01}^{-2})$ and $C_4=\arctan(C_{01}^2)$. Compare this with the linear setting in (A.7). The constant of proportionality, $\lambda_1\lambda_2K_2/2D_2$, is very similar, though the parts corresponding to the first stage, such as $K_1/2D_1$, now turn up inside the integral via C_1 because of the nonlinearity. The exponential decay in the solution persists via the term $\exp(-x/\lambda_2)$. Further simplification of the arctan terms is possible by recalling that the arctan function resembles a Hill function, though the upper and lower asymptotic limits are adjusted to $\pm\pi/2$, so for large $\pm x$ two of the arctan terms may be replaced by these constant limits.

A.3.2. The case of $\lambda_2=10\lambda_1$

As noted, it has been suggested that the length scale of the first system is much smaller than that of the second (Berezhkovskii et al., 2009). This corresponds to the choice $m=1/10$ in (A.2), but it is more convenient to change this definition and evaluate the integral instead via (A.6), as discussed. Integrals similar to $\int(1+u^{40})^{-1}\,du$ arise and have expressions in terms of simple trigonometric and log functions.

A.3.3. The case of $\lambda_2=\lambda_1$

Consider the case that the length scales are approximately the same, *i.e.* that $\lambda=\lambda_1=\lambda_2$, which corresponds to $m=1$ in (A.2). Again, it is possible to give an expression for the integral but it involves many terms. Some insight is obtained by considering the key integrals involved:

$$\begin{aligned}\int\frac{e^{z/\lambda}}{1+\left(\frac{C_0}{C_1e^{z/\lambda}}\right)^4}\,dz=\lambda e^{z/\lambda}-\frac{\sqrt{2}C_{01}}{8}\lambda\Bigg[&\ln\left(\frac{e^{2z/\lambda}+\sqrt{2}C_{01}e^{z/\lambda}+C_{01}^2}{e^{2z/\lambda}-\sqrt{2}C_{01}e^{z/\lambda}+C_{01}^2}\right)\\&+2\arctan\left(1+\frac{\sqrt{2}e^{z/\lambda}}{C_{01}}\right)-2\arctan\left(1-\frac{\sqrt{2}e^{z/\lambda}}{C_{01}}\right)\Bigg].\end{aligned} \tag{A.10}$$

The integral

$$\int \frac{e^{z/\lambda}}{1+\left(\frac{C_0}{C_1 e^{-z/\lambda}}\right)^4}\,dz,$$

is the same as (A.10) minus the first term ($\lambda \exp(z/\lambda)$), with C_{01} replaced by $1/C_{01}$ and with the sign of the remaining three terms swapped. We may substitute these expressions into (A.1) and compare with (A.8). Notice that for large $\pm z$ the logarithm term is approximately zero. Further simplification is possible for large negative z because the sum of the arctan terms is also approximately zero, while for large positive z their sum is approximately π.

A.4. Gradient at the end of the second stage

Applying (A.1), (A.6) and (A.4) in that order, the magnitude of the peak is

$$C_{im}^{(2)}(0) = \frac{g_h K_2}{D_2}\lambda_2^2 \int_0^1 (1+C_{01}^n u^{-mn})^{-1}\,du,$$
$$= \frac{g_h K_2}{D_2}\lambda_2^2 \left[u_2 H_1\left(\left[1,-\frac{1}{mn}\right],\left[1-\frac{1}{mn}\right],-(C_{01}^{1/m}u^{-1})^{mn}\right)\right]_0^1. \quad (A.11)$$

Some care with limits is required because this is an improper integral but this suggests how the framework may be extended to check, for example, feature (i) of Section 3.2 if required. Another nice aspect of the Green functions approach is that it provides a framework for devising numerical solutions, which, as we show in Fig. 7, is not an unreasonable way to investigate the behaviour of the gradient at the second stage.

Appendix B. Comparing the dynamics

For $x>0$,

$$h(x,\lambda_1) \equiv \frac{\partial f}{\partial x} = \frac{n}{\lambda_1}(f^2-f).$$

By symmetry we also obtain the derivative for $x<0$. As noted previously, the derivative at zero may not exist. However, the limit from one side does exist. Moreover, although the left- and right-sided limits may not agree at $x=0$, the magnitude will. Thus when evaluating h at $x=0$, bear in mind that we define this to be the right-sided limit of $\partial f/\partial x$. Notice $0<f^2<f<1$ so $h<0$. Also, $h\to 0$ as $\lambda_1\to\infty$ and as $\lambda_1\downarrow 0$. Together with the observation that h is not constant we deduce that, for each fixed x, h has at least one minimum.

In order to understand how changes in the characteristic length scale, λ_1, relate to changes in the sharpness of the curvature near the peak of the gradient, we differentiate with respect to λ_1, as if it is an independent variable. Bear in mind the following two issues. First, experimentally, direct perturbation is restricted to physical parameters, for example, the rates of uptake or diffusion. Indirectly, of course, this may then alter the characteristic length scale, λ_1. Second, in what follows $\hat{C}_{01}$ is treated as a constant. However, both $\hat{C}_{01}$ and λ_1 depend on physical parameters such as the diffusion constant so, in general, both change as a result of a change in one of the physical parameters. Thus the formulae derived below apply to the special case that $\hat{C}_{01}$ is held constant. One practical example to which this case applies is where k_m is experimentally perturbed.

