A Modular Computational Framework for the Rational Design and Exploration of Synthetic Development

[1,*]Calvin Lam, [1,2,†]Leonardo Morsut

[1]The Eli and Edythe Broad CIRM Center, Department of Stem Cell Biology and Regenerative Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033-9080, USA
[2]Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089-1111, USA
[*]Present address: College of Medicine, University of Nebraska Medical Center, Omaha NE, 68198-0001, USA
[†]Correspondence: Leonardo.Morsut@med.usc.edu

SUMMARY

Synthetic development is a nascent field of research that uses the tools of synthetic biology to design genetic programs directing cellular patterning and morphogenesis in higher eukaryotic cells, mainly mammalian ones. Current design methods of these genetic programs proceed inefficiently, relying on trial and error processes. By contrast, computational models can act as rapid testing platforms, revealing the networks, signals, and responses required for achieving robust target structures. We introduce a computational model where contact dependent cell-cell signaling networks and cellular responses can be chosen in a modular fashion, allowing *in silico* recreation of known synthetic morphogenic trajectories such as those resulting in multilayered synthetic spheroids. By altering the modular components, our model also allows for the exploration of new trajectories that can result in hollowed, elongated, and oscillatory structures. Our model functions as a testing ground illuminating how synthetic biology tools can be used to create particular structures. In addition, it can provide valuable insight into our understanding of both imagined and extant cellular morphologies.

KEYWORDS

Synthetic biology, self-organization, tissue engineering, computational modeling, cellular Potts, juxtacrine signaling, synNotch, patterning, morphogenesis

INTRODUCTION

During the development of multicellular organisms, cells self-organize into complex tissue and organ architectures essential for proper function. Self-organization is driven by genetically encoded programs that dictate interactions at multiple biological levels, from cellular to tissue and organ levels (Santorelli et al., 2019; Toda et al., 2019; Turner et al., 2016). Deconstructive approaches, such as gene knock-out or interference experiments have traditionally been used to investigate how self-organization arises from these levels (Bashor et al., 2010; Libby et al., 2018). More recently, reconstructive approaches are increasingly being employed (Elowitz and Leibler, 2000; Li et al., 2018; Santorelli et al., 2018; Tigges et al., 2009). These approaches emphasize identification of the universal principles that govern life (Elowitz and Lim, 2010).

In the field of developmental biology, the goal of reconstruction is to understand primary components and their role in self-organization (Bashor et al., 2010; Davies, 2017; Lim, 2010; Mukherji and van Oudenaarden, 2009). For instance, which genes or genetic circuits, in which cells, allow and control specific natural or non-natural tissue development? Synthetic development seeks to integrate synthetic genetic circuits into cells to control both the stimuli they sense and their subsequent behavior in order to achieve synthetic multicellular development (Santorelli et al., 2019; Toda et al., 2018, 2019).

For example, by adding synthetic signaling and adhesion circuits to the genome of non-assembling cells, e.g. mouse fibroblast cells in culture, it is possible to achieve self-assembling patterned spheroids (Toda et al., 2018). Similar genetic circuits could, in theory, be used to design other developmental trajectories in such non-developing cellular systems, however it remains unclear exactly how to predict which genetic circuits would be required for any given target structure (Briers et al., 2019; Santorelli et al., 2019).

Currently, synthetic development programs are designed intuitively, based on literature search and preliminary experiments in simplified setups, before implementation into living cells. As a result, reconstruction attempts move slowly, relying on a cycle of design, cell implementation, behavior monitoring, redesign, and so forth. This can be both time-consuming and costly, therefore limiting exploration of the genetic circuits, behaviors, and experimental parameters critical for achieving desired structures and features (Briers et al., 2019). Arguably, the most limiting step of reconstructive approaches is the first, design, which indicates our incomplete understanding of the underlying biology. A sensible initial guess could be shown to be erroneous only after months of cell implementation work. In these situations where not everything is known, even semi-predictive computational models could be of significant value, allowing rapid implementation of many designs *in silico*. These various designs could be tested for viability and pre-optimized before biological implementation, thereby saving time and cost (Briers et al., 2019; Santorelli et al., 2019). Additionally, such a framework could both facilitate the design of novel tissues of potential biomedical interest and delineate which structures might or might not be possible (Elowitz and Lim, 2010; Ollé-Vila et al., 2016).

While current computational models can successfully replicate various natural developmental processes (Belmonte et al., 2016; Hester et al., 2011; Hutson et al., 2017; Lakatos et al., 2018; Lambert et al., 2018; Lin et al., 2009; Marin-Riera et al., 2018; Shaya et al., 2017; Swat et al., 2015), they exist disparately, functioning as highly specialized tools tailored to the specific mechanisms/programs of the focal biological system. In the case of synthetic development, an effective computational framework requires generalizability and modularity. In order to maximize the designs that can be tested *in silico*, the model must reflect the broadness and modularity of tools employed in synthetic biology itself. Such a framework would not only enable great flexibility in design testing, it would also facilitate the exploration of numerous parameters and networks promoting robust development as well as elucidating novel designs that could yield as yet unimagined structures of interest.

Here we describe the development and use of a computational framework for synthetic development circuit design. Inspired by the use of synthetic Notch (synNotch) pathways to program multicellular morphogenesis (Toda et al., 2018), the framework allows the design of synthetic cell-cell contact-dependent signaling networks with user defined cellular inputs and responses. In this paper, we first describe a generalizable cell-cell contact dependent signaling model (named GJSM, generalized juxtacrine signaling model) that could be modified to capture different types of signaling inputs and dynamics. We describe then the GJSM implementation the model in a cellular Potts environment (Swat et al., 2012) and how we pair the signaling with different behavioral outputs to mimic the modularity of the synNotch protein tool in an extended framework (Fig.1 and methods). We describe then the use of this framework to recapitulate currently implemented synthetic structures like multi-layered spheroids (Fig. 2-4). Finally, we describe several design-implementation-test cycles for novel genetic programs for other multicellular behaviors such as hollowing, elongation, and oscillation (Fig. 5-7). We believe that this type of framework will enable the field to move towards a more deterministic and constructive understanding of developmental signaling logic.

RESULTS

Brief overview of the computational method (for more details see STAR Methods)

Our computational model captures contact-dependent signaling and links it to various changes in cell behaviors. Cell-cell signaling that relies on contact-dependence (aka juxtacrine) generally functions using similar logic: if cell (A) expressing a ligand is in contact with cell (B), and cell (B) expresses the cognate receptor for that ligand, then cell (B) will expresses the target genes (Bosenberg and Massagué,

1993; Massagué, 1990; Massagué and Pandiella, 1993) (Fig. 1a, Biological). When target gene products accumulate sufficiently, the behavior of the responding cell (B) changes (Fig. 1b, Biological).

To model this modality of signaling, we conceptually separated it in two parts: (i) signaling-dependent continuous changes in target gene expression in receiver cells, and (ii) gene expression dependent change in cell behavior (see STAR methods for full details). For the first part, we use differential equations to model input-dependent response. The strength of the response (target gene induction) in the receiver cell over time depends on several factors: the number of sender cells, the number of ligands on each sender cell, the number of receptors on the receiver cell, and the amount of contact between sender(s) and receiver. In a simplified two cell case with sender cell (A) and receiver cell (B), if the receptors on (B) are in excess, signaling depends primarily on the amount of ligand on cell (A) and the fraction of (A)'s surface contacting (B) (Fig. 1a, Biological). To capture this signaling, we define L as the number of ligands on cell (A)'s surface and Φ as the fraction of (A)'s surface in contact with (B)'s surface. We can then define the signal S that cell (B) receives as $S=\Phi*L$ (Fig 1a, Model). Increasing the fraction of shared surface induces a stronger response (Fig. 1a, right graph). This is one specific case of signaling, where ligand is limiting; for the generalized model, see STAR methods. This part of the model (Fig. 1a) accounts for the continuous changes in gene expression in receiver cells.

To model the target gene expression dependent change in cell behavior, we take a step-wise approximation where we define thresholds of gene activation that induce transition from a basal cell state to an activated state and vice versa. In this way, behavioral transitions are discrete instead of continuous (Anderson, 2005; Hester et al., 2011; Hutson et al., 2017) (Fig. 1b, right graph). Thresholds for transition from basal to active state and back from active to basal state can be different and are defined as model parameters. The state machine schematic for this type of network is proposed in Fig. 1b, Model.

The two parts of the model, continuous signaling and discrete response, are highly modular and can be designed independently of one another.

The Model Qualitatively Recapitulates Simple Synthetic Structures

Known synthetic development structures are currently derived from juxtacrine signaling dependent expression of different adhesion proteins (Toda et al., 2018). In this *in vitro* assays, cells are engineered with the appropriate signaling networks and responses, then between 20 and 200 cells are mixed together and grown for 2-3 days in ultra-low attachment plates, and their morphological and signaling evolution is followed via fluorescent microscopy. The first example is the two-layer spheroid (Fig. 2a); it is the result of a single forward network where CD19 on (A) blue cells bind to an αCD19 synNotch receptor on (B) gray cells to drive expression of green fluorescent protein (GFP) and E-cadherin, a homotypic adhesion protein (Toda et al., 2018). When approximately 100 (A) and 100 (B) cells are mixed together in a non-adhesive extracellular environment (ultra-low attachment, U-bottom plates), signaling induces (B) cells to become GFP-positive and preferentially adhere to one another. The (B) cells then deform to the extent where they cannot be individually distinguished from one another and, over time, form the inner layer of the spheroid. Blue (A) cells remain on the outer layer (Toda et al., 2018).

To replicate the *in vitro* experimental setup *in silico*, we defined cell signaling and response behaviors in L929 analogs that we call *in silico* L929 (ISL929); we then "mixed" these cells by initializing 20-200 cells in a 100x100x100 lattice containing an inert medium (to model ultra-low attachment plates), as a radially symmetric blob to maintain a consistent initial cell aggregate shape while also maintaining a similar cell total and ratio to that of the reference experiment (more info in the STAR methods, Computational Methods Details chapter). We then ran the simulation for 1000 monte carlo steps (mcs) per one hour of experiment time and follow the evolution of signaling and morphology (timescale

was determined by comparing the qualitative and quantitative space-time morphological evolution, and (A') and (B') activation rate, to that of the reference three-layer experiment, see below Fig. 3d-e).

To simulate the 2 layer spheroid in this setup, we programmed our *in silico* L929 (ISL929, see STAR Methods) cells with the same logic as the *in vitro* (Fig. 2b): (A) cells in contact with (B) cells send activation signals to (B) cells; (B) cells accumulate target genes over time if signaled to and with sufficient target gene level, excite to (B'). (B') cells adhere to other (B') cells and they can take a more compacted morphology, via relaxation of their spherical morphology constraints λ_{Sur} and λ_{Vol} (See STAR Methods for details). Starting with a mixture of approximately 100 (A) and 100 (B) cells, we consistently (n=3) obtained two-layer structures (Fig. 2b) qualitatively similar to that of the *in vitro* results (Fig. 2a). We noticed that this 2-layered structure could be generated for a number of values of adhesion matrix (not shown).

One way to quantitatively track sorting is to follow a homogeneity index (Flenner et al., 2008; Olimpio et al., 2018; Sun and Wang, 2013); we define a homogeneity index for a given cell type, based on the average percentage of surface area that the cells of that cell type share with either other of the same or of different cell types. This measure ranges from 0 to 1, with 1 indicating maximal homogeneity (see STAR Methods, Simulation Quantifications). We followed sorting over time by quantifying homogeneity index of cells (B)/(B') and (A) separately over time. The homogeneity index increases only slowly for cells of type A, reflecting the lack of active sorting for A cells, and instead increases dramatically for cells B/B', starting from around when the induction of transition B→B' happens (around 7000 mcs, Fig. 2b).

We noted that, from the *in vitro* system, it does seem that the reversion to the ground state is little or negligible. Computationally here we used an approximation where the active state B' never goes back to B. We compared, in the context of 3-layered structures below (Fig. 3c and S3), the results of simulations with this approximation with ones where the threshold for activation was the same as of induction, and we did not see statistic differences, which motivated us to keep this assumption of "irreversibility" for the positive-feedback based networks.

The Model Recapitulates Higher-order Synthetic Structures Qualitatively and Quantitatively

QUALITATIVE

While the two-layer structure is the result of a single forward network, we hypothesized that the model could easily be expanded to capture morphologies resulting from higher order signaling. We tested this hypothesis by extending the forward network to a back-and-forth network where (A) activates (B), then (B') activates (A). This network was used in vitro to generate both central-symmetric 3-layered structures (Fig. 3) and also non-central symmetric structures (Fig. 4), based on the choice of adhesion molecules.

One first central-symmetric 3-layered structure is the so-called 3-layered structure. *In vitro*, this is accomplished by modifying the forward network such that (B) cells express GFP-ligand that binds to αGFP-ligand synNotch on blue (A) cells, driving blue (A) cells' expression of mCherry and low E-cadherin. When around 200A cells are mixed with around 40B cells, the first step is the induction of B cells to B'; the B' cells are homotypically adhesive so they form the core, and they also signal to cells A to induce them to A'. (A') cells are moderately adhesive to (B') cells and weakly to (A') cells (Toda et al., 2018) (Fig. 3a, Development 0-20hr).

The implementation *in silico* follows the same design logic. The forward network used to generate the two-layer structure is modified such that the excited (B') state gains a ligand communication channel that can signal to (A) and (A') cells that express the cognate receptor. The adhesion matrix is modified such that the (A') cells are intermediately adherent to (B') and weakly to the rest of the cells (Fig. 3b). When we simulated the development of a system comprising around 200A and 50B cells, we

observed that there was first induction of B to B' cells that formed a core. Then the B' cells started to communicate to the A cells to turn red starting (Fig. 3b, Development). At the endpoint we observed structures similar to that of the *in vitro* results, a three-layer structure consisting of a green core surrounded by concentric shells of red, then blue (Fig. 3b and Movie S1). We also obtained similar structures when we started from different amount of cells A and B (Fig. S1 and Movie S2). In silico, we showed that signaling is necessary for three-layer formation: without excitation to (B') and (A'), no sorting occurred either qualitatively or quantitatively (Fig. S2). Of note, the three-layer structure has known regenerative capabilities *in vitro*, reforming the three-layer structure if bisected (Toda et al., 2018) (Fig. 3a). *In silico* three-layer structure also exhibited regeneration, but only partially (Fig. 3b and Movie S3). Reasons for this are addressed in the discussion.

At this point we carried out a fine-tuning of the parameters to qualitatively match the timelines of activation; we converged to model parameters for which 1000mcs = 1h by comparing the in silico and the in vitro timelapse stills. For example, the activation of green cells occur at around 7000mcs in silico, and the activation of red cells occur at 13,000mcs. This alignment was achieved by tuning the parameters.

Other, non-central-symmetric synthetic structures were generated with the back-and-forth network, differing in the signaling-induced adhesion proteins (Toda et al., 2018). Once we identified the signaling parameters that would best recapitulate the three layers structure, we moved to the other structures. We decided to keep the signaling parameters consistent between networks (same signaling parameters for all back-and-forth). Moreover, when E-cadherin was used again, we used the same numerical values across different simulations. For new cadherins, we estimated the value by best-guessing based on literature, and by fine tuning the similarity of the structures that we obtained to the in vitro structures. By modifying the adhesion matrix to reflect different cadherin-type adhesion and the corresponding spherical morphology constraints to reflect deformation in adhesive cells we were able to recapitulate the structures derived from the back-and-forth signaling network *in silico* with the correct timescale (Fig. 4a, Fig. S4-S5, Movies S3-7).

QUANTITATIVE

Quantifying morphogenesis is known to be relatively difficult. Nevertheless, we wanted to have a quantitative sense of whether our computational models were able to recapitulate the *in vitro* systems, apart from the qualitative observed likelihood of the images.

We decided to first focus on the target gene expression dynamics. We first measured the target gene induction over time in the in vitro system, from the timelapse movies; the % induction of GFP is the normalized amount of green fluorescence over time, and we can follow how that increases over time, starting from around 5-7h. For the *in silico* system, the quantification is on the activated cells, with the last time step normalized to 1. Fig. 3c shows the similarity between the induction dynamics in vitro and in silico, both for the green and the red channel. Interestingly, the similarities are not only in the induction time (which we obtained through parameter tuning), but also on the slope and shape of the curves, which have not been target of parameter optimization. This shows that our model can quantitatively recapitulate the time evolution of signaling over time of the in vitro system.

We wanted then to assess whether the morphological evolution over time in silico was similar to that in vitro. In vitro, we noticed from the published time-lapse of the three-layered structure development that the circularity of the structures in the 2D projection evolves over time to reach a steady state by the end of simulation; this was both true for the overall structure and of the cadherin-expressing cells (Fig. 3d). To quantify these features, we defined a circularity index in 2D and a sphericity index in 3D for the comparison of the in vitro and in silico structures (see STAR methods, Video Analysis). When we measured these morphology indexes over time, we found that they generated similar temporal evolution and convergence to a steady state by the end of simulations.

