Self-organized patterning by diffusible factors: roles of a community effect

Kirill Batmanov, Céline Kuttler, Cédric Lhoussaine *

Lille University, Laboratoire d'Informatique Fondamentale de Lille (LIFL, CNRS UMR 8022),

Cite Scientifique, Batiment M3, 59655 Villeneuve d'Ascq CEDEX, France

Yasushi Saka †

University of Aberdeen, School of Medical Sciences, Institute of Medical Sciences,

Foresterhill, Aberdeen AB25 2ZD, UK

Abstract. For decades, scientists have sought to elucidate self-organized patterning during development of higher organisms. It has been shown that cell interaction plays a key role in this process. One example is the community effect, an interaction among undifferentiated cells. The community effect allows cell population to forge a common identity, that is, coordinated and sustained tissue-specific gene expression.

The community effect was originally observed in muscle differentiation in *Xenopus* embryos, and is now thought to be a widespread phenomenon. From a modelling point of view, the community effect is the existence of a threshold size of cell populations, above which the probability of tissue-specific gene expression for a sustained period increases significantly. Below this threshold size, the cell population fails to maintain tissue-specific gene expression after the initial induction.

In this work, we examine the dynamics of a community effect in space and investigate its roles in two other processes of self-organized patterning by diffusible factors: Turing's reaction-diffusion system and embryonic induction by morphogens.

Our major results are the following. First, we show that, starting from a one-dimensional space model with the simplest possible feedback loop, a community effect spreads in an unlimited manner in space. Second, this unrestricted expansion of a community effect can be avoided by additional negative feedback. In Turing's reaction-diffusion system with a built-in community effect, if induction is localized, sustained activation also remains localized. Third, when a simple cross-repression gene circuitry is combined with a community effect loop, the system self-organizes. A gene expression pattern with a well-demarcated boundary appears in response to a transient morphogen gradient. Surprisingly, even when the morphogen distribution eventually becomes uniform, the system can maintain the pattern. The regulatory network thus confers *memory* of morphogen dynamics.
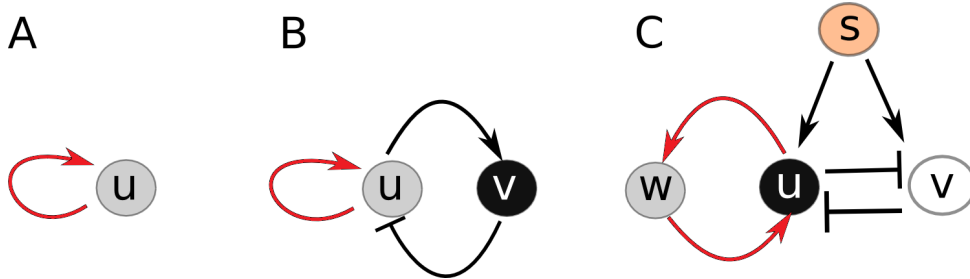
Figure 1.   This paper analyses three gene regulatory networks (GRNs), **(A)** minimal model of a community effect, **(B)** Turing pattern model, **(C)** dynamic morphogen gradient model. u and w are diffusible species for community effect (in grey), s is a morphogen species (in orange). Arrowheads: regular for activation, T-shape for inhibition. Red highlighting: direct or w-mediated feedback loops for community effects.

1.   Introduction

During animal development from a fertilized egg (a single cell) into a multicellular organism, embryos undergo complex series of transformations. By activating specific sets of genes, identical cells differentiate and are partitioned into distinct tissues with clear boundaries. Differentiating tissues are simultaneously patterned and moved to their final destinations. This biological process, in which an organism develops its shape, is called *morphogenesis* – a term coined in the late 19th century, having as origin the Greek words *morphe* (shape) and *genesis* (origin, creation, generation). During the self-organized, finely regulated process of morphogenesis, cells constantly create and exchange information by means of *diffusible factors*. These signalling molecules diffuse from a central source to other cells in the growing organism, forming a concentration gradient along an axis. Alan Turing introduced the term *morphogen* for diffusible factors in development, in his seminal paper *The chemical basis of morphogenesis* [49]. Turing's paper in 1952 shows the role of modelling in developmental biology. Quantitative frameworks from mathematics and computer science that allow the formulation and testing of hypotheses, have contributed to better understanding of development for decades [10, 14, 38, 41].

In the current paper, we investigate *gene regulatory networks* (GRNs) for animal development, illustrated as *directed graphs* in Fig.1. Similar graphical notations are widely used for knowledge representation in the biological literature [2, 14]. The graphs in Fig.1 summarize which genes control other genes. Each node of the graphs represents a *species*, indeed denoting both a gene, and the protein encoded by that gene. Directed edges represent *regulatory interactions*. The arrowhead style distinguishes between positive or negative regulation. Interaction can be direct between two species, or indirect through an intermediate sequence of molecular events that is not explicitly displayed as additional edges. It can also include intercellular communication by diffusible factors. Protein *decay* remains implicit in GRNs, unless it is due to inhibitory regulation between species.

From a computational perspective, graphical GRNs are useful representation to introduce and present models of chemical reactions. However, they have an expressiveness which is too weak to faithfully model the biological mechanisms of interest in this paper. We therefore refine graphical GRNs in two different semantics describing how chemical solutions evolve over time. On one hand, deterministic models are given by differential equations, which we solve numerically. On the other hand, stochastic

models are given in terms of continuous time Markov chains [33], that can be explored through Monte-Carlo simulations [21]. The latter models are actually defined by means of a formal ad-hoc rule-based programming language. This allows both for a compact and intelligible stochastic representation of GRNs. We do not directly address here theoretical questions of programming languages for biological modelling and design issues. It is rather a specific spatial biological aspect to which we apply the most accurate computational modelling tools currently available. We hope to give insights on how space and related questions may be best treated in biological models, and motivate future theoretical investigations of spatial properties of programs.

The community effect is a developmental phenomenon allowing to establish a *cell community with a common identity*. Upon reception of an inducing signal, a cell population starts co-ordinately expressing a specific set of genes. Whenever the population size exceeds a certain critical threshold, the cells maintain their common gene expression profile over an extended period of time. Upon terminal differentiation, the cell community has become a functional tissue, such as muscle.

A community effect was first identified in the African clawed toad *Xenopus laevis*, where it assists differentiation of muscle precursor cells [22, 25]. It has subsequently been observed in other organisms including the sea urchin [5] – and has been conjectured a widespread phenomenon [13]. Across organisms, the common denominator for a community effect is a *self-sustaining positive feedback loop*, with diffusible factors for intercellular communication [46].

The *minimal model of a community effect* shown in Fig.1A distils the essential properties of previous community effect models [5, 46]. It consists of the unique species u, which is exchanged between cells, and promotes its own production. The inducing signal for initial activation of u, which will be self-sustained by the positive feedback loop, is represented by injecting some u at the start of simulations.

The *space* in which community effects occur was disregarded in previous models. However, different communities within the embryo are spatially organized, with clearly demarcated boundaries. To examine how this happens and to study the properties of community effects in space, we adopt a *minimal model* of a community effect in which cells are arranged along a one-dimensional row. Diffusible factors for intercellular communication can migrate from one cell to the two directly neighbouring cells in this one dimensional grid. Although considering only one dimension might appear simplistic, compared to the three dimensions along which an embryo develops, we can obtain useful insights from it. We will show through stochastic simulations [20] that, in this configuration, a community effect spreads across the system.

The central question that arises is '*How can we limit the community effect in space?*'. Intuitively, the system needs some negative feedback control mechanism. To address this issue, we propose two approaches.

The Turing pattern model adds a second diffusible factor v with an inhibitory effect to the minimal model's self-sustaining species u. This second species is highlighted in black in Fig.1B. In order to close the loop of mutual dependencies, u activates not only itself, but also v. Intuitively, the effect of this seemingly small change in one dimensional space is not clear. We formulate this model by means of rules with parameters [12, 30], that represent sets of chemical reactions. Using stochastic simulations [20], we obtain the insight that *in this system, the spread of the community effects in space is restricted*.

Alan Turing's paper, *The Chemical Basis of Morphogenesis* [49], deals with the above network from

the perspective of pattern formation. Turing called it the reaction-diffusion system (*RD system*). When two diffusible factors interact in the manner shown in Fig.1B, the RD system may generate so-called *Turing patterns*. It has been shown that this system can generate many diverse patterns observed in development [36], for instance, colour stripes on fish skin [35, 48], colourings of sea shells [37], establishment of the left-right asymmetry in vertebrates [26] and plant growth [31]. The community effect, however, has not been considered in the context of the RD system.

The dynamic morphogen gradient model   sketched in Fig.1C introduces a diffusible factor s, which activates two mutually inhibitory species v and u. The latter participates in a self-activating feedback loop with w, another diffusible species. Fig.1C highlights the u-w community effect loop in red. Without the community effect loop and diffusion of the morphogen s in continuous space, this GRN was previously considered in [47].

Mutual repression between two species is widespread, and results in exclusive activation of either one species or the other [9]. It is often referred to as a genetic *toggle switch*, and controls various aspects of cell fate decision, e.g., the genetic switch of the bacteriophage lambda [1, 43], the gap gene network [28, 29, 42], and neural tube patterning [45].

In our model, the morphogen s spreads along one dimensional space from a localized source, and creates a *morphogen gradient*. It is postulated that embryonic cells learn their position by sensing the local morphogen level in a gradient for pattern formation [51]. This system is analogous to embryonic induction, where one tissue instructs and patterns the neighbouring tissue by diffusible factors.

Analysing the dynamic morphogen gradient model by means of partial differential equations (PDEs), we demonstrate that it efficiently limits community effects in continuous space (for an introduction to PDE modelling of morphogen gradient formation, see [50]). We also show that a *transient* morphogen gradient is sufficient for pattern formation in this system. This self-organizing property was previously unknown. Our analysis using stochastic simulations of the system indicates that community effects confer *robustness* to this patterning process. Interestingly, we found that even when the morphogen distribution eventually becomes uniform, the system can maintain the pattern. Thus our model confers *memory* to the system.

What has emerged from the analyses described in this paper is that diffusible factors for intercellular communication could provide many benefits to embryonic patterning, when interlinked with developmental GRNs.

Outline.   Section 2 defines the syntax and semantics of the rule-based language used in this paper. Section 3 introduces the *minimal model of a community effect* and the one dimensional model of diffusion. In Section 4, we discuss the analysis of the *Turing pattern model*, after reviewing Turing's RD theory briefly. In Section 5, we introduce and analyse the *dynamic morphogen gradient model*. We discuss the possible benefits of community effects to embryonic patterning, and other implications of our results in Section 6. In the appendix, we provide detailed model of the community effect in *Xenopus* with a self-activating feedback loop involving two species. We also list all parameter sets used in this paper. And finally we discuss briefly about numerical error in simulations of PDE models.

2. Computational modelling of biological systems

Inspired by control theory, the dynamics of gene regulatory systems are formalized traditionally as systems of ordinary differential equations (ODE). Variables of those systems track the average numbers of molecular species over time. Under the assumption of numbers of molecules being high enough, those variables track the continuous concentrations of molecules. When space comes into play, partial differential equations (PDE) are used [50], with variables which depend both on space and time. Those continuous, deterministic approximations are useful because they allow a formal analysis of models. However, they are often intractable analytically. Computational methods have thus been used to manipulate the equations symbolically and integrate numerically.

While deterministic ODE- and PDE-based models have been widely used to analyse the *average* behaviour of biological systems, they mask the stochastic noise that has a significant influence on biochemical reactions [32, 44]. One approach to take account of stochasticity are stochastic differential equations, another are Continuous Time Markov Chains (CTMC). Because CTMCs usually involve infinitely many states, their representation by ODEs (so called *Chemical Master Equation*) is not always useful.

Rule-based languages are a more practical (finite) approach, inspired by traditional representations of chemical reactions extended with programming concepts to cope with combinatorial explosion [12], model abstractions and refinements. New challenging questions thus arise in programming language theory. For instance, recent work investigates modelling languages in terms of not only their relative expressiveness [7, 30], their quantitative semantics (deterministic and/or stochastic) [6, 8], but also efficient implementations of related stochastic simulators [11]. Biological applications have also highlighted specific design issues not encountered so far in existing programming languages. The representation of space is of particular interest, and in our opinion, React(C) [30] has good capabilities in this area. In this paper, we use a simplified version of React(C) tailored for the specific models we consider. We now formally define this language and its stochastic semantics.

2.1. Molecules and solutions

We consider abstract biological "agents", that we call *molecules*, attributed with an integer value that represents the location of the molecule in a discretized one dimensional space. For instance, we denote by $u(3)$ a molecule of species u located at 3. Formally, given a set $Specs = \{u, v, w, s, \ldots\}$ of *species names*, a molecule $m \in Mols$ is thus a pair of a molecule species and an integer:

$$molecules \qquad m \in Mols = Specs \times \mathbb{N}$$

Molecular systems S, named *solutions*, are defined as multisets of molecules. We call *Sols* the set of solutions:

$$solutions \qquad S \in Sols = Mols \to \mathbb{N}$$

We only consider finite solutions, that is solutions S such that there are finitely many molecules m such that $S(m) > 0$. We define the multiset inclusion $\subseteq$, subtraction $\ominus$ and addition $\oplus$ as follows:

$$S \subseteq S' \text{ iff } \forall m \in Mols.S(m) \leq S'(m) \qquad S \ominus S'(m) = S(m) - S'(m) \qquad S \oplus S'(m) = S(m) + S'(m)$$

Note that $S \ominus S'$ is defined only if $S' \subseteq S$. As an example, the solution composed of two molecules of species u at location 3 and one molecule v at location 5 is written:

$$u(3) \oplus u(3) \oplus v(5)$$

A solution defines a state of our dynamic system.

2.2. Reactions and rules

We start with the definition of a language of *ground reactions*, then we define a language of *rules* (or reaction schemas). Rules abstract locations with quantified variables over finite domains. The idea is to design compact models by means of rules from which finitely many ground reactions are generated.

    Reactions can be seen as rewriting reduction relations over solutions. In addition, they are equipped with a real number, their *stochastic rate*, that quantifies how fast and how likely a reaction may apply concurrently with other reactions. Formally, a reaction is a triplet made of two solutions (i.e. the reactants and products) and a real number:

$$reactions \qquad R \in Reacts = Sols \times \mathbb{R} \times Sols$$

In the following, reactions (S_1, r, S_2), for $r \in \mathbb{R}$, are written $S_1 \xrightarrow{r} S_2$. For instance, the reaction

$$u(1) \oplus u(1) \quad \xrightarrow{2.5} \quad v(1) \tag{1}$$

can represent the dimerization of two molecules u at location 1 where v symbolizes the dimer, with (an arbitrary) rate 2.5. If such reaction applies at any location, which is very likely, one should write as many reactions as possible locations. Instead, we introduce *rules* that allow writing the reaction only once with a location abstracted with a quantified variable:

$$\forall x \in I, u(x) \oplus u(x) \quad \xrightarrow{2.5} \quad v(x) \tag{2}$$

where I is some finite set of integers and assumes a set of variables $\mathcal{V} = \{x, y, z, \ldots\}$.

    Usually, reactions such as (1) assume the *Mass Action Law* for chemical kinetics, that is the actual stochastic rate of (1) is 2.5 times the number of combinations of the reactants in the current solution. We don't apply this here because we want to use different kinds of kinetics (such as *Michaelis-Menten*). So, the actual rate of (1) is 2.5 irrespective of the current solution. Kinetics laws are explicitly specified by rules. Thus, the Mass Action version of (3) is written

$$\forall x \in I, u(x) \oplus u(x) \quad \xrightarrow{2.5 \frac{\sharp u(x)(\sharp u(x) - 1)}{2}} \quad v(x) \tag{3}$$

where $\sharp u(x)$ represents the number of u at location x in the current solution. Kinetic law expressions are arbitrary real valued arithmetic expressions collected in the set *Laws* that we formalize below.

    Finally, we need to express spatial *diffusion* of molecules, which is a molecule moving from one location to an adjacent one. Provided x does not index the last position, locations of x and $x + 1$ are adjacent to each other. Thus, using not only variables, but arithmetic expressions to attribute for molecules in rules, one can write diffusion as follows:

$$\forall x \in I_1, u(x) \quad \xrightarrow{D_u \sharp u(x)} \quad u(x + 1)$$

where D_u is the diffusion coefficient of molecule u, and I_1 an appropriate integer interval of diffusion. The other diffusion direction is written

$$\forall x \in I_2, u(x+1) \xrightarrow{D_u \sharp u(x+1)} u(x)$$

In this paper, it is indeed sufficient to consider expressions of location x and $x+1$, that is the least set *LocExprs* that satisfies the following:

$$\mathbb{N} \cup \mathcal{V} \subseteq \textit{LocExprs} \qquad \text{and} \qquad x \in \mathcal{V} \Rightarrow x+1 \in \textit{LocExprs}$$

$u(x)$ and $u(x+1)$ are called a *molecule schemas* and solutions of molecule schemas are *solution schemas*:

$$\begin{array}{ll} \textit{molecule schemas} & \mathbf{m} \in \textit{SMols} = \textit{Specs} \times \textit{LocExprs} \\ \textit{solution schemas} & \mathbf{S} \in \textit{SSols} = \textit{SMols} \to \mathbb{N} \end{array}$$

The set of rules *Rules* is formally defined by

$$\mathbf{R} \in \textit{Rules} = \textit{Domains} \times \textit{SSols} \times \textit{Laws} \times \textit{SSols}$$

where *Laws* is the least set of law expressions ℓ satisfying the following:

$$\mathbb{R} \cup \mathcal{V} \cup \{\sharp\mathbf{m}\} \in \textit{Laws} \qquad \text{and} \qquad \ell_1, \ldots, \ell_n \in \textit{Laws}, ar(f) = n \Rightarrow f(\ell_1, \ldots, \ell_n) \in \textit{Laws}$$

that assumes a set $\mathcal{F} = \{f, g, \ldots\}$ of function symbols with arities $ar(f) \geq 0$. As functions, we will make use of $+, -, \times$ and $/$ as well as exponentiation and *max*. The set *Domains* defines *variable domains* that is, for each variable, the integer values that the variable can be assigned to:

$$\textit{variable domains} \qquad \Delta \in \textit{Domains} = \mathcal{V} \to \mathcal{P}_{\textit{fin}}(\mathbb{N})$$

We denote $\textit{vars}(\Delta) = \{x \mid \Delta(x) \neq \emptyset\}$ and we only consider domains Δ such that $\textit{vars}(\Delta)$ is finite. Suppose that $\textit{vars}(\Delta) = \{x_1, \ldots, x_n\}$ and $\Delta = \{x_1 \mapsto I_1, \ldots, x_n \mapsto I_n\}$, then we denote Δ as $\forall x_1 \in I_1, \ldots, \forall x_n \in I_n$. Also, for the sake of readability, a rule $(\Delta, \mathbf{S}_1, \ell, \mathbf{S}_2)$ is denoted $\Delta, \mathbf{S}_1 \xrightarrow{\ell} \mathbf{S}_2$ and we assume that variables occurring in $\mathbf{S}_1, \mathbf{S}_2$ and ℓ also occur in $\textit{vars}(\Delta)$.

