A Modular Computational Framework for the Design of Multicellular Genetic Circuits of Morphogenesis

[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, such as mammalian cells. Synthetic genetic networks comprising cell-cell communications and morphogenesis effectors (e.g. adhesion) are generated and integrated into a cellular genome. Current design methods for these genetic programs rely on trial and error, which limit the number of possible circuits and parameter ranges that can be explored. By contrast, computational models act as rapid testing platforms, revealing the networks, signals, and responses required for achieving robust target structures. Here we introduce a computational framework, based on cellular Potts, where contact dependent cell-cell signaling networks and cellular responses can be chosen in a modular fashion. We represent and tune a number of recently described synthetic morphogenic trajectories *in silico*, such as those resulting in multilayered synthetic spheroids. Our parameters were tuned using a comparison with published *in vitro* experimental results. Our tuned parameters were then used to design and explore novel developmental trajectories for the formation of elongated and oscillatory structures. Here, multiple rounds of optimization suggested critical parameters for the successful implementation of these trajectories. The framework that we develop here could function as a testing ground to explore how synthetic biology tools can be used to create particular multicellular trajectories, as well as for understanding both imagined and extant developmental trajectories.

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 through 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. In the last two decades, synthetic biology has developed a variety of tools to reprogram and interrogate cellular behaviors in a reconstructive manner (Elowitz and Leibler, 2000; Li et al., 2018; Santorelli et al., 2018; Tigges et al., 2009). Thanks to advancements in synthetic biology, a new approach is now being proposed, called synthetic development. Synthetic development seeks to integrate synthetic genetic circuits into the cellular genome to control the stimuli they sense and their subsequent behavior to drive synthetic multicellular development (Davies, 2017; Ebrahimkhani and Ebisuya, 2019; Santorelli et al., 2019; Teague et al., 2016; Toda et al., 2018, 2019).

We recently showed an *in vitro* implementation of this idea in mouse fibroblast cells. We combined synthetic cell-cell contact dependent signaling with changes in cell adhesion. By adding synthetic signaling and adhesion circuits to the genome of non-assembling cells we achieved self-assembling patterned spheroids (Toda et al., 2018). As the use of cell-cell signaling interactions to conditionally induce morphological responses represents a common theme of a diverse set of developmental systems (Gilmour et al., 2017), genetic circuits based on similar abstractions could, in theory, be used to design other user-defined developmental trajectories. However, it remains unclear

exactly how to predict which genetic circuits are required for any given target structure. Currently, synthetic development programs are designed intuitively, based on literature search and preliminary experiments in simplified setups, before implementation in living cells. As a result, reconstruction attempts rely on a cycle of design, cell implementation, behavior monitoring, and redesign. This can be both time-consuming and costly, thereby limiting the exploration of circuits, behaviors, and experimental parameters critical for achieving desired structures and features. One of the most limiting steps of reconstructive approaches is design, which indicates our incomplete understanding of the underlying biology. A sensible initial guess could be shown to be erroneous after months of cell implementation work. In these situations where much is unknown, 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 or concurrently with, biological implementation, as was recently shown for simple genetic modifications and self-organization of stem cells (Briers et al., 2019). Here we develop a more general framework that can enable more complex genetic circuit engineering. Additionally, such a computational framework could both facilitate 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). Our work takes inspiration from the many examples of computational models developed and successfully used to 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), and brings them into the constructive viewpoint of design instead of understanding.

Here we describe the background and implementation of a computational framework for the design of synthetic development circuits. The framework allows for the design of synthetic cell-cell contact-dependent signaling networks with user-defined cellular inputs and responses. Our generalizable cell-cell contact dependent signaling model (named GJSM, generalized juxtacrine signaling model) can be modified to capture different types of signaling inputs and dynamics. We then describe GJSM implementation in a stochastic cellular Potts environment (Swat et al., 2012) and describe how signaling is paired with different behavioral outputs to mimic the modularity of the synNotch protein tool in an extended framework (Fig.1 and methods). We then describe the use of this framework to recapitulate synthetic structures like multi-layered spheroids (Fig. 2-4). Finally, we describe several *in silico* design-implementation-test cycles of novel synthetic development programs for other multicellular behaviors such as elongation and oscillation (Fig. 5-6). We believe that this type of framework will move the field towards a more deterministic and constructive understanding of developmental signaling logic.

RESULTS

Overview of the computational method (complete details in Methods)

Cell-cell signaling that relies on contact-dependence (also known as juxtacrine signaling) functions using the following 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) (Fig. 1a, Biological). When target gene products accumulate sufficiently, the behavior of responding cell (B) changes (Fig. 1c, Biological). Our computational model describes contact-dependent signaling and links it to various changes in cell behavior. This signaling framework is then embedded in the cellular Potts model to model stochastic cell behaviors evolution over time in 3D. To model juxtacrine communication, we conceptually separated it into two parts: (i) signaling-dependent continuous changes in target gene expression in receiver cells, and (ii) gene expression-dependent discrete

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

For modeling continuous changes in target gene expression, 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 contacting the receiver cell, 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). 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, *In silico*). Increasing the fraction of shared surface induces a stronger response (Fig. 1b). This part of the model accounts for the continuous changes in gene expression in receiver cells. For a more detailed explanation of the model, please see the "Generalized Juxtacrine Signaling Model (GJSM)" section of the Methods.

To model 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 (Fig. 1d). In this way, behavioral transitions are discrete instead of continuous, an approximation used by others, e.g. (Anderson, 2005; Hester et al., 2011; Hutson et al., 2017). Thresholds for state transitions from basal to active and from active to basal can be different and are defined as model parameters. The state machine schematic for this type of network is in Fig. 1c, *In silico*. For complete details, see "Quantized Signaling Inducible Output" paragraph of the Methods.

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

The Model Qualitatively Recapitulates Simple Synthetic Structures

We wanted first to show that our *in silico* model could recapitulate *in vitro* synthetic developmental trajectories such as that described in (Toda et al., 2018). In *in vitro* assays, mouse fibroblast cells (L929 line) are engineered with signaling networks and responses, then mixed together (20-200 cells) and grown for 2-3 days in ultra-low attachment U-bottom plates where their morphological and signaling evolution is followed via bright field and fluorescent microscopy.

To replicate this *in vitro* experimental setup *in silico*, we defined cell signaling and response behaviors in L929 analogs that we call *in silico* L929 (ISL929, specified in Methods, "*In Silico* L929 Cell Line Properties"). We then "mix" these cells by initializing 20-200 cells in a 100x100x100 lattice as a radially symmetric aggregate containing an inert medium to model ultra-low attachment plates (for more details see Methods "CompuCell3D and the cellular Potts Formalism"). We then ran the simulation for 1000 monte carlo steps (mcs) per hour of experimental time and followed the evolution of signaling and morphology by visualizing cell type and morphological evolution. Timescale correspondence between *in vitro* and *in silico* 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. 3c-d).

The first developmental trajectory that we modeled *in silico* is the two-layer spheroid (Fig. 2). *In vitro*, the two-layer spheroid is the result of a single forward network where CD19 on (A) blue cells binds to an anti-CD19 synNotch receptor on (B) gray cells to drive expression of green fluorescent protein (GFP) and E-cadherin, a homotypic adhesion protein (Fig. 2a). It should be noted that B cells, even in the basal state, express low levels of E-cadherin, due to leakiness. When approximately 100 (A) and 100 (B) cells are mixed together, signaling from (A) cells 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 core of the spheroid. Blue (A) cells remain distinct and form an outer layer (Fig. 2a, from (Toda et al., 2018)). No reversion of activated B cells to ground state is observed.

To simulate the two-layer spheroid *in silico*, we programmed our ISL929 cells with the same logic as the *in vitro* system (Fig. 2b) where (A) cells in contact with (B) cells send activation signals to (B) cells. In response, (B) cells accumulate target genes over time and with sufficient target gene level, excite to (B'). As in the *in vitro* system there is no observable reversion to the ground state, the active state (B') never goes back to (B). (Note: further discussion of reversion in the context of three-layer structure below (Fig. 3c-e and S3)). (B') cells are set to preferentially adhere to each other and to take a more compacted morphology. To model leakiness of cadherin expression in (B) cells, the homotypic adhesion between (B) cells is higher than that of (A) cells. Thus, adhesion levels are set to 0 for interactions between (A) cells and any other cell, to 1 for B-B and B-B' contacts, and to 3 for B'-B' contacts (Fig. 2b). (B') cells also have a higher deformability, achieved by setting λ_{Sur} and λ_{Vol} to 1.0 instead of 2.2. Details on the algorithm and parameters can be found in Methods, "*In Silico* L929 Cell Line Properties" and "CC3D and the cellular Potts Formalism" and full lists of simulation parameters can be found in Table S1-2 for adhesion and S3 for signaling.

With these parameters, and 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). One way to quantitatively track sorting is to follow a homogeneity index (Flenner et al., 2008; Olimpio et al., 2018; Sun and Wang, 2013). Here we define a homogeneity index Ψ for a given cell type based on the average percentage of surface area that that cell type shares with user-chosen cell types (Equation (10) in Methods, "Simulation quantifications section"). This measure ranges from 0 to 1, with 1 indicating maximal homogeneity. In this case, we followed sorting over time by quantifying the homogeneity index of cells (B) and (B') together (green line in Fig. 2c), and (A) (blue line). The homogeneity index increases slowly for cells of type (A), reflecting the lack of active sorting for (A) cells, but dramatically for cells (B)/(B') beginning with the transition of (B) to (B') (around 7000 mcs, Fig. 2c). From these data it can be noted that the slightly higher basal adhesion of (B) cells does not translate into dramatically higher homogeneity index (timesteps 0-7000 mcs).

The Model Recapitulates Higher-order Synthetic Structures Qualitatively and Quantitatively

The two-layer structure is the result of a single forward network, we next tested if the model could be expanded to capture morphologies resulting from higher-order signaling. We extended the forward network to a back-and-forth network where (A) activates (B) to an excited state (B'), which gains a signaling capacity to activate (A) to the excited state (A'). This network was used *in vitro* (Toda et al., 2018) to generate both central-symmetric three-layered structures (Fig. 3a) and also non-central symmetric structures (Fig. 4), based on the choice of adhesion molecule.

We first focused on the central-symmetric three-layered structure (Fig. 3). *In vitro*, the forward network of the two-layer structure is modified such that (B) cells also express a GFP-ligand that binds to anti-GFP synNotch on blue (A) cells, driving blue (A) cells' expression of mCherry and low E-cadherin. When around 200 (A) cells are mixed with around 40 (B) cells, (B) cells are induced to be homotypically adhesive and form a core, and also signal back to (A) cells inducing them to activate. Activated (A) cells are moderately adhesive to activated (B) cells and weakly to activated (A) cells (Toda et al., 2018) (Fig. 3a, Synthetic developmental trajectory *in vitro*, 0-20hr).

The implementation *in silico* follows the same circuit 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) cells that express the cognate receptor, so that (A) cells also can activate to an (A') induced state. 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 200 (A) and 50 (B) 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 (Fig. 3b). At later timepoints, 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 see also Movie S1).

As this was the best characterized structure *in vitro*, we used this synthetic developmental trajectory to carry out a fine-tuning of the computational system parameters. The goal was to match the timelines of activation such that 1,000 mcs = 1 hour. For example, the activation of green cells occurs at around 7 hours *in vitro* and we mapped it onto 7000 mcs *in silico* by comparing the *in silico* and the *in vitro* time-lapse stills and tuning the parameters, β, κ, ε, in the signaling equation (Equation (3) in the Methods, "Generalized Juxtacrine Signaling Model"). These parameters were used then for all the other simulations (including Fig. 2).

We next explored the robustness of the *in silico* development. First, we repeated the stochastic *in silico* development several times with similar initial conditions and obtained similar results also quantitatively following sorting with the homogeneity index (Fig. S1b). Second, we ran simulations starting from different amount of (A) and (B) cells, which showed morphologies similar to the *in vitro* ones (Fig. S1c and Movie S1). Third, we showed that signaling is necessary for three-layer formation: without signaling, there is no activation to (B') and (A'), no formation of core(s), and no sorting occurred either qualitatively or quantitatively *in silico* (Fig. S2a-b).

The three-layer structure has regenerative capabilities *in vitro*, reforming the three-layer structure if bisected ((Toda et al., 2018) and Fig. 3a, "Regeneration"). We tested capacity for regeneration *in silico* by removing half of the cells at t=20,000 mcs, and follow subsequent simulation behavior until 25,000 mcs. We observed regeneration of the spherical structure, and almost completely of the red lining, but only partially of the blue external lining (Fig. 3b and Movie S2). We noted that after bisection in the *in vitro* system some cells remain floating around the mass, we incorporated that observation into a new set of simulations and observed a more qualitative, three-layer appearance (Fig. S2c).

