

RESEARCH ARTICLE

A computational modeling approach for predicting multicell spheroid patterns based on signaling-induced differential adhesion

Nikita Sivakumar¹, Helen V. Warner¹, Shayn M. Peirce¹, Matthew J. Lazzara^{1,2*}

1 Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, United States of America, **2** Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia, United States of America

* mlazzara@virginia.edu



OPEN ACCESS

Citation: Sivakumar N, Warner HV, Peirce SM, Lazzara MJ (2022) A computational modeling approach for predicting multicell spheroid patterns based on signaling-induced differential adhesion. *PLoS Comput Biol* 18(11): e1010701. <https://doi.org/10.1371/journal.pcbi.1010701>

Editor: Philip K. Maini, Oxford, UNITED KINGDOM

Received: October 8, 2021

Accepted: November 1, 2022

Published: November 28, 2022

Copyright: © 2022 Sivakumar et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The code associated with this study can be accessed here: <https://github.com/lazzaralab/multicell-synthetic-cell-signaling-abm>.

Funding: This work was supported by funding from the National Science Foundation Grant No. 1700687 (MJL), University of Virginia Engineering in Medicine Program (MJL, SPC), University of Virginia Center for Advanced Biomanufacturing (MJL), and Arnold and Mabel Beckman Foundation (NS). The funders had no role in study design, data

Abstract

Physiological and pathological processes including embryogenesis and tumorigenesis rely on the ability of individual cells to work collectively to form multicell patterns. In these heterogeneous multicell systems, cell-cell signaling induces differential adhesion between cells that leads to tissue-level patterning. However, the sensitivity of pattern formation to changes in the strengths of signaling or cell adhesion processes is not well understood. Prior work has explored these issues using synthetically engineered heterogeneous multicell spheroid systems, in which cell subpopulations engage in bidirectional intercellular signaling to regulate the expression of different cadherins. While engineered cell systems provide excellent experimental tools to observe pattern formation in cell populations, computational models of these systems may be leveraged to explore more systematically how specific combinations of signaling and adhesion parameters can drive the emergence of unique patterns. We developed and validated two- and three-dimensional agent-based models (ABMs) of spheroid patterning for previously described cells engineered with a bidirectional signaling circuit that regulates N- and P-cadherin expression. Systematic exploration of model predictions, some of which were experimentally validated, revealed how cell seeding parameters, the order of signaling events, probabilities of induced cadherin expression, and homotypic adhesion strengths affect pattern formation. Unsupervised clustering was also used to map combinations of signaling and adhesion parameters to these unique spheroid patterns predicted by the ABM. Finally, we demonstrated how the model may be deployed to design new synthetic cell signaling circuits based on a desired final multicell pattern.

Author summary

The remarkable ability of cells to self-organize is critical for the assembly of functional tissues during embryogenesis and is impaired by the molecular aberrations that lead to tumorigenesis and metastasis. Learning the rules of cellular self-assembly will provide a new way to understand such physiological and pathophysiological processes and create an instruction manual for designing tissues from scratch for therapeutic applications.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Biologists have begun to learn the rules of self-assembly through reverse engineering—that is, through engineering biochemical circuits that control how cells adhere to one another. These approaches can yield simple, multi-cell structures that self-assemble into a core of one cell type surrounded by a shell of another cell type, for example. However, engineering more complex tissue patterns requires exploring a large domain of circuit structures and parameters. To facilitate this exploration in a guided and systematic manner, we created a computational model that predicts how mixtures of cells with different circuitry for expressing adhesion proteins will interact to form varied patterns in spheroidal tissues comprised of hundreds of cells. Our model enables the design of internal molecular signaling circuitry that permits cells to be used as building blocks for self-assembled tissues with specific structures, and ultimately functions.

Introduction

Multicell patterning processes in embryonic development and tumorigenesis are regulated by dynamic cell-cell signaling and differential adhesion between cell subpopulations [1,2]. Steinberg and colleagues [3] previously proposed the differential adhesion hypothesis to explain how patterns emerge at the multicell level in heterogeneous mixtures of cells. Specifically, they proposed that cell subpopulations self-organize to minimize the adhesive free energy between cells such that more adhesive cell subpopulations are surrounded by less adhesive ones [3–5]. The differential adhesion hypothesis has been validated experimentally [3–7] and incorporated into computational models of cell sorting and morphogenesis [8–11]. However, this conceptual model does not account for how dynamic cell-cell signaling processes, which regulate adhesion protein expression, may yield emergent patterns that differ from those predicted by purely equilibrium considerations (i.e., minimization of adhesive free energy).

The ability of a cell to adhere to another cell may be time- and space-dependent based on cell-cell signaling that alters expression of adhesion proteins. In embryonic development, for example, cell-cell interactions initiate transcriptional processes that regulate expression of specific adhesion proteins, leading to the segregation of cell subpopulations into organ-specific tissue layers [12–19]. Similarly, in many cancers, signaling cascades cause individual cells to downregulate expression of the cell adhesion protein E-cadherin, leading to epithelial-mesenchymal transition (EMT) [20–24]. EMT creates the opportunity for differential adhesion between epithelial cells and the mesenchymal counterparts they can generate, impacting the spatial configuration of these subpopulations throughout the tumor [25–28]. Emergent cell patterning within tumors can impact tumor growth, migration, and response to treatment [20,28–34]. Understanding the mechanisms by which multicell patterning occurs can inform drug targets for various congenital diseases and other pathologies that occur throughout adulthood.