Lastly, we wanted next to assess how much the computational system is able to quantitatively capture robustness of the development, which is an emergent property of the biological counterparts. Although the synthetic development of the in vitro structures does not happen identically 100% of the times, it is reported to form a similar structure with one core in the 57% of the times (Toda et al., 2018). We wanted to see if our computational system, based on the stochasticity of the cellular potts evolution algorithm, would recapitulate this feature. To measure reproducibility of the *in silico* system, we quantified the number of cores formed over repeated simulations (n=30 simulations). We saw that the majority of the simulations yielded a 1-core structure, some a 2-core, and a minority a non-core structure (Fig. 3c). We compared this distribution to the distribution of morphologies obtained in the biological system (Toda et al., 2018), and found them similar (Pearson χ^2=0.24, d.f.=2, P>0.89), indicating that our in silico system can recapitulate reproducibility features of the in vitro cellular systems.

We also tested whether these quantitative analysis were dependent on the reversibility of the induction. A chi-squared analysis did not reveal a significant difference from the *in vitro* core distribution, Pearson χ^2 =4.75, d.f.=2, P>0.09. Quantitatively, the sphericity and activation timescale were similar to that of *in vitro* circularity and activation (Fig. S3b).

The Model Recapitulates Synthetic Structures Generated by Lateral Inhibition Circuits Starting from Genetically Uniform Cell Populations

Our model can recapitulate synthetic structures resulting from single or multiple levels of activation juxtacrine signaling. However, patterning and morphogenesis can also result from inhibition juxtacrine signaling as in the classic example of checkerboard patterning from lateral inhibition (Cohen et al., 2010; Collier et al., 1996; Ghosh and Tomlin, 2001; Shaya et al., 2017; Simakov, S A David and Pismen, 2013; Sprinzak et al., 2010, 2011). To capture inhibition juxtacrine signaling, we modified our model such that, S→-S (S is signal) and β→-β (threshold of signaling) to yield high signaling to low reporter production and low signaling to high reporter production (See STAR Methods). We tested the inhibition version of the model on lateral inhibition by generating the following network: red (A) cells send inhibition signals to neighboring red (A) and (A') green cells (Fig. S6a). The inhibition signal then inhibits red color, representing red reporter fused ligand and activates green reporter. Red (A) cells with sufficient red inhibition/green activation excite to (A') green. In order to allow reporter inhibition as required in lateral inhibition, we set the reversion threshold equal to the activation threshold. The network is predicted, starting from genetically uniform cells, to generate cell states differentiation. To test if this was true in our setup, we simulated development starting from a static, regularly shaped monolayer of red (A) cells, and we were able to obtain the classic checkerboard pattern of lateral inhibition (Fig. S6, Model, top). We then tested this lateral-inhibition network on a disordered cell monolayer where cells move slightly around their position, and begin with different sizes, don't grow and don't divide. By adjusting the signaling parameters to reflect strong inhibition (del Álamo et al., 2011; Sprinzak et al., 2010, 2011), we obtained clean checkerboard patterning (Fig. S6b, model, bottom). Our model also accounted for fate bias due to cell size, a recently documented phenomenon of lateral inhibition (Shaya et al., 2017). We found that smaller cells were significantly more likely to be of the sender fate (high Delta) and large cells more likely to be of the receiver fate (high Notch) (n=10 lattices of ~400 cells each, Delta cells: 108.82±1.59 pixels^2, Notch cells: 122.53±1.89 pixels^2, two-tailed matched t-test, d.f.=9, t=14.90, P<0.0001).

Finally, we added adhesion of our inhibition model to test whether it could recapitulate, *in silico,* the multicellular synthetic 3D structure demonstrated *in vitro* using lateral inhibition and adhesion. The *in vitro* network consists of red (A) cells expressing mCherry fused CD19 and αCD19 synNotch (Toda et al., 2018). When αCD19 synNotch is bound to mCherry fused CD19 on a neighboring cell, mCherry

fused CD19 expression is inhibited, while GFP and E-cadherin expression is activated. This yields green cells that adhere to other green cells, thereby forming a two-layer structure with a shell of red cells surrounding a green core (Toda et al., 2018) (Fig. 4b left and Fig. S6c left).

To obtain this network in vitro, we modified the lateral inhibition network model such that green (A') cells adhere to other green (A') cells (Fig.4b, Model, Adhesion Matrix). When we run the simulation over time, starting from an aggregate of around 100 cells, we observe that first some of the red cells become green, and then that the green cells meet each-other in the center of the aggregate, similarly to what happens in vitro (Fig. 4b and S6c, *In silico*). Over time in the *in silico* system we observe that these green cells do sometime revert to red, leaving in the center an active dynamic of green cells turning red and being moved to the external layer (not shown).

Synthetic Hollowing can be Achieved by Apoptosis

With evidence of biological replicability of multiple structures, we then generated and tested a series of new programs for interesting morphologies. We started with hollow structures due to their developmental importance (e.g. the blastocyst (Watson, 1992), tubes (Nelson, 2003), and clinical relevance (e.g. polycystic kidney disease (Qian et al., 1996)).

We reasoned that the two-layer structure could provide an ideal starting point for generating a hollowed structure, if we fate the cells in the core to apoptosis. A mixture of A and B cells would initially engage in communication so that B cells would convert to B' adhesive cells and form a core. After the core is formed, the cells of the core would die to leave space for a cavity. Figure 5a shows our ideal target trajectory.

We thought that the implementation could start from the two-layer network (Fig. 2) for the formation of 2 layers. For the simulation of cell death, we converted cells to "medium" type cells to physically conserve cavity volume. Otherwise, it would behave as a vacuum rather than a medium-filled void. When we tried the simple implementation where the green cells are fated to apoptosis, we encountered a number of problems: the cells committed apoptosis before forming a core, and if more time was allowed for apoptosis, then the inner core of B' green cells would not receive the further signals for transitioning to media (not shown). To resolve these two issues, we introduced two modifications. (1) To have a controlled timing of response, with first increase of adhesion, and then commit to death, we introduced two thresholds for activation: one, lower, for activation of adhesion; a second, higher, for cell death. (2) To sustain signaling among B' cells, we equipped B' cells with signaling capacity on top of receiving capacity (Fig. 5b). Therefore, (B) cells with sufficient reporter activation become (B') and then may, with further reporter activation, convert to medium.

When we simulated this network, the formation of a hollow structure was at first incomplete, as the outer layer did not coat completely the inner core (not shown). In order to achieve complete coating of the cavity, we introduced a couple of variations: (B') cells neither grow nor divide, and (A) cells grow and divide slightly faster (see STAR methods). This transformed the two-layer structure into a hollow structure with a simple hollowed shell of blue (A) cells 50% of the time in a total of 10 runs, when starting from a mixture of 121.6±7.28 blue (A) and 57.4±7.28 gray (B) cells (Fig. 5c).

We propose that such network could be generated, *in vitro,* with an extension of the 2-layer network, where the response of B cells include: E-cadherin, ligand, G1 arrest; and, with lower efficiency, cell death (Fig. 5d). The *in silico* simulations use ISL929 (*in silico* L929) with slightly increased growth/division rates, which could be simulated by experimental conditions under low-dose mitogens insufficient to override (B') G1 arrest from p21 (Fig. 5d).

Elongation can be Achieved by Modulating Motility, Growth, and Activation Timescale

We next aimed to achieve elongated structures. Elongated structures have yet to be achieved synthetically *in vitro* but, like hollow structures, bear marked relevance in developmental processes. Inspired by the developmental processes of somitogenesis (Gossler and de Angelis, 1997; Hester et al., 2011; McGrew and Pourquié, 1998), bone growth (Kobayashi et al., 2005; Li and Dudley, 2009), wavefront activation and fluid-to-solid transition (Mongera et al., 2018), we developed the goal trajectory in Fig. 6a. We planned to start from a mixture of two cell types, (A) and (B), which would first form two poles via adhesion-mediated sorting (step 1). A and B cells would then signal to each other to induce the activated cell types, (A') and (B') at the interface (step 2). Activated cells A' and B' would stop proliferating and acquire a more "solid-like" features that inhibit further movement and sorting. Activated cells A' and B' would acquire also a signaling capacity for cells of the same type (A' towards A and A', and B' towards B and B') that would induce activation. In this way, at the interface between blue cells and red cells, the activated red cells A' are signaling to inactive A cells (blue) to become red. As the inactivated cells A and B keep proliferating, we reasoned that these areas on either end with reserve A and B cells could be poles able to provide an engine for growth.

To implement this synthetic developmental trajectory, we thought of describing a symmetric system where two cell types A and B work symmetrically. In the inactive state, they have homotypic adhesion preference, and low heterotypic adhesion. They also signal to each-other to activate to A' and B' respectively. The active state, for example A', gains different features: a signaling capacity towards same cell type (both A and A' have the receptor for this new signal), reduced motility and inhibited growth and division (Fig. 6b).

Seeding a mixture of approximately 30 (A) and 30 (B) cells with a first parameter set (Implementation 1 in Fig. 6c) led to the formation of short elongated structures in ~50% of the simulations. Elongation terminated between 20,000-50,000 mcs with 100% loss of both poles (only one run had (B) cells remaining as a small spheroid isolated from the elongated structure) (Fig. 6c). We wanted to see if we could optimize the growth to obtain continuous growth. We suspected that unreliable elongation was due to initial conditions that did not support pole formation (step 1 in Fig. 6a). We therefore hypothesized that robustness could be improved by seeding with separate spheroids of (A) and (B). This simple change in initial condition led to consistent, improved elongation overall (100% bidirectional elongation) and in both cell genotypes in all runs, even though only 40% of the structures retained both caps (Fig. 6d). This suggests that the incomplete elongation observed with Implementation 1 was due to an incomplete interface formation. It also suggested a way to improve elongation: increasing time delay between sorting (step 1) and activation (step 2). We therefore generated an Implementation 2 of the same network, where the activation threshold for A to A' and B to B' were higher, and so they took longer to happen. With this implementation we observed markedly improved elongation, both overall (Fig. 6f), and comparison to genotype from an initial mixture of cells, 68.67±9.27 pixels with Implementation 2 vs 48.95±8.28 pixels with Implementation 1 (significantly longer, two-tailed t-test, $t=5.01$, $P<0.0001$). Furthermore, 80% of the structures obtained with Implementation 2 retained the capability to elongate by retaining the (A) caps and (B) caps (Fig. 6e).

In vitro, this can be achieved by requiring higher amount of red and green reporter, relative to associated proteins, to achieve the phenotypic transition, a similar method to that employed in the hollowing network. This would allow to start from uniformly mixed cells, which is an easier to obtain initial condition.

Strongly Inhibiting Receptor Expression Allows Morphological Oscillation In Silico

The previous structures were inspired by existing biological structures, but synthetic biology can be expanded to potential structures as well (Elowitz and Lim, 2010). We wanted to see if we could

generate the networks for a "limit-cycle" attractor for the dynamic, i.e. a morphological oscillator. We tested our model's capability to explore possible structures by focusing on a novel morphological oscillator, oscillating between a highly mixed (A) and (B) structure (low homogeneity) and a poorly mixed (A') and (B') structure (high homogeneity) (Fig. 7a). Previous studies suggest that the low homogeneity structure can be achieved via heterotypic adhesion (Brodland and Chen, 2000; Glazier and Graner, 1993; Togashi et al., 2011) and simulations performed in this study suggest the high homogeneity structure can be achieved via homotypic adhesion, but the network required to link the two is unknown. We reasoned that the two states could be implemented by a switch-like behavior in the adhesion repertoire of the A and B genotypes. In their inactivated state, A and B cells would have heterotypic adhesion molecules (e.g. ephrins, nectins (Togashi et al., 2011), Het.Ad.1a and Het.Ad.1b in Fig. S8a) that would favor the checkerboard pattern. In their activated state, heterotypic adhesion would be replaced by two different homotypic adhesion molecules that would favor the formation of two poles (e.g. N-cadherin and P-cadherin, Homot.Ad.1 and Homot.Ad. 2 in Fig. S8a). The switch behavior relies on a receptor that can both inhibit and activate target genes. *In vitro* this could be achieved with expression of two synNotches with the same extracellular domain, and two different intracellular domain, one for activation one for repression (Morsut et al., 2016). We reasoned that some form of signaling could generate continuous oscillation between the two morphologies without reaching a stable endpoint.

We started by testing a simple version of the network where (A) and (A') cells and (B) and (B') cells have the same signaling capacity. A cells signal to B cells to switch their adhesion repertoire and vice versa: in the basal state cells have high heterotypic adhesion; in the activated state both the cell types have high homotypic adhesion. Seeding a mixture of 30 (A) and 27 (B) cells led to a multistep trajectory as follows: initial formation of a highly mixed structure, activation of A and B cells to A' and B' respectively, formation of 2 poles. At this point, although the cells far from the interface reverted to the basal sates, the interface became interlocked as A' cells are signaling to B' cells to maintain them in an activated state, and vice versa (Fig. 7b, network 1). This structure was stable and demonstrated no clear oscillatory behavior, showing that morphological separation alone was not enough to induce the oscillatory behavior because (A') and (B') still signal to each other, stabilizing the interface.

The above results suggest that morphological separation is not sufficient to bring the cells back to the initial state. We next tested the case where signaling capacity of the activated cell states is different from basal state. First, we tried ligand repression in the activated state in order to make the cells less capable of signaling to each other in the active state, thus favoring a return to ground state (Network 2). Seeding a mixture of 30 (A) and 27 (B) cells yielded overdamped morphological oscillation, with the endpoint being (B) cells locking (A) cells into the (A') red fate (Fig. 7c).

To prevent this locked endpoint, we adjusted the network such that the inhibition signal inhibited receptor expression instead of ligand expression. This led to quicker and sustained transition of the excited state to the ground state. In addition, this adjustment forced signaling coupling to generate robust morphological oscillations (Fig. 7b network 3) that lasted at least 13 cycles (Fig. S8b) with minimal dampening. This suggests that receptor expression can be used to generate stable intermediate states where sorting can be reverted.

DISCUSSION

Multicellular self-organization is a key facet of development and tissue formation. Numerous recent advances in synthetic biology(Baeumler et al., 2017; Barnea et al., 2008; Conklin et al., 2008; Daringer et al., 2014; Hartfield et al., 2017; Morsut et al., 2016; Qudrat and Truong, 2017, 2018; Scheller et al., 2018) have made it possible to control this process, facilitating synthetic reconstruction of native morphogenic processes and enabling custom tissue development. However, reconstruction and design often rely on a lengthy trial and error process. An initial design is implemented biologically and

subsequently modified for improvement in an iterative and lengthy process. Computational models can provide support by allowing rapid implementation of various designs *in silico*. In this way, robust formulas can be identified and selected before being implemented experimentally, effectively catalyzing reconstruction efforts and enabling rational design. Here we present the first of such models, focusing specifically on synthetic juxtacrine signaling, and demonstrate its ability to replicate synthetic morphogenesis and facilitate synthetic design. We began by recapitulating known synthetic structures along with their key features: morphology, self-organization, variability and dynamics. We then used the model to test, improve, and propose potential designs for yet to be achieved structures in synthetic biology: hollow, elongated, and oscillatory.

Our model demonstrates biological faithfulness, replicating numerous facets of currently known high complexity mammalian synthetic structures: back-and-forth, symmetrical, asymmetrical, and lateral inhibition. Nevertheless, the model can still be further improved, as evidenced by the incomplete regeneration of the three-layer structure. Several reasons are possible for the lack of blue (A) cells at the bisected area *in silico* (Fig. 3b), with the first, but unlikely, consideration being the (A) adhesion parameters. In the simulations, (A) cells have minimal adhesion, preferentially adhering to cells rather than medium, but with no preference between cell types (see Fig. 3b matrix or Table S1). Therefore, a simple remedy is to alter (A) cells to bear differential adhesion to (A') and (B'), driving them to the damaged area. However, this is unlikely as it contradicts other *in vitro* images of the same circuit; the lack of a smooth (A) layer, especially evident from the 3D reconstruction of a three-layer in the reference experiment, along with the retention of a rounded morphology even when contacting (B') or (A'), strongly indicates that (A) lacks differential adhesion to other cell types (Toda et al., 2018). A more plausible explanation is that the computational model is an idealized version of the experiment, performed with machine-like precision difficult to achieve humanly. Immediate post-bisection structures *in vitro* have noticeable imperfections such as latching cells post bisection and prominent proximal cells in suspension (Toda et al., 2018); these cells can easily reattach to aid regeneration. With the model demonstrated to be capable of powerful biological replication under controlled conditions, the foundation is established. The next step is to deliberately introduce "imperfections" to further improve similarity to realistic experimental setups. More importantly, this improvement raises an interesting question; can these "imperfections" be used to our advantage in rational design, for example, improving robustness?

Our model also generated interesting observations when we tried to design novel developmental trajectories that have not yet been implemented *in vitro*. Going through phases of design-test-learn-redesign can elucidate which parameter sets confer robustness to various trajectories and which instead perform poorly.