Quantifying morphogenesis is known to be relatively difficult. Nevertheless, we wanted to have a sense of whether our computational models were able to recapitulate the *in vitro* systems quantitatively as well as qualitatively. Initially, we decided to focus on target gene expression dynamics. We first measured target gene induction over time in the *in vitro* system from the time-lapse movies (Fig. 3c). We define an activation index *in vitro* as the normalized amount of fluorescence (See Methods, "Video analysis"). We also define and monitor an activation index in the *in silico* system where the activation index *in silico* is defined as the normalized ratio of activated cells over the total number of cells of the same type, i.e. #(A') / (#(A)+#(A')) and similarly for (B) and (B') cells. Figure 3c shows the similarity between the induction dynamics *in vitro* and *in silico*, both for (A)→(A') and (B)→(B') activations. Interestingly, the similarities are not only in the induction time (which we obtained through parameter tuning), but also in the slope and shape of the curves, which were not targets of parameter optimization. This shows that our model can recapitulate the time evolution of signaling in the *in vitro* system.

We then wanted to assess whether *in silico* morphological evolution over time was similar to *in vitro* morphological evolution. *In vitro,* we noticed from the published time-lapse of three-layered structure development that circularity of the structures in the 2D confocal slice evolves over time to reach a steady state by the end of the experiment. This was true for both the overall structure and for the cadherin-expressing cells. To quantify these features, we defined a circularity index in 2D that increases when the structure is more circular (see methods, "Video Analysis", Equation (11)), and quantified it over

time; results are in Figure 3d, blue and beige solid lines. To compare it with the *in silico* system, we defined a sphericity index in 3D for the comparison of the *in vitro* and *in silico* structures, which also increases with increasing sphericity. We measured this index over time, and plotted together with the *in vitro* results, and found that they generated similar temporal evolution (Fig. 3d).

Lastly, we wanted to assess how well the computational system was able to quantitatively capture robustness of 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 time, it is reported to form a similar structure with one core 57% of the time experiments were performed (n=28 experiments, (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). The majority of the simulations yielded a 1-core structure (53%), some a 2-core (33%), and a minority a non-core structure (13%) (Fig. 3e). 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 robustness features of the *in vitro* cellular systems.

Here we also tested whether these quantitative analyses were dependent on reversibility of induction. As discussed earlier, in the *in vitro* system we do not observe signs of reversion between active and ground states, but we reasoned that it may be hard to distinguish if it is truly zero. We simulated the development of spheroids *in silico* with reversion threshold set to equal the induction threshold, to simulate a situation where the activated cells would have a chance to get back to the ground state in absence of signaling. A chi-squared analysis did not reveal a significant difference from *in vitro* core distribution, Pearson χ^2 =4.75, d.f.=2, P>0.09. Qualitatively, the sphericity and activation timescale were similar to that of *in vitro* circularity and activation (Fig. S3b). This supported the use of reversion threshold set to zero for the networks discussed so far, i.e. ones based on positive feedback.

Other, non-central-symmetric synthetic structures were generated *in vitro* with the back-and-forth network by using different adhesion proteins (Toda et al., 2018) and Fig. 4a.1-4). With the parameter set that was able to qualitatively and quantitatively recapitulate three-layered structures, we tested our model on those other structures. We kept signaling parameters consistent between networks (same signaling parameters for all back-and-forth). Moreover, when E-cadherin was used *in vitro*, we used the same numerical values for the adhesion matrix across different simulations. For new cadherins, we estimated the value based on the *in vitro* results and comparison with E-cadherin values estimated previously. For the complete list of pair-wise adhesion parameters see Table S1. By modifying the adhesion matrix to reflect different cadherin-type adhesion, we were able to qualitatively recapitulate the structures derived from the back-and-forth signaling network *in silico* with the correct timescale (Fig. 4a, Fig. S4-S7, Movies S3-6).

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

Our *in silico* model could recapitulate synthetic structures resulting from single or multiple levels of activation juxtacrine signaling. During development, patterning and morphogenesis can also result from inhibition juxtacrine signaling. One classic example is the checkerboard patterning that results from lateral inhibition, where Notch receptor activation represses Delta ligand expression (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) see also Fig. S8a, biological. To capture inhibition juxtacrine signaling, we modified our signaling equations such that, $S \rightarrow -S$ (S is signal) and $\beta \rightarrow -\beta$ (threshold of signaling) to

yield an inverse relationship between signaling and reporter production, such that with low signaling receiver cells have high reporter production (See Methods, Equation (3) in "Generalized Juxtacrine Signaling Model (GJSM)"). We tested the inhibition version of the model on lateral inhibition by generating the following network: red (A) cells send and receive inhibition signals to/from neighboring (A) cells (Fig. S8a, *In silico*); the inhibition signal then inhibits red color and activates green reporter; (A) cells with sufficient red inhibition/green activation excite to (A') green. Starting from genetically uniform cells, the network is predicted to generate cell state differentiation. To test if this was true in our setup, we simulated development starting from a static, squared lattice monolayer of red (A) cells that neither move, grow, nor divide, and were able to obtain the classic checkerboard pattern of lateral inhibition (Fig. S8b), keeping all the parameters the same as for the simulations described in Fig. 2-4. When we tested this lateral-inhibition network on a monolayer of "disordered" cells where cells move slightly around their position and begin with different sizes we obtained cell types differentiation in red and green, but red cells where not completely isolated from each other as expected in a classic version of lateral inhibition and as observed in the biological counterparts. To see if we could optimize this patterning network in our computational system, we optimized parameters to introduce a stronger inhibition, which is described as being used by biological systems employing lateral inhibition patterning (del Álamo et al., 2011; Sprinzak et al., 2010, 2011); with these optimized parameters (see Table S3, "case of stronger inhibition"), we observed a checkerboard pattern in the disorganized lattice that reproduced the complete isolation of red cells (Fig. S8b). In this setup, we were able to observe fate bias due to cell size, a recently documented phenomenon of lateral inhibition (Shaya et al., 2017). In particular, we found that smaller cells were significantly more likely to be of the sender fate (A, red) and large cells more likely to be of the receiver fate (A', green) (Fig. S8b) A cells area: 108.82 ± 1.59 pixels^2, (A') cells area: 122.53 ± 1.89 pixels^2 ($P<0.0001$).

Finally, we added adhesion to our inhibition model to test whether it could recapitulate the multicellular synthetic 3D structure demonstrated *in vitro* using lateral inhibition and adhesion (Fig. 4b). The *in vitro* network consists of red (A) cells expressing mCherry alongside CD19 and anti-CD19 synNotch. When anti-CD19 synNotch is bound to mCherry/CD19 on a neighboring cell, mCherry/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 *In vitro*).

To obtain this network *in silico*, we used the lateral inhibition network parameters consistent with Fig. 2-4, the ones that generated the checkerboard pattern with non-completely isolated red cells, for consistency with the remaining of the other simulations, and to keep a close relationship with the cell type defined *in silico* earlier. The adhesion matrix is defined such that green (A') cells adhere to other green (A') cells (Fig.4b, *In silico*, Adhesion Matrix). Starting from a 3D aggregate of around 100 cells, we observe that initially some of the red cells become green, and then the green cells meet each-other in the center of the aggregate, as in the *in vitro* experiments (Fig. 4b, see also Fig. S8c and Movie S7). In the *in silico* system, we have observed that these green cells sometimes revert to red, leaving an active, dynamic, center of green cells turning red and moving to the external layer (not shown), such that the two-layer structure is the result of a dynamic equilibrium.

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

With a model that could recapitulate *in vitro* synthetic developmental trajectories, we then generated and tested a series of new programs for new morphologies *in silico*. We first aimed to achieve elongated structures. 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 posited that

elongation could be achieved via a “solid stem” of cells to structurally support the elongation and a “fluid” proliferating cap to provide the fuel for the elongation.

As we wanted our elongation to be able to self-organize, we started from a randomly mixed population of cells. We ideated the goal trajectory in Figure 5a. Starting from a random mixture of two cell types, (A) and (B), they 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” feature that would inhibit further movement and sorting. In Step 3 activated (A’) cells and (B’) cells would induce activation for cells of the same type through gaining signaling capacity to the basal state of the same genotype of cells (A’ to A and B’ to B). In this way, at the interface between (A) and (A’) cells the activated (A’) cells are signaling to inactive (A) cells to become activated, and similarly at the interface between (B) and (B’). As the inactivated cells (A) and (B) keep proliferating, we reasoned that these areas on either ends could provide an engine for continuous growth.

To implement this synthetic developmental trajectory, we envisioned the signaling network depicted in Fig. 5b. Two cell types (A) and (B) have high homotypic adhesion preference, and low heterotypic adhesion. They also have signaling capacity to each-other that leads to activation to (A’) and (B’) respectively. The active state gains different physical features: increased homotypic adhesion, reduced motility and inhibited growth and division to simulate a more “solid” feature. Moreover, activated cells, for example (A’), gain a signaling capacity towards same cell type (both (A) and (A’) have the receptor for this new signal, triangle in figure) (Fig. 5b).

Seeding a mixture of approximately 30 (A) and 30 (B) cells led to the formation of some elongated structures. Elongation was not robust and typically was not sustained after 100,000mcs and in general terminated around 20,000-50,000 mcs. Poles of inactivated cells were retained in only 10% of the simulations on both ends. We measured an average length of 48.95±8.28pixel for the stalk region (green + red cells) at t=100,000 mcs (n=10) (Fig. 5c,f and Movie S8; see Fig. S9 for length measures).

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 formation of poles that were separated enough (step 1 in Fig. 5a). We therefore hypothesized that robustness could be improved by seeding with separate spheroids of (A) and (B). This change in initial condition led to consistent, improved elongation overall (100% bidirectional elongation) in all runs, and also 40% of the structures retained both caps of inactivated (A) and (B) cells (Fig. 5d and Movie S9).

These results suggested to us that the incomplete elongation observed with Implementation 1 was due to an interface formation that was non-optimal for continuous growth. It also suggested a way to improve elongation by increasing time-delay between sorting (step 1) and activation (step 2). We therefore generated an Implementation 2 of the same network, where the activation thresholds for (A) to (A’) and (B) to (B’) were higher (See Table S3), and so activation took longer to happen. With this implementation we observed markedly improved elongation that generated continuously elongating structures that were of an average length of 68.67±9.27 pixels (Fig. 5e-f and Movie S10). Furthermore, we noted that 80% of the structures obtained with Implementation 2 retained the capability to elongate by retaining the (A) caps and (B) caps.

Inhibiting Receptor Expression Allows Morphological Oscillation In Silico

Next, we wanted to see if we could generate synthetic networks that would sustain a “morphological oscillator”. Following an analogy from nonlinear dynamics, this would be described by a “limit-cycle” as an attractor for the dynamics. We hypothesized that we could generate such a system with oscillations between a structure with checkerboard lattice-like arrangement of (A) and (B) cells, and

a structure with (A) and (B) cells forming two poles (Fig. 6a). Previous studies, as well as our work here, suggest that the checkerboard lattice could be achieved via heterotypic adhesion, and the two poles structure could be achieved via homotypic adhesion (Brodland and Chen, 2000; Glazier and Graner, 1993; Togashi et al., 2011). In order for the system to oscillate between the two states, the cells would need to switch their adhesion repertoire. In their inactivated state, (A) and (B) cells would have heterotypic adhesion molecules (Het.Ad.1a and Het.Ad.1b in Fig. S11a where we show a possible theoretical implementation) 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 (Homot.Ad.1 and Homot.Ad. 2 in Fig. S11a). An adhesion matrix supporting this is rather straightforward to define, the example that we used is shown in Fig. 6b. Less clear to us was how to generate a network that would the switch between the two states to be self-organizing and sustaining over time.

We explored different network topologies. To simplify the analysis, we performed the initial exploration with conditions where the cells neither grow nor divide. We started by testing a version of the network where the basal states and the induced states cells have the same signaling capacity. (A) cells signal to (B) cells to switch their adhesion repertoire and vice versa (Fig. 6c, Network 1). 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, and formation of two poles. At this point, although the cells far from the interface reverted to their basal states, the interface became locked as (A’) cells signal (B’) cells to maintain their activated state, and vice versa, blocking the reversion to a basal state thus locking the structure in a two-pole state (Fig. 6c, and Movie S11). This structure was stable and demonstrated no clear oscillatory behavior, showing that morphological separation alone was not enough to induce the oscillatory behavior due to (A’) and (B’) reciprocal signaling stabilizing the interface (Fig. S10a).