2.3. Stochastic semantics

We first define the semantics of ground reactions, then we formalize the interpretation of rules in terms of reactions thanks to *variable assignments*, and finally define the stochastic semantics of rules.

The semantics of a finite set of reactions $\mathcal{R} \subset \textit{Reacts}$ is given in terms of Continuous Time Markov Chains (CTMC) where states are solutions. The state transition relation $\to_{\mathcal{R}} \subset \textit{Sols} \times \mathbb{R} \times \textit{Sols}$ of a set of rules $\mathcal{R}$ is defined as follows:

$$S \xrightarrow{r}_{\mathcal{R}} S' \quad \text{iff} \quad r = \sum_{\{(r', R) \mid R = S_1 \xrightarrow{r'} S_2 \in \mathcal{R},\ S_1 \subseteq S \text{ and } S' = S \ominus S_1 \oplus S_2\}} r'$$

In other words, $S \xrightarrow{r}_{\mathcal{R}} S'$ means that there are reactions in $\mathcal{R}$ that can consume their reactants from the current state S and produce their products such that S' is the resulting solution. The rate r is the sum of

the rates of those reactions. As we said above, at the level of reactions, the computation of the rate r is independent of the current solution S, i.e. no specific kinetic law is assumed.

The interpretation of rules as reactions is based on variable assignments:

$$\textit{variable assignments} \qquad \alpha \in \textit{Assigns} = \mathcal{V} \to \mathbb{N}$$

We say that an assignment α satisfies a domain Δ, written $\alpha \in \Delta$, if $\alpha(x) \in \Delta(x)$ for any variable $x \in \textit{vars}(\Delta)$. Given an assignment α, the interpretation $[\![\cdot]\!]_\alpha$ of location expressions is straightforward:

$$[\![i]\!]_\alpha = i \text{ for } i \in \mathbb{N}, \qquad [\![x]\!]_\alpha = \alpha(x), \qquad [\![x+1]\!]_\alpha = \alpha(x) + 1$$

To interpret law expressions, in addition to an assignment, one needs an interpretation $[\![f]\!]$ of function symbols (for instance, $\times$ is interpreted as multiplication, etc.) and a solution. Indeed, solutions are needed to interpret the counting operator $\sharp$ of molecules. This interpretation $[\![\cdot]\!]_{\alpha,S}$ is defined as

$$[\![r]\!]_{\alpha,S} = r \text{ for } r \in \mathbb{R}, \quad [\![x]\!]_{\alpha,S} = \alpha(x), \quad [\![\sharp\mathbf{m}]\!]_{\alpha,S} = S([\![\mathbf{m}]\!]_\alpha),$$

$$[\![f(\ell_1, \ldots, \ell_n)]\!]_{\alpha,S} = [\![f]\!]([\![\ell_1]\!]_{\alpha,S}, \ldots, [\![\ell_n]\!]_{\alpha,S})$$

It follows the interpretation of molecule schemas and solutions schemas:

$$[\![u(e)]\!]_\alpha = u([\![e]\!]_\alpha) \qquad [\![\mathbf{S}]\!]_\alpha = \begin{cases} \textit{Mols} \to \mathbb{N} \\ m \mapsto \sum_{\{\mathbf{m} | [\![\mathbf{m}]\!]_\alpha = m\}} \mathbf{S}(\mathbf{m}) \end{cases}$$

Finally, the interpretation $[\![\cdot]\!]_S$ of a rule $\mathbf{R}$ for a given solution S, is the set of reactions that result from all assignments satisfying the domain of $\mathbf{R}$:

$$[\![\Delta, \mathbf{S}_1 \xrightarrow{\ell} \mathbf{S}_2]\!]_S = \{[\![\mathbf{S}_1]\!]_\alpha \xrightarrow{[\![\ell]\!]_{\alpha,S}} [\![\mathbf{S}_2]\!]_\alpha \mid \alpha \in \Delta\}$$

The stochastic semantics of a finite set of rules $\mathcal{R} \subset \textit{Rules}$ is given by the stochastic transition relation $\to_{\mathcal{R}} \subset \textit{Sols} \times \mathbb{R} \times \textit{Sols}$ that induces a CTMC and is defined by

$$S \xrightarrow{r}_{\mathcal{R}} S' \quad \text{iff} \quad S \xrightarrow{r}_\mathcal{R} S' \text{ where } \mathcal{R} = \cup_{\{\mathbf{R} \in \mathcal{R}\}} [\![\mathbf{R}]\!]_S.$$

Thus, at each reduction step, the rule-based model is interpreted as a "reaction-based" model that defines the possible transitions. Note that, from one step to another, the interpreted reactions only differ by their stochastic rate.

3.   Minimal model of a community effect in one dimension

In this section, we present a *minimal model* of a community effect in one dimensional space. Its GRN is minimal, in the sense that the positive feedback loop for a community effect loop is the smallest possible, with the single species u – illustrated in Fig.1A on p.2. Previous work considered a well-mixed system of cells, in which a factor could diffuse from one cell to any other with equal probability [46], or didn't include multiple cells explicitly at all [5]. However, in reality cells are arranged in a certain *geometry in space*. Therefore, the probability of cell-to-cell communication depends on the distance between
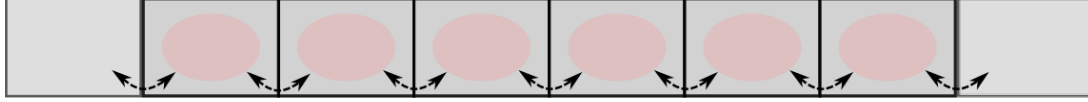
Figure 2.   Diffusion model scheme: row of N_a active cells, flanked by $2 \times N_p$ passive cells. Here $N_a = 6$, $N_p = 1$.

cells. The following model thus renders molecular diffusion in a more realistic way, considering a one-dimensional row of cells, in which we distinguish *active cells* (implementing the community effect GRN) flanked by *passive cells* (that merely support diffusion). With the restriction that molecules travel over longer distances by a random walk between immediate neighbours, we can see how a community effect behaves in space.

The contributions of this section are the following. As stochastic simulation shows, in this minimal model, a *community effect spreads in space*, conquering the entire population of active cells. Interestingly, without passive cells, we observe that the sustained activation becomes independent of the number of active cells. Finally, we formally *prove the necessity of passive cells for a community effect* in this model.

3.1.   Informal model of one-dimensional diffusion

Our diffusion model scheme is shown in Fig.2. Cells are arranged in a *row* of finite length. The row contains a domain of *active cells* in the middle, which is flanked by *passive cells*. Only active cells can respond to the inducing signal and produce the diffusible factors for a community effect. Passive cells do not respond to inducing signals and merely occupy the space on both sides of the active cells. N_p and N_a henceforth refer to the numbers of passive and active cells, respectively. Molecular *diffusion*, i.e. the exchange of a diffusible factor, occurs stepwise between adjacent positions. Molecules can move throughout the entire row by random walk.
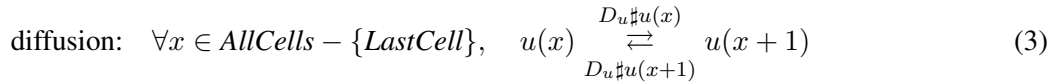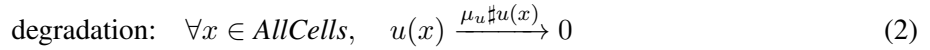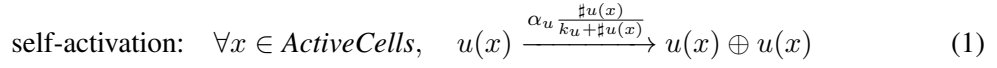
In our model, passive cells are required to obtain certain aspects of community effects. Without passive cells, those effects do not occur, regardless of initial conditions. We prove this later in Section 3.6 after formally defining the model.

The common structure of GRNs for community effects [14] is a self-sustaining feedback loop with diffusible factors. The GRN of our active cells illustrated in Fig.1A implements such a feedback loop in a minimal fashion: the diffusible factor u directly enhances its own production. Cells in the system are induced by a transient signal, applied either locally or globally. In our model, the inducing signal is represented by a number of diffusible factor molecules, which are deposited at some cells at time zero.

3.2.   Rule-based formalization of the minimal model

The minimal model is presented in Fig.3 in the rule-based language defined in the previous section.

Rule (1) describes u's self-activation, which is the GRN of this model. This rule applies to u at positions from the set of active cells, which is denoted $ActiveCells$. It creates an additional u molecule at the same position. The rate of the event described by this rule is computed from the expression $\alpha_u \frac{\sharp u(x)}{k_u + \sharp u(x)}$ which defines a Michaelis-Menten kinetic law. α_u and k_u are the rate parameter and the Michaelis constant, which are given regardless of the position. Michaelis-Menten law is a standard way

$$\text{self-activation:} \quad \forall x \in \textit{ActiveCells}, \quad u(x) \xrightarrow{\alpha_u \frac{\sharp u(x)}{k_u + \sharp u(x)}} u(x) \oplus u(x) \tag{1}$$

$$\text{degradation:} \quad \forall x \in \textit{AllCells}, \quad u(x) \xrightarrow{\mu_u \sharp u(x)} 0 \tag{2}$$

$$\text{diffusion:} \quad \forall x \in \textit{AllCells} - \{\textit{LastCell}\}, \quad u(x) \underset{D_u \sharp u(x+1)}{\overset{D_u \sharp u(x)}{\rightleftarrows}} u(x+1) \tag{3}$$

where $\textit{ActiveCells} = \{1 + N_p, \ldots, N_p + N_a\}$, $\textit{AllCells} = \{1, \ldots, \textit{LastCell}\}$, and $\textit{LastCell} = 2N_p + N_a$.

Figure 3.   Rules of the minimal model.

to model saturating reactions in biochemistry. We need the self-activation to be saturating, otherwise $\sharp u(x)$ may grow indefinitely.

Rule (2) describes degradation of u, which applies throughout the entire row, with mass action kinetics. The last rule (3) defines (reflexive) diffusion with diffusion rate D_u (which is also constant throughout the space).

The diffusion model described by these rules (Fig.3) is a standard approach of introducing space to models defined by chemical rules [40, 49]. It assumes that all cells have the same volume and geometry, and that the distance between all neighbours is the same. The diffusion rate D_u can be related to Fick's diffusion coefficient, which is invariant to geometry and discretization of space. The relationship can be derived by applying the finite volume method [18] to Fick's second law, which describes diffusion in continuous space. We refer the reader to [4] for the derivation.

3.3.   Simulation method

In this paper, we performed stochastic simulations of rule-based models using a variant of the Gibson and Bruck method [20] (itself being an optimization of the Gillespie algorithm [21]) which was implemented in the Scala language[1].

It is important to note that intercellular diffusion is modelled as two chemical reactions, one per direction. Each reaction transforms a molecule in one cell, into a molecule of the same type in the other cell (rule (3)). Typically, in order to accurately model diffusion using this approach, the rate of such reactions must be made significantly higher than the rates of real chemical interactions [4]. This, however, massively increases the simulation's computational cost, as the vast majority of generated events represent molecules moving back and forth between cells. We have chosen to reduce the diffusion rates of the stochastic systems to make the computations tractable; this is equivalent to making the space discretization coarser. Nonetheless the discretization is kept fine enough to demonstrate the phenomena we are interested in.
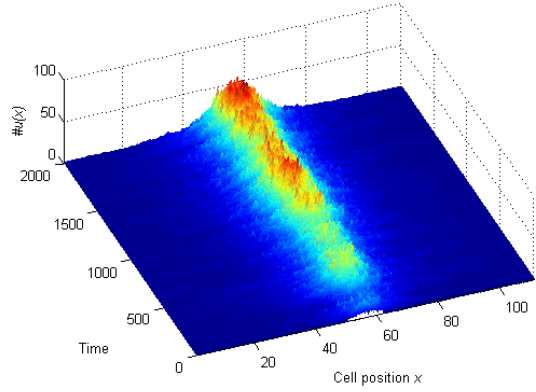
The parameter values for the simulations of the minimal model are listed in Table 4 on p.43 (where applicable), with the exception of D_u, which is set to 1.6 for this case. The choice of the parameters is made in such a way that the model can be simulated in a reasonable time and it shows the required properties.
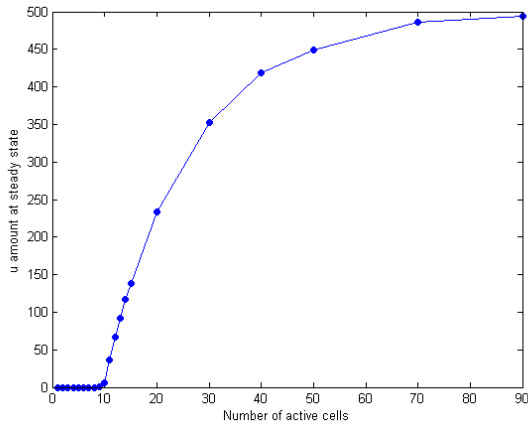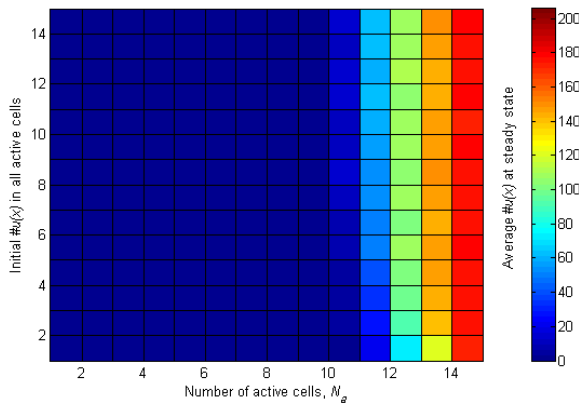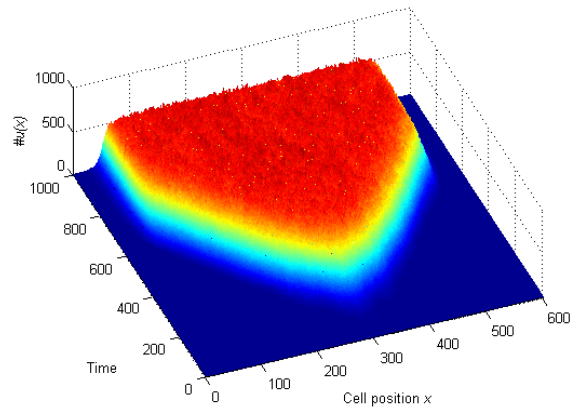
---

[1] http://www.scala-lang.org/
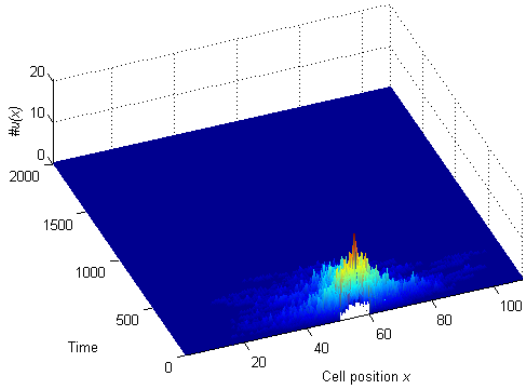
A. $N_a = 9$, $N_p = 50$.

B. $N_a = 11$, $N_p = 50$.

C. Number of u molecules in the middle cell at $t = 2000$, averaged across 100 runs. **D.** $N_a = 500$, $N_p = 50$. Initial induction localized at $x = 300$.

E. Effect of the initial u amount on the threshold. $N_p = 50$.
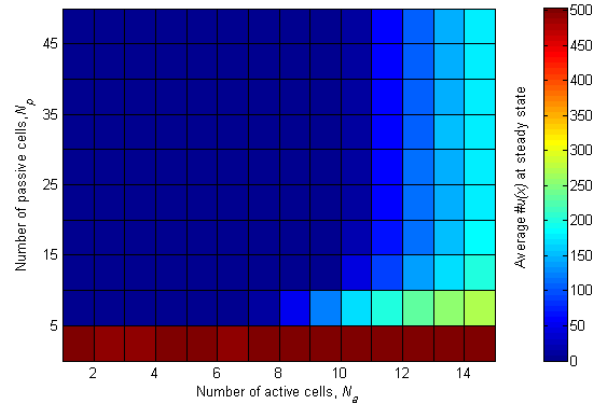
F. Effect of N_p on the threshold. Initial $\sharp u(x) = 10$.