In particular, when designing the elongating trajectories, we realized that without a fluid-to-solid transition we could not achieve directional elongation in our simulations. These types of transitions are shown to be at work during embryogenesis (Mongera et al., 2018). It would be powerful to understand the molecular underpinnings of these transitions and begin to control them *in vitro* in synthetic systems. We also observed that our first elongation network (network 1) was not very efficient in generating elongated structures; however, when we changed the initial conditions, the same genetic network did generate robust elongating structures. In and of themselves, boundary conditions could generate diversity, even without changes in the genetic program. During embryogenesis, initial conditions for a phase of morphogenesis are often dictated by a preceding phase of morphogenesis. For instance, the formation of two poles is the initial condition for subsequent elongation. By changing the dynamic of a previous phase of morphogenesis, we could affect a second phase without altering genetic program and the dynamic of the second phase directly.

As we attempted to create a trajectory leading to a stable oscillatory structure (Fig. 7), we learned valuable lessons from trajectory iterations that failed. We first thought that the strong homotypic adhesion within A' and B' cells by would be sufficient to separate the two group of cells so that they would not signal to each other. However, in the simulations we saw that the presence of an interface where A' cells and B' cells are still touching and signaling to each other prevented them from reverting back to the basal A and B states. We were able to overcome this when we had the communication between A and B cells change the network itself, so that A' and B' cells would have a reduction in signaling capacities. We made minimal progress when we removed the ligands from activated cells but saw more dramatic effects when we removed the receptors from activated A' and B' cells. This type of regulation seemed important to generate states that are basis for subsequent morphogenesis, and, together with the dynamics explored for the elongating structure, point to the need to pay attention to how the multistep developmental trajectories are linked together, and the dynamics of that as an engine of diversity, that might have been used during evolutionary times.

The model presented here enables rapid design development and testing for synthetic development, and is the first step in facilitating synthetic reconstruction (Santorelli et al., 2019). Further efficiency can be achieved by combining computational frameworks such as ours with machine learning algorithms (Briers et al., 2019). Algorithms could not only be trained to optimize parameters such as cell line, signaling network, and behavioral response, but could also incorporate subparameters such as: motility, proliferation, differentiability, juxtacrine and soluble morphogen signaling, mechanotransduction, adhesion, chemotaxis, and differentiation, to list a few. Such a framework would accelerate advances in synthetic biology on multiple fronts by expanding the breadth of testable designs, indicating directions for expansion within the current tool repertoire (e.g. lack of synthetic mechanotransduction), and revealing programs for targeted structures.

These frameworks and models will likely be employed as designed: for efficient rational design of a desired morphology. Therefore, "off-target' programs, programs that fail to yield a desired morphology, will inherently be viewed as less relevant compared to successful programs. Nonetheless, such "off-target" programs can still reveal unimagined structures alongside fundamental developmental rules. An interesting direction would be to utilize computational models to explore structures that can, but have yet to, exist. This was previously not possible due to the efforts required biologically, but is now feasible with computational models that can rapidly test programs and accurately reveal structures. In light of this possibility, we propose that computational models are not only methods for *identifying* potential programs for morphogenesis, but also methods for *proposing* programs for novel/unidentified structures.

We hope this is the first of many modular computational models and trust that with continued technological and biological advances each generation of model will improve computational performance.

MOVIES

FIGURE LEGENDS

Fig. 1. Concepts underlying the computational model (See STAR Methods for details and generalized model).

(a) On the left, representation of biological communication between cell pairs A and B. (A) cells express ligand (purple) and (B) cell express receptor (black). With contact (pink arrow), (B) cells receive signal (green) that triggers expression of the target gene. In the lower pair, the amount of ligands in cells A is higher, hence the signaling (green arrow) towards the target gene is stronger. On the right, the *in silico*

model shows a simplified representation of this process with parameters: ligand amount (L, purple boundary of in silico cells), surface area of contact (Φ, pink), and net signal (S, green arrow). In the schematic, *in silico* cells are multi-pixel objects with different levels of ligand and shared surface area. The cell pair at the bottom has a higher level of communication compared to the upper pair due to both a higher ligand level (L2>L1, shown as thicker border) and smaller surface area of contact (Φ2>Φ1, 2 pixels compared to 1). **(b)** Time evolution of target gene level in the receiving cell; cells A and B are first placed in contact for 100,000 steps of simulation to follow induction of target gene expression, and then moved far apart to stop signaling and follow decay of target gene. Two plots are shown for two different values of shared surface area Φ, with Φ2>Φ1; all other parameters are constant.
(c) Model representation of cell behavior state change. On the left, a sender cell (A) (purple) activates a receiver cell (B) (gray) to induce a target gene (green) that encodes for an effector protein. Over time, cell (B) accumulates target genes products, and at a certain threshold the effector gene product causes a cell state transition from (B) to (B'). To the right *in silico* representation of the state transition and communication relationship between cells (A), (B), and (B'). Orange curved arrows indicate state transitions. Corresponding ligand/receptor pairs indicate a communication channel from (A) to (B) that promotes the state change of (B) to (B'). **(d)** The graph shows the progression of target gene level over time for a (B) cell that is initially in contact with an (A) cell and is then isolated at 100,000 steps. Example thresholds for the excited state (5000 AU) and ground state transitions (2500AU) are shown as dotted horizontal lines. At the start, the (B) cell is in the basal state (black solid line), but when the target gene level passes the excited state threshold, (B) cell becomes a (B') cell. The (B') cell remains in the active state (green solid line) until target gene levels drop below the ground state threshold and reverts to (B) (line goes back to solid black).

Fig. 2. The *in silico* two-layer is similar to the *in vitro* two-layer.
(a) Biological implementation of two-layer circuit: diagram of (A) and (B) where (B) cells express GFP and E-cadherin in response to contact with (A). A cells constitutively express a blue fluorescent protein (BFP). Confocal fluorescent microscopy images, overlayed for green, blue and brightfield channels, at 1h after seeding of 100 (A) cells and 100 (B) cells and after 24h; images reproduced from (Toda et al., 2018).
(b) *In silico* implementation of the two-layer circuit. (A) cells, when physically contacting (B) or (B') cells, send activation signal to (B) and (B') cells, which induces target gene. In (B) cells, accumulation of sufficient target gene excites them to the (B') state. Adhesion matrix defines pair-wise adhesion preference on a scale 0-3 (0 minimal adhesion preference, 3 high adhesion preference); with this adhesion matrix indicated that (B') cells are strongly adhesive to other (B') cells but weakly to other (B) or (A) cells. On the right, shown are simulation renderings of a typical temporal evolution from t=1,000mcs to t=24,000mc. For this instance, initial conditions are a mixture of 92 (A) and 87 (B) cells. Still images of a midpoint cross section of the aggregate at t=1,000 and t=24,000mcs of representative simulation run are shown on the right. Scale bar is 17.5 pixels, approximately 100um. **(c)** Cell type homogeneity over time, as a measure of sorting. In green, homogeneity index for cell type B (both B and B'), and in blue for cell type A. Solid lines are the mean, shaded areas represent standard deviation (n=3). At around 7,000mcs the B cells accumulated enough target gene that allowed them to transition to B'. Time scale equivalence, 1,000 monte carlo steps (mcs) to 1 hour, and size equivalence, 17.5 pixels to ~100um, were obtained from images and movies of the *in vitro* three-layer structure (see Fig.3).

Fig. 3. The computational model captures qualitative and quantitative features of the synthetic developmental trajectory for formation of three-layered spheroids.
(a) Biological implementation of the back-and-forth network: cell A express constitutive BFP and CD19 ligand (round ligand), and conditional to anti-GFP synNotch also mCherry and low levels of E-cadherin.

Cell B express anti-CD19 synNotch that triggers expression of high levels of E-cadherin and GFP-ligand. The two inductions happen sequentially over time when A cells and B cells are mixed together, such that CD19-antiCD19 interaction happens first (Step 1), and GFPlig-antiGFP happens second (Step 2), only after GFP ligand is produced in Step 1.

For synthetic developmental trajectory, the time evolution of a mixture of 200 (A) cells and 40 (B) cells is shown with overlay pictures of bright field, blue green and red channel from confocal microscopy imaging, reproduced from (Toda et al., 2018). For the Regeneration, initial seeding is with 160 (A) with 80 (B); bisection was done with microguillotine, reproduced from (Toda et al., 2018).

(b) Computational implementation of back-and-forth network. (A) cells, when physically contacting (B) cells, send a signal (round ligand) to (B), which induces target gene. In (B) cells, accumulation of sufficient target gene excites them to (B') state. (B') cells become able to send a signal to cells (A) (square ligand), which initiate transition towards state A'. Adhesion matrix is shown on the right; it specifies that (A') cells are moderately adhesive to (B') green and weakly to other (A') red cells. All cells prefer binding to other cells than to media (black). Representative cross section of aggregates of an in silico synthetic developmental trajectory is shown below at the indicated time points; initial conditions are 200A cells and 50B cells. For the regeneration simulation, the initial condition is 24h timepoint of a synthetic development run started from 160A and 91B cells. Then, half of the cells were manually removed. Scale bar is 17.5 pixels (around 100um). Number of cells is approximate *in vitro*, and *in silico* is a feature that is not completely under control of the programming.

(c) Quantification of A' (red) and B' (green) activation *in silico* and *in vitro* followed over time for the duration of development of the three-layered structure. *In silico*, the activation index for A' cells is the number of activated cells over the total of A+A' cells normalized to be 100 at endpoint; similarly for activation index for B' cells. *In vitro* (solid lines), the activation index for A' cells is defined as the amount of green pixels, normalized to be 100 at the endpoint; similar for (see STAR methods, Video Analysis for details on thresholding) (n=30 simulations, n=1 for *in vitro*). We present mean±s.d. for the *in silico* results (dotted lines with standard deviations in the graph).

(d) Quantification of sphericity/circularity measures over the time development of synthetic and *in vitro* systems. In blue, all the cells are considered; in green only the activated (A') and (B') cells. Solid line is from *in* vitro measures; solid lines with shaded contours are from *in silico* measurements and represent mean and standard deviation interval respectively. *In silico* sphericity was rescaled to account for the cubic nature of the voxels (See STAR methods, Simulation quantifications for *in silico*, and Video Analysis for *in vitro* details). Vertical dashed line indicates time of (B') cells activation (n=30 simulations, n=1 for *in vitro*).

(e) Quantification of the number of cores formed over repeated simulations (n=30 simulations, n=28 for *in vitro*). Cores *in silico* are defined as contiguous assemblages of at least 5 cells.

Fig. 4. The model captures the formation of various synthetic structures.

(a) Gallery of different structures obtained *in vitro* (Toda et al., 2018) and the corresponding simulations. On top, the biological base framework (left) and the simulation network (right). They are the same for all 4 structures shown below. For the in vitro, con.gene # denotes a constitutively expressed transgene; tar.gene # are the target genes, induced upon signaling. For each structure shown are (from left to right): biological gene matrix that explains the specific genes present in the cells; biological confocal picture at endpoint reproduced from (Toda et al., 2018); model simulation section of the spheroid at the same endpoint; model adhesion matrix. Scale bar is 100um for *in vitro* and 17.5 pixles (approx. 100um) for *in silico*. The colors red and blue of the last row in the *in vitro* system have been switched when compared to the image published in (Toda et al., 2018), for keeping the color consistent with the *in silico* and the other trajectories. See Fig. S4 and S5 for more examples.

(b) Lateral inhibition with adhesion. On the left are the biological diagrams and cell behaviors over time, reproduced from (Toda et al., 2018). Starting from uniform population of red cells, a two-layered structure is obtained with the lateral-inhibition differentiation into red cells and green cells, plus adhesion-mediated sorting due to green cells expressing high Ecadherin.
On the right is the *in silico* version. (A) cells receive signal from neighboring (A) cells and send signal to neighboring (A) and (A') cells. The signals inhibit red color and activate green reporter. (A) cells, with sufficient red inhibition/green activation, excite to (A'). Schematics and adhesion matrix are shown above. Below, selected time points of a representative simulation run are shown. Scale bar is 17.5pixles =100um for *in silico*. See Fig. S6 for more details on lateral inhibition network.

Fig. 5. *In silico*, hollow structures can be achieved via two layers + apoptosis.
(a) Goal trajectory: start from an initial configuration of randomly mixed (B) gray and (A) blue cells, induction of (B) cells to (B') green, two-layer formation, and then cavitation.
(b) *In silico* network: the two-layer network is modified so that (B') cells can signal to (B) and (B') cells, facilitating the transition of (B) to (B') and (B') to medium M. Signal activates green color, representing green reporter associated to inducted ligand. A first threshold defines transition to (B'), a second, higher reporter threshold defines transition to medium M (see STAR methods for details). (B') cells neither grow nor divide. Adhesion matrix is shown on the right.
(c) Evolution over time of a representative simulation at the indicated time steps; shown are tilted cross section slices of the 3D aggregates.
(d) Proposed biological implementation of the *in silico* network. Compared to adhesion, ligand and growth arrest effectors, death effectors have a lower number of reporter sites (depicted as half-circles in the promoter cassette) and as such require higher threshold for induction.

Fig. 6. The *in silico* model identifies strategies for elongated structures.
(a) Goal trajectory: starting from a random mix of (A) gray and (B) blue cells, step 1 is formation of two poles, step 2 is activation of (A') red and (B') green cells at the interface, and step 3 is bidirectional elongation.
(b) *In silico* network. Signaling: (A) signals to (B) and vice versa via two independent channels (pointed and rounded ligands respectively); the signaling activates (A) to (A') red and (B) to (B') green after a threshold is reached. (A') red cells gain a new signaling capacity (triangular ligand) that can signal to other (A) cells as well. The same happens with (B') cells, which gain a squared ligand signaling capacity. Adhesion: (A) and (B) cells are homotypically adhesive and therefore do not adhere to each-other. (A') and (B') are highly adhesive both homotypically and to each-other, as expressed in the adhesion matrix. Other effector changes in cells (A') and (B') are: stopping cell division and growth, and decreasing motility, to simulate fluid-to-solid transition.
(c) Simulation result when starting from 40 A cells and 40 B cells with Implementation 1 (see text for details); representative image of endpoint result from a total of 10 runs (see (f) for quantification);
(d) Results of simulations with same parameters as Implementation 1, but controlled initial conditions. Initial conditions are as shown on the left: (A) and (B) cells are initialized as two separate spheroids of 36 inactivated cells each. Endpoint result of a simulation with the same parameters as Implementation 1, repeated n=10, average length=62.56±8.77 pixels
(e) Simulation results with Implementation 2 parameters, where activation threshold for transition A→A' and B→B' are higher. The elongation is slower but more robust.
(f) Quantification of length of simulated aggregates at 100,000 steps for Implementation 1 and Implementation 2. We present mean±s.d. (n=10). See Fig. S7 for details on how the length is measured.

Fig. 7. Oscillating morphologies can be achieved *in silico* by inhibiting receptor expression.
(a) Goal trajectory. Starting from randomly mixed A blue and B grey cells, a "limit cycle" is initiated: A and B cells are heterotypically adhesive, so they tend to maximize A-B contacts and form a checkerboard pattern; then (1), signaling between neighboring A and B cells induce them reciprocally to A' and B'; since A' and B' cells are homotypically adhesive, they sort to form two poles (2); at this point, signaling between A' and B' is minimized so they revert back to their basal states A and B (3); since basal states are heterotypically adhesive they should favor formation of a checkerboard pattern (4), thus re-initiating the cycle.
(b) Adhesion matrix common to all the attempted simulations. A and B are heterotypically adhesive whereas A' and B' are homotypically adhesive. Numerical parameters are set equal for the (A) genotype compared to (B) genotype. Cells neither grow nor divide here, and λ_{Sur} and λ_{Vol} are slightly increased compared to other simulations (see STAR methods, for details).
Network 1. (A) cells signal to (B) cells to induce them to B'; reciprocally, B signals to A to induce a transition to A'. Induced states A' and B' maintain all the signaling and receiving capacity of the basal states. Induced states change their adhesion repertoire (see adhesion matrix) to switch from heterotypic to homotypic adhesion.
Network 2, Signaling network is same as Network 1 but now activated states lose signaling capacity;
Network 3, Signaling network is same as Network 1 but now activated states lose signal reception capacity.
(c) Morphological evolution over time. Starting from 30 A cells and 27 B cells, the development is followed for 100,000mcs. Overall structure is followed by measuring the mean of the homogeneity indexes $\Psi_{B,B'}$ and $\Psi_{A,A'}$ over time; it reflects the quality of both separated structures and mixed structures when the total number of cells is kept relatively constant (n=1 for each attempt). Inserts are images from the simulated development at the corresponding timepoints.

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
CompuCell3D (CC3D) v3.7.8	(Swat et al., 2012)	RRID:SCR_003052
Mathematica v11.3.0.0	Wolfram Research	RRID:SCR_014448
ImageJ v1.52a	(Schneider et al., 2012)	RRID:SCR_003070
JMP Pro v14.0.0	SAS Institute	RRID:SCR_014242
Excel v1808	Microsoft	RRID:SCR_016137
General Juxtacrine Signaling Model (GJSM) in CC3D	This paper	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources or code should be directed to and will be fulfilled by the Lead Contact, Leonardo Morsut (Leonardo.Morsut@med.usc.edu).