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 (Fig. 6d, Network 2). Seeding a mixture of 30 (A) and 27 (B) cells yielded damped morphological oscillation, with (B) cells ultimately locking (A) cells into the (A’) fate (Fig. 6d, Movie S12 and Fig. S10b).

To prevent this locked endpoint, we adjusted the network such that the activated states had inhibited receptor expression instead of ligand expression (Fig. 6e, Network 3). 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. 6e, Network 3, Movie S13) that lasted at least 13 cycles (Fig. S10c) with minimal dampening. For Network 3, we ran simulations where the cells were allowed to grow, and still observed morphological oscillations (Fig. S11b).

DISCUSSION

Multicellular self-organization is a key facet of development and tissue formation. The identification of signaling and mechanical changes as core to multicellular development process, recently allowed us to design synthetic developmental trajectories *in vitro* (Toda et al., 2018). The design process of these trajectories relied on a trial-and-error process, where signaling networks and effectors are chosen from a literature search and implemented in cells. Computational models could provide support by allowing rapid implementation of various designs *in silico.* In this way, robust formulas for signaling and

effectors can be identified and selected before being implemented experimentally; effectively catalyzing reconstruction efforts and enabling rational design.

Here we present a first step towards that goal, with a computational framework focused on synthetic juxtacrine signaling and morphogenesis driven by changes in cell-cell adhesion. The combination of signaling and morphological effectors has been shown to be at the core of complex developmental transitions: tissues are a complex system where cell-cell signaling affects morphogenesis, and then morphogenesis feeds-back to influence signaling, to robustly generate complex multicellular structures (Gilmour et al., 2017). In facts, the combination of signaling and morphological effectors has been incorporated in several computational models and shown to be able to replicate complex morphogenesis of embryonic transitions (for example: (Belmonte et al., 2016; Marin-Riera et al., 2018). The modeling platform we put together here was generated not with a biological system in mind, but with the possibility of using it for design of novel trajectories We think part of the novelty of our work is that with relatively simple circuits and few parameters, we are able to capture a great deal of complexity. Our computational system replicated a range of different developmental trajectories of patterning and morphogenesis, and also allowed exploration of design principles of novel trajectories such as elongated and oscillatory structures through an iterative cycle of design.

Our *in silico* model demonstrates some degree of biological faithfulness, replicating numerous facets of currently known high complexity mammalian synthetic structures. We focused first on simple back-and-forward circuits, and on identifying parameters for the *in silico* model by tuning the timescales comparing to the *in vitro* counterpart. We observed that other emergent properties were also reproduced that were not explicitly used in the parameter estimation, in particular sphericity, cell activation dynamics and overall variability of final structures (Fig. 3). This prompted us to try keeping parameters for cell behaviors fixed while exploring other effectors (different cadherins) and other networks such as lateral inhibition, and we obtained qualitative replication of behaviors (Fig. 4).

In certain aspects, we note that the model can still be further improved, as evidenced by the incomplete regeneration of the three-layer structure upon 'injury' for example (Fig. 3a,b *Regeneration*). In particular, we noted lack of blue (A) cells in the regenerated half. One hypothesis is that (A) adhesion parameters could be improved. In implementation shown in Fig. 3, (A) cells have minimal adhesion, and no adhesion preference between cell types, so maybe increasing (A) cells adhesion towards (A') and (B') would potentially drive them to the damaged area. However, this seems unlikely to be a factor *in vitro* as (A) cells in the out layer retain their spherical morphology and do not seem to adhere strongly to the core of the structure. This is especially evident from the 3D reconstruction of a three-layer in the reference experiment (Toda et al., 2018). A more plausible explanation is that the *in silico* model is an idealized version of the experiment. Immediately post-bisection structures *in vitro* have noticeable imperfections such as latching cells post-bisection and prominent proximal cells in suspension (Toda et al., 2018) and these cells may reattach to aid regeneration. Our simulations with this modification (Fig. S2c) suggest that these latching cells do play a role in regeneration, but it remains unclear whether these cells are necessary for regeneration *in vitro*. It therefore remains uncertain whether complete regeneration is a shortcoming of the model or if there is an additional, unknown mechanism that drives complete regeneration *in vitro*. It is tempting to speculate that these "imperfections", like leftover cells, could be used to the designer advantage in rational design for improving robustness, for example.

One open challenge for the field is how to design novel developmental trajectories. In other, more mature fields, computational design is at the base of design of e.g. novel compounds (Gómez-Bombarelli et al., 2018). How to achieve that degree of predictability is still unmatched for biological systems based on complex artificial genetic circuits. With our model, we showed that we can design novel developmental trajectories, compare different designs, and identify the best ones. Going through phases of

design-test-learn-redesign *in silico* allowed us to identify which parameter sets confer robust target behavior to various trajectories and which ones perform poorly *in silico*.

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 (Implementation 1 in Fig. 5) was not very efficient in generating elongated structures. However, when we changed the initial conditions, the same genetic network did generate robust elongating structures, showing that 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 example, blastocyst compaction is the initial conditions for hollowing and internal cell mass differentiation during early mammalian development (Wennekamp et al., 2013). Our results suggest that one way to achieve diversity in complex morphogenesis is by changing parameters of an early transition, which would change the initial condition for subsequent transitions, without affecting the specific circuits of the second transition. This would provide a simpler way to generate diversity, one that could have been at play during evolution of animal body plans (e.g. Pre-Cambrian age). At the same time, it could be a design principle that could be used in the design of complex structures without increasing the number of control points. By changing the dynamics of a previous phase of morphogenesis, we could affect a second phase of morphogenesis without directly altering its genetic program or dynamics. These results also show that, in this system, the relative dynamics of signaling and morphological outputs are critical to achieve control of a the elongation trajectory.

Embryonic development is characterized by being initiated from a single cell. In our system, we always start from multiple cells and even sometimes with two cell types. We have shown a system that can recapitulate fate bifurcation starting from a single cell type, with the lateral-inhibition network. One possibility would be to string together lateral-inhibition for fate bifurcation, with the subsequent elongation network, to achieve elongation from a uniform genotype of cells.

As we attempted to create a trajectory leading to a stable oscillatory structure (Fig. 6), we went through a similar iteration of poorly performing designs. We first thought that the strong homotypic adhesion within (A’) and (B’) cells would be sufficient to separate the two groups such that they would not signal to each other. However, in the simulations, we saw that the presence of an interface where (A’) and (B’) cells are still touching and signaling to each other prevented them from reverting back to basal (A) and (B) states. We were able to overcome this when we altered communication between (A) and (B) and (A’) and (B’), in particular the excited states 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. Our results suggest that receptor expression can be used to generate stable intermediate states where sorting can be reverted. Interestingly, Notch receptor control via post-transcriptional mechanisms like microRNA have been shown to be used to stabilize different cell fates during choices of differentiation in human neuronal stem cells in particular in oscillating systems (Roese-Koerner et al., 2017).

This type of regulation seemed important for generating states that are stable enough to serve as the basis for subsequent morphogenesis. Again, together with the elongation trajectory results, this identifies the linking of multistep developmental trajectories as particularly important knob of control to achieve desired outcomes in our *in silico* system.

Future directions for these computational efforts could be the combination with artificial intelligence-based optimization either through evolutionary algorithms or machine learning. It has been

recently shown that these could be used to generate morphologies that can then be recapitulated *in vitro* (Briers et al., 2019; Kriegman 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. Incorporation of different modality of signaling and effectors is also a future direction of this study. Numerous other 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, 2017b, 2017a; Scheller et al., 2018) have made it possible to further control this process, facilitating synthetic reconstruction of complex native morphogenic processes towards enabling control over custom tissue development.

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. Critically, this and other computational systems, in order to prove helpful in the design of developmental trajectory, will have to integrate a full cycle in which the best designs identified *in silico* are then implemented in cells. We hope this is the first of modular computational models that will be incorporated in future rational design of multicellular structures in vitro, towards functional organs, both to study the design principles, and towards applications in regenerative medicine.

MOVIES

Movies files are attached to the submission, and can also be downloaded from:
https://www.dropbox.com/sh/6o7kh15fwfhrdk4/AAASs0Hija1gelnFviL-yLQBa?dl=0

S1. *In silico* formation of three-layer structure. First half of movie is with ~200 (A) cells, ~50 (B) cells. Second half of movie is with ~160 (A) and ~90 (B) cells. (A) cells are represented in blue, (B) in beige, (B’) in green, and (A’) in red. Logic as described in Figure 3b.1. When cells (A) and (B) are in contact, (A) cells send activation signals to (B) cells which, if a threshold is reached, excite to (B’) cells. (B’) cells preferentially adhere to each other and can signal back to (A) cells which, if a threshold is reached, excite to (A’) cells. (A’) cells are intermediately adherent to (B’) cells and weakly adherent to all other cell types. The end result is a layered structure with a green core followed by layers of red and blue.
For the rendering, still images were generated by overlaying the separate color fields (gray, blue, green, and red); the still images are taken in intervals of 500mcs; total run time is 20,000mcs for the first structure and 25,000mcs for the second structure. 3D renderings of a section of the structures are shown at the end of each developmental trajectory.

S2. *In silico* three-layer regeneration. Three layer network is used as in Movie S1, starting from with ~160 (A) cells and ~90 (B) cells. (A) cells are represented in blue, (B) in beige, (B’) in green, and (A’) in red. Logic as described in Figure 3b.1.
For the rendering, still images were generated by overlaying the separate color fields (gray, blue, green, and red); the still images are taken in intervals of 500mcs; total run time is 45,000mcs. Half of the cells are removed at 0:07 timepoint in movie (corresponding to 20,000mcs) to observe regeneration after bisection.

S3. *In silico* development of structure in Fig.4a.1. (A) cells are represented in blue, (B) in gray, (B’) in green, and (A’) in red. Logic is given in Figure 4a, and adhesion matrix is given in Figure 4.a.1. Starting from ~90 (A) cells and ~90 (B) cells. Images were generated by overlaying the separate color fields (gray,

blue, green, and red) and were taken in intervals of 500mcs over the course of 30,000mcs. Bisected 3D rendering is shown at the end of the movie.

S4. *In silico* development of structure in Fig.4a.2. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Logic is given in Figure 4a and adhesion matrix is given in Figure 4.a.2, starting from ~200 (A) cells, ~50 (B) cells. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 32,000mcs. Bisected 3D rendering is shown at the end of the movie.

S5. *In silico* development of structure in Fig.4a.3. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Logic as pictured in Figure 4a and adhesion matrix is given in Figure 4.a.3; first half of movie is with ~90 (A) cells, ~90 (B) cells. Second half of movie is with ~30 (A) and ~30 (B) cells. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 50,000mcs for each structure. Respective bisected 3D renderings are shown at the end of the first half and second half.

S6. *In silico* development of structure in Fig.4a.4. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Logic as pictured in Figure 4a and adhesion matrix is given in Figure 4.a.4; starting from ~160 (A) cells and ~90 (B) cells. Images were generated by overlaying the separate color fields (gray, blue, green, and red) were taken in intervals of 500mcs over the course of 34,000mcs. Bisected 3D rendering is shown at the end of the movie.

S7. *In silico* development of lateral inhibition with adhesion, as seen in Fig. 4b. (A) cells are represented in red and (A') in green. Logic and adhesion matrix are given in Figure 4b. Starting condition is 93 (A) cells. Images were generated by overlaying the separate color fields (green and red) and were taken in intervals of 500mcs over the course of 50,000mcs. Bisected 3D rendering is shown at the end of the movie.

S8. *In silico* development of elongation network, Implementation 1; refers to Fig. 5c. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Target trajectory as pictured in Figure 5a and proposed network along with adhesion matrix is given in Figure 5b. Starting conditions are 31 (A) and 26 (B) cells randomly mixed. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 100,000mcs. 3D rendering is shown at the end of the movie.

S9. *In silico* development of elongation network, Implementation 1, with controlled initial conditions; refers to Fig. 5d. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Target trajectory as pictured in Figure 5a and proposed network along with adhesion matrix is given in Fig. 5b. Starting conditions are ~36 (B) and ~36 (A) cells as separate spheroids. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 50,000mcs. 3D rendering is shown at the end of the movie.

S10. *In silico* development of elongation network, Implementation 2; refers to Fig. 5e. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Target trajectory as pictured in Figure 5a and proposed network along with adhesion matrix are given in Fig. 5b. Initial conditions are 36 (A) 21 (B) cells randomly mixed. Images were generated by overlaying the separate color fields (gray, blue,

green, and red) and were taken in intervals of 500mcs over the course of 100,000mcs. 3D rendering is shown at the end of the movie.