Figure 4.   Stochastic simulations of the minimal model of a community effect. (**A, B, D**): single simulation runs. (**C, E, F**): summary of many simulations.

3.4.   Stochastic simulations show a community effect with unlimited spread

We show typical simulation results in Figs.4A and B. The plots show the dynamics of u across the cell row over time with different numbers of active cells, below ($N_a = 9$, A) and above ($N_a = 11$, B) the community effect's activation threshold ($N_a = 10$). Both simulations start from an initial induction with $\sharp u(x) = 10$ molecules per active cell at time zero. The system can reach either a state with $\sharp u(x) = 0$ in all cells, the zero state, or enter the self-sustained activation state, where $\sharp u(x)$ stabilizes above zero. Since u can degrade at any moment, there is a non-zero probability for the system to collapse to the zero state at all time points. However, this probability becomes negligibly low for the cases studied here once the system reaches self-sustaining activation; therefore we call this outcome a steady state.

- **Fig.4A**. $N_a = 9$: the system is unable to sustain activation. The u molecules for induction at $t = 0$ are visible as a peak in all active cells. However, they quickly dissipate. This is a typical outcome of the simulations. Only 1 out of 100 simulations showed any residual activity at $t = 2000$.

- **Fig.4B**. $N_a = 11$: the system enters sustained activation in all 100 simulation runs.

To check if the system has reached self-sustaining activation, we simulate until $t > 2000$. This is long enough for the initial induction to degrade. Any residual activity in the system must be due to self-sustaining activation. We used this criterion to investigate systematically the activation threshold number of active cells:

- **Fig.4C** summarizes simulation results (2100 runs in total). Each point in this plot is the mean u amount at $t = 2000$ in the cell in the middle of the community, for 100 simulations with the given N_a. The threshold value of $N_a = 10$ clearly appears. The system enters sustained activation only when N_a is above this value. As N_a increases, the steady state amount of u approaches 500, its theoretical maximum when the feedback loop saturates.

Next, we examined the dynamics of the system in space and made an interesting observation.

- **Fig.4D**. With $N_a = 500$ that is well above the threshold, the system enters sustained activation with high probability even when inducing only *one* active cell at the centre of the row, with $\sharp u(x) = 10$ at time zero. Note that, in Fig.4A and B, the same amount was injected into *all* active cells. Significantly, *activation spreads over the entire row of active cells* from the location of the initial induction in the middle of the row.

The probability of sustained activation varies with the total amount of the initial induction when it is low. We found that initial activation by $\sharp u(x) = 10$ molecules per cell is sufficient to exclude this variation:

- **Fig.4E**. The effect of initial u amount on activation threshold N_a. Each square in the grid displays the colour-coded $\sharp u(x)$ in the middle of the community at steady state ($t = 1000$), averaged over 25 simulations, with a given combination of N_a and initial $\sharp u(x)$ per active cell. For low initial u amounts, the threshold N_a for self-sustaining activation increases. Beyond $\sharp u(x) = 10$, the threshold stabilizes. Note that $N_p = 50$ for all simulations.

- **Fig.4F**. The effect of the number of passive cells N_p on the threshold of active cell number N_a. The lower the N_p is, the higher the chance for u to re-enter the active area after being reflected back at the system's boundary. When $N_p = 0$, this boundary effect increases and effectively eliminate N_a threshold, down to zero (see also Section 3.6). With increasing N_p, the boundary effect become negligible, and the critical N_a threshold becomes independent of N_p.

Simulations indicate that the steady-state number of u, if the system ever reaches a non-zero steady state, is independent of the location, size (width), and the amplitude of the initial induction. Similar behaviour was observed in the spaceless model of the community effect in [46]. We generally observed that, for a self-sustaining activation, the number of active cells must exceed a critical threshold, which is one of the hallmarks of a community effect. We observed that the community effect spreads in an unlimited manner in the minimal model described in this section.

3.5. Biological relevance of the minimal model

The minimal model introduced in this section is abstract, but nonetheless useful because it is simple and it reproduces the essential features of community effects. We also modelled the community effect in *Xenopus* in greater detail, using reaction rules as well as partial differential equations (PDEs). See Section A in the appendix for the description of those detailed models and discussions.

The saturating self-activation reaction abstracts a process by which a diffusible factor enhances its own production. In real cells this involves binding of a diffusible factor to the cell surface, activation of the signalling pathway and the induction of the corresponding gene. In addition, the gene that produces the diffusible factors may be induced indirectly, by a cascade of inductions of intermediate genes, as happens in the community effect of *Xenopus* muscle precursor cells [46].

The diffusion scheme with the row of active cells between two layers of passive cells reflects the *sandwich*-like experimental setup of Gurdon *et al.* [24, 25], in which an inducer-soaked bead is placed between two slices of ectoderm tissue. To be more precise, our model's row corresponds to a vertical cut through the sandwich of cells. The inner layer of the ectoderm tissue corresponds to active cells and outer-most layer to passive cells.

3.6. Passive cells are required for the cell number-dependent community effect

We conclude the analysis of the minimal model by proving that the closed system without passive cells cannot have the threshold of N_a where its steady state changes. To formally analyse the steady states of the minimal model, we use its deterministic semantics assuming molecular diffusion in *continuous space*, clearly distinct from the discretized space which we used earlier. We transform the minimal model into the following PDE:

$$\frac{\partial u}{\partial t} = \frac{\alpha_u u}{k_u + u} - \mu_u u + \lambda_u \Delta u \tag{4}$$

where $u(x, t)$ is the average $\sharp u$ at space coordinate x and time t, which remain always implicit in PDEs. λ_u is a diffusion coefficient, which can be derived from D_u and the spatial configuration. The precise expression for it is not important for the following analysis. There are no passive cells in this PDE model because the activation term is applicable everywhere. All the parameters are assumed to be positive.

This model is using the Laplacian Δu to describe the diffusion of u, and therefore is applicable to n-dimensional case. This is the standard way to model diffusion in PDEs [17]. $\lambda_u \Delta u(x,t)$ defines the rate of u diffusing to point x at time t. Let $\Omega \subset \mathbb{R}^n$ be the spatial domain of the system, a compact, connected set, and $\overline{\Omega}$ its boundary. We consider a closed system with no passive cells, so one of the boundary conditions is

$$\nabla u(x,t) \cdot n(x) = 0, \forall x \in \overline{\Omega}, \forall t \geq 0 \tag{5}$$

where $n(x)$ is an outward unit normal of the boundary $\overline{\Omega}$ and ∇ is the gradient operator. This states that the flux of u across the boundary is zero at all points. Another boundary condition is the initial induction

$$u(x,0) = f(x) \geq 0, \forall x \in \Omega \tag{6}$$

Let $g(u) = \frac{\alpha_u u}{k_u + u} - \mu_u u$ and $u_c = \frac{\alpha_u - \mu_u k_u}{\mu_u}$. Note that $g(u)$ has two zero points: $g(0) = 0$ and $g(u_c) = 0$, and $g(u) > 0, \forall u : 0 < u < u_c$.

First, note that $\forall t \geq 0, \forall x \in \Omega, u(x,t) \geq 0$. It follows from the initial condition being nonnegative and $g(0) = 0$.

Second, note that if $f(x) = 0$ for some x, but $\int_\Omega f(x) \mathrm{d}\Omega > 0$ and $u_c > 0$, then $\forall t > 0, u(x,t) > 0$, because diffusion distributes concentration across space instantly. To show this, we consider initial induction to be a δ function and consider the heat equation $\frac{\partial u}{\partial t} = \lambda_u \Delta u$. It has an analytical fundamental solution which is positive for $t > 0$, but the solution of (4) is always greater or equal to the solution of the heat equation, because $g(u) \geq 0$.

Also $u(x,t)$ is continuous $\forall t > 0, \forall x \in \Omega$. So, without loss of generality, we consider the initial induction $f(x)$ to be continuous, and if $u_c > 0$ then also $f(x) > 0$.

We will introduce simple lower and upper bounds on the solutions, which can be easily analysed.

Lemma 3.1. (Lower bound)
If $u_c > 0$, then $\forall x \in \Omega, \forall t \geq 0, u(x,t) \geq \underline{u}(x,t)$, where $\underline{u}$ is the solution of (4)-(5) with initial condition $\underline{u}(x,0) = \min_\Omega f(x)/2$.

First, note that $\underline{u}$ is homogeneous in space at all times. The initial condition is homogeneous, so there is no diffusion inside the system; also by (5) there is no diffusion through the borders. So we can write $\underline{u}(x,t) = \underline{u}(t)$, which is the solution of the ODE

$$\begin{aligned} \frac{\mathrm{d}\underline{u}}{\mathrm{d}t} &= g(\underline{u}) \\ \underline{u}(0) &= \min_\Omega f(x)/2 \end{aligned} \tag{7}$$

The initial condition of $\underline{u}$ is chosen in this way to guarantee $0 < \underline{u}(0) < u(x,0), \forall x \in \Omega$.

The idea of the proof can be summarized as follows. We consider a point x_* where u crosses $\underline{u}$ the first time. Then it must be the minimum of u, and thus u will grow there at least as fast as $\underline{u}$, because u cannot diffuse out of the minimum point. Then it cannot go below $\underline{u}$ at this point.

Proof:
Suppose that $u(x,t) < \underline{u}(t)$ somewhere. Note that $u(x,0) > \underline{u}(0), \forall x \in \Omega$. Let $t_* = \inf\{t \geq 0 : \exists x \in \Omega : u(x,t) < \underline{u}(t)\}$. Note that

$$u(x,t_*) \geq \underline{u}(t_*), \forall x \in \Omega, \tag{8}$$

because the solutions are continuous. It follows that

$$\exists x_* : u(x_*, t_*) = \underline{u}(t_*), \tag{9}$$

and for any sufficiently small $\epsilon > 0 : u(x_*, t_* + \epsilon) < \underline{u}(t_* + \epsilon)$. Subtracting $u(x_*, t_*)$ from the left side and $\underline{u}(t_*)$ from the right side of the last inequality, dividing by ϵ, and taking $\lim_{\epsilon \to 0}$, we get

$$\frac{\partial u}{\partial t}(x_*, t_*) < \frac{\mathrm{d}\underline{u}}{\mathrm{d}t}(t_*) \tag{10}$$

From (8) and (9), $u(x_*, t_*) \leq u(x, t_*), \forall x \in \Omega$, i.e. $u(x_*, t_*)$ is a local minimum in space. Then $\Delta u(x_*, t_*) \geq 0$, and we have

$$\frac{\partial u}{\partial t}(x_*, t_*) = g(u(x_*, t_*)) + \Delta u(x_*, t_*) \geq g(u(x_*, t_*)) = g(\underline{u}(t_*)) = \frac{\mathrm{d}\underline{u}}{\mathrm{d}t}(t_*),$$

but this contradicts (10). □

Lemma 3.2. (Upper bound)
$\forall x \in \Omega, \forall t \geq 0, u(x, t) \leq \overline{u}(x, t)$, where $\overline{u}$ is the solution of (4)-(5) with initial condition $\overline{u}(x, 0) = \max_{\Omega} f(x) \cdot 2$.

The proof is analogous to lemma 3.1, exchanging $\max$ and $\min$ and reversing inequalities as appropriate.

Theorem 3.1. If $u_c > 0$, then $\lim_{t \to \infty} u(x, t) = u_c, \forall x \in \Omega$.

Proof:
From the lemmas we have $\underline{u}(t) \leq u(x, t) \leq \overline{u}(t), \forall t \geq 0, \forall x \in \Omega$. Since $u_c > 0$, the ODEs for $\underline{u}$ and $\overline{u}$ have an attractor at u_c, to which they will converge asymptotically starting from any point greater than zero. Therefore, $\lim_{t \to \infty} \underline{u}(t) = \lim_{t \to \infty} \overline{u}(t) = u_c$, from which follows the statement. □

Theorem 3.2. If $u_c \leq 0$, then $\lim_{t \to \infty} u(x, t) = 0, \forall x \in \Omega$.

Proof:
From the lemmas we have $0 \leq u(x, t) \leq \overline{u}(t), \forall t \geq 0, \forall x \in \Omega$. Since $u_c \leq 0$, $g(u) \leq 0$ in the ODE for $\overline{u}$, and it has an attractor at $u = 0$, to which it will converge asymptotically starting from any point greater than zero. Therefore, $\lim_{t \to \infty} \overline{u}(t) = 0$, from which follows the statement. □

From the theorems it follows that for the closed system, where all cells are active, the outcome doesn't depend on the geometry of space and the initial conditions; it depends solely on the parameters of the GRN. Therefore we cannot observe the property that there exists a threshold on the number of cells below which there is no sustained activation. Number of cells in this model corresponds to the volume of Ω. We see that for any spatial configuration, the system will either always enter the sustained activation, or, if $u_c \leq 0$, the activity will always fall to zero. For the parameters we used in the simulations, $u_c = 500$.

If we introduce the passive cells, in terms of the PDE the areas where $g(u) = -\mu_u u$, the lemmas still seem to hold, but the PDEs for $\underline{u}$ and $\overline{u}$ can no longer be reduced to ODEs, since the solutions will not be homogeneous anymore. We do not investigate the system with passive cells further, because its PDE contains discontinuous functions and thus its analysis is tedious.

4. Regulation of a community effect by a Turing pattern mechanism

The community effect spreads in an unlimited manner over the entire range of active cells in Section 3's minimal model, contradicting experimental observations [5, 16, 25]. We address the issue of unlimited spread by two different models in the remainder of this paper. In this section, we consider a system with *two interacting diffusible factors*: the self-enhancing species for a community effect, and a second species which negatively regulates it. We review this combination, which was introduced in Turing's reaction-diffusion theory [49], in Section 4.1. In Section 4.2, we show that it allows limiting community effects in space. Establishing this connection between the community effect and Turing's RD theory constitutes one of the core contributions of the present paper.

4.1. Turing's reaction-diffusion theory

Turing investigated systems with two interactive and diffusing species in general, regardless of the type of their interaction, called *reaction-diffusion systems* (RD system).

The following set of PDEs describing the dynamics of two interacting and diffusing species u and v, is the simplest deterministic model of an RD system [34]:

$$\frac{\partial u}{\partial t} = c_1\,u + c_2\,v + c_3 + \lambda_u\,\Delta u - \mu_u\,u,$$
$$\frac{\partial v}{\partial t} = c_4\,u + c_5\,v + c_6 + \lambda_v\,\Delta v - \mu_v\,v,$$

where $u(x,t)$ and $v(x,t)$ are concentrations of the two species at space point x and time t. The continuous space is of arbitrary dimension, i.e. $x \in \mathbb{R}^n$, similar to the model used in Section 3.6. The interactions between u and v are specified by the parameters $c_1, \ldots, c_6$. By taking positive or negative values, these can represent activation or inhibition, exerted by either species on itself (c_1, c_5), on the other (c_2, c_4), or basal production (c_3, c_6). We use diffusion coefficients λ_u and λ_v, and the Laplace operator Δ to define diffusion in continuous space, following Fick's second law. Finally, μ_u and μ_v are degradation rates.

Turing showed that, starting from homogeneous initial distributions of the two species in space, with small random perturbations, this system eventually converges to a steady state. It can reach six different states, depending on the values of the parameters [49].

Turing patterns are the most interesting among the six possible stable states of RD systems. In two dimensional space, they can generate a nearly endless variety of spatial patterns. For the simpler case of one-dimensional space that we consider in this work, Turing patterns correspond to *stable periodic waves*. Fig.5A shows a typical example. u and v concentrations are plotted against the space coordinate x, they oscillate in space. Note that the pattern is stable in time, i.e. it is a steady state of the system. Generally, the following *conditions* must hold on the two species of an RD system, in order for Turing patterns to appear [39]:

- One of the species, called *activator*, must have a positive effect on both itself and the other. The second species, called *inhibitor*, must have a negative effect on the activator. Fig.5B illustrates this GRN, with activator u and inhibitor v. We refer to this GRN as the *Turing pattern model*.

- The inhibitor must diffuse faster than the activator.

A. A typical Turing pattern in one dimensional space.　　**B.** GRN of Turing pattern model.
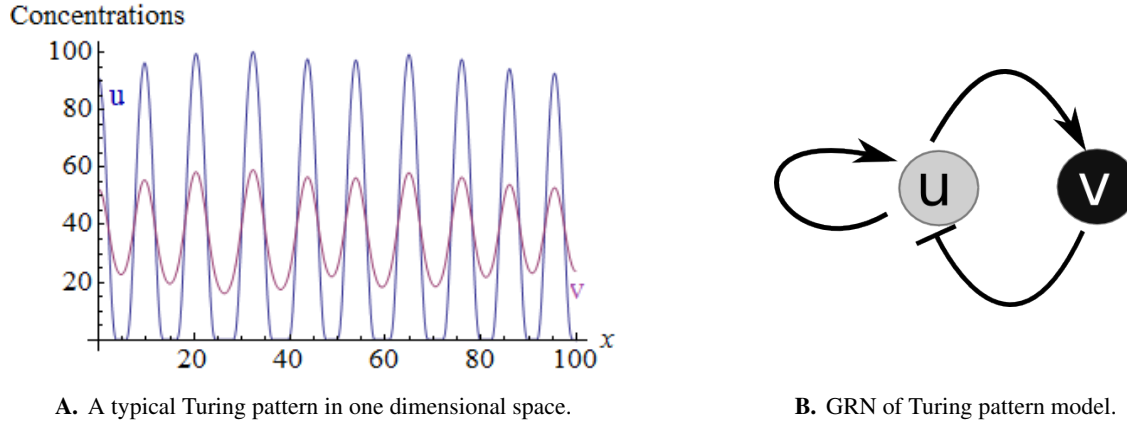
Figure 5.　A Turing pattern and the corresponding GRN.