COMPUTATIONAL METHOD DETAILS

742 CompuCell3D and the cellular Potts Formalism
743
744 We implemented our model in CompuCell3D (CC3D) v.3.7.8 (Swat et al., 2012), a modeling
745 software that allows simulation of cells and their behaviors using the cellular Potts formalism. By itself,
746 CC3D contains numerous built-in features for replicating *in vitro* cell behavior, several of which we
747 utilized either directly or adjusted via CC3D Python v.2.7.13 scripting according to manual v3.7.9. In our
748 model, we incorporated default features from CC3D such as surface area constraint, volume constraint,
749 cell division, adhesion, cell-cell surface contact, and cell types. We implemented custom cell motility, cell
750 growth, and cell signaling, as described below and in subsequent sections.
751 We defined cells as multi-pixel entities in 3D that physically act by performing "pixel copy
752 attempts" over simulation time steps (monte carlo steps, mcs). Performing "pixel copy attempts"
753 effectively moves and changes both cell geometry and position over time. These pixel copy attempts
754 succeed probabilistically, determined by the Boltzmann acceptance function, $P=e^{-\Delta H/T}$, where P is
755 probability of attempt success, ΔH is change in total effective energy of the system from all attempted
756 pixel copy attempts at the mcs t, and T is the cell motility (Swat et al., 2012).
757 Effective energy (H). Because we incorporated surface area constraint, volume constraint, and
758 adhesion, our total effective energy H at a given mcs t therefore takes the form,

$$H = \sum_{i,j} J_{\sigma(i),\sigma(j)} \left(1 - \delta_{\sigma(i),\sigma(j)}\right) + \sum_{\sigma} (\lambda_{Sur}(\sigma)\left(Sur(\sigma) - Sur_{Tar}(\sigma)\right)^2 + \lambda_{Vol}(\sigma)\left(Vol(\sigma) - Vol_{Tar}(\sigma)\right)^2)$$

759 as described in (Hester et al., 2011). The terms σ(i) and σ(j) denote the identity of the cells occupying
760 pixel sites i and j separately, with the Kronecker Delta limiting inclusion to only the cell interface. J is a
761 matrix that contains the contact energy per pixel of the boundaries while λ_{Sur} and λ_{Vol} constrain deviations
762 of a cell from the ideal surface area Sur_{Tar} and Vol_{Tar}, hereafter referred to as target surface area and target
763 volume, respectively.
764 J controls adhesion in cellular Potts. J represents a stability index: lower J makes for a more stable
765 state, which is then how you achieve stronger adhesion. Conversely, a higher J leads to weaker adhesion.
766 Throughout the manuscript (mainly in the figures), we use a grouped representation of the adhesion for
767 presentation simplicity; the exact values of J are shown in Table S2.
768 Cell motility (T)
769 Cell adhesion to environment is complexly linked to cell motility, and adhesion effects on
770 motility vary widely between different adhesion proteins and cell types (Gumbiner, 1996; Nieman et al.,
771 1999; Takeichi, 2011). In general, although clearly not all-encompassing, the adhesion abstraction is that
772 strong cell adhesion to environment tends to decrease cell motility (Alberts et al., 2002; Gumbiner, 1996;
773 Takeichi, 2011). We therefore defined motility as a function of a cell's environment (neighboring cells
774 and medium); so different cells can have different motiliy. Each cell's individual motility T_σ is:
775

$$T_\sigma = T_0 + \zeta \frac{\sum_{\sigma(i)_i,j} J_{\sigma(i),\sigma(j)} \left(1 - \delta_{\sigma(i),\sigma(j)}\right)}{Sur(\sigma)}$$

776 This formula iterates over each neighboring cell pixel and medium uniquely, but ultimately only the focal
777 cell type, neighboring cell types, and total contact with medium determine motility because J differs only
778 between types and is constant to medium. Categorizing environment by cell types and medium instead,
779 accomplished in CC3D via cell-cell surface contact feature and cell type index, we obtained a
780 computationally simpler approximate formula,

$$T_\sigma = T_0 + \zeta \sum_{k} \frac{J_{type(\sigma),k} * total\ contact\ surface\ area\ with\ \ k}{Sur(\sigma)}$$

T_0 is a constant representing basal cell motility, ζ a constant representing how effectual adhesion is at attenuating, and k denotes cell type (including medium here). This T allows each cell to sense its local environmental adhesiveness, decreasing motility if adhesive to neighbors and restoring motility when exposed to non-adhesive conditions.

In Silico L929 Cell Line Properties

In our model, each *in silico* L929 (ISL929) cell consists of multiple pixels and starts with a target radius (TR) randomly chosen using a Gaussian distribution (μ=3.0 pixels, σ=0.5 pixels). This TR is then used to calculate the target surface area ($4\pi r^2$) and target volume ($4\pi r^3/3$) for each cell, as *in vitro* L929 cells adopt a spherical shape when in suspension (Toda et al., 2018). Each cell then undergoes growth by experiencing net positive increase in TR from small positively skewed uniformly distributed fluctuations in TR. Target surface area and target volume thus increase slowly over time. Upon reaching a threshold volume, $2*4\pi\mu^3/3$, the cell then undergoes division, resulting in the original cell and a new cell. The original cell is subsequently reassigned a new TR from the above Gaussian distribution and both target surface area and target volume are recalculated. The new cell is assigned the same post-division parameters as the original cell.

I*n vitro* L929 mouse fibroblasts weakly adhere to one another under ultra-low attachment suspension conditions (Toda et al., 2018) thus we designate our ISL929 cells to have a relatively high J to one another and a slightly higher J to the medium, resulting in the formation of weak aggregates in medium. As a result, these ISL929 cells also bear high motility, again similar to *in vitro* L929 (Persson et al., 2010; Toda et al., 2018).

With the Potts model and growth model, we generated ISL929 cells with reasonable resemblance to *in vitro* L929 cells, favoring a rounded morphology, growing, roughly doubling in 24000 mcs (24 hours as estimated in (Toda et al., 2018)) being highly motile, and, when non-adhesive, forming non-compact aggregates with one another in suspension (Toda et al., 2018). Additionally, due to our stochastic implementation of growth, cell death occurs as well, incorporating yet another behavior of *in vitro* L929 cells. We also observed from the reference experiments that strongly adhesive L929 cells tightly cluster, deform markedly, and lose their rounded morphology (Toda et al., 2018). To roughly mimic this characteristic of adhesive L929 cells in our simulations, we relaxed the spherical morphology constraint such that for cells with an adhesion matrix value of at least 2 (see Table S2), λ_{Sur} and λ_{Vol} were set to 1.0. Other cells had λ_{Sur} and λ_{Vol} set to 2.2.

In our hollowing, elongation, and oscillation simulations, we modify some of these basic parameters to incorporate new behaviors and if so, give the changes in the respective sections.

Generalized Juxtacrine Signaling Model (GJSM)

Juxtacrine signaling is the method employed to achieve the known synthetic structures. For a generic signaling ligand whose expression was constitutive, constant, and unaffected by signaling, we describe the total ligand level, L, on a cell's surface by the equation

$$General: L = \frac{\gamma}{\eta + e^{-(t-\theta)/\xi}} \quad (1a)$$

$$Model\ Simplified: L = \frac{\gamma}{1 + e^{-t/\xi}} \quad (1b)$$

where t is the given time in mcs, while γ, η, θ, and ξ are constants. We chose this equation because of its generalizability. It can represent steady state ligand level on a cell's surface, recovery of surface ligand level from trypsinization, and experimental conditions such as ligand induction via tetracycline from a drug-controlled promoter (e.g. Tet On). Here we also give the simplified form used in our simulations.

Then, a receiver cell in contact with the sender cell would change its reporter level, R, by the differential equation

$$General: \frac{dR}{dt} = \frac{\tau}{\alpha + e^{-(S-\beta)/\varepsilon}} - \frac{R}{\kappa} \quad (2a)$$

$$Model\ Simplified: \frac{dR}{dt} = \frac{1}{1 + e^{-(S-\beta)/\varepsilon}} - \frac{R}{\kappa} \quad (2b)$$

where τ, α, β, ε, and κ are constants and S is signal strength. We chose this form for several reasons. First, parameters have intuitive interpretations: τ and α control maximal reporter synthesis due to S, β controls sensitivity to S, ε modulates magnitude of S and β, and κ represents the standard linear protein decay rate constant commonly employed in biological models. Secondly, these parameters have kinetic/biological interpretations, due to the logistic function's intrinsic relation to the Hill function (Reeve and Turner, 2013). Lastly, the logistic function is easily tunable and well behaved, due to its monotonicity from negative infinity to positive infinity and bound between 0 and τ. This tunability is not as easily achievable with the Hill function, where odd or fractional hill constants lead to the existence of singularities.

Signal strength S is affected by four primary factors in juxtacrine signaling: the number of receptors on the receiver cell, the number of ligands on each sender cell, the surface contact area between the sender and receiver cell, and the number of sender cells in contact. For a receiver cell σ with receptor level Ω and signaling neighbors SN this allows S to be defined as,

$$General: S_\sigma = \sum_{SN} Min[\Phi_{SN} * L\ , \Phi_\sigma * \Omega] \quad (3)$$

where Φ and L/Ω are the separate components reflecting surface contact fraction and ligand/receptor level, respectively. The Min function takes the minimum of the two values. Then, we define

$$\Phi_{SN} = \frac{contact\ surface\ area\ with\ SN}{Sur\ (SN)} \quad (4a)$$

$$\Phi_\sigma = \frac{contact\ surface\ area\ with\ SN}{Sur\ (\sigma)} \quad (4b)$$

The receiver cell thus receives a signal strength S as a function of both the ligand amount (L) present on a sender cell, the number of receptors (Ω) present on the receiver cell, and the surface area shared Φ with each sender cell, iterated over all neighboring sender cells. The model assumes homogeneity of ligand and receptor on the cell's surface.

In our biological replication simulations, we assumed the receptors are non-limiting (i.e. ligand induction or signaling mediated ligand inhibition), and in our exploration simulations, we utilized the case where receptor is limiting (i.e. signaling mediated receptor inhibition). We thus obtain

$$Ligand\ Limiting\ Model\ Simplified: S_\sigma = \sum_{SN} \Phi_{SN} * L \quad (5a)$$

$$Receptor\ Limiting\ Model\ Simplified: S_\sigma = \sum_{SN}\ \Phi_\sigma * \Omega \quad (5b)$$

Because these factors evolve over time, S is therefore a morphological dependent and time dependent function that evolves according to structure's spatial organization.

Some cell types send and receive signals. These cells have reporters that also function as, or are associated with, signaling ligands. For these cells, we set

$$General\ and\ Model\ Simplified: L = R \quad (6)$$

By assigning different combinations of these equations to *in silico* cells, we can generate different genotypes of cells. In CC3D, these different genotypes are coded as different cell types. To match the biological ligand receptor pair specificity, we set different types of cells to receive signal only from designated types of cells, reflecting the lack of cross-activation by synNotch (Toda et al., 2018). Thus, a neighboring cell that expresses a ligand L on its surface sends signal only to a neighboring cell with the cognate receptor. Likewise, a cell with receptor Ω can only receive signal from a cognate ligand expressing cell.

Quantized Signaling Inducible Output

To implement signaling inducible behavioral response, we further classify our cell signaling genotypes into states. We borrow notation from physics; cells of each genotype, if excitable, bear a ground state and an excited state or even multiple higher order excited states. Cells that pass the threshold would enter the excited state, with the excited state bearing different properties, such as color change, being adhesive, capacity to deform, or even different signaling/reception capacity (Fig. 1b). This quantized representation of cell behavior has been applied to great effect, though not with this notation, in other models (Anderson, 2005; Hester et al., 2011; Hutson et al., 2017).

Because the reference experiments primarily focus on signaling inducible adhesion with reporter, we utilize two states per genotype, ground and excited, in the biological replication simulations. The excited state bears a different color from the ground state, reflecting signaling induced reporter expression. J changes depending on adhesive strength and binding specificity that the cadherin types in the *in vitro* counterpart express upon sufficient signaling (see Table S1). It is also possible for a cell to fall from the excited state to the ground state due to loss of signaling, leading to the existence of a reversion threshold, in contrast to the activation threshold. Falling under the reversion threshold transitions an excited state cell to the ground state, reverting color and excited properties. It is of interest to note that for the reference experiments, the activation threshold need not necessarily equal the reversion threshold, as the adhesion protein and reporter can have different degradation rates since they are not fused. We tested both cases where the reversion threshold is zero and activation threshold equals the reversion threshold. We did not detect a clear difference between the two methods (See Results, Fig. 3, and Fig. S3). We therefore based our decision on the reference results, where no notable deactivation and loss of adhesion occurred (Toda et al., 2018) and thus set our reversion threshold to be zero in our replication simulations. In our hollowing, elongation, and oscillation simulations, we conjugated signaling to other types of output, such as changes in motility, growth, and tested the effect of different thresholds. This led to additional behavior differences between genotypes, their ground states, and their excited states. We give these changes in the respective sections.

Simulation Conditions

Our simulations employed two genotypes, usually both excitable, with the following notation: (A) as ground state of the first genotype, (A') as excited state of the first genotype, (B) as ground state of the second genotype, and (B') as excited state of the second genotype. We generated these genotypes in our replication simulations by programming ISL929 with the appropriate signaling network and behavioral response when excited (adhesion and/or color change), reflecting that of the *in vitro* counterpart. At the center of a 100x100x100 lattice, we seeded a mixture of (A) and (B) cells as a radially symmetric blob to maintain a consistent initial cell aggregate shape while also maintaining a similar cell total and ratio to that of the reference experiment. We then ran the simulation according to the timescale, 1000 monte carlo steps (mcs) per one hour of experiment time. Our timescale was determined by comparing the qualitative and quantitative space-time morphological evolution, and (A') and (B') activation rate, to that of the reference three-layer experiment (Fig. 3d-e).

We ran lateral inhibition patterning on a 100x100x5 pixel cell monolayer (~400 cells) for 20,000 mcs. Cells in inhibition networks (Fig. 4b, Fig.7, Fig. S6, and Fig. S8) began with a nonzero reporter as per reference experiment or as stipulated according to designed circuit.

QUANTIFICATION AND STATISTICAL ANALYSES

Simulation Quantifications

To determine how our simulations quantitatively compared to experimental runs for the three-layer structure, we quantified three measurements: core distribution, sphericity, and activation timescale. (B') green cells were visualized in 3D to determine core amounts and counted for each simulation at the endpoint. Sphericity was measured over time, both for excited states and over all states (Fig. 3d), using the formula (Cruz-Matías et al., 2019; Wadell, 1932)

$$Sphericity = \frac{\pi^{\frac{1}{3}}(6*Structure\ Volume)^{\frac{2}{3}}}{Structure\ Surface\ Area}$$

We roughly rescaled the sphericity by dividing by 0.48 to compensate for the cubic nature of the voxels. We measured activation timescale by measuring the number of (B') and (A') cells present per timestep and normalized each to 1 maximum.

We were also interested in the detailed spatial morphogenesis of these structures over time, thus we developed and quantified homogeneity degree Ψ per cell type x, calculated according to the formula below

$$\Psi_x = \frac{\sum_{\sigma x} \frac{contact\ surface\ area\ with\ x}{Sur\ (\sigma)}}{total\ cells\ of\ x\ in\ contact\ with\ x}$$

This measure ranges from 0 to 1, with 1 indicating maximal homogeneity, and is similar to sorting measurements employed in other studies (Flenner et al., 2008; Olimpio et al., 2018; Sun and Wang, 2013). Our measure has several additional desirable properties, the first being the ability to generalize beyond two cell types. In our simulations, we separate over each genotype (A vs B) and measure $\Psi_{A,A'}$ and $\Psi_{B,B'}$ to distinguish quality of structural homogeneity due to each genotype. If desired, this measure can be simply extended to the ground and excited states of each genotype as well, Ψ_A, Ψ_B, $\Psi_{A'}$, $\Psi_{B'}$, or condensed as desired, $\Psi_{A,A',B,B'}$, making it possible to distinguish the effects of different behaviors on morphogenesis. Second, our measure can be applied to many different morphologies, beyond fixed lattices (Olimpio et al., 2018) and spherical, their spatial organization over time, and is insensitive to single cells isolated from the focal structure.

Finally, to compare our elongated structures, we additionally measured the elongation length by estimating each length of the linear segments comprising the structure. Each length was calculated as the distance between the endpoints of each linear segments. Then, the sum of all linear segments per structure represents the elongation of each structure (Fig. S7).