Synthetic biology approaches have been used to impart self-organizing properties to multicellular systems [2,35–41]. These systems provide an ideal experimental setting to explore pattern formation resulting from dynamic cell-cell signaling because the cell interactions are clearly defined and programmable. A recent example of interest is the bidirectional signaling system engineered by Toda et al. [2], in which sender cells drive expression of a particular cadherin in receiver cells that then reciprocally drive expression of a different cadherin in the initial sender cells (Fig 1A). The authors observed that the relative homotypic and heterotypic cadherin interaction strengths directly impacted observed spheroid patterns. When the network was engineered with N-cadherin (Ncad) and P-cadherin (Pcad), which both have high

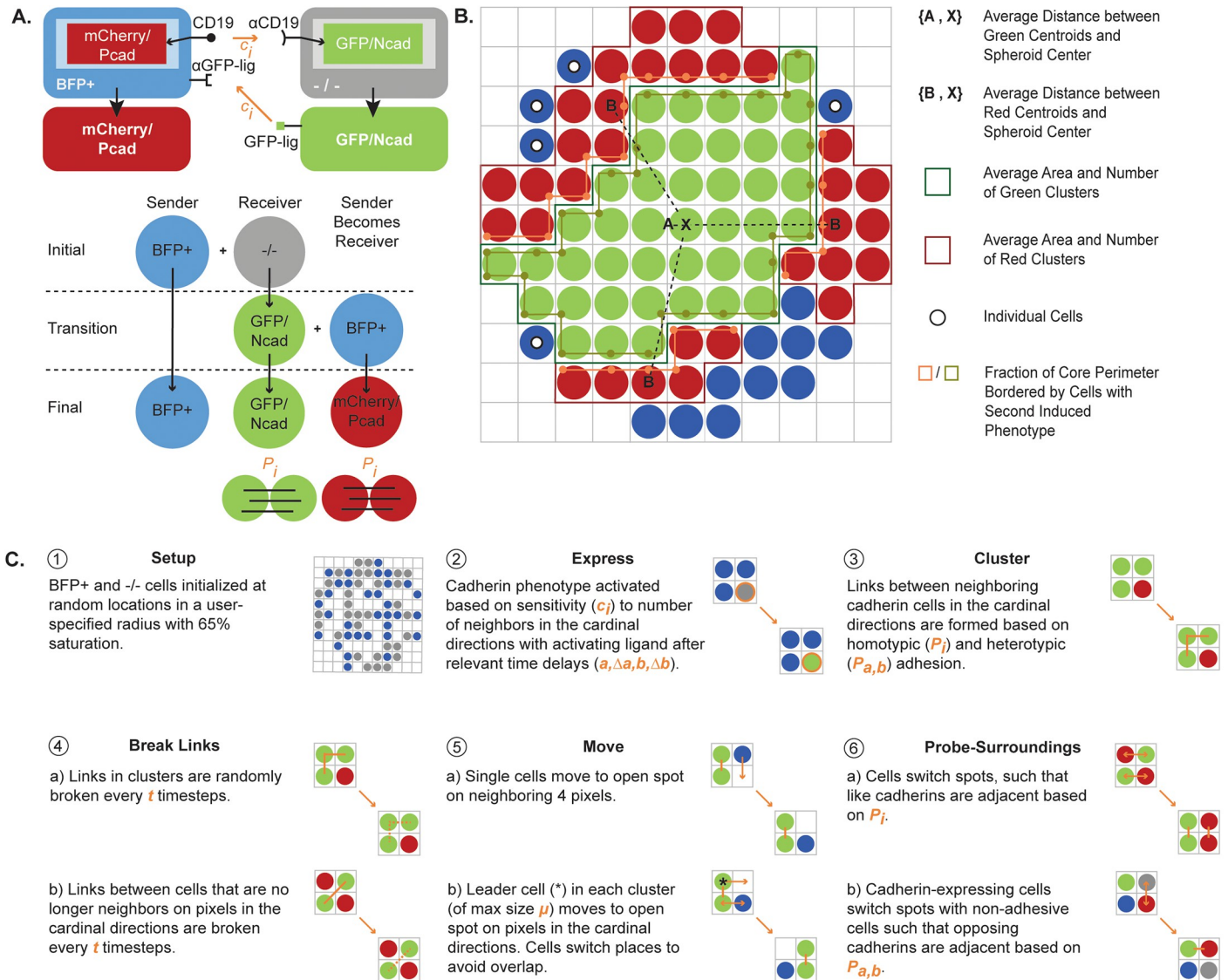


Fig 1. Synthetic biology functionality can be encoded in an agent-based model of cell-cell signaling within a spheroid. (A) A schematic of the synthetic bidirectional signaling circuit engineered by Toda et al. [2] and implemented in the agent-based model (ABM) is shown. Spheroids are seeded with non-adhesive “sender” BFP+ cells and “receiver” -/- cells (labeled -/- to indicate no fluorescent protein expression). Sender BFP+ cells can induce GFP/Ncad expression in -/- receiver cells, and newly formed GFP/Ncad cells behave as sender cells to induce mCherry/Pcad expression reciprocally in BFP+ cells (at which point BFP+ cells behave as receiver cells). The final spheroid composition can include GFP/Ncad, mCherry/Pcad, or BFP+ cells (that never contacted a GFP/Ncad cell and thus failed to transition to the mCherry/Pcad phenotype). In this ruleset, the model cadherin induction constant (c_i) controls the probability of GFP/Ncad induction in -/- cells based on neighboring BFP+ cells and the probability of mCherry/Pcad induction in BFP+ cells based on neighboring GFP/Ncad cells. The model homotypic adhesion strength (P_i) controls the probability of forming junctions between cells expressing like cadherins on adjacent grid spaces. (B) Metrics used to analyze ABM outputs are represented schematically. Regions of a given color comprising three or more cells are defined as cell clusters. These metrics are further clarified in Table 1. (C) The rules executed at each ABM timestep are depicted. Parameters, indicated in bold italics, are defined in Table 2.

<https://doi.org/10.1371/journal.pcbi.1010701.g001>

homotypic and low heterotypic adhesion strengths, a “core/pole” pattern arose. Specifically, Ncad-expressing receiver cells formed a core decorated by smaller clusters, or poles, of Pcad-expressing sender cells and surrounding non-adhesive sender cells [2]. When the system was instead engineered with high and low expression of E-cadherin (Ecad), creating cells with high (but differential) homotypic and heterotypic adhesion, a “core/shell” pattern arose.

Specifically, high-Ecad receiver cells formed a core encapsulated by a shell of low-Ecad sender cells [2]. In both systems, the number of poles and core/shell spheroids formed varied across trials with the same initial conditions [2], but the factors that led to this variable outcome were not systematically investigated. These results open several questions about how combinations of cell-cell signaling and differential adhesion events affect multicell patterning. The multicell patterns that form in these systems may depend on the initial ratio of cell subpopulations, the order of signaling events, and/or the induced adhesion strengths between individual cells, but more work is needed to explore such potential dependencies and how combinations of these perturbations may affect the robustness of pattern formation.