S11. *In silico* development of oscillation Network 1. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Adhesion matrix is given in Figure 6b; signaling network is in Fig. 6c. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 100,000mcs. See Supplementary Figure 10a for extended time course.

S12. *In silico* development of oscillation Network 2. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Adhesion matrix is given in Figure 6b; signaling network is in Fig. 6d. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 100,000mcs. See Supplementary Figure 10b for extended time course.

S13. Oscillation with receptor inhibited, first 100,000 mcs. (A) cells are represented in blue, (B) in gray, (B') in green, and (A') in red. Adhesion matrix is given in Figure 6b; signaling network is in Fig. 6e. Images were generated by overlaying the separate color fields (gray, blue, green, and red) and were taken in intervals of 500mcs over the course of 100,000mcs. See Supplementary Figure 10c for extended time course.

FIGURE LEGENDS

Fig. 1. Concepts underlying the computational model (See 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 contact surface is higher compared to the upper pair, hence the signaling (green arrow) towards the target gene is stronger.

On the right, the *in silico* model shows a simplified representation of the cell-cell signaling process with parameters: ligand amount (L1), surface area of contact (Φ, pink), and net signal (S, green arrows). Cells (A) are the sender of the ligand, cells (B) are the receiver. In the schematic, *in silico* cells are objects of 9 pixels. The cell pair at the bottom has a higher level of signaling compared to the upper pair due to a larger surface area of contact (Φ2>Φ1, 2 pixels compared to 1).

(b) Time evolution of target gene level in the receiving cell (B); cells (A) and (B) are placed in contact at time zero and kept in the same configuration for 100,000 steps of simulation, and then moved far apart to stop signaling. Target gene levels is followed over time for two different values of shared surface area **Φ**, with Φ2>Φ1, with all other parameters kept identical.

(c) Model representation of cell behavior state change.

On the left, representation of biological effector activation: 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. Matching 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. Thresholds for the excited state (5000 AU) and ground state transitions (2500 AU) 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. (B) cells express anti-CD19 synthetic receptors, and response elements for induction of GFP and E-cadherin in response to receptor activation; (A) cells constitutively express CD19 ligands and a blue fluorescent protein (BFP). Confocal fluorescent microscopy images, overlayed for green, blue and brightfield channels, at 1h after seeding of approximately 100 (A) cells and 100 (B) cells, and after 24h are shown; microscope images are reproduced from (Toda et al., 2018), reprinted with permission from AAAS.

(b) *In silico* implementation of the two-layer circuit. Signaling network: (A) cells, when physically contacting (B) cells, send activation signal. In (B) cells, accumulation of sufficient target gene excites them to the (B') state. Cell behavior changes: adhesion matrix defines pair-wise adhesion preference on a scale 0-3 (0 minimal adhesion preference, 3 high adhesion preference). This adhesion matrix indicates that (B') cells are strongly adhesive to other (B') cells but weakly to other (B) or (A) cells.

Indicated time points (t=1 is equal to 1,000 simulation steps) of a midpoint cross section of the simulated aggregate are shown for a representative simulation with initial conditions a mixture of 92 (A) and 87 (B) cells. 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', hence the homogeneity of the B cell types starts to increase at around that time point. 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 and text for details).

Fig. 3. The computational model captures qualitative and quantitative features of the synthetic developmental trajectory for formation of three-layered spheroids.

(a.1) *In vitro* implementation of the back-and-forth network: cell A expresses constitutive BFP and CD19 ligand (round ligand), and, conditional to anti-GFP synNotch activation, also mCherry and low levels of E-cadherin. Cell B expresses 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.

(a.2) The time evolution of a mixture of approximately 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). **(a.3)** Initial seeding for Regeneration experiment is with approximately 160 (A) and 80 (B). Microscope images are reproduced from (Toda et al., 2018), reprinted with permission from AAAS.

(b.1) Computational implementation of back-and-forth network. Signaling 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'. Cell behavior changes:

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). **(b.2)** Representative cross section of aggregates of an *in silico* synthetic developmental trajectory is shown below at the indicated time points; initial conditions are approximately 200A cells and 50B cells. For the regeneration simulation **(b.3)**, 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).

(c) Temporal quantification of cell activation index (see methods for details on definition of activation indexes). Activation of A cell to A' (red) and B cells to B' (green) *in silico* and *in vitro* are shown for the duration of development of the three-layered structure. *In silico* (shaded lines), 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 B' (see methods, "Video Analysis" for details on thresholding) (n=30 simulations *in silico*, 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 indexes 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. (See 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*). *In vitro* results data are from (Toda et al., 2018). See Methods, "Quantification and Statistical Analyses; Simulation Quantifications", for more details on cores definitions *in silico*.

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 cross section of the aggregate at the same endpoint; *in silico* adhesion matrix. Scale bar is 100um for *in vitro* and 17.5 pixels (approx. 100um) for *in silico*. See Fig. S4 and S5 for more examples. Microscope images are reproduced from (Toda et al., 2018), reprinted with permission from AAAS. The colors red and blue of the structure (a.2) and (a.4) 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; modified and reprinted with permission from AAAS.

(b) Lateral inhibition with adhesion.

On the left are the *in vitro* diagrams (top) and cell behaviors (bottom) over time. Initially light red cells bifurcate to either dark red or green types. Starting from a uniform population of approximately 100 light 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 E-cadherin (Toda et al., 2018).

On the right is the *in silico* network. (A) cells receive and signal to/from neighboring (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 on the right. Below, cross sections of the simulated aggregate at the indicated 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. Microscope images are reproduced from (Toda et al., 2018), reprinted with permission from AAAS.

Fig. 5. *in silico* 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 "solid" 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') and (B) to (B') after a threshold is reached. (A') 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 but only to themselves. (A') and (B') are highly adhesive both homotypically and to each-other. Other effector changes in cells (A') and (B') are: stopping cell division and growth, and decreasing motility (to simulate fluid-to-solid transition).

(c) 3D rendering of simulation result when starting from 40 A cells and 40 B cells with Implementation 1 (see text for details); representative image of endpoint result (see (f) for quantification). t=100 means 100,000 mcs.

(d) 3D rendering 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. t=50 means 50,000mcs.

(e) 3D rendering of simulations with Implementation 2 parameters, where activation threshold for transition A→A' and B→B' are higher. Images at the indicated timepoints of 3D rendering of the simulations are shown.

(f) Quantification of length of simulated aggregates at 100,000 steps for Implementation 1 and Implementation 2. We present mean±s.d. (n=10); significantly longer, two-tailed t-test, t=5.01, P<0.0001. See Fig. S7 for details on how the length is measured.

Fig. 6. 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 signaling between neighboring A and B cells (1) induces 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 signaling parameters are set equal for the (A) genotype compared to (B) genotype (see Table S3).

(c) 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. Below, 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 (more on this index in Methods, "Simulation Quantification"); inserts are 3D renderings of the simulated aggregates at the corresponding timepoints. See also Fig. S10a for longer simulation results.

(d) Network 2, signaling network is same as Network 1 but now activated states lose signaling capacity, indicated by lack of rounded and squared ligand on A' and B' cells, and corresponding lack of pink dotted arrows. Below, 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; inserts are 3D renderings of the simulated aggregates at the corresponding timepoints. See also Fig. S10b for longer runs.

(e) Network 3, Signaling network is same as Network 1 but now activated states lose signal reception capacity, indicated by lack of receptors on activated cells. Below, 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; inserts are 3D renderings of the simulated aggregates at the corresponding timepoints. See also Fig. S10c for longer runs.

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

CompuCell3D and the cellular Potts Formalism

We implemented our model in CompuCell3D (CC3D) v.3.7.8 (Swat et al., 2012), a modeling software that allows simulation of cells and their behaviors using the cellular Potts formalism. By itself, CC3D contains numerous built-in features for replicating *in vitro* cell behavior, several of which we utilized either directly or adjusted via CC3D Python v.2.7.13 scripting according to manual v3.7.9. In our model, we incorporated default features from CC3D such as surface area constraint, volume constraint,

cell division, adhesion, cell-cell surface contact, and cell types. We implemented custom cell motility, cell growth, and cell signaling, as described below and in subsequent sections.

We defined cells as multi-pixel entities in 3D that physically act by performing "pixel copy attempts" over simulation time steps (monte carlo steps, mcs). Performing "pixel copy attempts" effectively moves and changes both cell geometry and position over time. These pixel copy attempts succeed probabilistically, determined by the Boltzmann acceptance function, $P=e^{-\Delta H/T}$, where P is probability of attempt success, ΔH is change in total effective energy of the system from all attempted pixel copy attempts at the mcs t, and T is the cell motility (Swat et al., 2012).

Effective energy (H). Because we incorporated surface area constraint, volume constraint, and adhesion, our total effective energy H at a given mcs t therefore takes the form,

$$Eq.(1)$$

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

as described in (Hester et al., 2011). The terms σ(i) and σ(j) denote the identity of the cells occupying pixel sites i and j separately, with the Kronecker Delta limiting inclusion to only the cell interface. J is a matrix that contains the contact energy between cells of different identity. λ_{Sur} and λ_{Vol} constrain deviations of a cell from the ideal surface area Sur_{Tar} and Vol_{Tar}, hereafter referred to as target surface area and target volume, respectively.

The numerical values in J control adhesion in cellular Potts. The values in J represent a stability index: lower J makes for a more stable state, which is then how you achieve stronger adhesion. Conversely, a higher J leads to weaker adhesion. Throughout the manuscript (mainly in the figures), we use a grouped representation of the adhesion for presentation simplicity; the exact values of J are shown in Tables S1-2.

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

$$L = \frac{\gamma}{1+e^{-\frac{t}{\xi}}} \tag{2}$$

where t is the given time in mcs, while γ, and ξ are constants. We chose this equation because of its simplicity. It could be generalized to 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).

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

$$\frac{dR}{dt} = \frac{1}{1+e^{-\frac{S-\beta}{\varepsilon}}} - \frac{R}{\kappa} \tag{3}$$

where β, ε, and κ are constants, whereas S is the signal strength. We chose this form for several reasons. First, parameters have intuitive interpretations: 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, this form of the logistic function is easily tunable and well behaved, due to its monotonicity from negative infinity to positive infinity and

bound between 0 and 1. This tunability is not as easily achievable with the Hill function, where odd or fractional Hill constants lead to the existence of singularities.

In the case where the target gene is a ligand itself, we use equation (2) to calculate ligand levels.

The time-dependent evolution of the reporter, apart from the parameters, depends on **signal strength** S; this reflects the biological fact that the promoter of the target gene is under the control of the receptor in juxacrine signaling. Signal strength is itself affected by: the number of receptors on the receiver cell (Ω), the number of ligands on contacting neighboring cells (L), the surface contact area between receiver cell and its neighbors (Φ). Because these factors evolve over time, S is therefore a morphological dependent and time dependent function that evolves according to structure's spatial organization.

The following describes how we take into consideration the shared surface area to compute S. We consider a single receiver cell σ; first we need to identify which of its neighbors can signal with it. The different cells are assigned different types, according to whether they can signal (ligand expressing) or can receive (receptor expressing) or both. These cell types are indicated as A, A', B and B' in the text, in CC3D. For example, we have a cell σ of type A, that express receptor r_A, that can be activated by ligand l_A. This allows us to identify the neighbors of sigma, by looking through the list of all neighbors of sigma and identifying the ones that are of a type that bear ligand l_A.

Then each different ligand/receptor interaction is treated identically regardless of the specific mechanism (e.g. if it models anti-GFP/GFP or anti-CD19/CD19).

Receiver cells has receptors on its membrane, quantified by Ω_σ. The neighbors have ligands on their membrane, quantified by L_i. To compute the amount of receptor-ligand interactions that can happen when receiver + cognate sender cells are in contact, we need to calculate the amount of receptor and ligand that are present on the surface of contact.

To do that we first define the portion of contact surface for sender (SNi) and receiver (sigma):

$$\Phi_{SNi(\sigma)} = \frac{surface\ area\ of\ contact\ between\ SNi\ and\ sigma}{Sur\ (SN)} \tag{4a}$$

$$\Phi_{\sigma(i)} = \frac{surface\ area\ of\ contact\ between\ SNi\ and\ sigma}{Sur\ (\sigma)} \tag{4b}$$

These are now multiplied for the total amount of ligand (or receptor), to obtain the amount of ligand (or receptor) that is available at the area of contact.