Video Analysis

In vitro data was either provided in the reference paper or obtained by analyzing the supplementary video for the counterpart structure from the reference experiments (Toda et al., 2018). The video was split into constituent frames using Mathematica v11.3.0.0, then circularity analyzed by drawing a region of interest around the structure using ImageJ v1.52a, both in bright field (all cells) and merged color field (activated cells only), and data collated in Microsoft Excel v1808. Circularity was then calculated using the classic equation

$$Circularity = \frac{4\,\pi\,Area}{Perimeter^2}$$

To estimate how fast cells activated over time, we color separated the green, red, and black merged image portion of each frame by green and red to generate two sets of frames, one for green and one for red, representing respectively the activated cells of (B) and (A). We then converted these frames into binary images using the MorphologicalBinarize function in Mathematica, replacing pixels with an intensity above 0.1 with pixels of intensity 1. This threshold value was minimally low to remove non-cellular background fluorescence and prevent biasing activated cell detection. Binarization additionally facilitated comparison by splitting *in vitro* cells into discrete states. Totaling the pixel intensity for each

frame of each set estimates activation per timepoint for (B) and (A). Cellular background fluorescence, due to a few cells beginning with some green/red (Toda et al., 2018) was removed by subtracting the minimum background fluorescence of the time series. Using the minimum helped negate cellular background fluorescence with again minimal biasing of activated cell detection. This yielded two estimated activation curves over time, one for green and one for red, and each was then normalized to 1 maximum.

Statistical Analyses

Sample sizes are given in the text and/or figure caption. Statistical tests were performed in JMP PRO v14.0.0 with a significance level of 0.05. We performed a chi-squared analysis for our core distribution analyses (Fig. 3c and Fig. S3b). We performed a matched t-test between average (A) cell surface area and average (A') cell surface area per lattice, with 10 lattice replicates, for our inhibition signaling cell monolayer patterning (Fig. S6b bottom right). We performed a standard two-tailed t-test for comparing the lengths of elongation for Fig. 6f. Appropriate test was chosen according to data type and assumptions tested by residuals analysis. We report and show mean ±s.d. for all measures.

DATA AND SOFTWARE AVAILABILITY

All simulations were performed in CompuCell3D v3.7.8 with custom scripts coded in Python v2.7.13.

SUPPLEMENTAL INFORMATION

Fig. S1. Additional three-layer structures and their sorting dynamics.
(a) Biological back-and-forth network that leads to the formation of the three-layer structures and model implementation.
(b) *In vitro* structure from 200 (A) and 40 (B) as reference (Toda et al., 2018). Two additional structures from replicate simulations starting from a mixture of 202.2±4.33 (A) and 48.8±4.33 (B) are given. $\Psi_{B,B'}$ and $\Psi_{A,A'}$ are used to measure system sorting. We present mean±s.d.. (n=30 simulations).
(c) *In vitro* structure from 160 (A) and 80 (B) as reference (Toda et al., 2018). Two additional structures from replicate simulations starting from a mixture of 162.8±7.54 (A) and 87.2±7.54 (B) is given. $\Psi_{B,B'}$ and $\Psi_{A,A'}$ are used to measure system sorting. We present mean±s.d.. (n=10 simulations).

Fig. S2. Activation from signaling is required for three-layer formation.
(a) Network is disrupted so that (A) and (B) cells no longer signal and thus cannot excite one another to their respective activated state. Adhesion matrix remains unchanged, only signaling is disrupted, and is analogous to treatment with DAPT, which inhibits synNotch mediated signaling (Toda et al., 2018).
(b) Lack of structure formation starting from 201.33±6.02 (A) and 49.67±6.02 (B) due to loss of signaling that drives cadherin mediated sorting. Quantitatively, no cores form, no activation dynamics occur, no sorting occurs, but the structure still adopts an overall spherical shape, similar to the spheroids obtained *in vitro* under DAPT treatment (Toda et al., 2018). We present mean±s.d.. Two structures from replicate simulations shown. (n=30 simulations, n=1 *in vitro* experiment).

Fig. S3. There is no detectable difference in setting the ground state/deactivation threshold equal to the excited state/activation threshold and zero.

A and B cells initial communicate via ligand-receptor binding. They communicate back and forth until a threshold (same for both A and B?) is reached and they become either A' or B' cells, respectively. In the absence of ligand-receptor binding, A' and/or B' cells can revert to their previous state.
(a) Same code to obtain the three-layer structure in Fig. S1b and Fig. 3a. The only change is the deactivation/ground state threshold is set equal to the activation/excited state threshold.
(b) No detectable quantitative difference. We obtain a nonsignificant difference in core distribution from the *in vitro* core distribution (Toda et al., 2018) and obtain dynamics similar to *in vitro* dynamics as well. Structures are also qualitatively indistinguishable from simulated structures using the primary parameter set. We present mean±s.d.. Two structures from replicate simulations shown alongside reference (Toda et al., 2018) (n=30 simulations with initial mixture of 201.9±4.95 (A) and 49.1±4.95 (B), n=1 *in vitro* experiment).

Fig. S4. Additional structures for the symmetric structures (first two) in Fig. 4 gallery.
(a) Biological base framework and the simulation network common to the structures below.
(b) Two additional structures from replicate simulations of the first structure in Fig. 4 gallery. We also give the sorting dynamics as well as the time development. We present mean±s.d.. (n=10 simulations with initial mixture of 89.7±10.78 (A) and 89.3±10.78 (B), *in vitro* 100 (A) and 100 (B) (Toda et al., 2018)).
(c) Two additional structures from replicate simulations of the second structure in Fig. 4 gallery. We also give the sorting dynamics as well as the time development. We present mean±s.d.. (n=10 simulations with initial mixture of 202.2±4.49 (A) and 48.8±4.49 (B), *in vitro* 200 (A) and 40 (B) (Toda et al., 2018)).

Fig. S5. Additional structures for the asymmetrical structures (last two) in Fig. 4 gallery.
(a) Biological base framework and the simulation network common to the structures below.
(b) Two additional structures from replicate simulations of the third structure from 100 (A) and 100 (B) in Fig. 4 gallery (Toda et al., 2018). We also give the sorting dynamics as well as the time development. We present mean±s.d.. (n=10 simulations with initial mixture of 89.7±10.78 (A) and 89.3±10.78 (B)).
(c) Smaller biological structures from the third structure of Fig. 4 gallery, but with 30 (A) and 30 (B) (Toda et al., 2018), and the analogous simulated structures starting from 27.1±3.07 (A) and 29.9±3.07 (B). Our model captures variability in structure formation in more than just the three-layer structure. We present mean±s.d.. (n=10 simulations).
(d) Two additional structures from replicate simulations of the fourth structure in Fig. 4 gallery. We also give the sorting dynamics as well as the time development. We present mean±s.d.. (n=10 simulations with initial mixture of 163.4±7.63 (A) and 87.6±7.63 (B), *in vitro* 160 (A) and 80 (B) (Toda et al., 2018).

Fig. S6. The model captures inhibitory signaling to recapitulate well known patterns.
(a) Synthetic implementation of lateral inhibition (Toda et al., 2018).
(b) Checkerboard patterning. The model can capture the classic checkerboard patterning resulting from lateral inhibition, resulting in a typical red interspersed with green pattern (left drawing). Beginning with a fixed (A) cell monolayer (no differential adhesion, cells are constant in morphology, no growth, division, and motion), fates bifurcate to yield the classic checkerboard patterning (top right). On a disorganized cell (A) monolayer, no differential adhesion, cells still do not grow nor divide, but can slightly move and vary in size naturally (before fate bifurcation occurs). We found that smaller cells are more likely to take the red (Notch) fate while larger cells tend to be (Delta) green fate. (n=10 cell monolayers, one shown on bottom right).
(c) Two additional structures from replicate simulations of the lateral inhibition structure in Fig. 4b. We give the sorting dynamics and *in vitro* structure from 100 (A) as reference (Toda et al., 2018). We present mean±s.d.. (n=10 simulations with initial mixture of 93 (A)).

Fig. S7. Measuring length of structure for various elongation programs. We measured the length of the elongated structure by dissecting the structure into multiple linear portions, then estimated coordinates of the endpoints of the line traversing through the center of each portion. In this example, we split the structure into two linear portions with a line each. The coordinates of the endpoints of the line can be found via the cross-sectional slices, requiring 3 slices (a,b,c) in this example. Summing the length of each line estimates the overall length of the elongated structure. We focused only on the elongation from activated (A') and (B') cells.

Fig. S8. Potential biological versions of the oscillation structures.
(a) Possible biological implementation for each network.
(b) Sorting dynamics. Measuring mean of $\Psi_{B,B'}$ and $\Psi_{A,A'}$ indicates the oscillation consists of high mixing and low mixing structural switching as designed and remains stable for at least 300,000 simulation steps.

	AC Hi.Ecad	BL Hi.Ecad	AC Ecad	BL Ecad	AC Lo.Ecad	BL Lo.Ecad	AC Ncad	BL Ncad	AC Pcad	BL Pcad	C.Pcad
AC Hi.Ecad	20	40			35	49					40
BL Hi.Ecad		45			40	49					42
AC Ecad			25	42					33	45	
BL Ecad				47					45	49	
AC Lo.Ecad					40	49					
BL Lo.Ecad						49					
AC Ncad							35	42	49	49	49
BL Ncad								47	49	49	49
AC Pcad									35	42	
BL Pcad										47	
C.Pcad											43

Table S1: Adhesion parameters utilized in the model. Due to the definition of contact energy in CC3D, the larger the value, the lower the adhesion interaction. Adhesion parameters are related to the type of cadherin the cell expresses and two parameters exist per cadherin due to the step-wise approximation (ground state and excited state) in the model. Notation AC denotes increased cadherin levels due to activation and BL denotes baseline cadherin levels due to basal promoter activity. Hi. denote high expression, Lo. denotes low expression, and C. denotes constitutive expression per (Toda et al., 2018).

Adhesion Value (appears in matrix J)	Adhesion value (appers in Figures)	
49	0	Lowest adhesion
48-40	1	
39-30	2	
29-20	3	
19-10	4	
9-0	5	Highest adhesion

Table S2: Representation of adhesion values in the adhesion matrix. Adhesion values are grouped according the range above to simplify representation. 0 is the lowest, representing lack of cadherin mediated adhesion and 5 is the highest, representing very strong adhesion.

ACKNOWLEDGEMENTS

The authors are grateful to the developers of CompuCell3D and the users of the help forum for their assistance in learning the program. The authors would also like to thank Dr. Matt Thomson, members of the Thomson lab, members of the Morsut lab, and members of USC's stem cell department for feedback that improved the model. This project was supported by a National Institute of Biomedical Imaging and Bioengineering R00 to LM (4R00EB021030-03) along with a USC Department of Stem Cell Biology and Regenerative Medicine Startup Fund.

AUTHOR CONTRIBUTIONS

Conceptualization: C.L., L.M.; Methodology: C.L., L.M.; Software: C.L.; Validation: C.L.; Formal Analysis: C.L.; Investigation: C.L.; Resources: L.M.; Data Curation: C.L.; Writing - Original Draft: C.L., L.M.; Writing – Review and Editing: C.L., L.M.; Visualization: C.L., L.M.; Supervision: L.M.; Project Administration: L.M.; Funding Acquisition: L.M.

DECLARATION OF INTERESTS

L.M. is a co-inventor of synNotch, which was licensed to Cell Design Labs (acquired by Gilead), and receives royalty payments for this from UCSF.

REFERENCES

del Álamo, D., Rouault, H., and Schweisguth, F. (2011). Mechanism and Significance of cis-Inhibition in Notch Signalling. Curr. Biol. *21*, R40–R47.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Molecular Biology of the Cell (Garland Science).

Anderson, A.R.A. (2005). A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. Math. Med. Biol. A J. IMA *22*, 163–186.

Baeumler, T.A., Ahmed, A.A., and Fulga, T.A. (2017). Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. Cell Rep. *20*, 2639–2653.

Barnea, G., Strapps, W., Herrada, G., Berman, Y., Ong, J., Kloss, B., Axel, R., and Lee, K.J. (2008). The genetic design of signaling cascades to record receptor activation. Proc. Natl. Acad. Sci. *105*, 64 LP – 69.

Bashor, C.J., Horwitz, A.A., Peisajovich, S.G., and Lim, W.A. (2010). Rewiring Cells: Synthetic Biology as a Tool to Interrogate the Organizational Principles of Living Systems. Annu. Rev. Biophys. *39*, 515–537.

Belmonte, J.M., Clendenon, S.G., Oliveira, G.M., Swat, M.H., Greene, E. V, Jeyaraman, S., Glazier, J.A., Bacallao, R.L., and Edelstein-Keshet, L. (2016). Virtual-tissue computer simulations define the roles of cell adhesion and proliferation in the onset of kidney cystic disease. Mol. Biol. Cell *27*, 3673–3685.

Bosenberg, M.W., and Massagué, J. (1993). Juxtacrine cell signaling molecules. Curr. Opin. Cell Biol. *5*, 832–838.

1157 Briers, D., Ashley R.G., L., Haghighi, I., Joy, D.A., Conklin, B.R., Belta, C., and McDevitt, T.C. (2019).
1158 Self-Organized Pluripotent Stem Cell Patterning by Automated Design. Cell Syst. Sneak Peek.

1159 Brodland, G.W., and Chen, H.H. (2000). The Mechanics of Heterotypic Cell Aggregates: Insights From
1160 Computer Simulations . J. Biomech. Eng. *122*, 402–407.

1161 Cohen, M., Georgiou, M., Stevenson, N.L., Miodownik, M., and Baum, B. (2010). Dynamic Filopodia
1162 Transmit Intermittent Delta-Notch Signaling to Drive Pattern Refinement during Lateral Inhibition. Dev.
1163 Cell *19*, 78–89.

1164 Collier, J.R., Monk, N.A.M., Maini, P.K., and Lewis, J.H. (1996). Pattern Formation by Lateral Inhibition
1165 with Feedback: a Mathematical Model of Delta-Notch Intercellular Signalling. J. Theor. Biol. *183*, 429–
1166 446.

1167 Conklin, B.R., Hsiao, E.C., Claeysen, S., Dumuis, A., Srinivasan, S., Forsayeth, J.R., Guettier, J.-M.,
1168 Chang, W.C., Pei, Y., McCarthy, K.D., et al. (2008). Engineering GPCR signaling pathways with
1169 RASSLs. Nat. Methods *5*, 673–678.

1170 Cruz-Matías, I., Ayala, D., Hiller, D., Gutsch, S., Zacharias, M., Estradé, S., and Peiró, F. (2019).
1171 Sphericity and roundness computation for particles using the extreme vertices model. J. Comput. Sci. *30*,
1172 28–40.

1173 Daringer, N.M., Dudek, R.M., Schwarz, K.A., and Leonard, J.N. (2014). Modular Extracellular Sensor
1174 Architecture for Engineering Mammalian Cell-based Devices. ACS Synth. Biol. *3*, 892–902.

1175 Davies, J. (2017). Using synthetic biology to explore principles of development. Development *144*, 1146
1176 LP – 1158.

1177 Elowitz, M., and Lim, W.A. (2010). Build life to understand it. Nature *468*, 889–890.

1178 Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators.
1179 Nature *403*, 335–338.

1180 Flenner, E., Marga, F., Neagu, A., Kosztin, I., and Forgacs, G. (2008). Relating Biophysical Properties
1181 Across Scales. In Multiscale Modeling of Developmental Systems, S. Schnell, P.K. Maini, S.A. Newman,
1182 and T.J.B.T.-C.T. in D.B. Newman, eds. (Academic Press), pp. 461–483.

1183 Ghosh, R., and Tomlin, C.J. (2001). Lateral Inhibition through Delta-Notch Signaling: A Piecewise
1184 Affine Hybrid Model BT - Hybrid Systems: Computation and Control. M.D. Di Benedetto, and A.
1185 Sangiovanni-Vincentelli, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 232–246.

1186 Glazier, J.A., and Graner, F. (1993). Simulation of the differential adhesion driven rearrangement of
1187 biological cells. Phys. Rev. E *47*, 2128–2154.

1188 Gossler, A., and de Angelis, M.H. (1997). 6 Somitogenesis. R.A. Pedersen, and G.P.B.T.-C.T. in D.B.
1189 Schatten, eds. (Academic Press), pp. 225–287.

1190 Gumbiner, B.M. (1996). Cell Adhesion: The Molecular Basis of Tissue Architecture and Morphogenesis.
1191 Cell *84*, 345–357.

1192 Hartfield, R.M., Schwarz, K.A., Muldoon, J.J., Bagheri, N., and Leonard, J.N. (2017). Multiplexing
1193 Engineered Receptors for Multiparametric Evaluation of Environmental Ligands. ACS Synth. Biol. *6*,
1194 2042–2055.

1195 Hester, S.D., Belmonte, J.M., Gens, J.S., Clendenon, S.G., and Glazier, J.A. (2011). A Multi-cell, Multi-
1196 scale Model of Vertebrate Segmentation and Somite Formation. PLOS Comput. Biol. *7*, e1002155.

1197 Hutson, M.S., Leung, M.C.K., Baker, N.C., Spencer, R.M., and Knudsen, T.B. (2017). Computational
1198 Model of Secondary Palate Fusion and Disruption. Chem. Res. Toxicol. *30*, 965–979.