While it would be impossible to explore the effects of perturbing different system parameters exhaustively through experiments, computational models provide an efficient method for predicting unique multicellular patterns for different parameter combinations [42–46]. ABMs are a particularly useful computational approach for modeling heterogeneous cell types and spheroid patterning because ABMs can represent individual cells within a population and predict how cell type-specific behaviors produce spatial patterns at the multicell level [8–10,28,47–54]. Here, we developed two-dimensional (2D) and three-dimensional (3D) ABMs of the Ncad/Pcad bidirectional signaling system engineered by Toda et al. [2]. Cells were modeled as agents expressing different cadherins with a specified time delay in response to heterotypic receptor-ligand interactions and as differentially adhering to like or unlike cells based on the homotypic and heterotypic adhesion properties of the cadherins. Model parameters were tuned to fit the ABM to experimental observations we made for a 1:1 ratio of the seeded cell types, and the fitted model was experimentally validated against experiments with unequal cell seeding densities. Unsupervised machine learning was then employed to map different combinations of perturbed model parameters (e.g., characterizing cadherin induction and adhesion strengths) to specific spheroid patterns predicted by the ABM, including the previously described core/pole and core/shell patterns and others resembling a soccer ball, bull's eye, or stripes. The ABM was also used to design synthetic cell-cell signaling circuits that encode specified multicellular patterns, demonstrating a potential practical application of this modeling approach.

Results

2D and 3D ABMs predict the impacts of dynamically activated differential adhesion on spheroid patterning

2D and 3D ABMs were developed based on the synthetic Pcad/Ncad circuit engineered by Toda et al. [2] (Fig 1A). The models were seeded with “sender” and “receiver” cells. Sender cells express blue fluorescent protein (BFP+), CD19 ligand, and green fluorescent protein (GFP) receptor. Receiver cells express the CD19 receptor (-/-). In response to CD19 receptor-ligand binding between the two cell types, receiver -/- cells express Ncad and GFP to become GFP/Ncad cells. GFP/Ncad cells then behave as sender cells and reciprocally promote Pcad and mCherry expression in BFP+ cells (that then behave as receiver cells) via GFP ligand-receptor binding, forming mCherry/Pcad cells. Hence, in this synthetically engineered multicell system, both cell types can function as either sender or receiver cells based on receptor-ligand binding interactions with adjacent cells, and the terms “sender” and “receiver” are used to clarify a cell's role during specific steps in bidirectional signaling. Final spheroid patterns were quantified using metrics such as total fraction of spheroid area occupied by each cell type, average areas of homotypic cell clusters, and average distances between homotypic cell clusters and the spheroid center (Fig 1B, Table 1). Homotypic cell “clusters” were defined as three or more cells of the same cell type adjoined by cell-cell junctions on adjacent grid spaces in the

Table 1. Metrics used to quantify multicell patterns in 2D ABMs.

Metric	Definition
Total fractional area	Number of cells of a specific color (blue, green, or red) as a fraction of the total number of cells of that color in the spheroid.
Average cluster fractional area	Average of the areas of all clusters of cells of the same color as a fraction of the total spheroid area. Note that a “cluster” is defined as three or more neighboring cells in the cardinal directions of the same color.
Average fractional distance from cluster centroids to spheroid center	Average of the distances between the centroids of individual clusters of a specific color to the center of the spheroid as a fraction of the radius of the spheroid.
Number of clusters	Total number of clusters of a given cell type/color.
Number of individual cells	Total number of cells of a given cell type/color that are not a part of a cluster.
Cluster circularity	Average of the ratios of width to height for all clusters of a given cell type/color.
Small cluster circularity	Ratio of cluster width to height for the smallest cluster in the spheroid.
Contiguous area fraction	Fraction of identically colored cells on the perimeter of the core cell type bordered by the other-colored, adhesive cells in the system.