For real biological systems, it is hard to prove experimentally that these conditions are fulfilled. However, RD systems explain many observed patterns. One well-studied case is the formation of the oral ectoderm in sea urchin embryos. There is significant evidence [16] that its area is established by an RD system with Nodal (activator) and Lefty (inhibitor). In this system, a community effect is suggested to occur with Nodal as diffusible factor for cell-to-cell communication [5]. Thus, it may not seem too speculative to expect some connection between the RD system and the community effect.

One of our important insights is that *Turing patterns can explain the spatial restriction of the community effect* in our one-dimensional model. We will show that while in the minimal model for a community effect, u's activity spreads over the entire available space, embedding the minimal model into the Turing pattern model confines u's activity to a single wave of limited width.

4.2. Stochastic simulation of Turing pattern model shows limited spread of the community effect

We present a stochastic model of the community effect as a list of reaction rules in Fig.6. It covers *two diffusible species* in a one-dimensional row of cells. The diffusible activator at the xth position is $u(x)$, and the diffusible inhibitor is $v(x)$. Note that the rules (1)-(3) are equal to those of the minimal model of a community effect with the diffusible factor u (Fig.3 on p.10). The behaviour of the second diffusible factor v is described by four additional rules. Rules (4) and (5) represent activation of the inhibitor and inhibition of the activator required for Turing patterns, with Michaelis-Menten kinetic laws. Rules (6) and (7) define degradation and diffusion of v, similar to that of u.

Figs.7A,B show typical stochastic simulation results of the system, using the same visualization as in Fig.4. The parameter values for these simulations are given in Table 4 on p.43. The system is very sensitive to the choice of the parameters, and they were tuned in a way that allows demonstration of the patterns while keeping simulation time reasonable. Only u is shown, as the amount of v is strongly correlated with it, see Fig.5A. As with the minimal model, the system is initialized with ten u molecules in some cells at time zero, which represent the transient signal that starts the differentiation in biological systems. The main difference between the two cases we show originates from *how many active cells* are initialized with u at time zero: either all of them (A), or just one (B). *The community effect is present in*

$$\text{self-activation of } u: \quad \forall x \in \textit{ActiveCells}, \quad u(x) \xrightarrow{\alpha_u \frac{\sharp u(x)}{k_u + \sharp u(x)}} u(x) \oplus u(x) \tag{1}$$

$$\text{degradation of } u: \quad \forall x \in \textit{AllCells}, \quad u(x) \xrightarrow{\mu_u \sharp u(x)} 0 \tag{2}$$

$$\text{diffusion of } u: \quad \forall x \in \textit{AllCells} - \{\textit{LastCell}\}, \quad u(x) \underset{D_u \sharp u(x+1)}{\overset{D_u \sharp u(x)}{\rightleftarrows}} u(x+1) \tag{3}$$

$$\text{activation of } v: \quad \forall x \in \textit{ActiveCells}, \quad u(x) \xrightarrow{\alpha_v \frac{\sharp u(x)}{k_v + \sharp u(x)}} u(x) \oplus v(x) \tag{4}$$

$$\text{inhibition of } u: \quad \forall x \in \textit{ActiveCells}, \quad u(x) \oplus v(x) \xrightarrow{\alpha_i \sharp v(x) \frac{\sharp u(x)}{k_i + \sharp u(x)}} v(x) \tag{5}$$

$$\text{degradation of } v: \quad \forall x \in \textit{AllCells}, \quad v(x) \xrightarrow{\mu_v \sharp v(x)} 0 \tag{6}$$

$$\text{diffusion of } v: \quad \forall x \in \textit{AllCells} - \{\textit{LastCell}\}, \quad v(x) \underset{D_v \sharp v(x+1)}{\overset{D_v \sharp v(x)}{\rightleftarrows}} v(x+1) \tag{7}$$

where $\textit{ActiveCells} = \{1 + N_p, \ldots, N_p + N_a\}, \textit{AllCells} = \{1, \ldots, \textit{LastCell}\}, \textit{LastCell} = 2N_p + N_a$.

Figure 6.   Rules of Turing pattern model.

either case: with $N_a = 5$, 13 out of 100 simulations show activity at $t = 1000$, while with $N_a = 10$ all 100 simulations enter sustained activation. Thus the critical number for the used parameter set is between 5 and 10.

With homogeneous initial conditions in all active cells (Fig.7A), the system exhibits a one-dimensional Turing pattern, i.e., stable periodic waves of high u levels over time. This occurs because the stochastic noise in $\sharp u$ and $\sharp v$ introduces an initial random asymmetry, which is later amplified by the system to larger scales. Note that this is impossible in the deterministic model, where starting from a homogeneous initial condition will always give a homogeneous solution.

The following observation is essential: if the initial signal is localized to a narrower area than the width of one wave of the pattern (e.g., a single cell as in Fig.7B), the sustained activation also remains localized. It appears as an isolated wave that is stable over time. The reader might want to refer back to Fig.4D: when no negative feedback is present, a community effect spreads over the entire range of active cells from an initial localized stimulation. In this model, the activated area is centred on the initial signal's location and it is about 20 cells wide.

Figs.7C and D show the outcomes of the pattern formation under different initial conditions. Each row is a snapshot of one simulation of a system at $t = 1000$, with $N_a = 200$ and $N_p = 50$ for all cases.

Fig.7C shows the effect of varying the initial u amount, while localizing it at the leftmost active cell. Localization at the edge of the active area is biologically plausible, since the induction signal usually comes from another tissue; also the resulting pattern is better aligned in this case. Generally, a stronger initial signal yields a more stable pattern. If present, the width of the pattern is mostly the same. In rarer cases, secondary patterns can be seen - in real embryos those could be suppressed by other, less precise mechanisms.
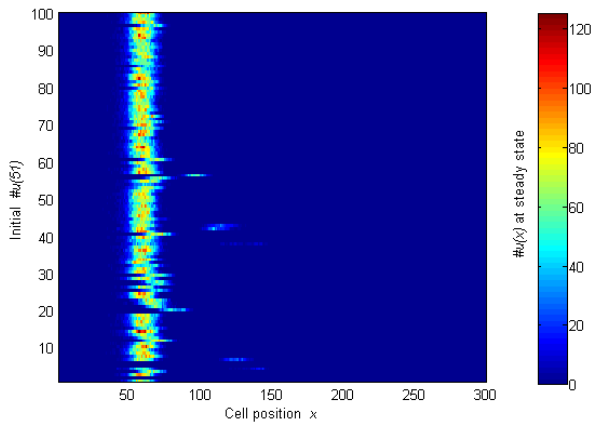
Fig.7D shows the effect of variable initial induction area on the pattern. The initial induction signal
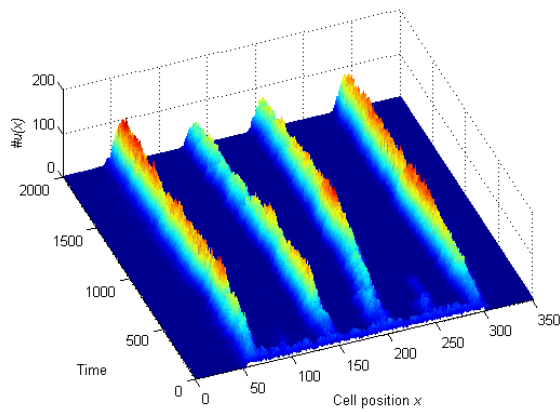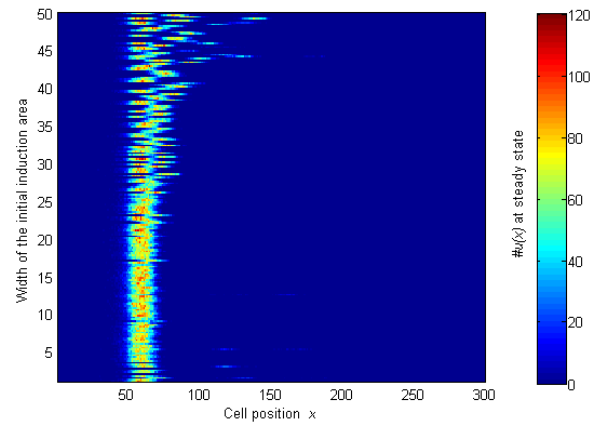
A. Initial distribution of u is uniform across all active cells.
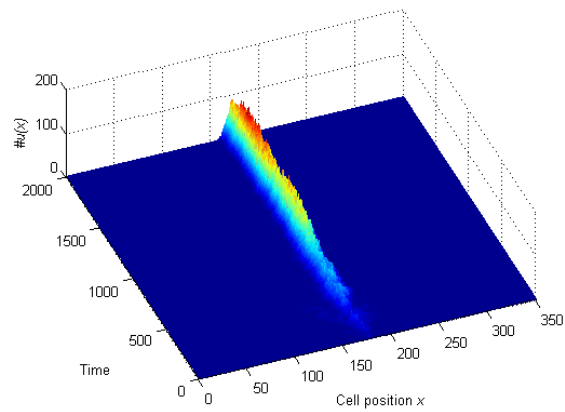
B. Initial distribution of u is localized at cell 175.
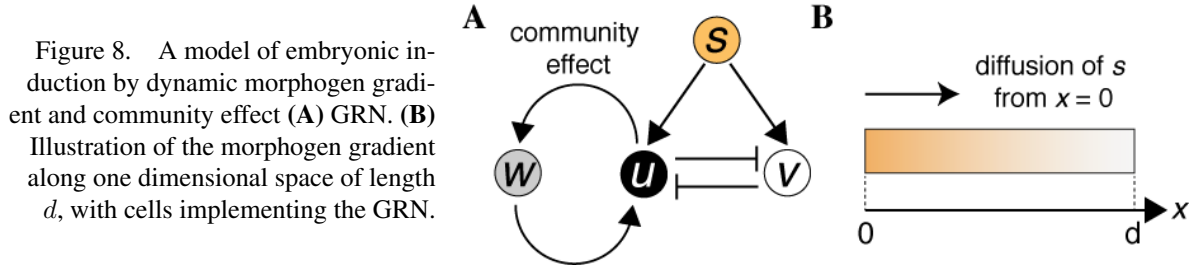

C. Varying initial u amount is localized at cell 51.

D. Fixed initial amount of $u = 10$ is spread over an area of varying width starting at near cell 51.

Figure 7. Stochastic simulations of Turing pattern model, $N_a = 200$, $N_p = 50$. **(A, B)**: single runs, number of u molecules against time and space coordinates. **(C, D)**: Impact of initial conditions at $t = 1000$.

Figure 8. A model of embryonic induction by dynamic morphogen gradient and community effect **(A)** GRN. **(B)** Illustration of the morphogen gradient along one dimensional space of length d, with cells implementing the GRN.

is 10 u molecules in every cell in the induction area, which is again located on the left of the active cells row. The width of the generated pattern does not depend on the width of the initial signal, provided that the initial signal is smaller than the pattern itself.

Therefore, the system is organized into a robust stable pattern, which is not fully specified by the initial signal. Rather, it is defined by the properties of the gene regulatory network. In summary, *the self-regulation imposed by this RD system constrains the community effect within an area with well-defined boundaries.*

5. Pattern formation by a dynamic morphogen gradient and a community effect

We now introduce a second model that prevents the spread of a community effect in one-dimensional space. It is based on a gene regulatory network controlling pattern formation in response to a morphogen gradient in *Xenopus* embryos [47]. Fig.8A shows this GRN we refer to as the *dynamic morphogen gradient model*. First, it contains a community effect loop between two mutually activating genes u and w. The latter is a diffusible factor for cell-to-cell communication, which controls the community effect. Second, it has a genetic toggle switch [19] – a well-known GRN motif – formed by the mutual repression between u and v. Third, both u and v are positively controlled by s. Fourth, and importantly, s is a morphogen - forming a gradient by diffusion from a source in space, as illustrated in Fig.8B.

We formulate the model in terms of partial differential equations (PDEs). We also provide a rule-based model of the morphogen gradient model, and use it for stochastic simulations.

The insights of this section are the following:

- We show that a pattern of gene expression emerges with a sharp boundary in space, and that the community effect orients this spatial pattern.

- We examine the system's behaviour as the morphogen saturates the entire space. Surprisingly, the established pattern is maintained in the community with uniform morphogen concentration. *This shows that positional information is not provided by morphogen concentration alone, but is encoded in the morphogen dynamics.*

- A community effect is dispensable for pattern formation. However, it is essential for the refinement of gene expression boundary and confers *robustness* to patterning processes.

5.1. PDE model

We have previously shown [47] that this network, without the positive feedback mediated by w, can convert a graded morphogen signal into a binary output (u ON v OFF, or vice versa). We now explicitly take account of molecular diffusion of s and w. We consider a gradient of the morphogen s, that arises by diffusion from the origin at position $x = 0$ into one-dimensional space. At the start of a simulation ($t = 0$), the levels of the other species u, v and w are set to zero. This corresponds to an undifferentiated state, before the morphogen s starts to induce the system.

The formulae of the model are as follows:

$$
\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{k_1\, s + k_2 \left(\frac{w^n}{w^n+1}\right)}{v^n + 1} - \mu_1\, u \\
\frac{\partial v}{\partial t} &= \frac{k_3\, s}{u^n + 1} - \mu_2\, v \\
\frac{\partial w}{\partial t} &= D_1\, \Delta\, w + k_4 \left(\frac{u^n}{u^n + 1}\right) - \mu_3\, w \\
\frac{\partial s}{\partial t} &= D_2\, \Delta\, s - \mu_4\, s,
\end{aligned}
\tag{1}
$$

where D_1 and D_2 are diffusion coefficients, k_1 to k_4 are synthesis rates and μ_1 to μ_4 are degradation rates. The synthesis term for u is inversely proportional to $v^n + 1$ to model repression of u by v. Repression of v by u is implemented in a similar way. The Hill coefficient n represents cooperativity, which introduces non-linearity to the network [47]. The boundary conditions are:

$$
s = 1 - e^{-\tau t}, \quad x = 0
\tag{2}
$$

$$
\frac{\partial s}{\partial x} = 0, \qquad x = d
\tag{3}
$$

$$
\frac{\partial w}{\partial x} = 0, \qquad x = \begin{cases} 0 \\ d\,. \end{cases}
\tag{4}
$$

We assume that the morphogen s is produced outside of our one-dimensional space, left of the origin ($x < 0$), and that its concentration at $x = 0$ is defined by (2). The speed of injection of the morphogen s into the one dimensional space varies according to the control parameter τ. For this deterministic model, the system is closed, meaning that except for the supply of s from the boundary $x = 0$, no diffusion of any molecule into or out of the system occurs. This boundary condition corresponds to the equations (3)-(4) on this page.

5.2. Numerical simulations of the PDE model

We performed numerical simulations to analyse the system's dynamics, fixing the length of one-dimensional space to $d = 100$ for all simulations. Exploring the parameter space of the system, we found three categories of *patterns of gene expression at steady state* along the one dimensional space (beyond the trivial case of $u = v = w = 0$):
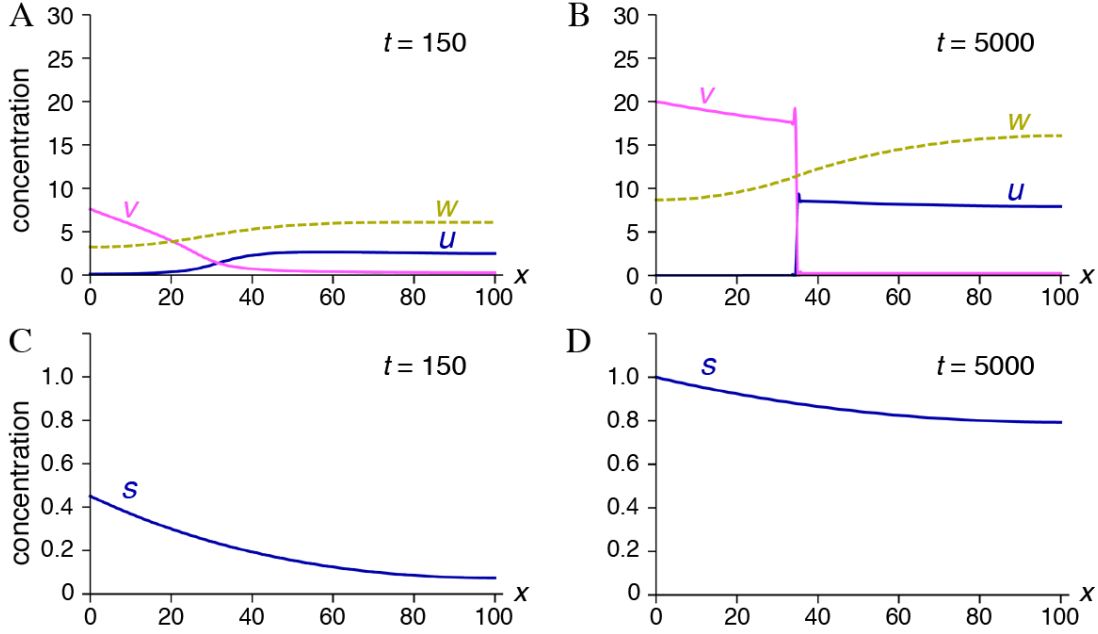
- uniform expression of u, denoted as $[u]$,

Figure 9.   Pattern formation by a dynamic morphogen gradient and a community effect.