1199 Kobayashi, T., Soegiarto, D.W., Yang, Y., Lanske, B., Schipani, E., McMahon, A.P., and Kronenberg,
1200 H.M. (2005). Indian hedgehog stimulates periarticular chondrocyte differentiation to regulate growth
1201 plate length independently of PTHrP. J. Clin. Invest. *115*, 1734–1742.

1202 Lakatos, D., Somfai, E., Méhes, E., and Czirók, A. (2018). Soluble VEGFR1 signaling guides vascular
1203 patterns into dense branching morphologies. J. Theor. Biol. *456*, 261–278.

1204 Lambert, B., MacLean, A.L., Fletcher, A.G., Combes, A.N., Little, M.H., and Byrne, H.M. (2018).
1205 Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis. J. Math.
1206 Biol. *76*, 1673–1697.

1207 Li, Y., and Dudley, A.T. (2009). Noncanonical frizzled signaling regulates cell polarity of growth plate
1208 chondrocytes. Development *136*, 1083 LP – 1092.

1209 Li, P., Markson, J.S., Wang, S., Chen, S., Vachharajani, V., and Elowitz, M.B. (2018). Morphogen
1210 gradient reconstitution reveals Hedgehog pathway design principles. Science (80-.).

1211 Libby, A.R.G., Joy, D.A., So, P.-L., Mandegar, M.A., Muncie, J.M., Mendoza-Camacho, F.N., Weaver,
1212 V.M., Conklin, B.R., and McDevitt, T.C. (2018). Spatiotemporal mosaic self-patterning of pluripotent
1213 stem cells using CRISPR interference. Elife *7*, e36045.

1214 Lim, W.A. (2010). Designing customized cell signalling circuits. Nat. Rev. Mol. Cell Biol. *11*, 393–403.

1215 Lin, C.-M., Jiang, T.X., Baker, R.E., Maini, P.K., Widelitz, R.B., and Chuong, C.-M. (2009). Spots and
1216 stripes: Pleomorphic patterning of stem cells via p-ERK-dependent cell chemotaxis shown by feather
1217 morphogenesis and mathematical simulation. Dev. Biol. *334*, 369–382.

1218 Marin-Riera, M., Moustakas-Verho, J., Savriama, Y., Jernvall, J., and Salazar-Ciudad, I. (2018).
1219 Differential tissue growth and cell adhesion alone drive early tooth morphogenesis: An ex vivo and in
1220 silico study. PLOS Comput. Biol. *14*, e1005981.

1221 Massagué, J. (1990). Transforming growth factor-alpha. A model for membrane-anchored growth factors.
1222 J. Biol. Chem. *265*, 21393–21396.

1223 Massagué, J., and Pandiella, A. (1993). MEMBRANE-ANCHORED GROWTH FACTORS. Annu. Rev.
1224 Biochem. *62*, 515–541.

1225 McGrew, M.J., and Pourquié, O. (1998). Somitogenesis: segmenting a vertebrate. Curr. Opin. Genet.
1226 Dev. *8*, 487–493.

1227 Mongera, A., Rowghanian, P., Gustafson, H.J., Shelton, E., Kealhofer, D.A., Carn, E.K., Serwane, F.,
1228 Lucio, A.A., Giammona, J., and Campàs, O. (2018). A fluid-to-solid jamming transition underlies
1229 vertebrate body axis elongation. Nature *561*, 401–405.

1230 Morsut, L., Roybal, K.T., Xiong, X., Gordley, R.M., Coyle, S.M., Thomson, M., and Lim, W.A. (2016).
1231 Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. Cell
1232 *164*, 780–791.

1233 Mukherji, S., and van Oudenaarden, A. (2009). Synthetic biology: understanding biological design from
1234 synthetic circuits. Nat. Rev. Genet. *10*, 859–871.

1235 Nelson, W.J. (2003). Tube morphogenesis: closure, but many openings remain. Trends Cell Biol. *13*,
1236 615–621.

1237 Nieman, M.T., Prudoff, R.S., Johnson, K.R., and Wheelock, M.J. (1999). N-Cadherin Promotes Motility
1238 in Human Breast Cancer Cells Regardless of Their E-Cadherin Expression. J. Cell Biol. *147*, 631 LP –
1239 644.

1240 Olimpio, E.P., Dang, Y., and Youk, H. (2018). Statistical Dynamics of Spatial-Order Formation by
1241 Communicating Cells. IScience *2*, 27–40.

1242 Ollé-Vila, A., Duran-Nebreda, S., Conde-Pueyo, N., Montañez, R., and Solé, R. (2016). A morphospace
1243 for synthetic organs and organoids: the possible and the actual. Integr. Biol. *8*, 485–503.

1244 Persson, J., Mölder, A.L., Pettersson, S.-G., and Alm, K. (2010). Cell motility studies using digital
1245 holographic microscopy. p.

1246 Qian, F., Watnick, T.J., Onuchic, L.F., and Germino, G.G. (1996). The Molecular Basis of Focal Cyst
1247 Formation in Human Autosomal Dominant Polycystic Kidney Disease Type I. Cell *87*, 979–987.

1248 Qudrat, A., and Truong, K. (2017). Engineering Synthetic Proteins to Generate Ca2+ Signals in
1249 Mammalian Cells. ACS Synth. Biol. *6*, 582–590.

1250 Qudrat, A., and Truong, K. (2018). Antibody-Based Fusion Proteins Allow Ca2+ Rewiring to Most
1251 Extracellular Ligands. ACS Synth. Biol. *7*, 531–539.

1252 Reeve, R., and Turner, J.R. (2013). Pharmacodynamic Models: Parameterizing the Hill Equation,
1253 Michaelis-Menten, the Logistic Curve, and Relationships Among These Models. J. Biopharm. Stat. *23*,
1254 648–661.

1255 Santorelli, M., Perna, D., Isomura, A., Garzilli, I., Annunziata, F., Postiglione, L., Tumaini, B.,
1256 Kageyama, R., and di Bernardo, D. (2018). Reconstitution of an Ultradian Oscillator in Mammalian Cells
1257 by a Synthetic Biology Approach. ACS Synth. Biol. *7*, 1447–1455.

1258 Santorelli, M., Lam, C., and Morsut, L. (2019). Synthetic development: building mammalian multicellular
1259 structures with artificial genetic programs. Curr. Opin. Biotechnol. *59*, 130–140.

1260 Scheller, L., Strittmatter, T., Fuchs, D., Bojar, D., and Fussenegger, M. (2018). Generalized extracellular
1261 molecule sensor platform for programming cellular behavior. Nat. Chem. Biol. *14*, 723–729.

1262 Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image
1263 analysis. Nat. Methods *9*, 671–675.

1264 Shaya, O., Binshtok, U., Hersch, M., Rivkin, D., Weinreb, S., Amir-Zilberstein, L., Khamaisi, B.,
1265 Oppenheim, O., Desai, R.A., Goodyear, R.J., et al. (2017). Cell-Cell Contact Area Affects Notch
1266 Signaling and Notch-Dependent Patterning. Dev. Cell *40*, 505-511.e6.

1267 Simakov, S A David, and Pismen, L.M. (2013). Discrete model of periodic pattern formation through a
1268 combined autocrine–juxtacrine cell signaling. Phys. Biol. *10*, 46001.

1269 Sprinzak, D., Lakhanpal, A., LeBon, L., Santat, L.A., Fontes, M.E., Anderson, G.A., Garcia-Ojalvo, J.,
1270 and Elowitz, M.B. (2010). Cis-interactions between Notch and Delta generate mutually exclusive
1271 signalling states. Nature *465*, 86–90.

1272 Sprinzak, D., Lakhanpal, A., LeBon, L., Garcia-Ojalvo, J., and Elowitz, M.B. (2011). Mutual Inactivation
1273 of Notch Receptors and Ligands Facilitates Developmental Patterning. PLOS Comput. Biol. *7*, e1002069.

1274 Sun, Y., and Wang, Q. (2013). Modeling and simulations of multicellular aggregate self-assembly in
1275 biofabrication using kinetic Monte Carlo methods. Soft Matter *9*, 2172–2186.

1276 Swat, M.H., Thomas, G.L., Belmonte, J.M., Shirinifard, A., Hmeljak, D., and Glazier, J.A. (2012).

Chapter 13 - Multi-Scale Modeling of Tissues Using CompuCell3D. In Computational Methods in Cell Biology, A.R. Asthagiri, and A.P.B.T.-M. in C.B. Arkin, eds. (Academic Press), pp. 325–366.

Swat, M.H., Thomas, G.L., Shirinifard, A., Clendenon, S.G., and Glazier, J.A. (2015). Emergent Stratification in Solid Tumors Selects for Reduced Cohesion of Tumor Cells: A Multi-Cell, Virtual-Tissue Model of Tumor Evolution Using CompuCell3D. PLoS One *10*, e0127972.

Takeichi, M. (2011). Self-Organization of Animal Tissues: Cadherin-Mediated Processes. Dev. Cell *21*, 24–26.

Tigges, M., Marquez-Lago, T.T., Stelling, J., and Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. Nature *457*, 309–312.

Toda, S., Blauch, L.R., Tang, S.K.Y., Morsut, L., and Lim, W.A. (2018). Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science (80-.). *361*, 156 LP – 162.

Toda, S., Brunger, J.M., and Lim, W.A. (2019). Synthetic development: learning to program multicellular self-organization. Curr. Opin. Syst. Biol. *14*, 41–49.

Togashi, H., Kominami, K., Waseda, M., Komura, H., Miyoshi, J., Takeichi, M., and Takai, Y. (2011). Nectins Establish a Checkerboard-Like Cellular Pattern in the Auditory Epithelium. Science (80-.). *333*, 1144 LP – 1147.

Turner, D.A., Baillie-Johnson, P., and Martinez Arias, A. (2016). Organoids and the genetically encoded self-assembly of embryonic stem cells. BioEssays *38*, 181–191.

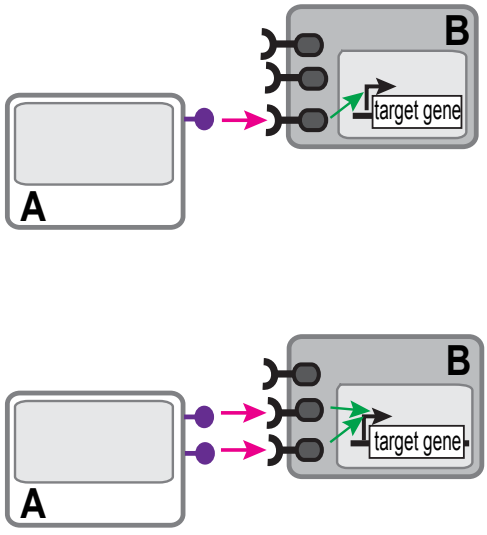
Wadell, H. (1932). Volume, Shape, and Roundness of Rock Particles. J. Geol. *40*, 443–451.

Watson, A.J. (1992). The cell biology of blastocyst development. Mol. Reprod. Dev. *33*, 492–504.

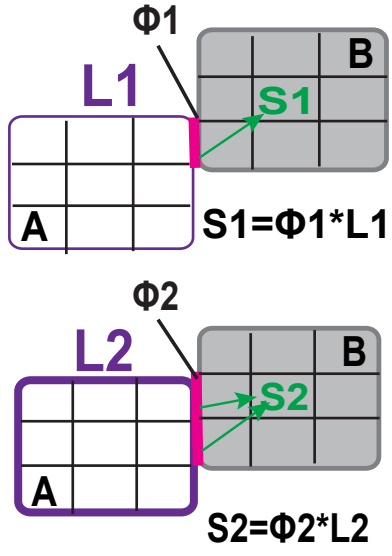
Fig. 1

(a)

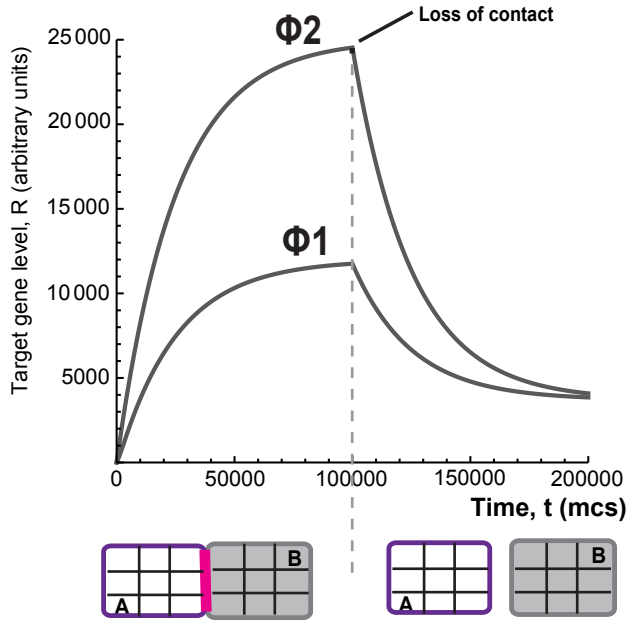
Biological



In silico

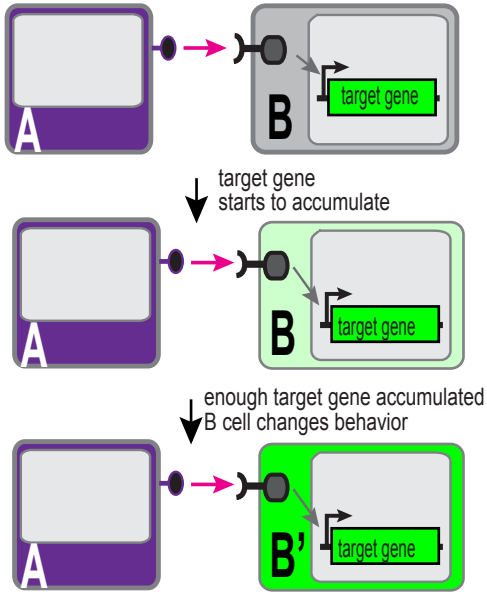


(b) Target gene induction over time

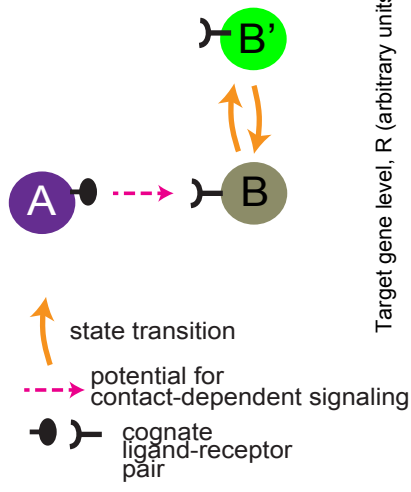


(c)