<https://doi.org/10.1371/journal.pcbi.1010701.t001>

cardinal directions. The “core” cell type of a spheroid was defined as the cell type of the homotypic cluster with minimum distance to all other homotypic clusters in the spheroid.

Interactions among cells in the ABM were encoded as six rules incorporating nine model parameters that execute on each timestep of the simulation, where one timestep represents 30 min (Fig 1C, Table 2). Cadherin phenotypes were induced in the model based on induction constants (c_i) that control the probability of cadherin induction in a receiver cell based on the number of neighboring sender cells. Over the time course of the model, neighboring cadherin-expressing cells form junctions with each other based on homotypic (P_i) and heterotypic ($P_{a,b}$) adhesion strengths encoded as probabilities of adhesion in the model. To recapitulate the

Table 2. ABM parameter values.

Parameter	Value	Definition	Justification for Value
a	10 ticks	Delay for GFP/Ncad phenotype to be induced in the system.	Estimated from Toda et al. (2018); set such that GFP/Ncad cells start to appear at ~26 ticks in the model.
Δa	10 ticks	Time to express Ncad.	
b	25 ticks	Delay for mCherry/Pcad to be induced in the system.	Estimated from Toda et al. (2018); set such that mCherry/Pcad cells start to appear at ~42 ticks in the model.
Δb	10 ticks	Time to express Pcad.	
μ	5 cells	Maximum movable cluster size (in number of cells).	Manually set such that smaller homotypic clusters can move throughout the spheroid to improve likelihood of homotypic clusters merging to form larger homotypic clusters.
t	3 ticks	Delay allotted after links are randomly broken in clusters to be reformed.	Manually set such that sufficient time is allowed for individual cell agents to move after a link is broken, so that new clusters can be formed in the spheroid.
c_i	0.20 (unitless)	Induction constant used in function that controls the probability of a cadherin phenotype being induced in a receiver cell based on the number of neighboring sender cells. c_a and c_b are used to refer to induction constants for GFP/Ncad and mCherry/Pcad phenotypes, respectively, in Figs 6–10.	Identified by the evolutionary algorithm.
P_i	0.95 (unitless)	Probability of homotypic adhesion in induced cadherin cell types. P_a and P_b are used to refer to individual homotypic adhesion strengths of GFP/Ncad and mCherry/Pcad, respectively, in Figs 6–10.	
$P_{a,b}$	0.05 (unitless)	Probability of heterotypic adhesion between GFP/Ncad and mCherry/Pcad cells.	

<https://doi.org/10.1371/journal.pcbi.1010701.t002>

Ncad/Pcad spheroids observed in Toda et al. [2], we assumed that both cadherin cell types have an equal induction constant (c_i) and equal homotypic adhesion strength (P_i). In later sections, we explore how spheroid patterning is impacted by assigning unequal values of c_i and P_i to the GFP/Ncad phenotype (c_a, P_a) and to the mCherry/Pcad phenotype (c_b, P_b). All simulations that were compared to *in vitro* experiments were run for 100 timesteps, where one timestep represents 30 min, because *in vitro* spheroids were observed for 50 hr. Additionally, because published data in Toda et al. [2] indicated that cells roughly doubled in number over the course of 50 hr, ABMs were seeded with twice the number of cell agents as the corresponding *in vitro* experiment (i.e., an *in vitro* experiment seeded with 200 cells corresponded to an ABM seeded with 400 cell agents). All other simulations were run based on pilot simulations to identify runtimes beyond which no changes were observed in the spheroid pattern, and the specific model runtime for each set of simulations is denoted in corresponding figure captions.

Published data and an evolutionary algorithm were used to parameterize the ABM

Parameters in the 2D ABM were either estimated based on published experimental data in Toda et al. [2] or computationally tuned with an evolutionary algorithm (EA) (Table 2). The time delay for cadherin phenotypes to be induced (a for GFP/Ncad and b for mCherry/Pcad) and the times to express each cadherin ($\Delta a, \Delta b$) were manually adjusted in the 2D ABM according to published data in Toda et al. [2]. Specifically, these parameters were set such that in simulations of experiments seeded with 100 BFP+ and 100 -/- cells, GFP/Ncad cells formed after 13 hr and mCherry/Pcad cells developed adjacent to GFP/Ncad cells after 21 hr, the respective times reported by Toda et al. [2]. Because the remaining five parameters ($c_i, P_i, P_{a,b}, \mu, t$) could not be determined from experimental data, these parameters were first manually approximated in the 2D ABM as values yielding outputs qualitatively similar to published images of spheroids from the Ncad/Pcad bidirectional signaling system in Toda et al. [2] (i.e., characterized by a GFP/Ncad central core flanked by mCherry/Pcad clusters and individual BFP+ cells). Then, a univariate parameter sensitivity analysis was implemented for the 2D ABM to determine which of these parameters had the largest relative impact on spheroid patterning. Each of the five parameters was increased or decreased by 10% from its base value, and the effects of these perturbations on the numbers and average areas of green and red clusters were determined (Fig 2A–2D). This analysis identified the cadherin induction constant (c_i) and the homotypic adhesion strength of both cadherin cell types (P_i) as the two parameters whose variation produced the largest changes in all four measured outputs. In this univariate sensitivity analysis, we assumed both cadherin cell types have an equal induction constant (c_i) and equal homotypic adhesion strength (P_i).