- uniform expression of v, denoted as $[v]$,

- $[v, u]$, denoting: v on and u off close to the origin $x = 0$, and further away, v off and u on.

Fig.9A and B show concentration profiles of u, v and the diffusible molecule w over the entire range of cells, at time points $t = 150$ (A) and $t = 5000$ (B), obtained by simulation of the PDE model, with the parameter values in Table 5 on p.44. Concentration profiles of the morphogen s at the same time points are shown in Fig.9C and D.

At the stable state ($t = 1500$), simulation yields a sharp boundary of gene expression for the pattern $[v, u]$. This gene expression pattern is self-organising and is refined spatially in response to the dynamic gradient of the morphogen s. Small bumps at the $[v, u]$ boundary in (B) are artefacts of numerical integrations (see Section D in the appendix). Interestingly, once established, the $[v, u]$ boundary is maintained even when the morphogen concentration along space becomes uniform at steady state, i.e., even when the gradient is transient (Fig.10). This observation indicates that *a transient morphogen gradient is sufficient*, and that maintenance of the morphogen gradient at steady state is not a prerequisite for pattern formation in our model.

We found that the steady-state pattern is either $[u]$, $[v]$ or $[v, u]$, over a wide range of parameter values, as shown by phase diagrams of the parameter plane in Fig.11. We will later show that the pattern $[u, v]$ only appears in the absence of the community effect and positive feedback loop. Therefore it seems that the *positive feedback of community effect introduces an asymmetry to the system, and orients the spatial pattern of gene expression*. We next examined how the gene expression pattern is affected by the dynamics of the diffusible factors s (morphogen) and w (community effect factor).
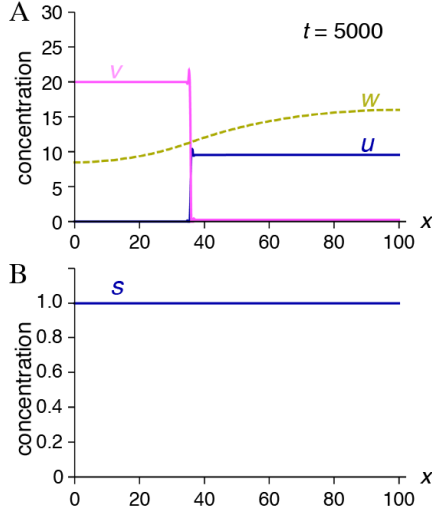
Figure 10. Pattern formation in a uniform concentration of morphogen along one dimensional space at steady state. A simulation snapshot at $t = 5000$ of the deterministic model is shown. Uniform distribution of the morphogen at steady state is established with $\mu_4 = 0$. The simulation result at a quasi-steady state ($t = 5000$) is shown. Parameter values used for the simulation are listed in Table 5 except $\mu_4 = 0$.

Disabling the cell-to-cell communication mediated by w. First, we tested the system's behaviour without cell-to-cell communication mediated by w (by setting $D_1 = 0$) – while preserving the mutual activation between w and u within the same cell ($k_2, k_4 > 0$). Numerical simulation of the PDE model shows that the $[v, u]$ pattern of gene expression can be established without cell-to-cell communication (Fig.11A). Fig.11B, C shows phase diagrams in the parameter plane (k_1, k_3) obtained from the PDE model with cell-to-cell communication (B), versus without (C). They were obtained by leaving w's diffusion rate on ($D_1 = 20$) and switching it off ($D_1 = 0$). We recall that k_1 is the rate for u's activation by the morphogen s, and k_3 is the rate of v's activation by the morphogen. The parameter pair (k_1, k_3) for the simulation in Fig.11A is marked by a blue dot in Fig.11C. As the phase diagrams show, the parameter range for a $[v, u]$ pattern widens in the presence of cell-to-cell communication.

Pattern formation is sensitive to the dynamics of morphogen gradient. We next examined how the gene expression pattern is affected by the dynamics of the morphogen s and the community effect factor w. Fig.12A shows a plot of the position of the boundary between v and u as a function of τ ($\log \tau$), the parameter controlling the supply of s to the system. The plot demonstrates that the boundary shifts as τ changes: the faster the supply of s to the system, the broader the expression domain of v is at steady state. Fig.12B is a phase diagram in the parameter plane (D_1, D_2). The diagram indicates that faster diffusion of the morphogen, i.e., larger diffusion coefficient D_2 shifts the expression boundary away from the origin. This is consistent with the above results shown in Fig.12A. In contrast, as shown in the same diagram, the boundary of $[v, u]$ is insensitive to the change of D_1, the diffusion rate of w, especially when D_2 value is low. This is also evident in the stochastic simulations, following in Section 5.4.

Pattern formation without both positive feedback and community effect. Finally we consider the system without both positive feedback and a community effect (Fig.13). This gene regulatory network without w is effectively identical to the one described in [47], with the addition of the diffusion term of s. In contrast to the network with w, this system still form the $[v, u]$ pattern at steady state, but also the previously unseen pattern $[u, v]$, in which u is on in a domain closer to the origin, while v is on further away. The pattern depends on the parameter values (Fig.14). However, a phase diagram in
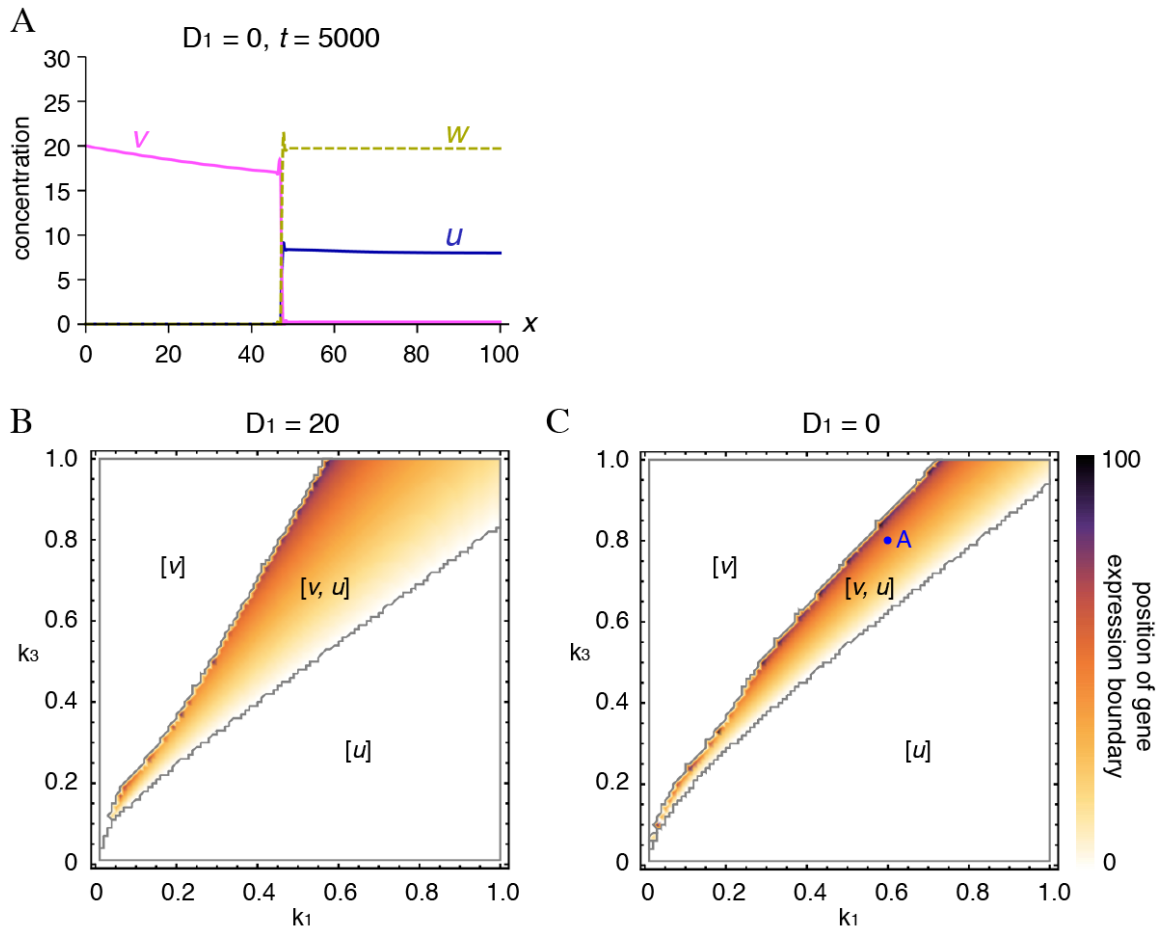
Figure 11.   Simulations of the PDE model show that cell-to-cell communication by the diffusible factor w is dispensable for pattern formation, but is essential for community effects. **(A)** Steady state profile. **(B, C)** Phase diagrams in the parameter plane (k_1, k_3) with $D_1 = 20$ (B) and $D_1 = 0$ (C). The position of the expression boundary for $[v, u]$ pattern is colour-coded as indicated. Borders between the patterns $[u]$, $[v]$ and $[v, u]$ are indicated by grey lines. Simulation parameters for (B,C) are listed in Table 5, except k_1, k_3 and D_1 as indicated in the panels.
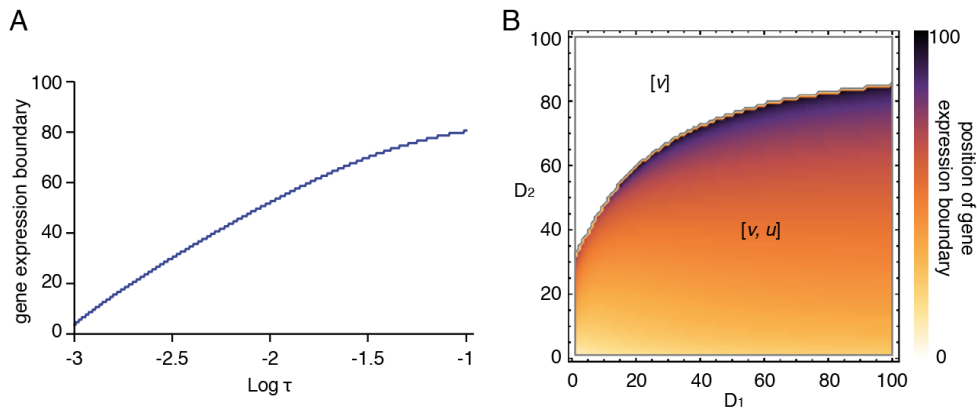
Figure 12. The position of the gene expression boundary depends on the dynamics of the morphogen s. The figures are derived from simulations of the *PDE model*. **(A)** Position of gene expression boundary is plotted as a function of τ (in log scale). τ defines the speed of morphogen supply (the larger τ is, the faster the supply and s gradient reaches a steady state). The parameter values used are the same as Fig.9, except τ. **(B)** Phase diagram in the parameter plane (D_1, D_2). The parameter values used can be found in Table 5, except D_1 and D_2. The expression boundary is relatively insensitive to the fluctuation of D_1, the diffusion rate of w. By contrast, the boundary position is sensitive to the dynamics of the morphogen s and shifts by changes in τ or D_2 value.
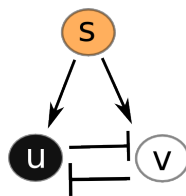

Figure 13. GRN without positive feedback and a community effect, analysed in Fig.14

the parameter plane (k_1, k_3) shows that the $[u, v]$ and $[v, u]$ *patterns only appear for narrow parameter ranges* (Fig.14A,C). In fact, stochastic simulations confirm that the system mostly ends up with no clear pattern, and a heterogeneous population of u-ON and v-ON cells at steady state (data not shown). These results highlight an important role of the community effect that *confers robustness to this patterning system*.

5.3.   Rule based model

We rewrote the PDEs model from p.21 into rules. We hereby used the discretized space of Sections 3 and 4, i.e., a row of active cells implementing the GRN, and passive cells merely supporting diffusion. This allows *stochastic simulations* of the dynamic morphogen gradient model. Due to the discretization of space and concentrations, the presence of passive cells, and stochastic effects, the precise numerical correspondence between the stochastic and the deterministic model cannot be established. However, the behaviour of the stochastic model is qualitatively similar to that of the PDE model. The stochastic model is useful to study robustness of pattern formation in the presence of noise.

The rules are given in Fig.15. Because they directly reflect our PDE model, certain rules (e.g., for inhibition) diverge from those for Turing pattern model in Section 4.2. Rule (1) defines the activation of u by s, inhibited by v. δ is the scaling coefficient which relates concentrations in the deterministic model to number of molecules in the rules; this is set to $1/100$, corresponding to 100 molecules per one unit of concentration. Rule (2) defines the activation of v by s, inhibited by u. Rule (3) defines u's activation by w with saturation, which is also inhibited by v. Rule (4) defines w's saturating activation by u. Rules (5)-(8) define the degradation of all species. Rules (9), (10) define the diffusion for w and s, following the diffusion scheme of Section 3, with one unit of space in the PDE model corresponding to one cell in the rules. Rule (11) describes the creation of s at the leftmost active cell, approximating the boundary condition of the PDEs (2) on p.21. Since we did not perform simulations with varying τ, τ is not present in the reaction rate. In the presence of only rule (11), the mean number of $\sharp s(N_p + 1)$ at time t would be $(1 - e^{-t})\delta^{-1}$, which corresponds to the boundary condition with $\tau = 1$. Because $\sharp s(N_p + 1)$ is also affected by rules (8) and (10), its actual number will be lower in general. Its maximum, for example, is about 55 instead of 100. This is the main source of numerical differences between the stochastic and the deterministic models.

5.4.   Stochastic simulation

We carried out stochastic simulations of this system, results of which are shown in Fig.16. We paid particular attention to the roles of the community effect, and observed the impact of enabling or disabling it. Disabling the community effect can be achieved in two technical ways in the rule based model, either by setting w's diffusion rate to zero, following the PDE model, or equivalently by withdrawing (9) from our rule based model in Fig.15.

Fig.16 shows snapshots at $t = 990$ (quasi-steady state) for 70 simulations with community effect (left column) and without community effect (right column). The panels in the first row show the active cell's s level, averaged over all simulations. The second and third row in Fig.16 show the number of u and w molecules, respectively, in active cells at $t = 990$. Each plot contains 70 lines, each corresponding to the outcome of one stochastic simulation. We do not show the distribution of v molecules across active cells; it is opposite to that of u (middle panels). Simulation parameters are the same as those of the PDE
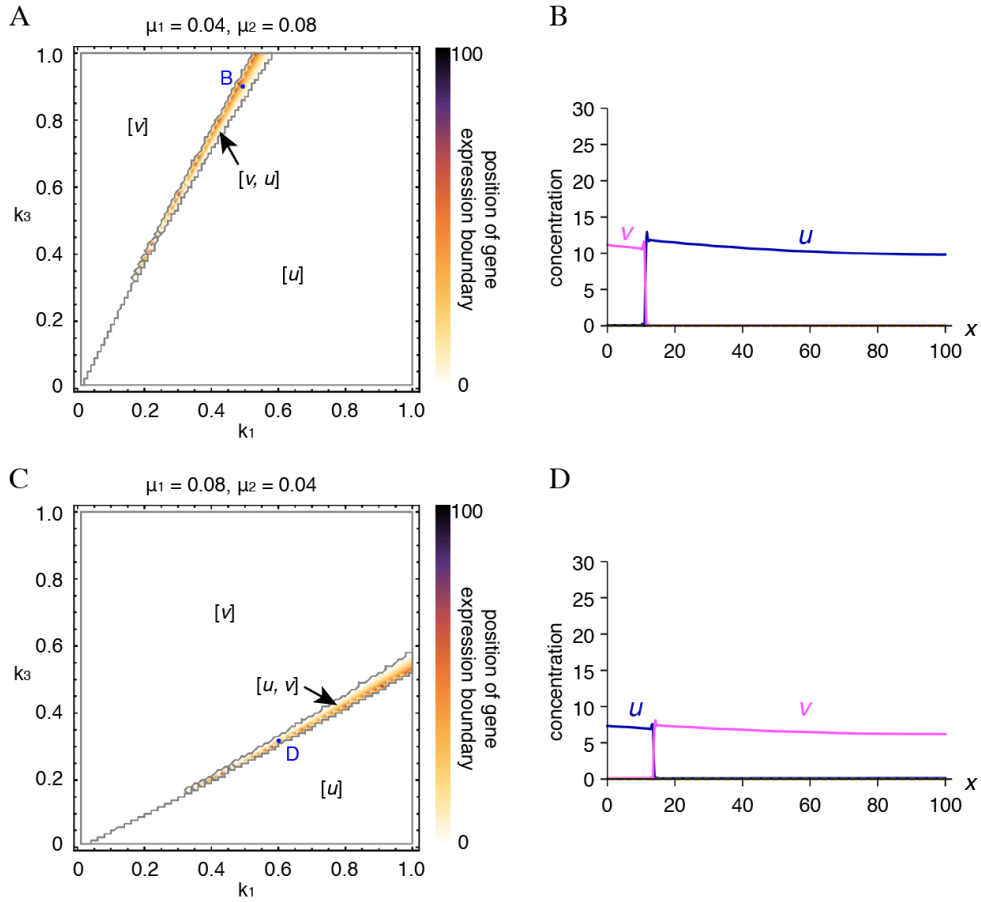
Figure 14.   Pattern formation without positive feedback and a community effect ($k_2 = k_4 = 0, D_1 = 0$) in the deterministic model. **(A)** Phase diagram in parameter plane (k_1, k_3) for $(\mu_1, \mu_2) = (0.04, 0.08)$. **(B)** Steady-state profile ($t = 5000$) for the parameter values indicated as a blue dot in (A). **(C)** Phase diagram for $(\mu_1, \mu_2) = (0.08, 0.04)$. **(D)** Steady-state profile ($t = 5000$) for the parameter values indicated as a blue dot in (C). The parameter space for $[v, u]$ (A) or $[u, v]$ pattern (B) (i.e., without w) is much smaller than in the system with positive feedback and a community effect. Without w, both $[v, u]$ and $[u, v]$ patterns are possible. Except for those indicated in the figure, the parameter values for simulations are listed in Table 5.