Biological



In silico



(d) Cell States Over Time

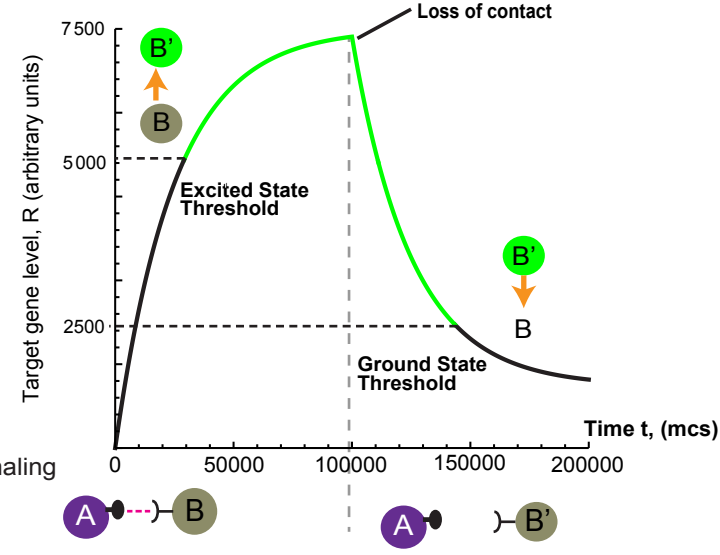


Fig. 2

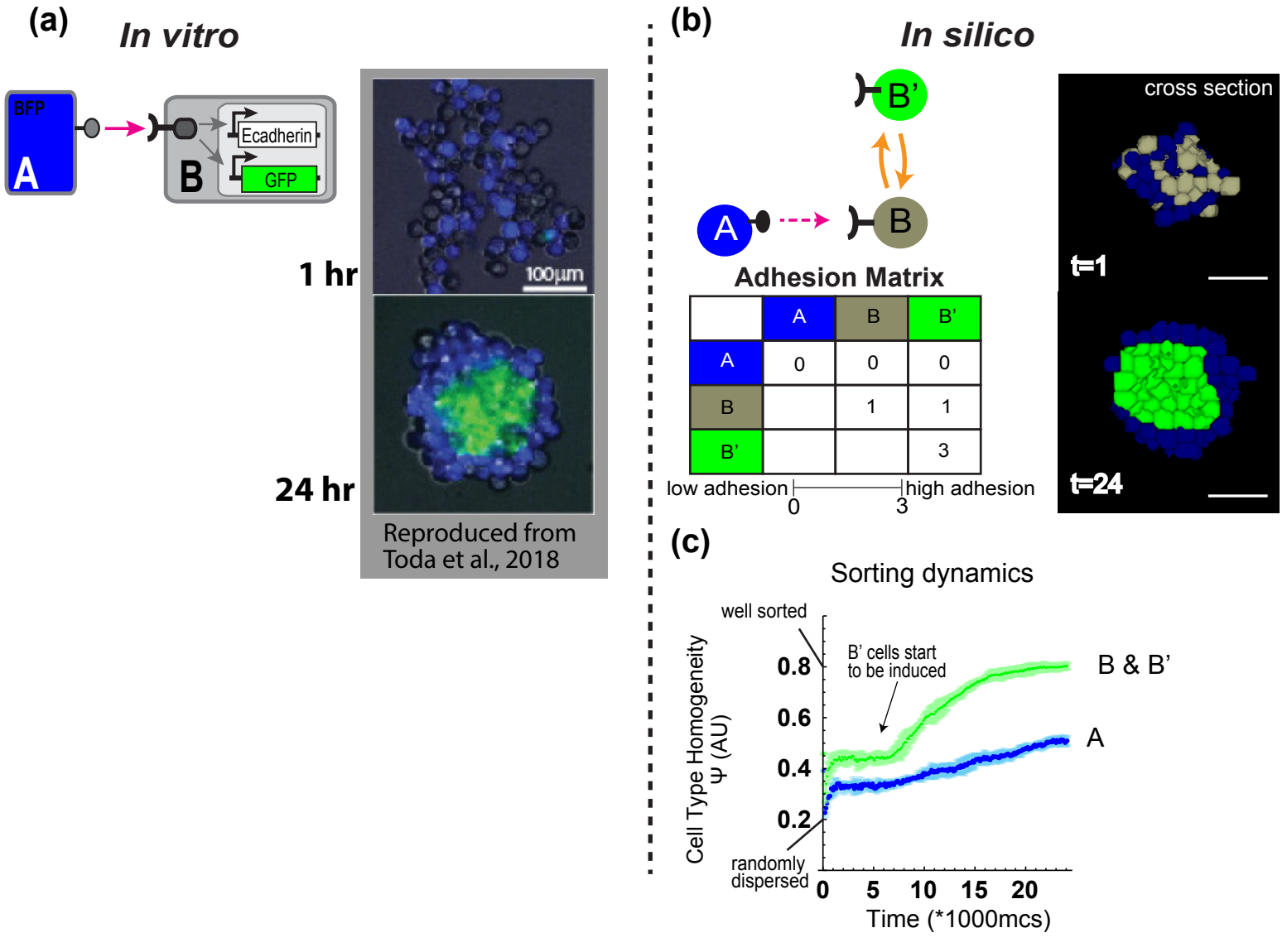
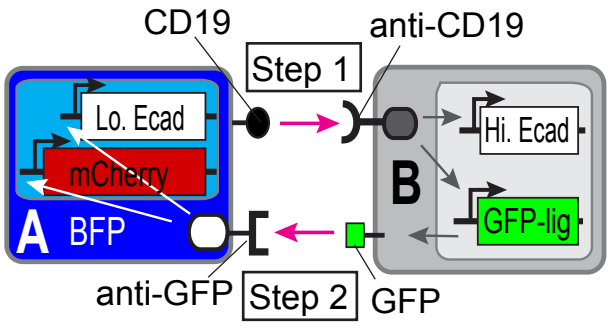
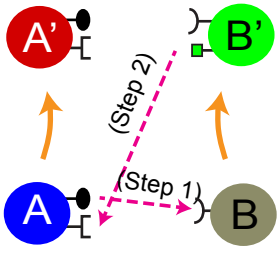

(a)
In vitro
BFP
A
B
Ecadherin
GFP
1 hr
24 hr
100μm
Reproduced from
Toda et al., 2018
(b)
In silico
B'
A
B
Adhesion Matrix
A
B
B'
A
0
0
0
B
1
1
B'
3
low adhesion
high adhesion
0
3
cross section
t=1
t=24
(c)
Sorting dynamics
well sorted
B' cells start
to be induced
randomly
dispersed
Cell Type Homogeneity
Ψ (AU)
0.8
0.6
0.4
0.2
0
5
10
15
20
Time (*1000mcs)
B & B'
A

Fig. 3

(a) *In vitro*



(b) *In silico*



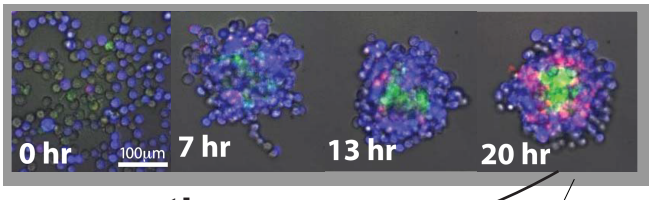
Adhesion Matrix

	A	A'	B	B'
A	0	0	0	0
A'		1	1	2
B			1	1
B'				3

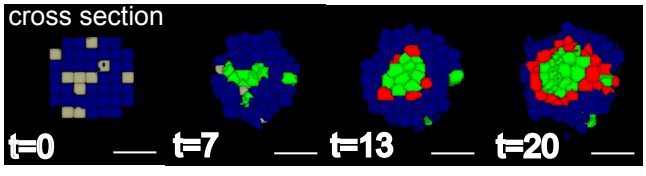
low adhesion 0 — 3 high adhesion

Synthetic developmental trajectory

In vitro



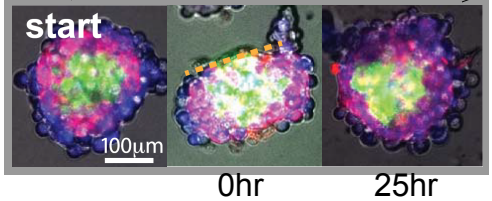
In silico

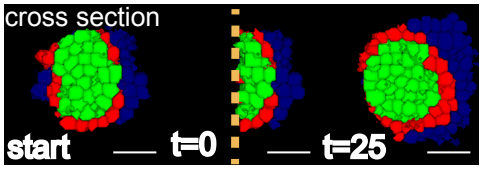


Regeneration

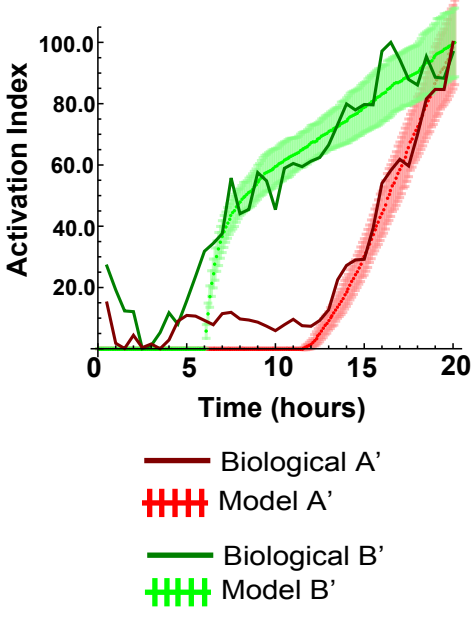
Reproduced from Toda et al., 2018

cut in half

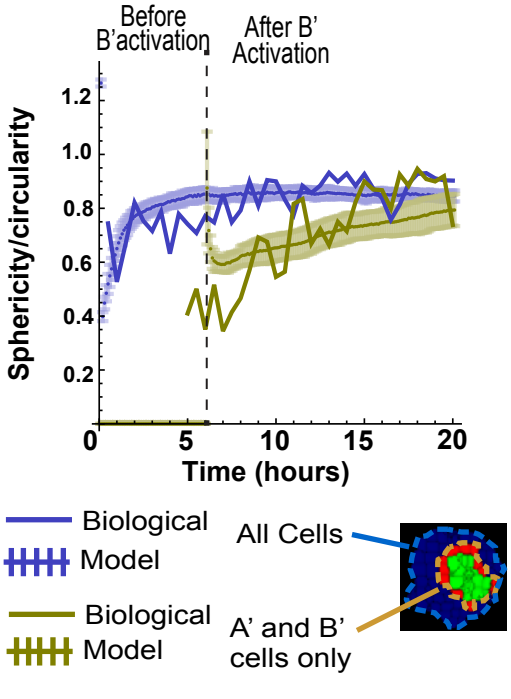




(c) **Cell activation dynamics**

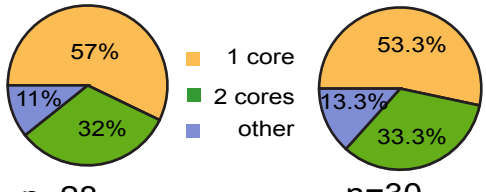


(d) **Morphology dynamics**



(e) **Core(s) Distribution**

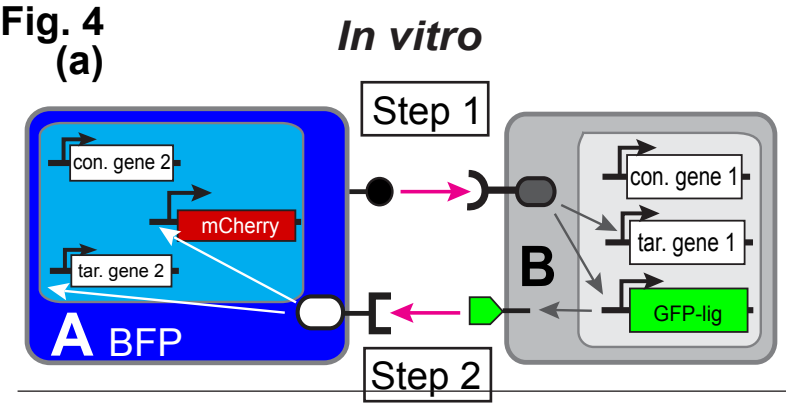
In vitro *In silico*

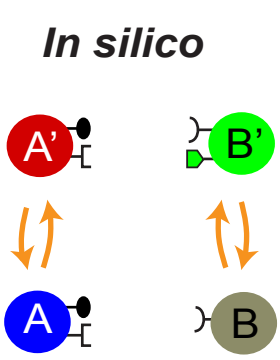


n=28 n=30

Fig. 4

(a)

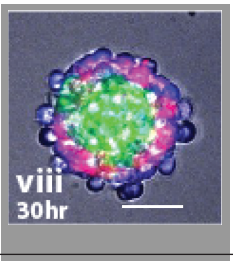


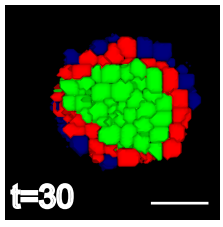


cross sections

Gene Matrix

cell B	con. gene 1	none
	tar. gene 1	Ecad
cell A	con. gene 2	none
	tar. gene 2	Pcad



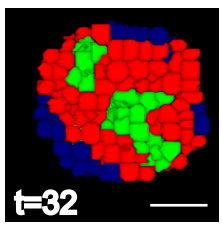


Adhesion Matrix

	A	A'	B	B'
A	1	1	0	1
A'		2	1	2
B			1	1
B'				3

Gene Matrix

cell B	con. gene 1	none
	tar. gene 1	Hi. Ecad
cell A	con. gene 2	Pcad
	tar. gene 2	none

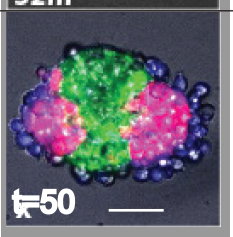


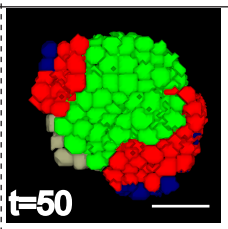
Adhesion Matrix

	A	A'	B	B'
A	1	1	1	1
A'		1	1	1
B			1	1
B'				3

Gene Matrix

cell B	con. gene 1	none
	tar. gene 1	Ncad
cell A	con. gene 2	none
	tar. gene 2	Pcad



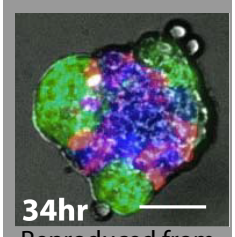


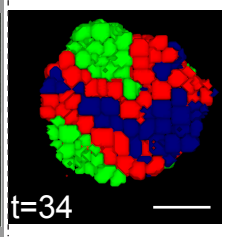
Adhesion Matrix

	A	A'	B	B'
A	1	1	0	0
A'		2	0	0
B			1	1
B'				2

Gene Matrix

cell B	con. gene 1	none
	tar. gene 1	Ncad
cell A	con. gene 2	Pcad
	tar. gene 2	none





Adhesion Matrix

	A	A'	B	B'
A	1	1	0	0
A'		1	0	0
B			1	1
B'				2

low adhesion — high adhesion

0 — 3

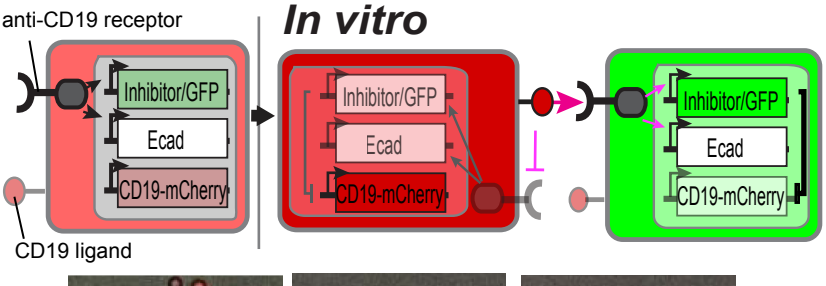
Estimated in vitro adhesion preferences:

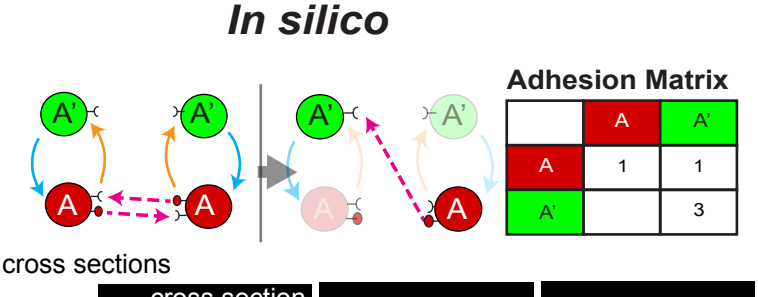
Pcad:Pcad = Ncad:Ncad > Pcad:Ncad

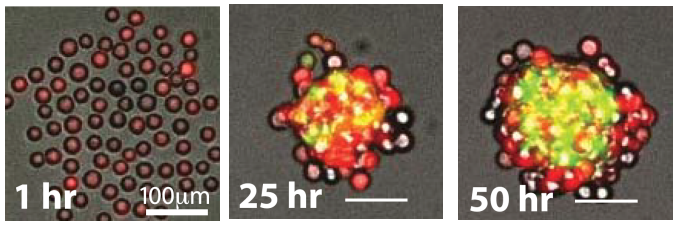
Hi. Ecad: Hi.Ecad > Hi. Ecad: Lo. Ecad > Lo. Ecad: Lo. Ecad

Reproduced from Toda et al., 2018

(b)