Values for c_i and P_i were determined using an evolutionary algorithm (EA) that optimized the 2D ABM prediction of patterns that were experimentally observed in spheroids seeded with 200 cells at a 1:1 ratio of BFP+ and -/- cells. A detailed description of the EA implementation is provided in *Materials and Methods*. A plot of 30 EA runs over 100 generations demonstrates that the EA accomplished a substantial reduction in error (difference between model prediction and data) to identify values for c_i and P_i (Fig 2E). The parameter values optimized for the 2D ABM were also used in the 3D ABM.

ABMs predict spheroid patterns across a range of different initial conditions that match *in vitro* results

2D and 3D ABMs were used to predict spheroid patterns when the BFP+: -/- cell seeding ratio was varied (1:9, 1:4, 1:1, 4:1, 9:1) for a fixed total number of seeded cells (60 or 200 cells)

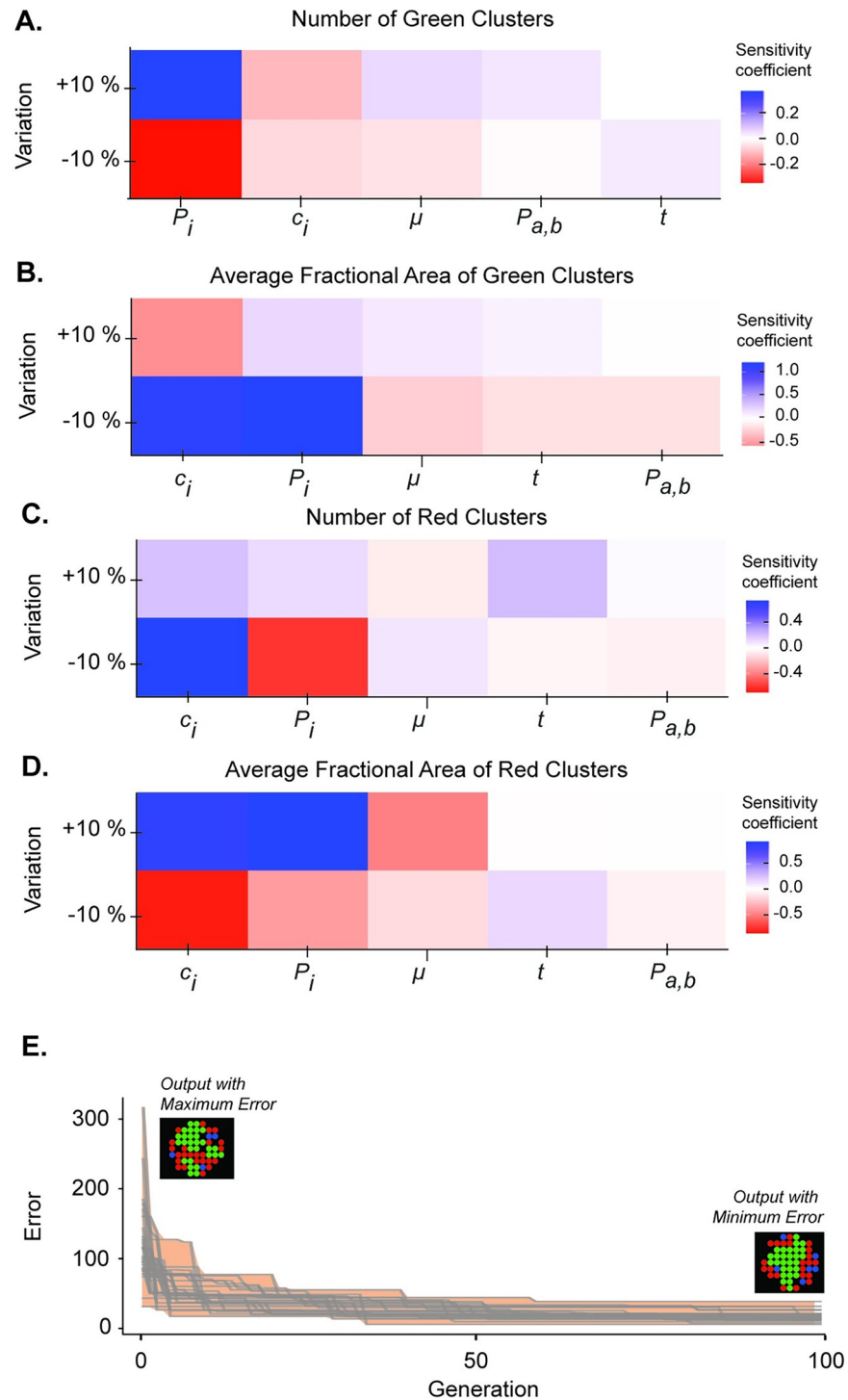


Fig 2. Parameter sensitivity analysis and optimization improves model fit to *in vitro* data. Parameters that could not be estimated based on experimental data (c_i , P_i , $P_{a,b}$, μ , t) were perturbed by a factor 0.9 and 1.1, and the sensitivity coefficient was quantified for the following outputs: (A) number of green clusters, (B) average fractional area of green clusters, (C) number of red clusters, and (D) average fractional area of red clusters. Sensitivity coefficients are reported as the average of 100 model runs, each run for 100 timesteps. The x -axis is ordered from greatest to least absolute-value sensitivity coefficient for each panel. (E) An evolutionary algorithm (EA) was used to tune parameters c_i and P_i by minimizing the error between model outputs and experimental images of 200-cell spheroids seeded with a 1:1 ratio of BFP+ and -/- cells. The EA was