$$\text{activation of } u \text{ by } s: \quad \forall x \in \textit{ActiveCells}, \quad s(x) \xrightarrow{k_1 \frac{\sharp s(x)}{(\delta \sharp v(x))^n + 1}} s(x) \oplus u(x) \tag{1}$$

$$\text{activation of } v \text{ by } s: \quad \forall x \in \textit{ActiveCells}, \quad s(x) \xrightarrow{k_3 \frac{\sharp s(x)}{(\delta \sharp u(x))^n + 1}} s(x) \oplus v(x) \tag{2}$$

$$\text{activation of } u \text{ by } w: \quad \forall x \in \textit{ActiveCells}, \quad w(x) \xrightarrow{k_2 \frac{(\delta \sharp w(x))^n}{\delta((\delta \sharp v(x))^n + 1)((\delta \sharp w(x))^n + 1)}} w(x) \oplus u(x) \tag{3}$$

$$\text{activation of } w \text{ by } u: \quad \forall x \in \textit{ActiveCells}, \quad u(x) \xrightarrow{k_4 \frac{(\delta \sharp u(x))^n}{\delta((\delta \sharp u(x))^n + 1)}} u(x) \oplus w(x) \tag{4}$$

$$\text{degradation of } u: \quad \forall x \in \textit{ActiveCells}, \quad u(x) \xrightarrow{\mu_1 \sharp u(x)} 0 \tag{5}$$

$$\text{degradation of } v: \quad \forall x \in \textit{ActiveCells}, \quad v(x) \xrightarrow{\mu_2 \sharp v(x)} 0 \tag{6}$$

$$\text{degradation of } w: \quad \forall x \in \textit{AllCells}, \quad w(x) \xrightarrow{\mu_3 \sharp w(x)} 0 \tag{7}$$

$$\text{degradation of } s: \quad \forall x \in \textit{AllCells}, \quad s(x) \xrightarrow{\mu_4 \sharp s(x)} 0 \tag{8}$$

$$\text{diffusion of } w: \quad \forall x \in \textit{AllCells} - \{\textit{LastCell}\}, \quad w(x) \mathrel{\substack{D_1 \sharp w(x) \\ \rightleftarrows \\ D_1 \sharp w(x+1)}} w(x + 1) \tag{9}$$

$$\text{diffusion of } s: \quad \forall x \in \textit{AllCells} - \{\textit{LastCell}\}, \quad s(x) \mathrel{\substack{D_2 \sharp s(x) \\ \rightleftarrows \\ D_2 \sharp s(x+1)}} s(x + 1) \tag{10}$$

$$\text{external inflow of } s: \quad 0 \xrightarrow{max(\delta^{-1} - \sharp s(N_p+1), 0)} s(N_p + 1) \tag{11}$$

where $\textit{ActiveCells} = \{1 + N_p, \ldots, N_p + N_a\}$, $\textit{AllCells} = \{1, \ldots, \textit{LastCell}\}$, and $\textit{LastCell} = 2N_p + N_a$.

Figure 15.   Rules of dynamic morphogen gradient model.

model (Table 5 on p.44), except $k_2 = 0.3$ that represents u's activation by w.

The community effect being turned on (left column), the result of the stochastic simulations are similar to those of the deterministic model, except stochastic noise. In stochastic simulations, although a boundary is formed around $x = 30$, minor populations of *rogue* cells with opposite gene expression pattern are evident in both u-ON and v-ON domains around the boundary in approximately 80% of simulations.

Interestingly, simulations revealed that stochasticity near the expression boundary is more pronounced in the absence of a community effect with more *rogue* cells (Fig.16, right column). This result demonstrates that *a community effect sharpens the gene expression boundary in our model*, although rogue cells are not completely eliminated. Without the community effect the boundary is blurred. It has revealed an interesting role of the community effect for refining gene expression boundaries during patterning, in addition to its known role for a coordinated gene expression in a group of cells.

In conclusion, stochastic simulations showed that the dynamic morphogen gradient model has more robust pattern formation when including a community effect. Although a community effect is dispensable for patterning process *per se* (see Fig.11 in Section 5.2), it plays a crucial role for refining gene expression boundary.

6. Conclusion

The main focus of this paper is how diffusible factors in extracellular space control embryonic pattern formation. In particular, we have been interested in community effects, which are mediated by information exchange between cells by molecular diffusion, and their interaction with gene regulatory networks (GRNs) in animal development. What has emerged through our theoretical work is the strong *self-organising* property of the patterning by diffusible factors, when combined with particular GRNs. It has been suggested that there probably exists a small number of sub-circuits in the GRNs that function for embryogenesis [14]. Such sub-circuits include community effect loops and genetic toggle switches, consisting of a pair of mutually repressive genes.

Our previous and present work has highlighted one of the hallmarks of community effect, that is, the coordination of cellular behaviour within a cell group. Our minimal model of community effect illuminated another property of community effects, that is, its uniform unrestricted expansion throughout the cell population. This property is obviously inconvenient for pattern formation during embryogenesis. To resolve this issue, we have proposed two different models that limit the spread of a community effect within a cell population: the Turing pattern model and the dynamic morphogen gradient model. These not only presented mechanisms of regulating community effects, but also revealed several interesting insights into embryonic patterning. Our findings are summarised as follows:

1. A community effect spreads out without any negative feedback (Fig.4D).

2. With negative feedback, a community effect does not spread unrestrictedly. We show that localised induction leads to localised gene expression in our Turing pattern model (Fig.7B).

3. A modified form of genetic toggle switch combined with a community effect loop converts a dynamic morphogen gradient into a spatially-asymmetric pattern of gene expression (Fig.9).
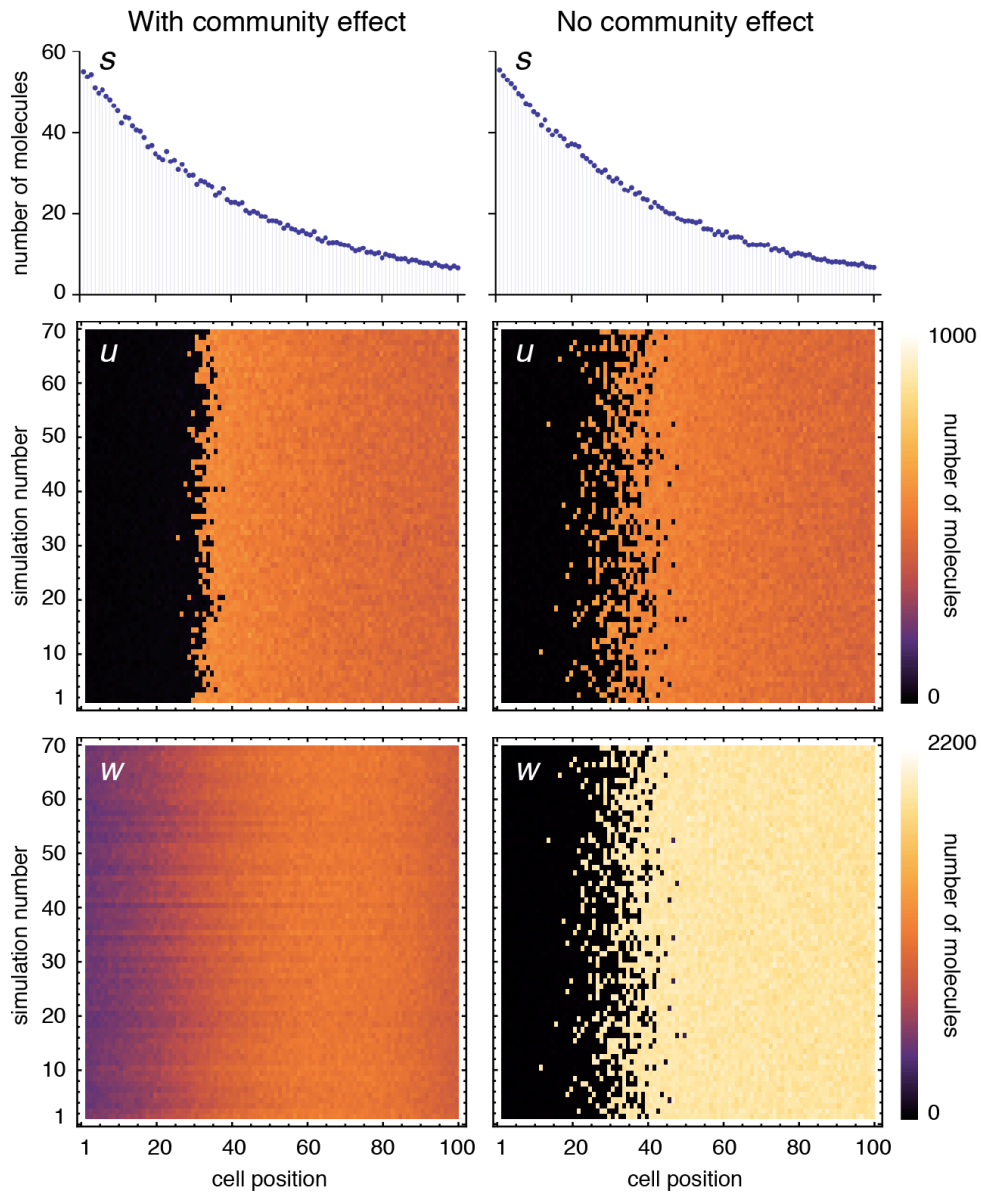
Figure 16.   Stochastic simulations of the morphogen gradient model, enabling a community effect (left column) or disabling it (right column).

4. A transient and dynamic morphogen gradient is sufficient for patterning. Moreover, a uniform morphogen distribution across the induced tissue can maintain the established pattern with a well-demarcated boundary (Fig.10).

An inconvenient truth of the community effect. We have shown that a community effect spreads across the system, without any interlocked negative feedback. Although this is indeed an inconvenient property for patterning of the embryo, it has turned into a control mechanism for coordinated gene expression in nature. Our two different models demonstrated that it can play another role as a control sub-circuit for pattern formation.

We have also shown that, when the spatial dimension is taken into account, the community effect must be treated as not a simple cell-number dependent phenomenon, but also as tissue size and cell-density dependent (Supp. Section A). This is relevant to embryonic patterning, because embryos undergo dynamic morphogenetic movement. It raises a question: *How does tissue geometry affect patterning?* We plan to address this issue in the future.

Roles of community effect in pattern formation by Turing mechanism. Turing patterns emerge from a uniform signal with noise as shown in Fig.7A. Embryonic induction, by contrast, is initiated by diffusion of signalling molecules from a localised source. Our simulation results suggested that the Turing pattern mechanism can be exploited as a mechanism to restrict the effective range of a community effect, in response to a localised inducing signal (Fig.7B). Such an isolated area is maintained as a stable wave. A community effect may also coordinate the gene expression within the area.

Roles of community effect in dynamic morphogen gradient interpretation. Dynamic assessment of a morphogen gradient has been suggested previously based on experimental evidence [15, 23, 27, 28]. Surprisingly, our simulations indicate that a morphogen gradient at steady state is not a prerequisite for pattern formation. It turned out that a transient gradient is sufficient for patterning and the steady-state morphogen distribution can be uniform (Fig.10).

In our dynamic morphogen gradient model, therefore, positional information is encoded in the dynamics of diffusing morphogens, rather than by the concentration in a stable gradient. In other words, the system confers memory of morphogen dynamics. The dynamic morphogen gradient model thus provides an alternative to the so-called French flag model and the concept of positional information for embryonic patterning [51].

The network of a mutually-repressive pair of genes has been implicated in interpretation of positional information in the gradients of maternal proteins, which work as morphogens [28] in the segmentation of *Drosophila* embryos. However, because it involves a complex GRN, its exact mechanism has remained unresolved. We found that, when being combined with a community effect, this cross repression sub-circuit (genetic toggle switch) turned into a robust patterning GRN with coordinated gene expression. The community effect plays a role in refining patterns by sharpening gene expression boundaries.

Our simulation results have also shown that a community effect helps orient gene expression patterns in space (Fig.9). This is one of the roles of a community effect not recognised previously. In the context of animal development, the community effect and the GRN depicted in Fig.1C is a robust system to establish an asymmetric pattern in embryonic tissues. We have not fully understood the origin of

this asymmetry of spatial gene expression introduced by a community effect, which awaits further investigation (see Section B). These results indicate that diffusible factors for cell-to-cell communication, either with positive or negative influence from one cell to the other, may provide myriad mechanisms for embryonic patterning when interlinked with a sub-circuitry of developmental GRNs.

References

[1] Ackers, G. K., Johnson, A. D., Shea, M. A.: Quantitative Model for Gene Regulation by λ Phage Repressor, *Proceedings of the National Academy of Sciences USA*, **79**(4), February 1982, 1129–1133.

[2] Alon, U.: Network motifs: theory and experimental approaches., *Nature reviews Genetics*, **8**(6), June 2007, 450–61.

[3] Berg, H. C.: *Random Walks in Biology*, (Princeton University Press, Princeton, 1993).

[4] Bernstein, D.: Simulating mesoscopic reaction-diffusion systems using the Gillespie algorithm., *Physical review. E, Statistical, nonlinear, and soft matter physics*, **71**(4 Pt 1), April 2005, 041103.

[5] Bolouri, H., Davidson, E. H.: The gene regulatory network basis of the "community effect," and analysis of a sea urchin embryo example, *Developmental Biology*, **340**(2), 2010, 170–8.

[6] Cardelli, L.: Artificial biochemistry, *Algorithmic Bioprocesses, LNCS*, Springer, 2008.

[7] Cardelli, L., Zavattaro, G.: On the Computational Power of Biochemistry, *Proceedings of the 3rd international conference on Algebraic Biology*, Springer-Verlag, Berlin, Heidelberg, 2008.

[8] Chabrier-Rivier, N., Fages, F., Soliman, S.: The Biochemical Abstract Machine BioCham, *Proceedings of CMSB 2004*, 3082, 2005.

[9] Cherry, J. L., Adler, F. R.: How to make a biological switch, *Journal of Theoretical Biology*, **203**(2), March 2000, 117–33.

[10] Crick, F.: Diffusion in embryogenesis, *Nature*, **225**(5233), February 1970, 671.

[11] Danos, V., Feret, J., Fontana, W., Krivine, J.: Scalable simulation of cellular signaling networks, *Proceedings of the 5th Asian conference on Programming languages and systems*, APLAS'07, Springer-Verlag, 2007.

[12] Danos, V., Laneve, C.: Formal Molecular Biology, *Theoretical Computer Science*, **325**(1), 2004, 69–110.

[13] Davidson, E. H.: *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*, Academic Press, 2006.

[14] Davidson, E. H.: Emerging properties of animal gene regulatory networks, *Nature*, **468**(7326), December 2010, 911–920.

[15] Dessaud, E., Ribes, V., Balaskas, N., Yang, L. L., Pierani, A., Kicheva, A., Novitch, B. G., Briscoe, J., Sasai, N.: Dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen sonic hedgehog, *PLoS Biology*, **8**(6), January 2010, e1000382.

[16] Duboc, V., Lapraz, F., Besnardeau, L., Lepage, T.: Lefty acts as an essential modulator of Nodal activity during sea urchin oral-aboral axis formation, *Developmental Biology*, **320**(1), 2008, 49 – 59.

[17] Evans, L.: *Partial Differential Equations*, American Mathematical Society, Providence, 1998, ISBN 0821807722.

[18] Eymard, R., Gallouet, T., Herbin, R.: *Finite volume methods*, vol. 7 of *Handbook of Numerical Analysis*, Elsevier, 2000.

[19] Gardner, T. S., Cantor, C. R., Collins, J. J.: Construction of a genetic toggle switch in Escherichia coli., *Nature*, **403**(6767), January 2000, 339–42.

[20] Gibson, M., Bruck, J.: Efficient exact stochastic simulation of chemical systems with many species and many channels, *J. Phys. Chem. A*, **104**(9), 2000, 1876–1889.

[21] Gillespie, D. T.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *Journal of Computational Physics*, **22**, 1976, 403–434.

[22] Gurdon, J. B.: A community effect in animal development, *Nature*, **336**(6201), 1988, 772–4.

[23] Gurdon, J. B., Mitchell, A., Mahony, D.: Direct and continuous assessment by cells of their position in a morphogen gradient, *Nature*, **376**(6540), Aug 1995, 520–521.

[24] Gurdon, J. B., Mitchell, A., Ryan, K.: An experimental system for analyzing response to a morphogen gradient, *Proc Natl Acad Sci U S A*, **93**(18), Sep 1996, 9334–8.

[25] Gurdon, J. B., Tiller, E., Roberts, J., Kato, K.: A community effect in muscle development, *Current Biology*, **3**(1), 1993, 1–11.

[26] Hamada, H., Meno, C., Watanabe, D., Saijoh, Y.: Establishment of vertebrate left-right asymmetry, *Nature reviews Genetics*, **3**(2), February 2002, 103–13.

[27] Harvey, S. a., Smith, J. C.: Visualisation and quantification of morphogen gradient formation in the zebrafish, *PLoS Biology*, **7**(5), May 2009, e1000101.