Available ligand = $\Phi_{SNi(\sigma)} * L_i$

Available receptor = $\Phi_{\sigma(i)} * \Omega_\sigma$

With these two values we can calculate the value of S_σ as follows:

$$General: S_\sigma = \sum_{i=1}^{n} Min[(\Phi_{SNi(\sigma)} * L_i), (\Phi_{\sigma(i)} * \Omega_\sigma)] \tag{5}$$

where n is the total number of cells that are currently in contact with sigma and that can engage in signaling with sigma, i.e. produce the ligand for which cell sigma produces the receptor.

This results from assumption of:

a 1-1 stochiometry of 1 ligand activating 1 receptor, given the biochemistry of the signaling;

homogeneity of ligand and receptor on the cell's surface.

This is in general; in the majority of the simulations that we describe in the results, we employed a simplified version where the receptor is considered to be in excess, as we do not have evidence to think otherwise. In this case, the only part of (5) that is determining the amount of signal S is given by the ligand, so S takes the form of:

$$Ligand\ Limiting: \quad S_\sigma = \sum_i \Phi_{SNi(\sigma)} * L_i \tag{6a}$$

In one case in the results, we use a simplification where the receptor is considered limiting, mainly for the results described in Fig. 6e. The signaling is calculated with the receptor limiting form

$$Receptor\ Limiting:\ S_\sigma = \sum_i \Phi_{\sigma(i)} * \Omega_\sigma \qquad (6b)$$

Quantized Signaling Inducible Output

To implement signaling-inducible behavioral response, we use a state-transition model. 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. The transition between states is regulated by the target gene activation, which is regulated by the cell-cell signaling (see above Generalized Juxtacrine Signaling Model (GJSM)). Cells in different states can have different properties such as color/state, adhesive properties, deformability properties, signaling/reception capacity (Fig. 1c). 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). In this way, the signaling can induce behavioral changes in the cells that receive the signal, generating a highly non-linear system of interacting agents.

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. Adhesion matrix J can be defined for the different states, to mirror 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 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.

In Silico L929 Cell Line Properties

Cell division. *In silico* L929 (ISL929) cells consist 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 at the beginning of experiments (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, modelling completely symmetric cell division. All the parameters and state variables are inherited by the two daughter cells from the parent cell. These choices result in roughly doubling time of 24,000 mcs (equivalent to 24 hours, the estimated doubling time of L929 cells (Toda et al., 2018)). We note that, due to stochastic nature of growth, cell death is also possible within this model.

Cell adhesion. *In vitro* L929 mouse fibroblasts weakly adhere to one another under ultra-low attachment suspension conditions (Toda et al., 2018) thus we designate our basal, parental 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).

Cell deformation. Cells that express adhesion proteins and adhere to each-other in vitro, deform markedly, and lose their rounded morphology (Toda et al., 2018). In CCC3D, a way to change deformability of cells is through modulating parameters λ_{Sur} and λ_{Vol} in Eq (1) for H, with lower values corresponding to higher deformability. Cells with an adhesion matrix value of at least 39 (i.e. 0-39 range) (see Table S1 and S3), λ_{Sur} and λ_{Vol} were set to 1.0. Other cells had λ_{Sur} and λ_{Vol} set to 2.2.

For the morphological oscillation, we had simulations without growth and cell division. Since removal of growth in motile cells leads to an artifact of unrealistic warped morphologies, we therefore used λ_{Sur} and λ_{Vol} equal to 2.0 each (instead of 1.0) to minimize unrealistic morphologies.

Cell motility. Cell motility is defined in CC3D via the parameter T. Biologically it is know that cell adhesion to environment is complexly linked to cell motility, and adhesion effects on motility vary widely between different adhesion proteins and cell types (Gumbiner, 1996; Nieman et al., 1999; Takeichi, 2011). In general, although clearly not all-encompassing, the adhesion abstraction is that strong cell adhesion to environment tends to decrease cell motility (Alberts et al., 2002; Gumbiner, 1996; Takeichi, 2011).

For our purposes, in the *in vitro* L929 system, we noticed that the cell motility is different between adherent cells (low motility) compared with non-adherent cells (high motility). We therefore defined motility as a function of a cell's environment (neighboring cells and medium), so that higher adhesion results in lower motility; in this way, different cells can have different motility. Each cell's individual motility T_σ is:

$$T_\sigma = T_0 + \zeta \frac{\sum_{\sigma(i)_i,j} J_{\sigma(i),\sigma(j)} \left(1 - \delta_{\sigma(i),\sigma(j)}\right)}{Sur(\sigma)} \tag{7}$$

This formula iterates over each neighboring cell pixel and medium uniquely, but ultimately T is determined only by the type of the focal cell, the types of the neighbors, and total contact with medium. Categorizing environment by cell types and medium instead, accomplished in CC3D via cell-cell surface contact feature and cell type index, we obtained a computationally simpler approximate formula, which is the one that we use in our simulations:

$$T_\sigma = T_0 + \zeta \sum_k \frac{J_{type(\sigma),k} \; * \, total\ contact\ surface\ area\ with\ \ k}{Sur(\sigma)} \tag{8}$$

T_0 is a constant representing basal cell motility, ζ a constant representing how effectual adhesion is at attenuating motility, and k denotes "cell type" (can be either a cell type or medium). This T allows each cell to sense its adhesivity to local environment, decreasing motility if adherent to neighbors and restoring motility when exposed to non-adhesive conditions.

T_0 is kept constant throughout except for the elongation; there, to model fluid-solid transition, the activated cells A' and B' acquire a lower T_0 (see Table S3).

<u>Simulation Conditions</u>

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, 1,000 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

Sphericity index. (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)

$$Sphericity = \frac{\pi^{\frac{1}{3}}(6*Structure\ Volume)^{\frac{2}{3}}}{Structure\ Surface\ Area} \tag{9}$$

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.

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

$$\Psi_x = \frac{\sum_{i=1}^{n} \frac{surface\ area\ of\ \sigma_i\ that\ is\ in\ contact\ with\ cells\ of\ type\ X}{Sur\ (\sigma_i)}}{n} \tag{10}$$

Where the sum is taken only for cells of type X that are in contact with other cells of type X, and n is the total number of cells of type X in contact with cells of type 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). By focusing only on cells that have neighbors of the same kind, this measure focus on clusters and not on isolated cells.

This measure can be generalized for more than one cell type, by considering for example two cell types together. For example, in our simulations, when we have a genotype B that gives rise to two cell types (B basal and B' excited), we may be interested in measuring homogeneity index for B and B' combined. To do so we calculated $\Psi_{B,B'}$ by defining type X as {type B or type B'} in the above formula. 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. This measure can be applied to many different morphologies, beyond fixed lattices (Olimpio et al., 2018) and spherical morphologies.

Core Distribution. We counted the number of cores per structure by visualizing only (B') green cells in 3D at the endpoint of each simulation. To determine whether a core was a single core or double core, we visualized the endpoint simulation (B') green cells from multiple perspectives to prevent viewpoint bias. This allowed us to determine whether there was truly a single core or multiple cores of similar sizes, with the latter generating 2-cores or 3+ cores (counted as "other"in Fig 3e.). Small cell stripe connections between cores were negligible and therefore counted based on the cores. Structures that did not appear core-like (i.e. large cell stripes) were counted in the "other" category. See Fig. S12-13 for examples of core counting.

Elongation length. We measured the elongation length of the "solid stem" portion, i.e. the red+green cells. We did that by decomposing the structure in linear fragments, and then estimating each length of the linear segments. 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 (See Fig. S9 for an example).

Activation index. the activation index *in silico* is defined as the normalized ratio of activated cells over the total number of cells of the same type, i.e. #(A') / (#(A)+#(A')) and similarly for (B) and (B') cells.

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).

Circularity index. 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

$$Circularity = \frac{4\,\pi\,Area}{Perimeter^2} \tag{11}$$

This measure ranges between 0 and 1 and the closer to 1 is indicates it is closer to a perfect circle.

Activation index. 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 (Fig. 3c).

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. 5f. 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. General source code and specific codes for the structures presented in this study are available at [https://github.com/lmorsut/Lam_Morsut_GJSM]. Dataset for the structures are available upon request.

SUPPLEMENTAL INFORMATION

Fig. S1. Three-layer structures obtained from different initial conditions (Related to Main Fig. 3).
(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) at the top; two independent t=20h endpoint structures from replicate simulations are shown at the bottom. The simulations where repeated 30 times with slightly different initial conditions; the sorting dynamics are shown on the right. The initial number of cells across the 30 replicates were 202.2±4.33 (A) and 48.8±4.33 (B). $\Psi_{B,B'}$ and $\Psi_{A,A'}$ are used to measure system sorting. Scale bar is 17.5 pixels, approximately 100um.
We present mean±s.d.. (n=30 simulations). Microscope images of *in vitro* system are reproduced from (Toda et al., 2018), reprinted with permission from AAAS.
(c) *In vitro* structure from 160 (A) and 80 (B) as reference (Toda et al., 2018) is shown at the top. Two additional structures from replicate simulations starting from a mixture of 162.8±7.54 (A) and 87.2±7.54 (B) are shown at the bottom. $\Psi_{B,B'}$ and $\Psi_{A,A'}$ are used to measure system sorting. We present mean±s.d.. (n=10 simulations). Microscope images of *in vitro* system are reproduced from (Toda et al., 2018), reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S2. Activation from signaling is required for three-layer formation (Related to Main Fig. 3).
(a) Biological schematic of a back-and-forth network with DAPT treatment; DAPT inhibit synNotch signaling, so that (A) and (B) cells no longer signal and thus cannot excite one another to their respective activated state (Toda et al., 2018). *In silico* representation is on the right, with corresponding adhesion matrix, where absence of signaling is described with absence of ligands.
(b) *In silico* results from the network depicted in (a). Two distinct structures at t=20h are shown on the left. Number of starting cells for simulations is 201.33±6.02 (A) and 49.67±6.02 (B). Sphericity and circularity index for in silico structures are shown in the Morphology dynamics graph; the *in silico* index for A' and B' cells only lies at the bottom as there is no generation of A' and B' cells. (Toda et al., 2018). We present mean±s.d.. Two structures from replicate simulations are shown on the left. (n=30 simulations, n=1 *in vitro* experiment). Scale bar is 17.5 pixels, approximately 100um.
(c) Initial condition of a simulation to test regeneration with leftovers. Network and adhesion matrix is as in Main Fig. 3. The "start" timepoint is 24h from beginning. At indicated t=0, the shown perturbation is done where half of the cells are removed but some are left behind. Snapshot of t=25h of that same structure after 25,000mcs is shown. Scale bar is 17.5 pixels, approximately 100um.

Fig. S3. Results of simulation of three-layer structure formation networks under reversible activation conditions are similar to the ones with irreversible activations (Related to Main Fig. 3).
(a) *In vitro* and *in silico* networks and adhesion matrix. It is identical to Fig. S1b and Fig. 3a. The difference is on deactivation threshold is set equal to the activation threshold here instead of zero, so that the activation reactions are reversible.
(b) Quantification of cell activation dynamics, morphology dynamics and cores distribution as in Fig. 3a. We present mean±s.d.. Two structures from replicate simulations are 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). Microscope image of *in vitro* system is reproduced from (Toda et al., 2018), reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S4. Additional structures for the symmetric structures in Fig. 4a.1.
(a) *In vitro* base framework and *in silico* representation common to the structures below.
(b) Two additional structures from replicate simulations of the structure in Fig. 4a.1. Additionally, aggregate sorting dynamics as well as the time development are presented. 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* approximately 100 (A) and 100 (B) (Toda et al., 2018)). Microscope images of *in vitro* system are reproduced from (Toda et al., 2018), reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S5. Additional structures for the symmetric structures in Fig. 4a.2.
(a) *In vitro* base framework and *in silico* representation common to the structures below.
(b) Two additional structures from replicate simulations of the structure in Fig. 4a.2. Additionally, aggregate sorting dynamics as well as the time development are presented. 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* approximately 200 (A) and 40 (B) (Toda et al., 2018)).
Microscope images of *in vitro* system are reproduced from (Toda et al., 2018), reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S6. Additional structures for the asymmetrical structures in Fig. 4a.3.
(a) *In vitro* base framework and *in silico* representation common to the structures below.
(b) Two additional structures from replicate simulations of the third structure from approximately 100 (A) and 100 (B) in Fig. 4a.3. Additionally, aggregate sorting dynamics as well as the time development are presented. 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 structure of Fig. 4a.3 generated starting from approximately 30 (A) and 30 (B) *in vitro* from (Toda et al., 2018), and the analogous simulated structures starting from 27.1±3.07 (A) and 29.9±3.07 (B) *in silico*. We present mean±s.d. for the aggregate sorting dynamics (n=10 simulations). Microscope images of *in vitro* system are modified from (Toda et al., 2018) (the colors red and blue in the *in vitro* system have been switched) for consistency with *in silico* framework and reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S7. Additional structures for the asymmetrical structures in Fig. 4a.4.
(a) *In vitro* base framework and *in silico* representation common to the structures below.
(b) Two additional structures from replicate simulations of the structure in Fig. 4a.4. Additionally, aggregate sorting dynamics as well as the time development are presented. 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* approximately 160 (A) and 80 (B) (Toda et al., 2018).
Microscope images of *in vitro* system are modified from (Toda et al., 2018) (the colors red and blue in the *in vitro* system have been switched) for consistency with *in silico* framework and reprinted with permission from AAAS. Scale bar is 17.5 pixels, approximately 100um.