run 30 times for a population of 20 model combinations over 100 generations. Within each generation, the error for each parameter combination was averaged over 10 runs for 100 timesteps. Sample model graphical outputs generated from the parameter combinations with maximum or minimum error are shown.

<https://doi.org/10.1371/journal.pcbi.1010701.g002>

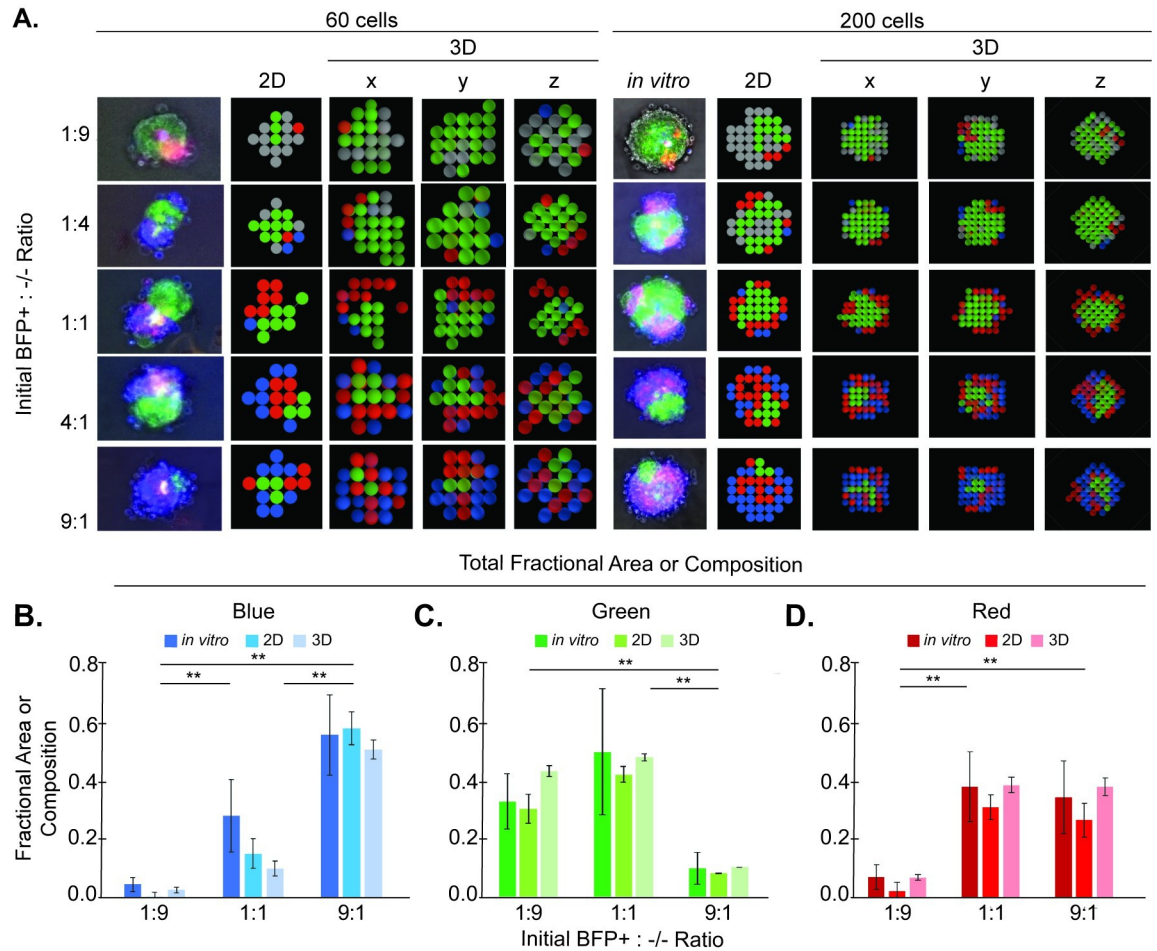


Fig 3. The ABM qualitatively and quantitatively recapitulates *in vitro* pattern formation. (A) Comparisons are shown of representative cross-sectional images of *in vitro* spheroids with 2D or 3D ABM outputs. 3D ABM *x*-, *y*-, and *z*-cross-sections provide 2D views of the spheroid cores. *In vitro* cross-sectional images are microscopy images of a single focal plane, without deconvolution, in the approximate center of the spheroid. Comparisons of (B) blue, (C) green, and (D) red fractional areas or compositions between *in vitro*, 2D ABM, and 3D ABM images for varying initial BFP+/-/- seeding ratios (1:9, 1:1, 9:1) and 200 cells. For 3D ABM outputs, a fractional composition, rather than fractional area, was computed as the number of cell agents of a given color channel divided by the total number of cell agents in the spheroid. ABM-predicted fractional areas or compositions for the 2D and 3D models, respectively, were calculated as an average of 100 model runs, each run for 100 timesteps, for each experimental condition. Fractional areas from *in vitro* images were calculated from a sample of 9, 10, and 10 images for the 1:9, 1:1, and 9:1 ratios, respectively. A one-way Kruskal-Wallis test followed by multiple-comparisons testing was used to compare fractional areas or compositions between all *in vitro*, 2D, and 3D ratios; ** indicates $p < 0.001$. Error bars indicate within-group standard deviation.