[28] Jaeger, J., Blagov, M., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Surkova, S., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., Reinitz, J.: Dynamical analysis of regulatory interactions in the gap gene system of Drosophila melanogaster, *Genetics*, **167**(4), August 2004, 1721–37.

[29] Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., Reinitz, J.: Dynamic control of positional information in the early Drosophila embryo, *Nature*, **430**(6997), July 2004, 368–71.

[30] John, M., Lhoussaine, C., Niehren, J., Versari, C.: Biochemical Reaction Rules with Constraints., *European Symposium On Programming (ESOP)* (G. Barthe, Ed.), 6602, Springer, 2011.

[31] Jonsson, H., Heisler, M., Reddy, G. V., Agrawal, V., Gor, V., Shapiro, B. E., Mjolsness, E., Meyerowitz, E. M.: Modeling the organization of the Wuschel expression domain in the shoot apical meristem, *Bioinformatics*, **21**(suppl 1), 2005, i232–i240.

[32] Kaern, M., Elston, T., Blake, W., Collins, J.: Stochasticity in gene expression: from theories to phenotypes, *Nature Reviews Genetics*, **6**(6), 2005, 451–467.

[33] Kampen, N. V.: *Stochastic Processes in Physics and Chemistry*, North Holland, 2001.

[34] Kondo, S., Asai, R.: A reaction-diffusion wave on the skin of the marine angelfish Pomacanthus, *Nature*, **376**(6543), 1995, 765–768.

[35] Kondo, S., Miura, T.: Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation, *Science*, **329**(5999), September 2010, 1616–1620.

[36] Meinhardt, H.: *Models of biological pattern formation*, vol. 6, Academic Press, 1982.

[37] Meinhardt, H.: *The algorithmic beauty of sea shells.*, The virtual laboratory, Springer, 1995.

[38] Meinhardt, H.: Models for the generation and interpretation of gradients, *Cold Spring Harbor Perspectives in Biology*, **1**(4), October 2009, a001362.

[39] Meinhardt, H., Gierer, a.: Applications of a theory of biological pattern formation based on lateral inhibition, *Journal of Cell Science*, **15**(2), July 1974, 321–46.

[40] Nicolis, G., Prigogine, I.: *Self-organisation in non-equilibrium systems*, Wiley International, 1977.

[41] Oates, A. C., Gorfinkiel, N., González-Gaitán, M., Heisenberg, C.-P.: Quantitative approaches in developmental biology, *Nature reviews Genetics*, **10**(8), August 2009, 517–30.

[42] Papatsenko, D., Levine, M.: The Drosophila Gap Gene Network Is Composed of Two Parallel Toggle Switches, *PLoS ONE*, **6**(7), 07 2011, e21145.

[43] Ptashne, M.: *A genetic switch: phage lambda revisited*, 3rd ed edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2004.

[44] Raj, A., van Oudenaarden, A.: Nature, nurture, or chance: stochastic gene expression and its consequences, *Cell*, **135**(2), October 2008, 216–26.

[45] Ribes, V., Briscoe, J.: Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback, *Cold Spring Harbor Perspectives in Biology*, **1**(2), August 2009, a002014.

[46] Saka, Y., Lhoussaine, C., Kuttler, C., Ullner, E., Thiel, M.: Theoretical basis of the community effect in development, *BMC Systems Biology*, **5**, March 2011, 54.

[47] Saka, Y., Smith, J. C.: A mechanism for the sharp transition of morphogen gradient interpretation in Xenopus, *BMC Developmental Biology*, **7**, 2007, 47.

[48] Sanderson, A., Kirby, R., Johnson, C., Yang, L.: Advanced Reaction-Diffusion Models for Texture Synthesis, *Journal of Graphics Tools*, **11**(3), 2006, 47–71.

[49] Turing, A. M.: The Chemical Basis of Morphogenesis, *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **237**(641), 1952, 37–72.

[50] Wartlick, O., Kicheva, A., González-Gaitán, M.: Morphogen gradient formation, *Cold Spring Harbor Perspectives in Biology*, **1**(3), September 2009, a001255.

[51] Wolpert, L.: Positional information and the spatial pattern of cellular differentiation., *Journal of Theoretical Biology*, **25**(1), October 1969, 1–47.
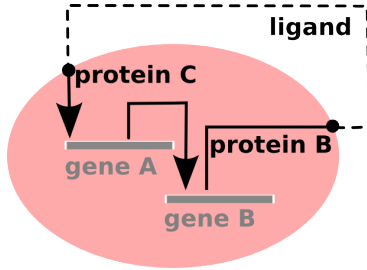
Figure 17. The GRN for the community effect in an active cell with genes A and B. Gene A activates gene B, protein B is expelled from the cell, transformed into a diffusible factor, here called *ligand*, that diffuses to adjacent cells. This signal's reception activates protein C, which in turns activates gene A.

Appendix

A. Model of a community effect in Xenopus in space

In this section we present a model of the community effect in *Xenopus*. It can be seen as an expansion of our minimal model in Section 3 on two levels. First, the community effect loop here involves two genes and their protein products, instead of one, in intercellular communication (Fig.17). Second, instead of collapsing a gene expression into one reaction, we consider separately transcription of genes into messenger RNA, and its translation into proteins. This detailed reaction network was previously investigated in [46]. The new contribution of this section is to consider the same detailed network of a community effect *in space*. Based on a rule-based formalization (Section A.2), we observe an unlimited spread of a community effect in stochastic simulation (Section A.3). This confirms our observations based on the minimal model in Section 3. To confirm our results observed in *discrete space*, we investigate a deterministic PDE model with explicit diffusion in *continuous* one-dimensional space in Section A.4. The findings confirm our previous conclusions in [46] and strengthen the results we have presented in Section 3.
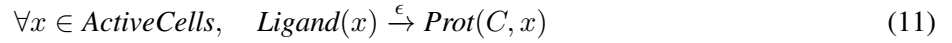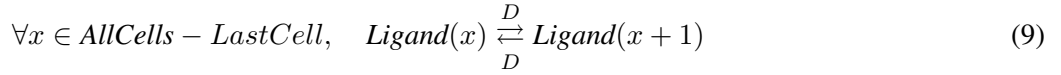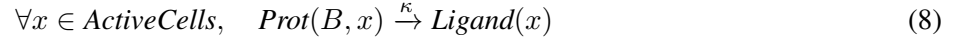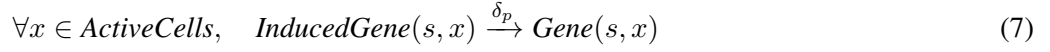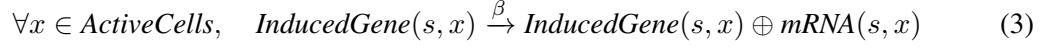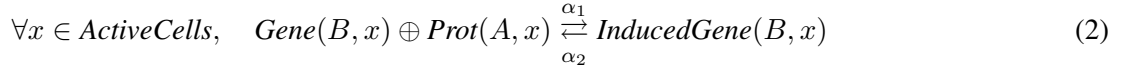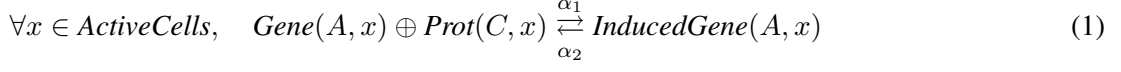
A.1. Informal description of the GRN with one-dimensional diffusion

Due to the greater level of details, we resort to a different graphical notation for the *GRN of active cells* in Fig.17. Each cell contains two genes A and B, where A *activates* B. In turn, B *activates* A over a feedback loop mediated by inter-cellular communication: Gene B expresses a protein, which is then exported out of the cell and turned into a diffusible factor, named $ligand$. This $ligand$ either degrades, or *diffuses* to a directly neighbouring cell. Reception of a ligand by an active cell results in an intra-cellular signal, which creates a protein C. This protein C then activates gene A.

A.2. Rule-based model of the detailed reaction network in one dimensional space

The detailed reaction network including the intermediate steps of transcription and translation is listed in Fig.18. Note that terms such as $Gene(A, x)$ include the species variables $\in \{A, B, C\}$ as the first parameter, and the position variable x as the second one.

Rule (1) denotes the activation of the gene A in an active cell at position x, represented by the term $Gene(A, x)$, by the protein $Prot(C, x)$. It results in an induced gene, represented by $InducedGene(A, x)$. This is indeed a complex between the protein C and the gene A's promoter. α_1 and α_2 are the rate parameters for the forward and reverse reactions, respectively. These parameters define the mass action

$$\forall x \in ActiveCells, \quad Gene(A, x) \oplus Prot(C, x) \overset{\alpha_1}{\underset{\alpha_2}{\rightleftarrows}} InducedGene(A, x) \tag{1}$$

$$\forall x \in ActiveCells, \quad Gene(B, x) \oplus Prot(A, x) \overset{\alpha_1}{\underset{\alpha_2}{\rightleftarrows}} InducedGene(B, x) \tag{2}$$

$$\forall x \in ActiveCells, \quad InducedGene(s, x) \overset{\beta}{\rightarrow} InducedGene(s, x) \oplus mRNA(s, x) \tag{3}$$

$$\forall x \in ActiveCells, \quad mRNA(s, x) \overset{\gamma}{\rightarrow} mRNA(s, x) \oplus Prot(s, x) \tag{4}$$

$$\forall x \in ActiveCells, \quad mRNA(s, x) \overset{\delta_m}{\longrightarrow} \mathbf{0} \tag{5}$$

$$\forall x \in ActiveCells, \quad Prot(s', x) \overset{\delta_p}{\longrightarrow} \mathbf{0} \tag{6}$$

$$\forall x \in ActiveCells, \quad InducedGene(s, x) \overset{\delta_p}{\longrightarrow} Gene(s, x) \tag{7}$$

$$\forall x \in ActiveCells, \quad Prot(B, x) \overset{\kappa}{\rightarrow} Ligand(x) \tag{8}$$

$$\forall x \in AllCells - LastCell, \quad Ligand(x) \overset{D}{\underset{D}{\rightleftarrows}} Ligand(x + 1) \tag{9}$$

$$\forall x \in AllCells, \quad Ligand(x) \overset{\delta_\ell}{\longrightarrow} \mathbf{0} \tag{10}$$

$$\forall x \in ActiveCells, \quad Ligand(x) \overset{\epsilon}{\rightarrow} Prot(C, x) \tag{11}$$

where $s \in \{A, B\}, s' \in \{A, B, C\}$, *ActiveCells* $= \{1 + N_p, \dots, N_p + N_a\}$, *AllCells* $= \{1, \dots, 2N_p + N_a\}$, and $LastCell = \{2N_p + N_a\}$.

Figure 18.   Detailed rule-based model of a community effect with diffusion in one-dimensional space.

rate of this event. In this model, we only write the rate parameters, since all the rate expressions are defined by the mass action rule. Likewise, the rule (2) describes the activation of gene B by protein A. Let us briefly explain the remaining rules.

Transcription and translation of a gene are defined by rule (3) and (4). Transcription requires an induced gene. Both mRNA and protein degrade according to the rules (5) and (6). Proteins may also degrade while they are bound to a gene, as stated in rule (7). The rate of this degradation is the same as the degradation rate of the unbound protein, δ_p.

Diffusion is implemented in a more fine-grained manner than the one described previously. Rule (8) defines the export of a B protein from the xth cell into the extracellular space and its conversion into a ligand. The position of a ligand is specified by an integer $x \in AllCells$, the same coordinate system as the cell. Note the distinction between active cells (pink ovals) and its neighbourhood (grey box without ovals) in Fig.2 on p.9. A secreted ligand hops from one position to the next by random walk, which approximates its diffusion in the entire extracellular space via (9). In any position, the ligand can degrade (see rule (10)). Finally, the binding of a ligand to a cellular receptor of an active cell and subsequent production of protein C is governed by rule (11).

Parameter	Description	Value
α_1	binding of a transcription factor to a gene	1.93×10^{-4}
α_2	dissociation of a transcription factor from a gene	3.47×10^{-2}
β	transcription of a gene to mRNA	1.16×10^{-2}
γ	translation of mRNA, creating a protein	2.31×10^{-2}
δ_m	degradation of mRNAs	1.16×10^{-3}
δ_p	degradation of proteins	3.47×10^{-4}
δ_l	degradation of Ligand	1.4×10^{-2}
κ	export of Prot(B) outside of a cell, creating a Ligand	3.85×10^{-4}
ϵ	binding of Ligand to an active cell, creating a Prot(C)	5.78×10^{-5}
D	diffusion of Ligand between two adjacent cells	5.205×10^{-2}

Table 1. Parameter values for stochastic simulations are the same as in [46], except those concerning Ligand.

A.3. Stochastic simulations

We simulated the dynamics of the one-dimensional cell row over time by stochastic simulations similar to those in Sections 3-5. Again, we first consider typical single simulation runs in Fig.19. Panels (A,C,E) show the dynamics of the $Ligand$ number across the cell row over time. Panels (B,D,F) are snapshots of $Prot(A)$ taken at the end of those simulations. All simulations are initiated by the initial induction with 100 $Ligand$ molecules in the leftmost active cell at time zero. We tested three different combinations of (N_a, N_p) for simulations. See Table 1 for the parameter values used in simulations.

We observe essentially the same behaviour as in the minimal model in Section 3:

- $N_a = 25$, $N_p = 100$: below the threshold number of active cells, the system is unable to sustain activation (Fig.19A,B). Only 16 of 100 simulations showed any residual activity at $t = 6 \times 10^5$.

- $N_a = 50$, $N_p = 100$: the system enters sustained activation in all of 100 simulation runs, as shown for a typical run in Fig. 19C. Thus the activation threshold lies in the range between 25 and 50 active cells. At the end of the simulation, $Prot(A)$ is present in all active cells (Fig.19D), indicating a coordinated gene expression in the community.

- $N_a = 500$, $N_p = 100$: With N_a well above the threshold, all simulations showed sustained activation, as in the run in Fig.19E and F, where all active cells reached a self-sustaining gene expression at a significant level. We observe that, as in the minimal model, a community effect spreads all over the active cells, similar to Fig.4D on p.11.

To better understand our observation in stochastic simulations and the relationship between N_p, N_a and the community effect, we constructed a deterministic model, which we describe below.
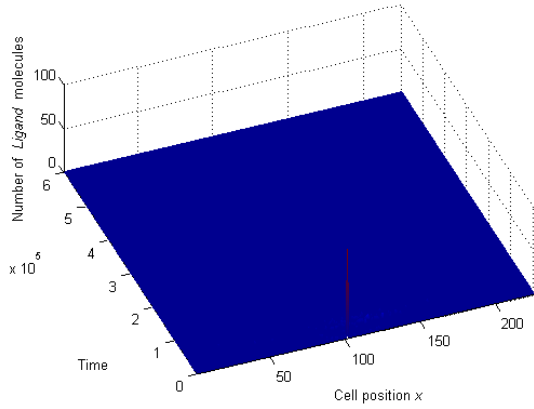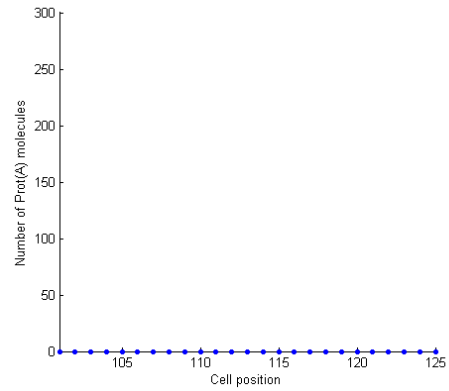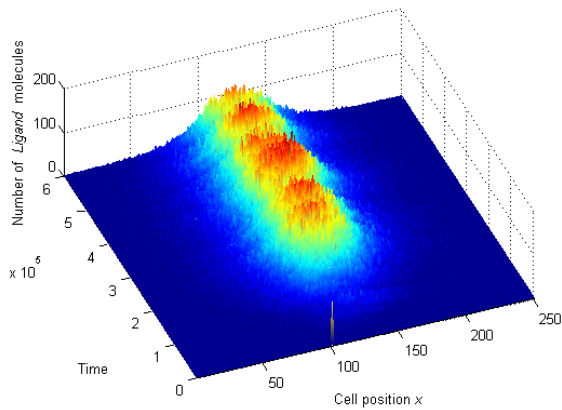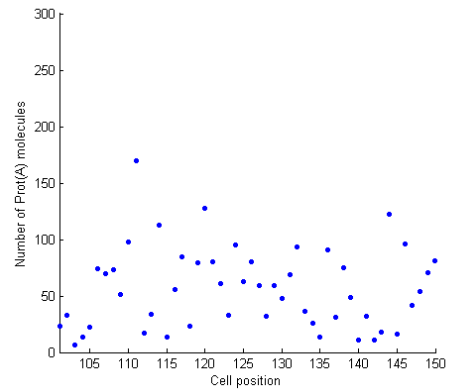
A. Ligand: $N_a = 25$, $N_p = 100$.