Thus, consider $\partial(|h(0,\lambda_1)|)/\partial\lambda_1 = 0$ to find critical points. Letting $f_0 \equiv f(0)$ critical points, λ_1^*, are solutions of

$$0 = 2n\hat{C}_{01}\lambda_1^{-n} f_0^2 - (n\hat{C}_{01}\lambda_1^{-n}+1)f_0 + 1.$$

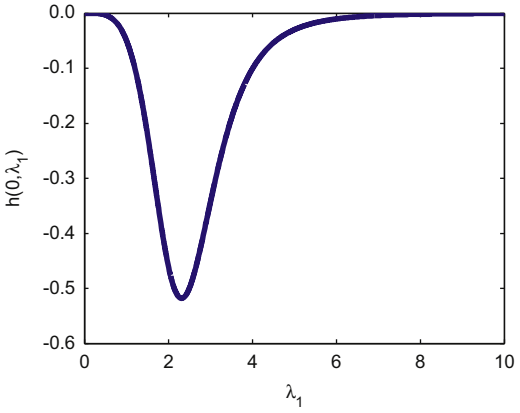

Fig. B1. A unique minimum at $2.3 \approx \lambda_1^* < \sqrt[n]{\hat{C}_{01}} \approx 2.5$. Parameters are as follows: $n=5$, and $\hat{C}_{01}=100$.

There exists a solution by our previous observation but we can be more precise. Introducing the change of variables $\lambda_1 = \sqrt[n]{\hat{C}_{01}/\gamma}$ reduces the equality to a quadratic in γ, of which only one root is admissible: $\gamma^* = (n+1)/(n-1)$. The unique solution is

$$\lambda_1^* = \sqrt[n]{\hat{C}_{01}/\gamma^*}.$$

Fig. B1 shows a plot of $h(0,\lambda_1)$, with the unique solution just below $\lambda_1^* \approx \sqrt[n]{\hat{C}_{01}}$. Fig. B1 also shows that the peak will not sharpen if λ_1 is too small.

References

Berezhkovskii, A.M., Coppey, M., Shvartsman, S.Y., 2009. Signaling gradients in cascades of two-state reaction–diffusion systems. Proc. Nat. Acad. Sci. USA 106 (4), 1087–1092.
Coppey, M., Berezhkovskii, A.M., Kim, Y., Boettiger, A.N., Shvartsman, S.Y., 2007. Modeling the bicoid gradient: diffusion and reversible nuclear trapping of a stable protein. Dev. Biol. 312 (2), 623–630.
Coppey, M., Boettiger, A.N., Berezhkovskii, A.M., Shvartsman, S.Y., 2008. Nuclear trapping shapes the terminal gradient in the *Drosophila* embryo. Curr. Biol. 18 (12), 915–919.
Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., Sethna, J.P., 2007. Universally sloppy parameter sensitivities in systems biology models. PLoS Comput. Biol. 3 (10), e189.
Huang, C.Y., Ferrell, J.E., 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Nat. Acad. Sci. USA 93 (19), 10078–10083.
Kanodia, J.S., Rikhy, R., Kim, Y., Lund, V.K., DeLotto, R., Lippincott-Schwartz, J., Shvartsman, S.Y., 2009. Dynamics of the dorsal morphogen gradient. Proc. Nat. Acad. Sci. USA 106 (51), 21707–21712.
Mackeigan, J.P., Murphy, L.O., Dimitri, C.A., Blenis, J., 2005. Graded mitogen-activated protein kinase activity precedes switch-like c-fos induction in mammalian cells. Mol. Biol. Cell 25 (11), 4676–4682.
Qiao, L., Nachbar, R.B., Kevrekidis, I.G., Shvartsman, S.Y., 2007. Bistability and oscillations in the Huang–Ferrell model of MAPK signaling. PLoS Comput. Biol. 3 (9), 1819–1826.
Sample, C., Shvartsman, S.Y., 2010. Multiscale modeling of diffusion in the early drosophila embryo. Proc. Nat. Acad. Sci. USA 107 (22), 10092–10096.
Shvartsman, S.Y., Coppey, M., Berezhkovskii, A.M., 2008. Dynamics of maternal morphogen gradients in drosophila. Curr. Opin. Genet. Dev. 18 (4), 342–347.
Shvartsman, S.Y., Coppey, M., Berezhkovskii, A.M., 2009. MAPK signaling in equations and embryos. Fly (Austin) 3 (1), 62–67.
Takahashi, K., Tănase-Nicola, S., Ten Wolde, P.R., 2010. Spatio-temporal correlations can drastically change the response of a MAPK pathway. Proc. Nat. Acad. Sci. USA 107 (6), 2473–2478.
Wolpert, L., Beddington, R., Jessell, T., Lawrence, P., Meyerowitz, E., Smith, J., 2002. Principles of Development. Oxford University Press.