<https://doi.org/10.1371/journal.pcbi.1010701.g003>

(Fig 3A). For the 1:1 ratio with 200 cells, the model predicted a GFP/Ncad core surrounded by mCherry/Pcad poles and individual BFP+ cells. For the 1:1 ratio with 60 cells, the model predicted two neighboring poles, one of GFP/Ncad cells and one of mCherry/Pcad cells, with surrounding individual BFP+ cells. In spheroids with a minority of seeded BFP+ cells, fewer mCherry/Pcad cells were generated, leading to fewer and smaller mCherry poles. Conversely, in spheroids with a majority of seeded BFP+ cells, fewer GFP/Ncad cells were generated, leading to smaller GFP/Ncad clusters that were typically located farther from the spheroid center. Model predictions aligned well with observations from *in vitro* experiments we performed independently using the cell lines engineered by Toda et al. [2]. The average fractional blue, red, and green areas of the *in vitro* spheroids and the model predictions were statistically similar to each other for the 1:9, 1:1, and 9:1 ratios (Fig 3B–3D). Moreover, 2D and 3D model

predictions were statistically similar, suggesting that the 2D model sufficiently represented the synthetic cell signaling system.

Comparison of fractional areas of induced phenotypes among the 1:9, 1:1, and 9:1 ratios indicates that a small number of sender cells is required to induce cadherin phenotypes in the majority of receiver cells. For example, despite there being fewer BFP+ cells in the 1:9 ratio, the total GFP/Ncad fractional area was not significantly different from that of the 1:1 ratio (Fig 3C). Similarly, despite the total fractional area of GFP/Ncad cells being significantly lower in the 9:1 ratio than in the 1:1 ratio (Fig 3C), the total mCherry/Pcad fractional area was not significantly different from the 1:1 ratio (Fig 3D). These results suggest that sender BFP+ cells in the 1:9 ratio and sender GFP/Ncad cells in the 9:1 ratio disperse sufficiently over the model's time course such that sender BFP+ and sender GFP/Ncad cells experience enough interactions with receiver cells that the total fractional area of phenotypes induced by these sender cells is similar to that in the 1:1 ratio. This hypothesis was tested by quantifying the positions and movements of GFP/Ncad cells over time in the 2D ABM for 1:1, 4:1, and 9:1 BFP+ $-/-$ seeding ratios. For ratios with BFP+ cells in the majority (4:1 and 9:1), GFP/Ncad clusters were significantly farther from the spheroid center than for the 1:1 ratio and, thus, more dispersed throughout the spheroid (Fig 4A). Additionally, after the GFP/Ncad phenotype had been induced in the majority of $-/-$ cells (~13 hr or 26 timesteps in the model), the fraction of GFP/Ncad cells that moved per timestep in the 9:1 ratio was larger than that in the 4:1 and 1:1 ratios (Fig 4B). This result suggests that simulated cell clusters composed of fewer cadherin-expressing cells were more mobile due to fewer movement-restricting intercellular interactions. Increased mobility permits more interactions of GFP/Ncad cells with BFP+ cells, ultimately producing as many mCherry/Pcad cells for the 9:1 condition as for the 1:1 condition.

The model also predicted that the total number of seeded cells affects final spheroid patterns (Fig 5A). As the total number of cells in the spheroid increased, the number of green clusters formed also increased (Fig 5B). Simultaneously, the average fractional area of green clusters in the spheroid decreased, and the average fractional distance of green clusters to the center of the spheroid increased (Fig 5C and 5D). These trends suggest that as the overall cell number increases, the central GFP/Ncad core partitions into smaller pieces that become interspersed throughout the spheroid. This result is further explored in Fig 7.