B. Prot(A): $N_a = 25$, $N_p = 100$, $t = 6 \times 10^5$

C. Ligand: $N_a = 50$, $N_p = 100$.

D. Prot(A): $N_a = 50$, $N_p = 100$, $t = 6 \times 10^5$
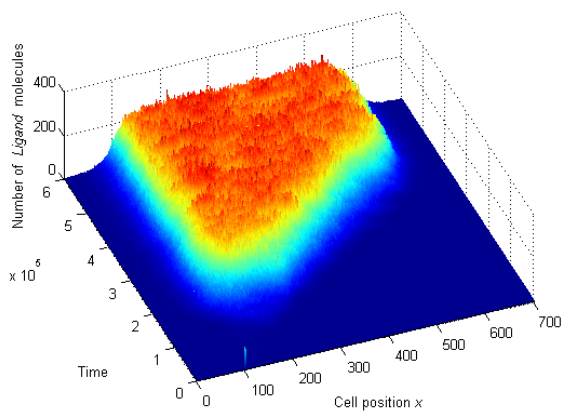
E. Ligand: $N_a = 500$, $N_p = 100$.
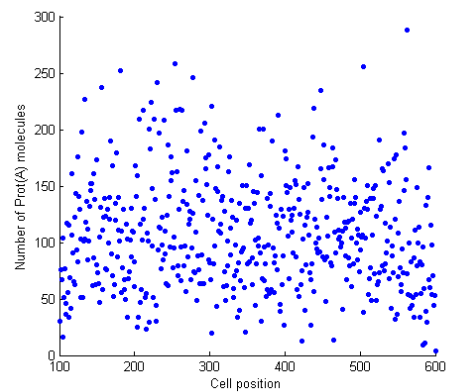
F. Prot(A): $N_a = 500$, $N_p = 100$, $t = 6 \times 10^5$

Figure 19.   Single runs of stochastic simulation of the detailed model of a community effect. Left column: time course of $Ligand$ over the entire cell row. Right column: at the end of the simulation to the left, level of protein A in active cells.

Molecules	PDE variables	Rule-based terms
Gene A	Ag	$Gene(A, x)$
Gene B	Bg	$Gene(B, x)$
Induced gene A	Agi	$InducedGene(A, x)$
Induced gene B	Bgi	$InducedGene(B, x)$
Transcripts (mRNA) of gene A	Ar	$mRNA(A, x)$
Transcripts (mRNA) of gene B	Br	$mRNA(B, x)$
Protein product of gene A	Ap	$Prot(A, x)$
Protein product of gene B	Bp	$Prot(B, x)$
Protein C	Cp	$Prot(C, x)$
The diffusible factor	Lp	$Ligand(x)$

Table 2. Terms of the stochastic model and corresponding variables of the deterministic model. Variables in the PDE model, briefly referred to as e.g. Ap, are indeed functions of (continuous) space x and time t, e.g. $Ap(x, t)$.

A.4. Deterministic model

Using the level of detail as in the rule-based model, we have constructed a deterministic model with molecular diffusion in *continuous* one-dimensional space. To avoid any confusion, we use different nomenclatures for terms in the rule-based model (representing *numbers* of molecules) and variables in this deterministic model (which stand for *concentrations*), as summarized in Table 2.

The model is described as follows: Because the number of gene copies per cell ($= 1$ in our model) is conserved, in the region occupied by active cells,

$$Ag + Agi = \rho$$
$$Bg + Bgi = \rho, \tag{1}$$

where ρ is the cell density. The dynamics of the variables are described by a set of partial differential equations (PDEs):

$$\frac{\partial Agi}{\partial t} = \alpha_1 \, Ag \, Cp - (\alpha_2 + \delta_p) \, Agi, \tag{2}$$

$$\frac{\partial Bgi}{\partial t} = \alpha_1 \, Bg \, Ap - (\alpha_2 + \delta_p) \, Bgi, \tag{3}$$

$$\frac{\partial Ar}{\partial t} = \beta \, Agi - \delta_m \, Ar, \tag{4}$$

$$\frac{\partial Br}{\partial t} = \beta \, Bgi - \delta_m \, Br, \tag{5}$$

$$\frac{\partial Ap}{\partial t} = \gamma \, Ar + \alpha_2 \, Bgi - \alpha_1 \, Bg \, Ap - \delta_p \, Ap, \tag{6}$$

$$\frac{\partial Bp}{\partial t} = \gamma \, Br - (\kappa + \delta_p) \, Bp, \tag{7}$$

$$\frac{\partial Cp}{\partial t} = \epsilon \, Lp + \alpha_2 \, Agi - \alpha_1 \, Ag \, Cp - \delta_p \, Cp, \tag{8}$$

$$\frac{\partial Lp}{\partial t} = \lambda \, \Delta Lp - \delta_\ell \, Lp + J + I, \tag{9}$$

where the functions J and I in (9) are defined as

$$J(x,t) = \kappa \, Bp(x,t) - \epsilon \, Lp(x,t), \tag{10}$$
$$I(t) = \eta \, \theta(\tau - t). \tag{11}$$

Here, θ is the unit step function and J is the net flow of Lp in and out of the cell. The function I is the initial induction of amplitude η and duration τ. The term $\lambda \, \Delta Lp$ describes diffusion of Lp. It follows Fick's second law [3], where λ is the diffusion coefficient. The boundary condition is

$$Ag(x,0) = Bg(x,0) = \rho,$$

and all other variables are 0 at $t = 0$. We consider the system in $0 \leq x \leq d$ with closed boundaries, therefore

$$\frac{\partial Lp}{\partial x} = 0, \; x = \begin{cases} 0 \\ d \end{cases}. \tag{12}$$

As in the rule-based model, the active cells occupy the middle domain of width r_a, and are flanked by a passive cell region of size r_p on each side. Therefore,

$$d = r_a + 2r_p.$$

Parameter	Description	Value	Range
α_1	binding of Cp to Ag	1.93×10^{-4}	fixed value
α_2	dissociation of Cp from Ag	3.47×10^{-2}	fixed value
β	transcription	1.16×10^{-2}	fixed value
γ	translation w	2.31×10^{-2}	fixed value
δ_m	degradation of mRNAs	1.16×10^{-3}	fixed value
δ_p	degradation of proteins	5.78×10^{-4}	fixed value
δ_ℓ	degradation of Lp	5×10^{-4}	fixed value
ϵ	import and conversion of Lp into Cp	10^{-5}	n.a.
κ	export and conversion of Bp into Lp	3.85×10^{-4}	fixed value
λ	diffusion rate of Lp	20	fixed value
ρ	cell density	see Fig.20	$0 \leq \rho \leq 20$
η	amplitude of induction	2	fixed value
τ	duration of induction	200	fixed value
r_a	domain size of active cells (community size)	see Fig.20	$0 \leq r_a \leq 400$
r_p	domain size of passive cells	see Fig.20	see Fig.20

Table 3.   Parameter values for numerical simulations of a community effect in Section A.4. Used for simulations, unless defined explicitly in the main text. Parameter ranges in the last column are used for phase diagrams.
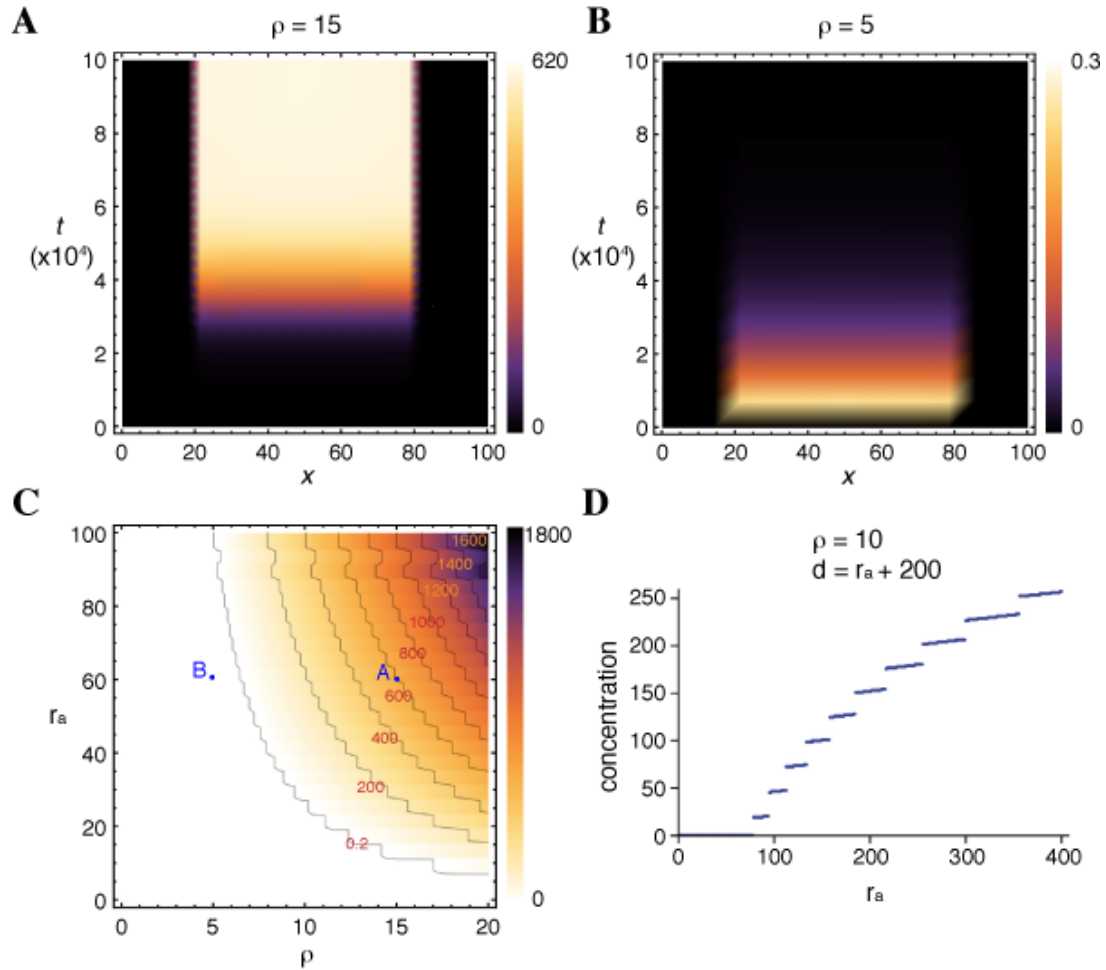
Figure 20. Simulations of the deterministic model of a community effect. **(A, B)** The dynamics of Ap in 1D space over time, for two different densities. The concentration is colour-coded as indicated. **(A)** At high density ($\rho = 15$) the community effect spreads rapidly once initiated. **(B)** With low density ($\rho = 5$), gene expression is transient. Note the different scale of colour-coding for gene expression level. Parameter values for induction are $\eta = 2, \tau = 200$ and the width of induction is 1. **(C)** Phase diagram in the parameter plane (ρ, r_a). Initial induction is applied uniformly across the community. Ap concentration at quasi-steady state ($t = 4 \times 10^5$) is colour-coded as indicated. Contours are also shown. Gene expression depends on the cell density ρ and the size of the community r_2. The blue dots correspond to the parameter sets in panels A and B. **(D)** Ap concentration plotted as a function of the community size r_a. Simulations are performed with the constant $r_p = 100$ and $\rho = 10$. Initial induction is applied uniformly across the community. The plot shows many discontinuous jumps in concentration, which is most likely an artefact of the numerical integration.

A.5. Community effects are size- and density-dependent in the one-dimensional model

The simulation results of the PDE model are qualitatively similar to what we observed in stochastic simulations (Fig.20). The most notable observation is that self-sustained gene expression depends on the cell density and the size of the community relative to that of the whole system. This indicates that community effects are influenced by the spatial arrangement of cells. Let us now explain the simulation results of the PDE model.

Fig.20A and B show the typical dynamics of Ap concentration in one dimensional space with different cell densities. In these simulations, $d = 100$ and $r_a = 60$ ($r_p = 20$). Although only the narrow middle region of width one is induced for a brief period, the entire domain of active cells starts expressing Ag after a time lag. This seemingly dormant period before the surge of gene expression was previously observed with a spaceless model of a community effect [46]. Such expression surge happens almost simultaneous across the community in the simulations. The speed of the expansion of a community effect obviously depends on the diffusion rate of the ligand δ_ℓ. With lower diffusion rates, it becomes slower and is qualitatively similar to the spread of a community effect shown in Fig.4D and Fig.19E. In the simulation in Fig.20A, the ligand concentration becomes almost uniform across the whole system at steady state (data not shown).

The system reproduces a community effect: once the size and density of the community exceeds certain thresholds the group can maintain gene expression after a transient induction. We next asked what is the *relationship between the size of the community and the density of the cell group*. Fig.20C is a phase diagram plotted for the parameter pair (ρ, r_a), with a constant system size $d = 100$. The diagram indicates that the condition for a self-sustained gene expression is strongly correlated with the cell density and the community size. The critical *contour* for $Ap = 0.2$ (an arbitrary number close to 0) shown in the panel does not fit to $r_a\rho$ (number of cells) $= constant$, indicating that the observed community effects are not a simple *cell number*-dependent phenomena, but are influenced by the cell density and the community's size.

We also performed simulations with a constant cell density, constant passive region $r_p = 100$ and varying r_a, that is, $d = r_a + 200$: this scenario resembles that of the stochastic rule-based model of Section A.3. The results are consistent with the stochastic simulation results, showing the clear minimal critical r_a that is required for a community effect (Fig.20D). The observed behaviour is qualitatively similar to the minimal model of the community effect, see Fig.4C.

B. Asymmetry in the dynamic morphogen gradient model

Without positive feedback and a community effect (Fig.13), (1) on p.21 is reduced to

$$\frac{\partial u}{\partial t} = \frac{k_1\, s}{v^n + 1} - \mu_1\, u, \tag{13}$$

$$\frac{\partial v}{\partial t} = \frac{k_3\, s}{u^n + 1} - \mu_2\, v, \tag{14}$$

$$\frac{\partial s}{\partial t} = D_2\, \Delta\, s - \mu_4\, s. \tag{15}$$

(13) and (14) are invariant after a symmetric transformation,

Parameter	Description	Value
α_u	u self-activation maximum rate	33.25
k_u	u self-activation Michaelis constant	450
α_v	activation of v by u maximum rate	60
k_v	activation of v by u Michaelis constant	450
α_i	inhibition of u by v maximum rate	0.04
k_i	inhibition of u by v Michaelis constant	0.1
μ_u	degradation of u	0.035
μ_v	degradation of v	0.1
D_u	diffusion of u	0.4
D_v	diffusion of v	15

Table 4.   Parameter values for stochastic simulations in Sections 3.4 and 4.2.

$$
\begin{bmatrix} u \\ k_1 \\ \mu_1 \end{bmatrix} \leftrightarrow \begin{bmatrix} v \\ k_3 \\ \mu_2 \end{bmatrix}. \tag{16}
$$

This is reflected in the symmetry around the diagonal $k_1 = k_3$ in Fig.14A and C. The symmetry is broken by introducing a community effect (see (1)), which makes the GRN *topologically* asymmetric. The topological asymmetry of the GRN seems to favour one pattern ($[v, u]$) over the other ($[u, v]$) in response to a dynamic morphogen gradient (Section 5). This bias of pattern formation was observed with wide range of parameter values in the system with a community effect, but not in the symmetric one (13)-(15). Asymmetries by topology of GRNs are thus distinct from those introduced by unbalanced parameter values. However, the fundamental, non-trivial question remains unanswered: how does the topological asymmetry of the GRN bring about the bias in pattern formation? Although out of scope in this paper, this is one of the main subjects of our future work.

C.   Parameter sets

Table 4 provides the parameter sets for stochastic simulations in Sections 3.4 and 4.2, and Table 5 for numerical simulations in Section 5.2.

D.   Numerical error in simulations of PDE models

We have used Mathematica for numerical integration of deterministic models. Mathematica's numerical integrator (the function `NDSolve`) returns an interpolating function. Therefore there are two sources of numerical errors: one is by the numerical integration itself and another by interpolation. We examined how much error is introduced by a numerical integration. The error can be calculated by plugging numerical data in $lhs - rhs$ of (1) on p.21, for example,

Parameter	Description	Value	Range
D_1	diffusion of w	20	$0 \leq D_1 \leq 100$
D_2	diffusion of s	20	$0 \leq D_2 \leq 100$
k_1	activation of u by s	0.6	$0 \leq k_1 \leq 1$
k_2	activation of u by w	0.2	n.a.
k_3	activation of v by s	0.8	$0 \leq k_3 \leq 1$
k_4	activation of w by u	0.2	n.a.
μ_1	degradation of u	0.08	n.a.
μ_2	degradation of v	0.04	n.a.
μ_3	degradation of w	0.01	n.a.
μ_4	degradation of s	0.001	n.a.
τ	control parameter for $s(0,t)$	0.004	$10^{-3} \leq \tau \leq 10^{-1}$
d	responding tissue size	100	n.a.

Table 5.   Parameter values for numerical simulations in Section 5. Used for simulations, unless defined explicitly in the main text. Parameter ranges in the last column are used for phase diagrams. n.a., not applicable.

$$Error(u) = \frac{\partial u}{\partial t} - \left(\frac{k_1\, s + k_2 \left(\frac{w^{\mathrm{n}}}{w^{\mathrm{n}}+1} \right)}{v^{\mathrm{n}} + 1} - \mu_1\, u \right).$$

The numerical data should be as close as $Error(u) = 0$. Fig.21 is an example of numerical error for u in the simulation shown in Fig.9A-D on p.22. $Error(u)$ is plotted for numerical integrations with high precision (dark blue) and low precision (magenta). The error is negligible except near the point of discontinuity (e.g., gene expression boundary). The oscillatory errors at the edge are a typical phenomenon due to finite order approximation by Fourier series.

Analytical solutions of our PDE models presented in this paper are not trivial to derive, if they exist at all. Instead we rely on numerical simulations for the analysis of the system with the aid of phase diagrams.

Figure 21.   Numerical error in simulations of PDE models **(A)** Concentration profile of u. **(B)** Error(u) using the numerical data. Both (A) and (B) are plotted using numerical simulation data at $t = 5000$. Two sets of data, with high precision (dark blue) and low precision (magenta) for numerical integration, are shown.