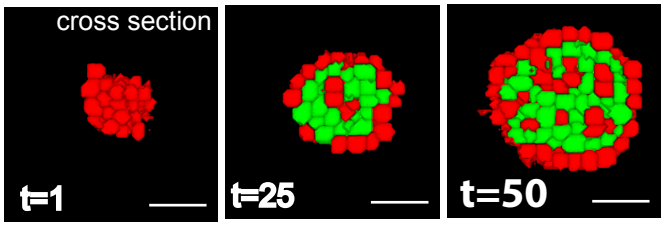
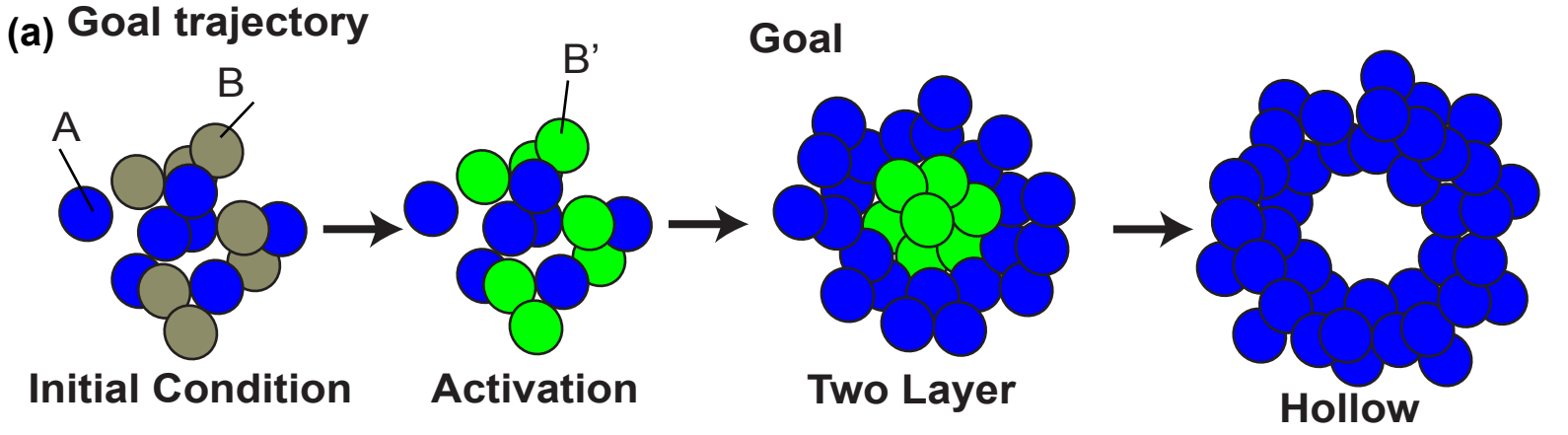
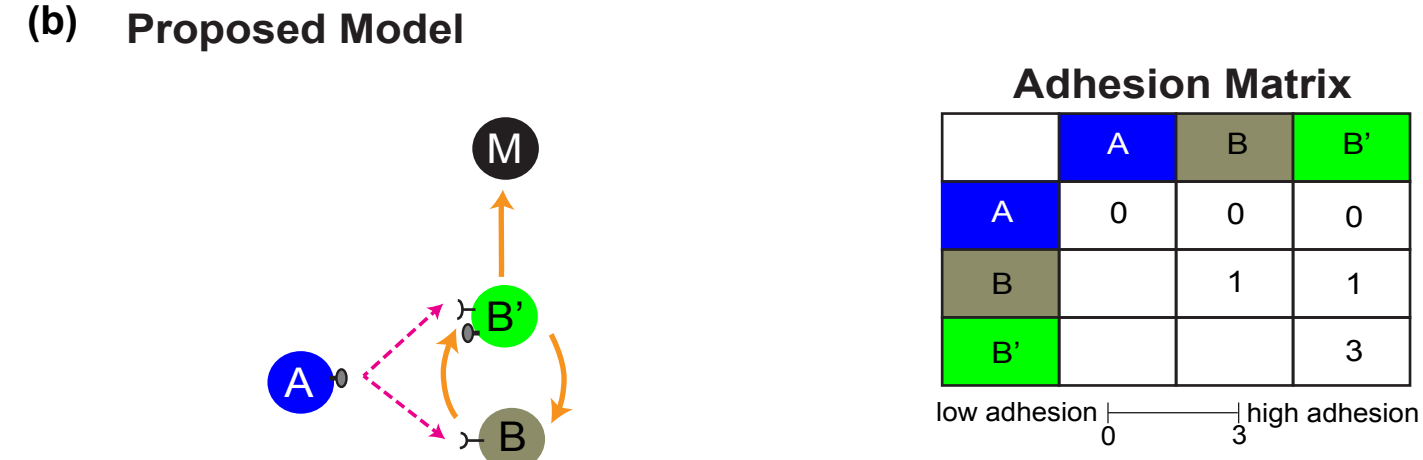
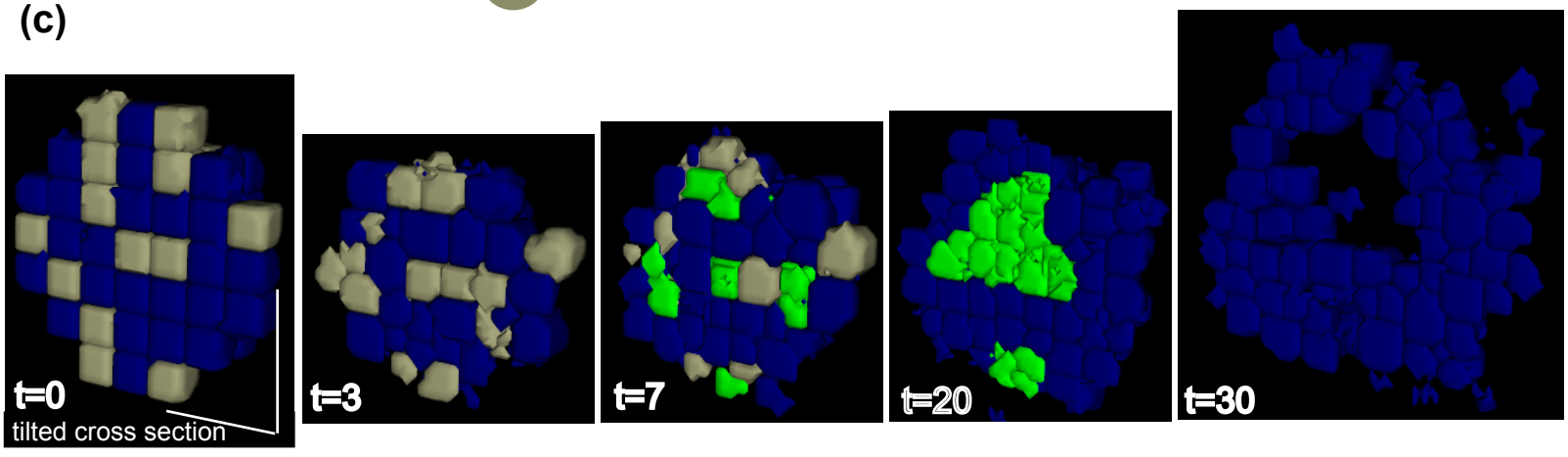


Fig. 5







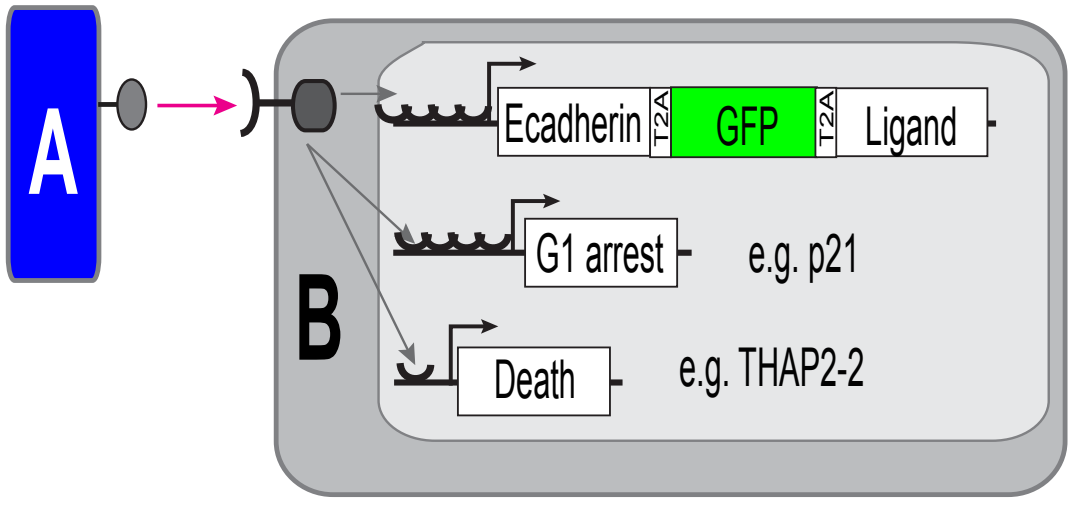
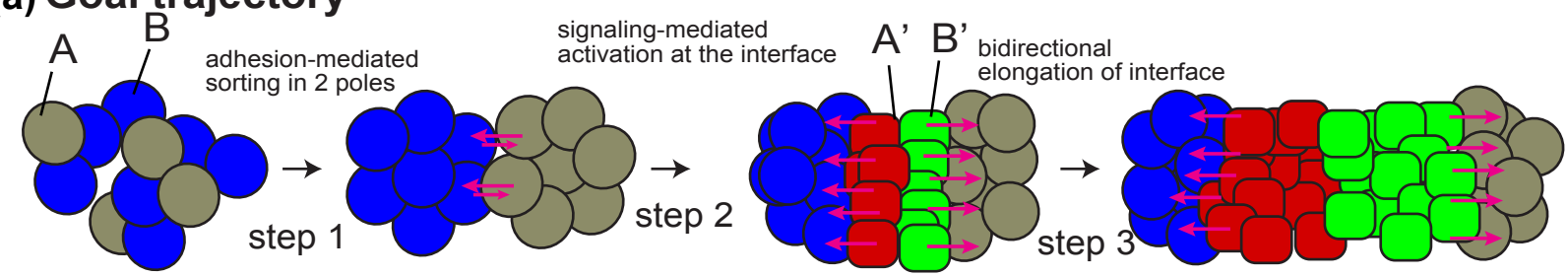
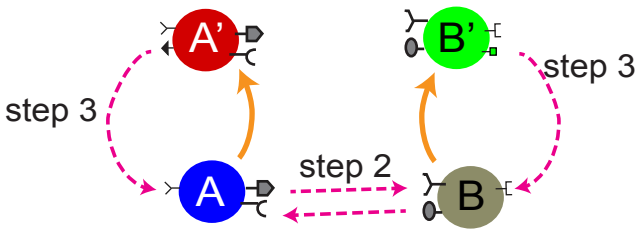


Fig. 6

(a) Goal trajectory



(b) Proposed Network

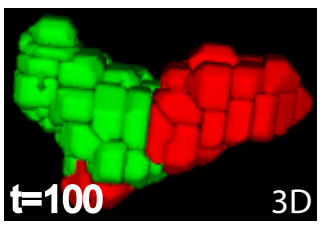


Adhesion Matrix

	A	A'	B	B'
A	2	0	0	0
A'		5	0	5
B			2	0
B'				5

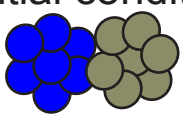
low adhesion — 0 — 2 — 5 — ultra high adhesion
high adhesion

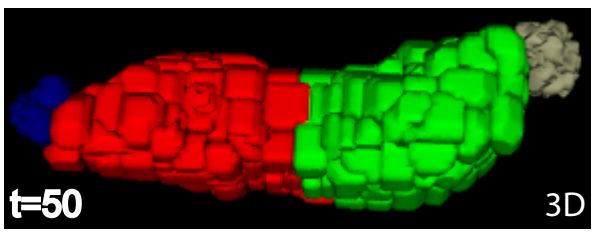
(c) Implementation 1



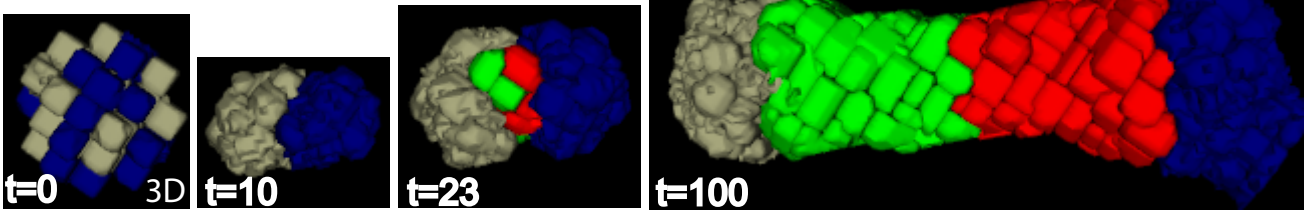
(d)

Implementation 1 with controlled initial condition

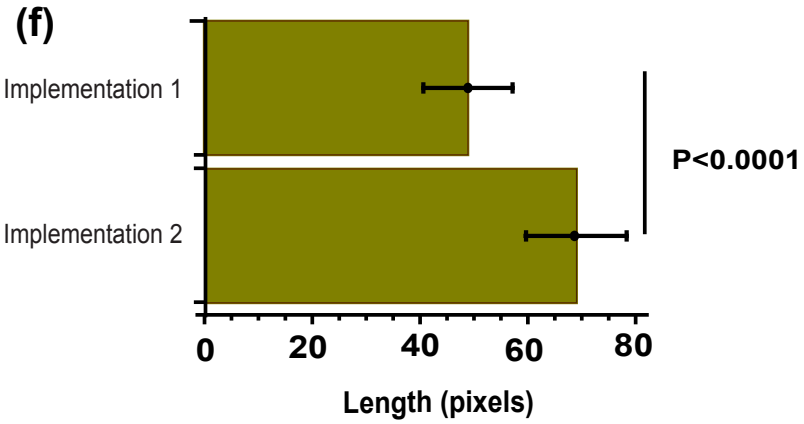


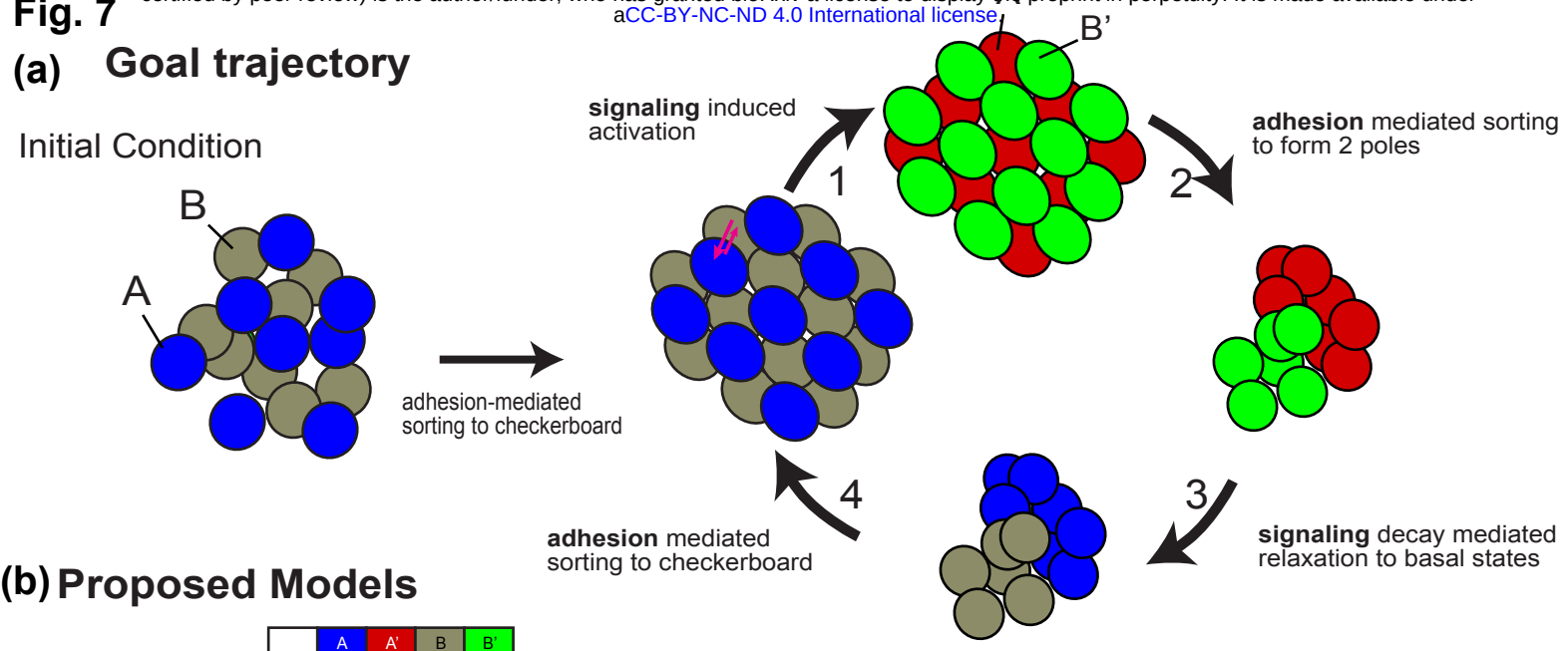


(e) Implementation 2



(f)





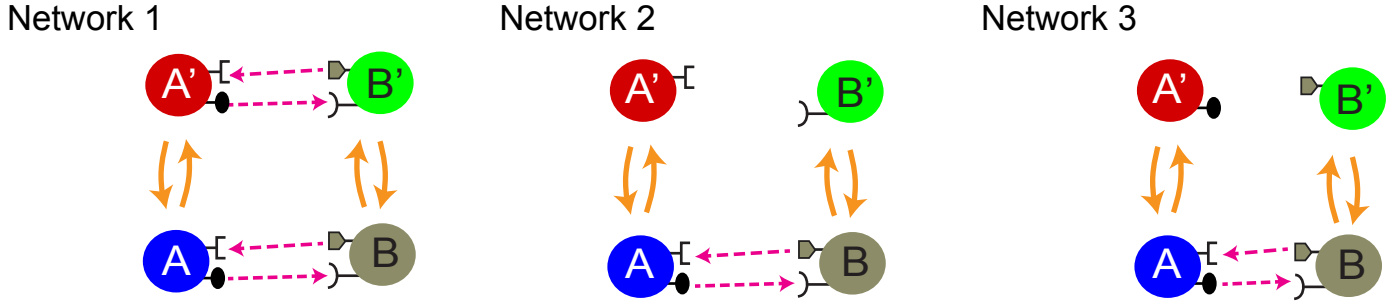


Fig. 7

(a) Goal trajectory

(b) Proposed Models

Adhesion Matrix

	A	A'	B	B'
A	0	0	2	0
A'		2	0	0
B			0	0
B'				2

(c)