Fig. S8. The model captures inhibitory signaling to recapitulate well known patterns (Related to Main Fig. 4b).
(a) *In vitro* schematics and *in silico* representation of lateral inhibition network. In the *in silico* network, A is the sender state, A' is the receiver state. **(b.1)** Endpoint at 20,000mcs of a simulation starting from a fixed (A) cell monolayer (left) and on a disorganized cell (A) monolayer (right). **(b.2)** Endpoint of simulation after 20,000 mcs from a disorganized cell (A) monolayer with signaling parameters for strong inhibition (see Table S3). On the right is quantification quantification of cell area for cells A and cells A'; n=10 lattices of ~400 cells each, A cells area: 108.82±1.59 pixels^2, (A') cells area: 122.53±1.89 pixels^2, two-tailed matched t-test, d.f.=9, t=14.90, P<0.0001.
(c) Two additional structures from replicate simulations of the lateral inhibition structure in Fig. 4b. We give, additionally, the aggregate sorting dynamics. We present mean±s.d. (n=10 simulations *in silico* with initial condition of 93 (A) cells). Scale bar is 17.5 pixels, approximately 100um.

Fig. S9. Measuring length of structure for various elongation programs (Related to Main Fig. 5).
Graphical description of the process to calculate length of red+green cell stem of elongated structures.
3D structure and spatial cross section reconstruction of one representative elongation run are shown. In each, 3 cross sections are identified (white dotted squares), and the reference internal core is shown (orange dotted line).

Fig. S10 (Related to Main Fig. 6).
(a) Extended run time, 3000,000 steps, of simulation based on Network 1 of Fig. 6c. The mean of the homogeneity indexes $\Psi_{B,B'}$ and $\Psi_{A,A'}$ is shown over time.
(b) Extended run time, 3000,000 steps, of simulation based on Network 2 of Fig. 6c. The mean of the homogeneity indexes $\Psi_{B,B'}$ and $\Psi_{A,A'}$ is shown over time.
(c) Extended run time, 3000,000 steps, of simulation based on Network 3 of Fig. 6c. The mean of the homogeneity indexes $\Psi_{B,B'}$ and $\Psi_{A,A'}$ is shown over time.

Fig. S11. Potential biological versions of the oscillation structures (Related to Main Fig. 6).
(a) Possible biological implementation for each network. Het.Ad.1a and Het.Ad.1b are pairs of a heterotypic adhesion molecules. Homot. Ad. 1 and 2 are a homotypic adhesion pair.
(b) Sorting dynamics. Mean of sorting indexes $\Psi_{B,B'}$ and $\Psi_{A,A'}$ is shown over time for a Network 3 implementation as shown in Main Fig. 6c. The difference here is that cells are allowed to proliferate. The oscillation of morphology is still visible.

Table S1

	Induced **Hi.Ecad**	Basal **Hi.Ecad**	Induced **Ecad**	Basal **Ecad**	Induced **Lo.Ecad**	Basal **Lo.Ecad**	Induced **Ncad**	Basal **Ncad**	Induced **Pcad**	Basal **Pcad**	**C.Pcad**
Induced **Hi.Ecad**	20	40			35	49					40
Basal **Hi.Ecad**		45			40	49					42
Induced **Ecad**			25	42					33	45	
Basal **Ecad**				47					45	49	
Induced **Lo.Ecad**					40	49					
Basal **Lo.Ecad**						49					
Induced **Ncad**							35	42	49	49	49
Basal **Ncad**								47	49	49	49
Induced **Pcad**									35	42	
Basal **Pcad**										47	
C.Pcad											43

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. Basal values are before activation (due to promoter leakiness); induced values are after activation. Hi. denotes high expression, Lo. denotes low expression, and C. denotes constitutive expression per (Toda et al., 2018).

Note:

Cell-cell = 49 when at least one of the two is a parental cell

Cell-media = 52 always

Table S2

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

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 Marion Johnson and the other members of the Morsut lab, and members of USC stem cell department, Dr. Matt Thomson, members of the Thomson lab at Caltech 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. *105*.

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.

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

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

Cohen, M., Georgiou, M., Stevenson, N.L., Miodownik, M., and Baum, B. (2010). Dynamic Filopodia Transmit Intermittent Delta-Notch Signaling to Drive Pattern Refinement during Lateral Inhibition. Dev.

Cell *19*, 78–89.

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

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

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

Daringer, N.M., Dudek, R.M., Schwarz, K.A., and Leonard, J.N. (2014). Modular Extracellular sensor architecture for engineering mammalian cell-based devices. ACS Synth. Biol.

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

Ebrahimkhani, M.R., and Ebisuya, M. (2019). Synthetic developmental biology: build and control multicellular systems. Curr. Opin. Chem. Biol. *52*, 9–15.

Elowitz, M., and Lim, W.A. (2010). Build life to understand it. Nature.

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

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

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

Gilmour, D., Rembold, M., and Leptin, M. (2017). From morphogen to morphogenesis and back. Nature *541*, 311–320.

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

Gómez-Bombarelli, R., Wei, J.N., Duvenaud, D., Hernández-Lobato, J.M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T.D., Adams, R.P., and Aspuru-Guzik, A. (2018). Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. ACS Cent. Sci.

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

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

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

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

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

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

Kriegman, S., Blackiston, D., Levin, M., and Bongard, J. (2019). A scalable pipeline for designing reconfigurable organisms.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Qudrat, A., and Truong, K. (2017b). Antibody-based fusion proteins allow Ca2 + rewiring to most extracellular ligands.

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

Roese-Koerner, B., Stappert, L., and Brüstle, O. (2017). Notch/Hes signaling and miR-9 engage in complex feedback interactions controlling neural progenitor cell proliferation and differentiation. Neurogenesis.

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

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

Scheller, L., Strittmatter, T., Fuchs, D., Bojar, D., and Fussenegger, M. (2018). Generalized extracellular molecule sensor platform for programming cellular behavior Leo. Nat. Chem. Biol.

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

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

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

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

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

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

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.

Teague, B.P., Guye, P., and Weiss, R. (2016). Synthetic Morphogenesis. Cold Spring Harb. Perspect. Biol. *8*, a023929.

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*, 156–162.

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.

Wennekamp, S., Mesecke, S., Nédélec, F., and Hiiragi, T. (2013). A self-organization framework for symmetry breaking in the mammalian embryo. Nat. Rev. Mol. Cell Biol.

Fig. 1

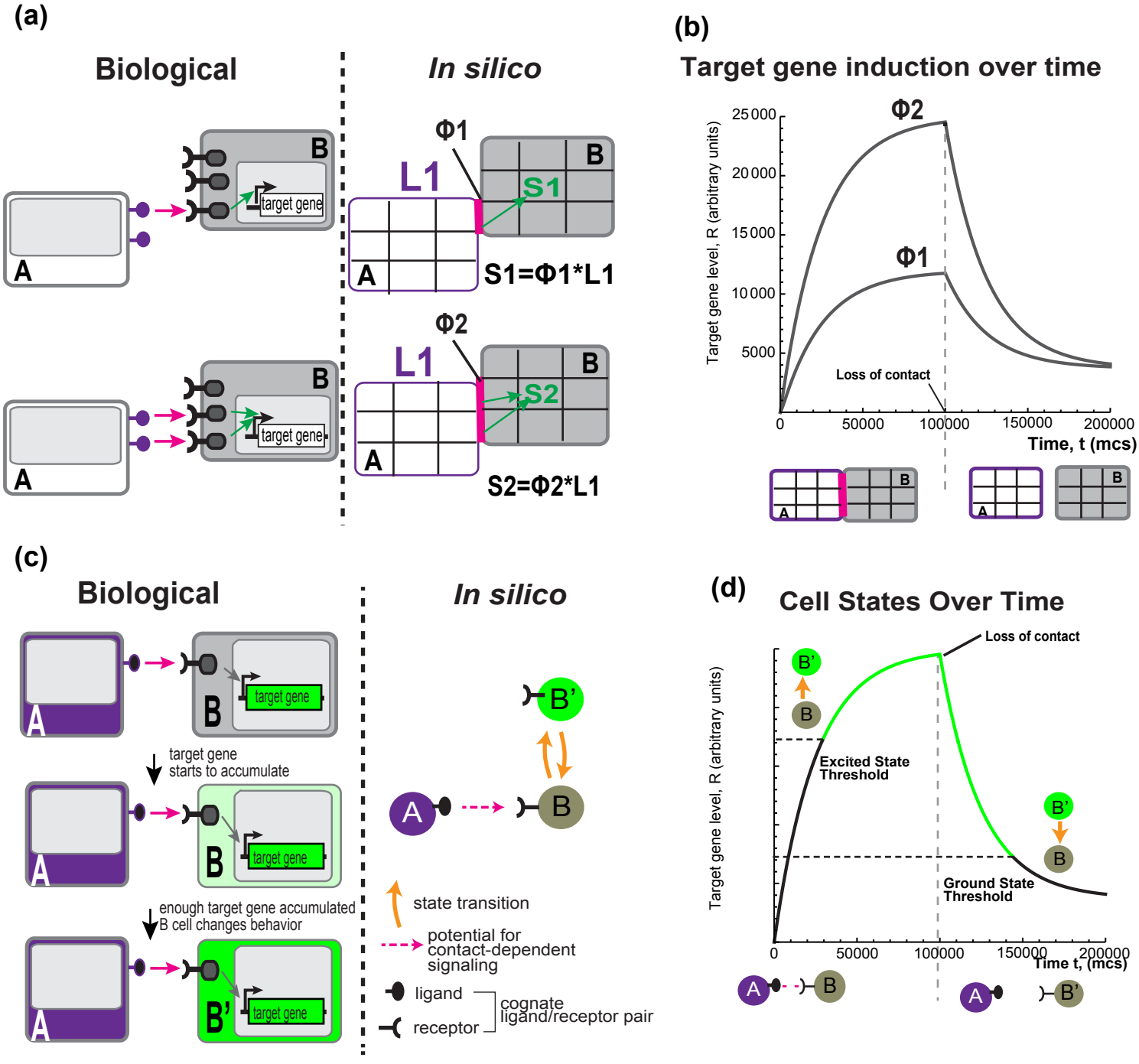

(a)
Biological
In silico
A
B
target gene
Φ1
L1
S1
S1=Φ1*L1
Φ2
S2
S2=Φ2*L1
(b)
Target gene induction over time
Target gene level, R (arbitrary units)
25000
20000
15000
10000
5000
Loss of contact
0
50000
100000
150000
200000
Time, t (mcs)
(c)
Biological
In silico
target gene
starts to accumulate
enough target gene accumulated
B cell changes behavior
B'
state transition
potential for
contact-dependent
signaling
ligand
receptor
cognate
ligand/receptor pair
(d)
Cell States Over Time
Loss of contact
Excited State
Threshold
Ground State
Threshold
Target gene level, R (arbitrary units)
0
50000
100000
150000
200000
Time t, (mcs)

Fig. 2

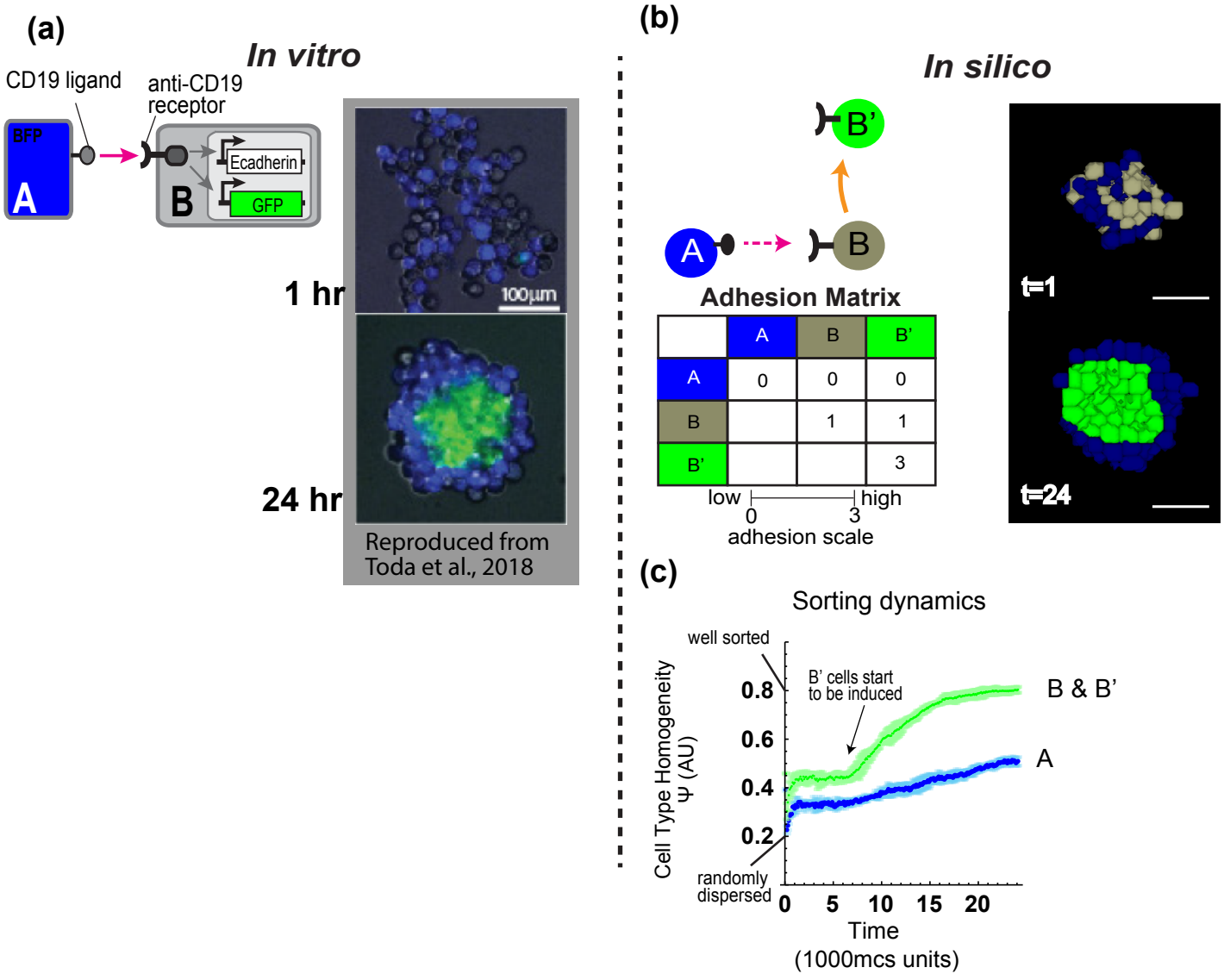
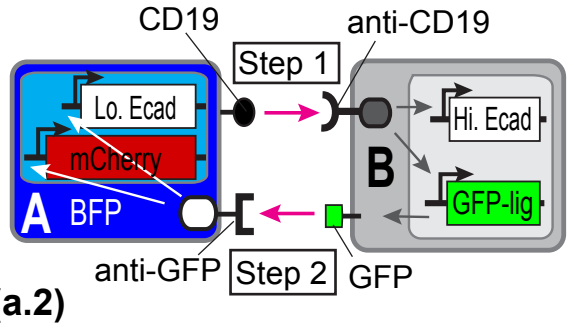
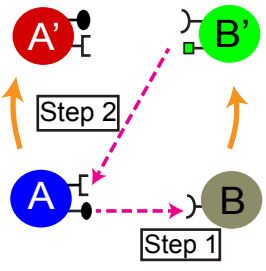

(a)
In vitro
CD19 ligand
anti-CD19 receptor
BFP
A
B
Ecadherin
GFP
1 hr
24 hr
100µm
Reproduced from Toda et al., 2018
(b)
In silico
B'
A
B
Adhesion Matrix
A B B'
A 0 0 0
B 1 1
B' 3
low high
0 3
adhesion scale
t=1
t=24
(c)
Sorting dynamics
well sorted
randomly dispersed
B' cells start to be induced
B & B'
A
Cell Type Homogeneity Ψ (AU)
0.8
0.6
0.4
0.2
0 5 10 15 20
Time
(1000mcs units)

Fig. 3

(a.1) *In vitro*



(b.1) *In silico*



Adhesion Matrix

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

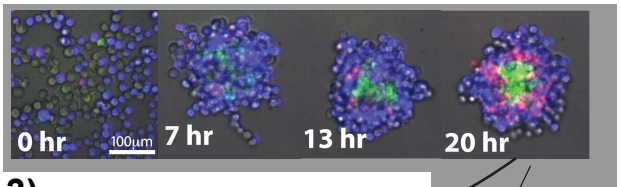
low — high
0 — 3
adhesion scale

(a.2)

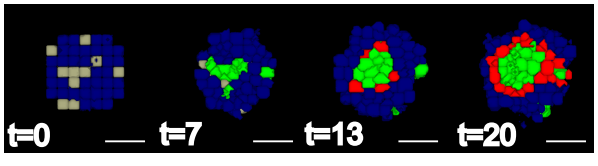
Synthetic developmental trajectory

(b.2)

In vitro



In silico

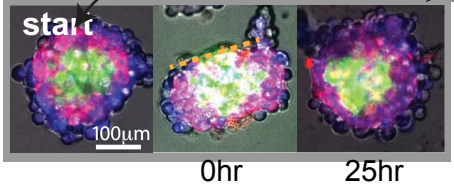


(a.3)

Regeneration

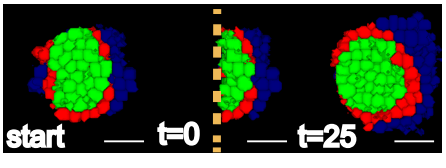
Bisected

Reproduced from Toda et al., 2018



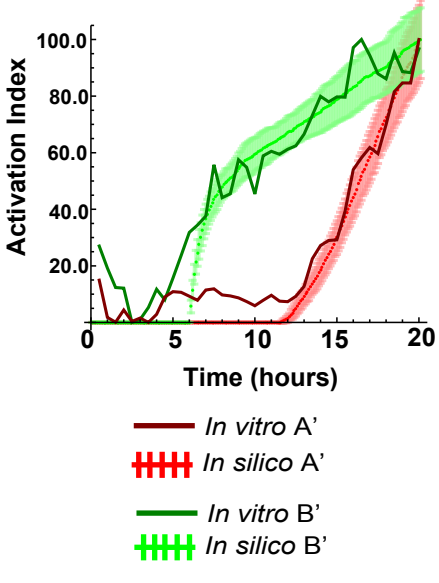
0hr 25hr

(b.3)



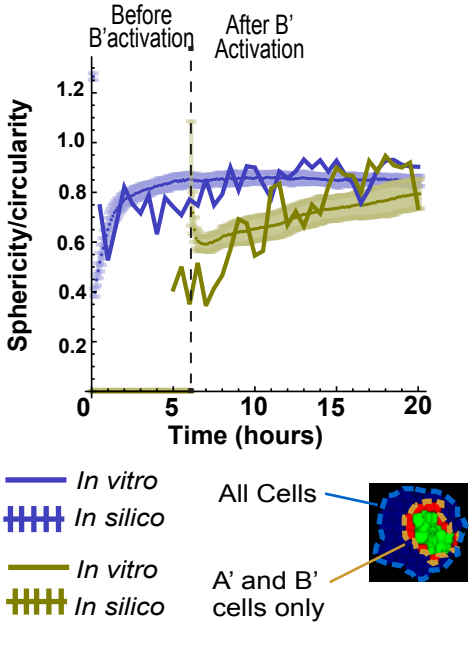
(c)

Cell activation dynamics



(d)

Morphology dynamics



(e)

Core Distribution

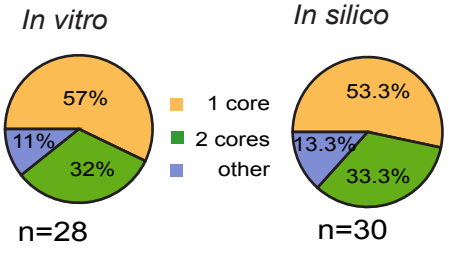
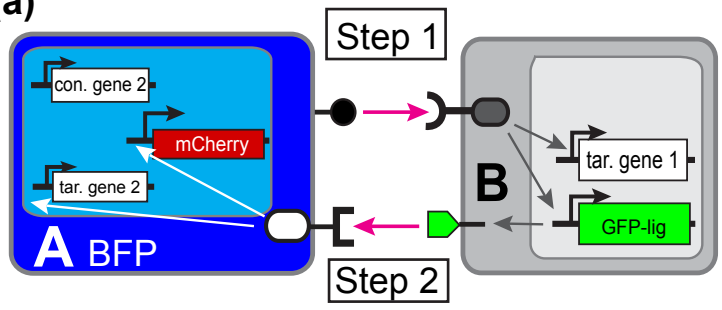


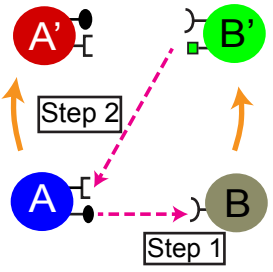
Fig. 4

(a)

In vitro



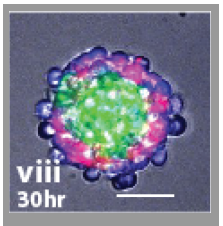
In silico

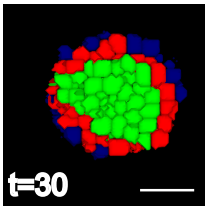


(a.1)

Gene Matrix

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





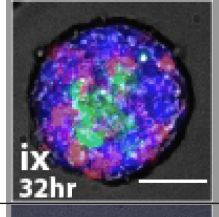
Adhesion Matrix

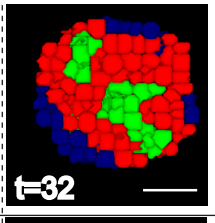
	A	A'	B	B'
A	1	1	0	1
A'		2	1	2
B			1	1
B'				3

(a.2)

Gene Matrix

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





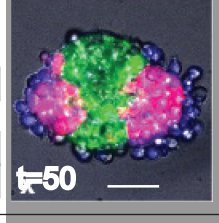
Adhesion Matrix

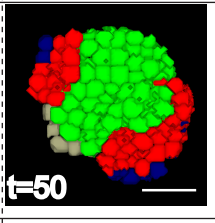
	A	A'	B	B'
A	1	1	1	1
A'		1	1	1
B			1	1
B'				3

(a.3)

Gene Matrix

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





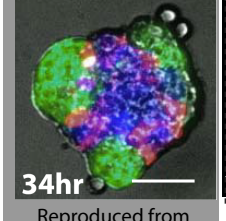
Adhesion Matrix

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

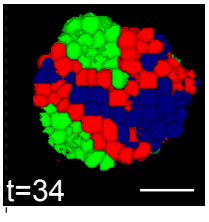
(a.4)

Gene Matrix

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



Reproduced from Toda et al., 2018



Adhesion Matrix

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

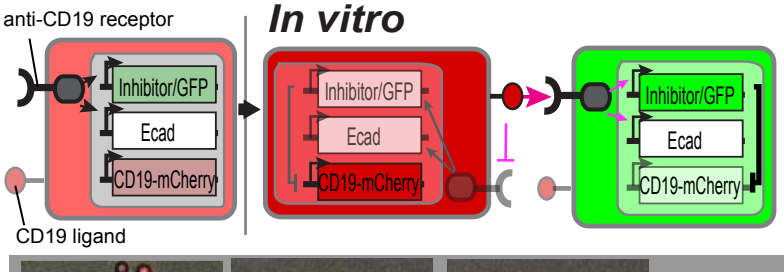
low — high
0 — 3
adhesion scale

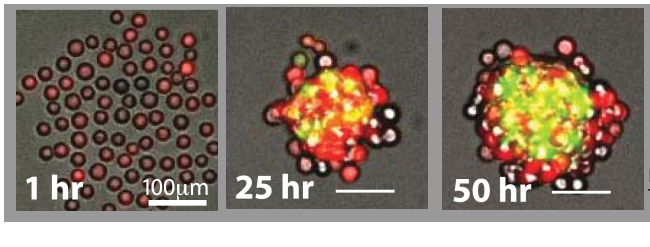
Estimated *in vitro* adhesion preferences:

Pcad:Pcad = Ncad:Ncad > Pcad:Ncad

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

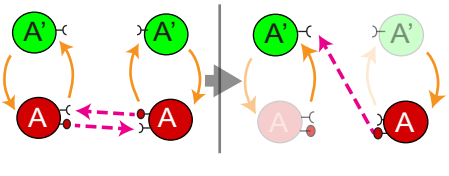
(b)





Reproduced from Toda et al., 2018

In silico



Adhesion Matrix

	A	A'
A	1	1
A'		3

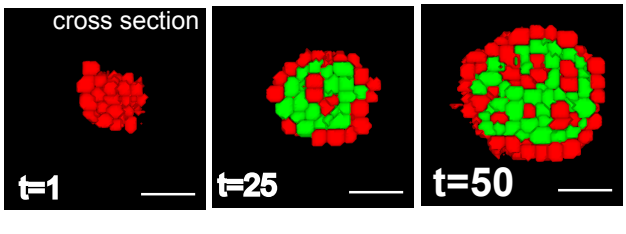
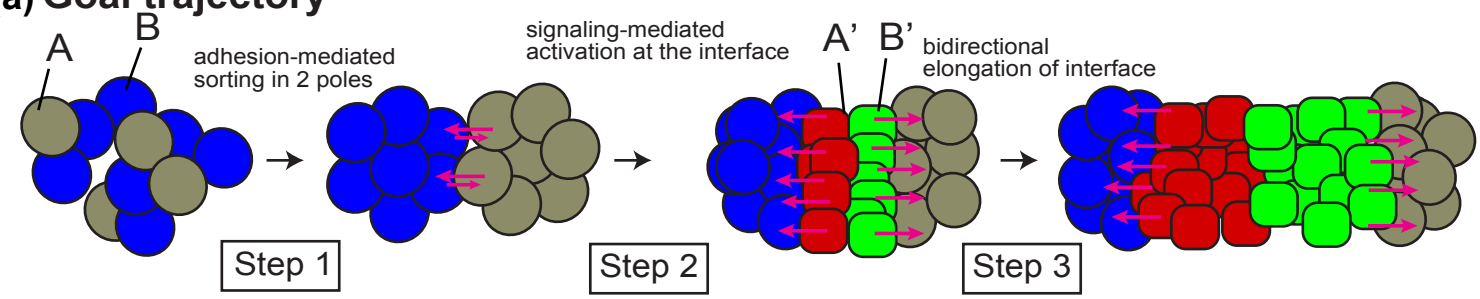
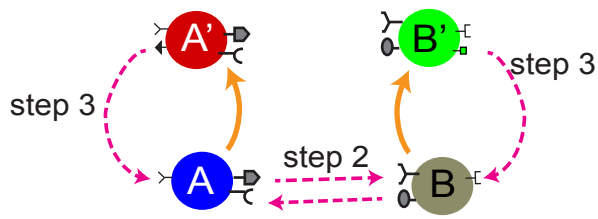


Fig. 5

(a) Goal trajectory



(b) Proposed Network

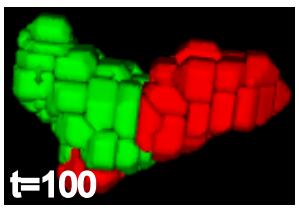


Adhesion Matrix

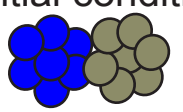
	A	A'	B	B'
A	2	0	0	0
A'		5	0	5
B			2	0
B'				5

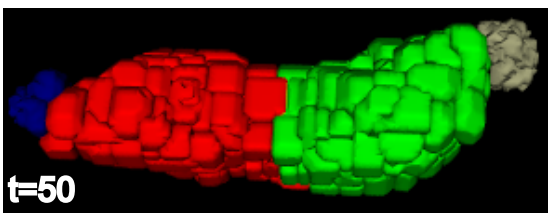
adhesion scale: 0 low, 2 high, 5 highest

(c) Implementation 1

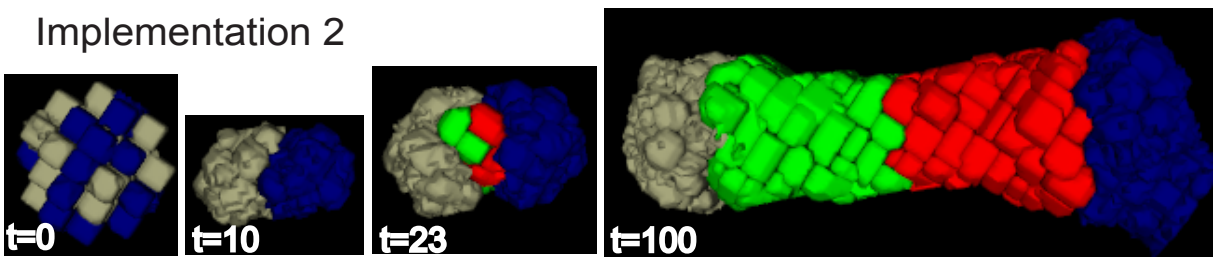


(d) Implementation 1 with controlled initial condition





(e) Implementation 2



(f)

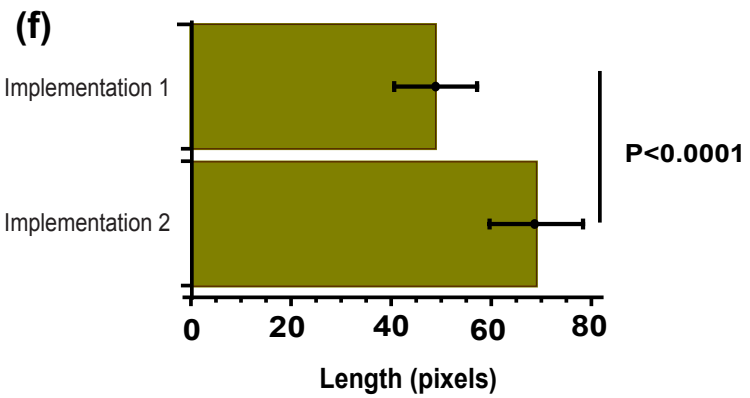
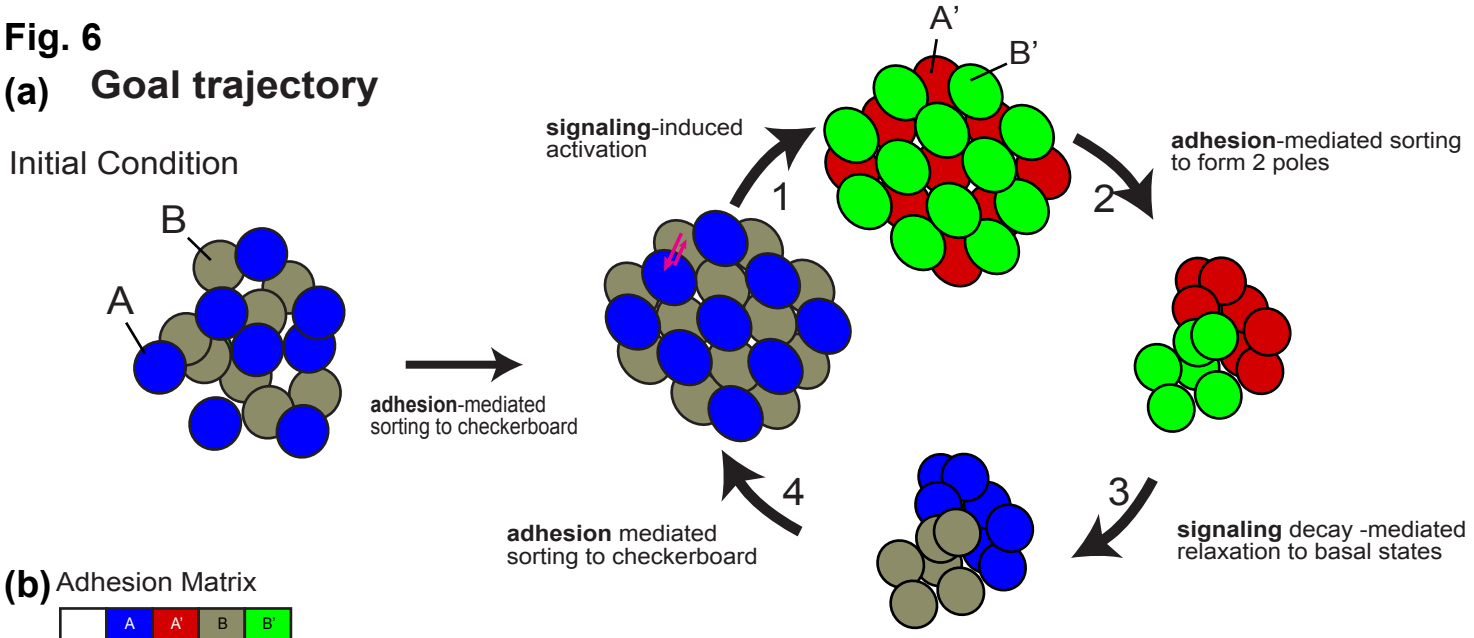


Fig. 6

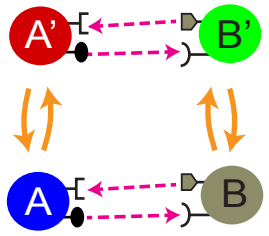
(a) Goal trajectory

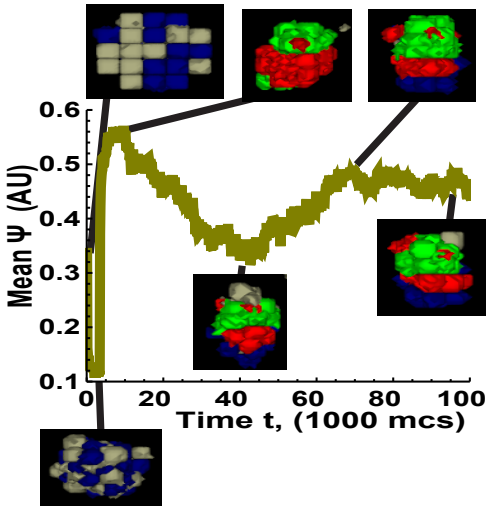


(b) Adhesion Matrix

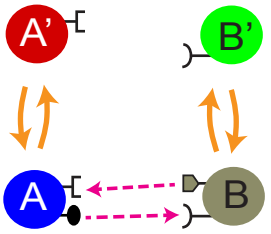
	A	A'	B	B'
A	0	0	2	0
A'		2	0	0
B			0	0
B'				2

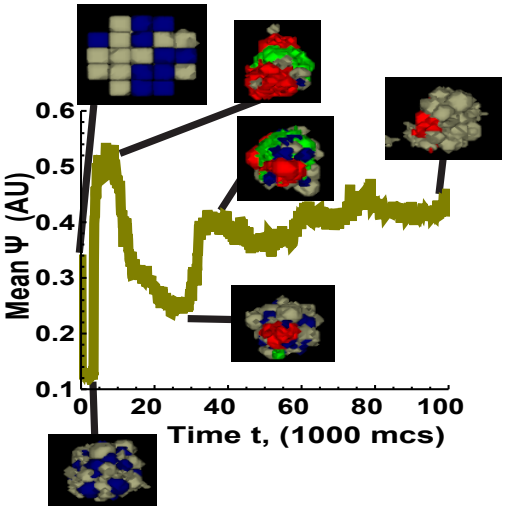
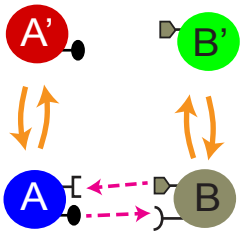
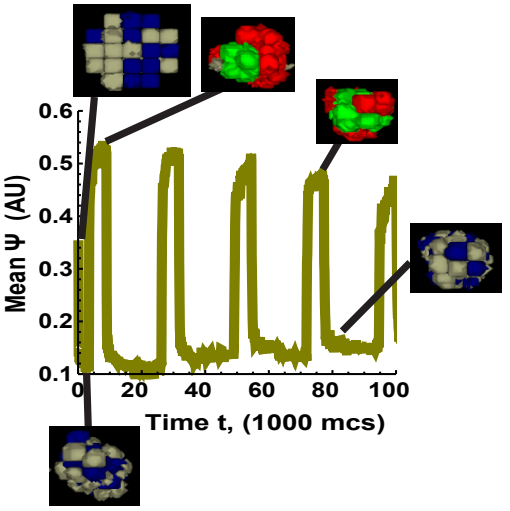
(c) Network 1





(d) Network 2





(e) Network 3