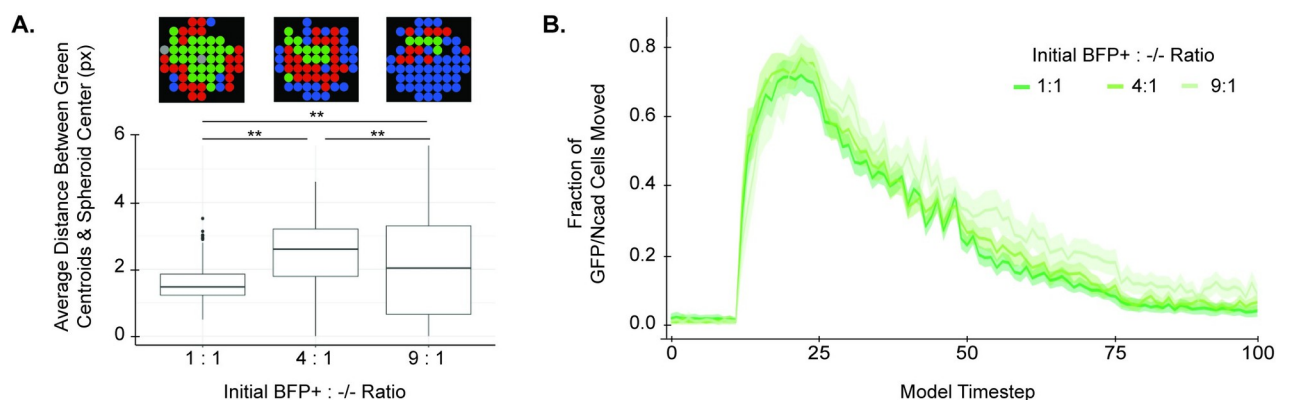


Fig 4. The relative fraction of seeded BFP+ sender cells impacts the size and mobility of GFP/Ncad clusters. (A) The average distance between green cluster centroids and spheroid centers was quantified for increasing ratios of initial BFP+ seeded cells in 200-cell spheroids. Data are shown as a boxplot distribution for 100 model runs, each run for 100 timesteps. The groups were compared using a one-way Kruskal-Wallis test followed by individual Dunn's tests for posthoc testing; ** indicates $p < 0.001$. (B) From timestep 26 onwards, the number of green cells that move per tick is generally higher in spheroids with a 9:1 ratio of initial BFP+ cells to $-/-$ cells. Metrics were calculated as the average from 100 model runs, each for 100 timesteps. Ribbons in the line plot represent 95% confidence intervals calculated using the standard error of the mean.

<https://doi.org/10.1371/journal.pcbi.1010701.g004>

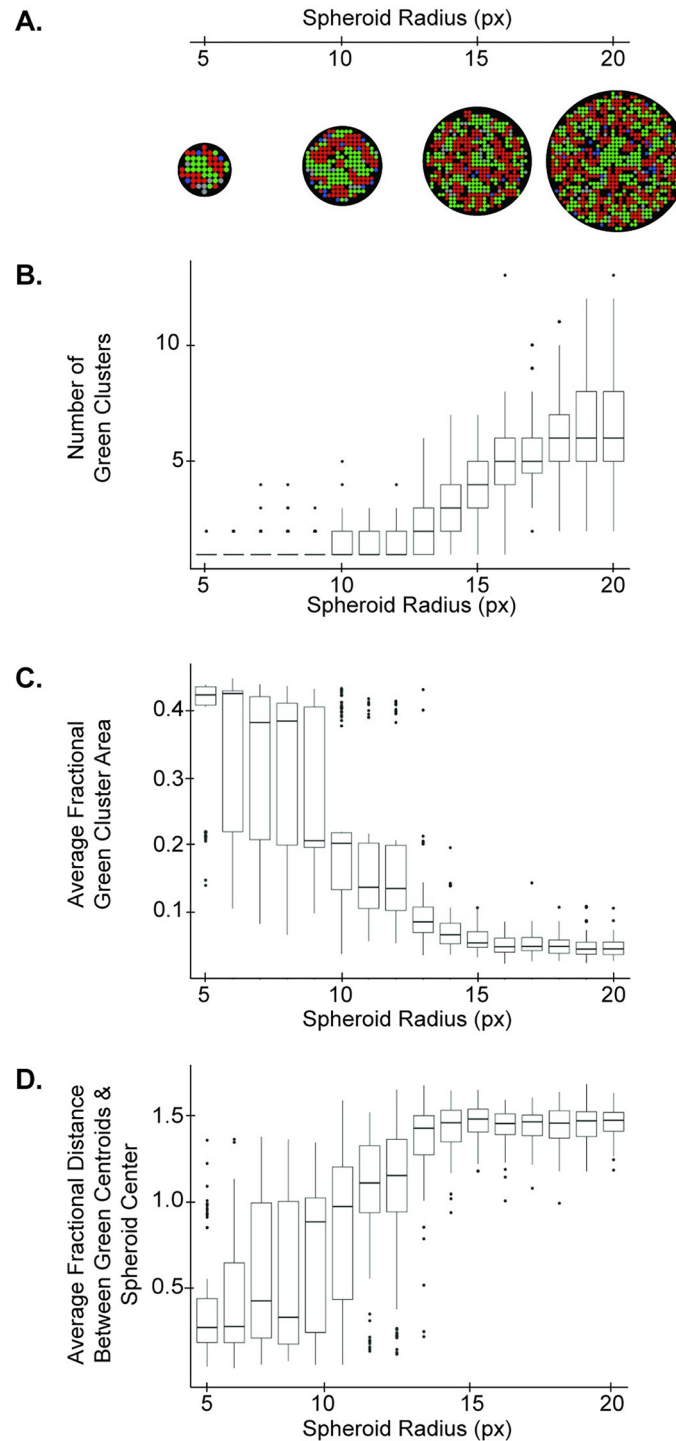


Fig 5. Increasing the radius of spheroids with a 1:1 BFP+/- seeding ratio leads to deviation from the core/pole configuration. The (A) number of green clusters, (B) average fractional area of green clusters, and (C) average fractional distance between green cluster centroids and spheroid center were quantified as a function of model spheroid radius (reported in units of pixels, px, where 1px is the approximate diameter of a cell, 10 μm). Data are shown as boxplots of values for 100 model runs, each run for 300 timesteps, to accommodate the aggregation of larger spheroids.

<https://doi.org/10.1371/journal.pcbi.1010701.g005>

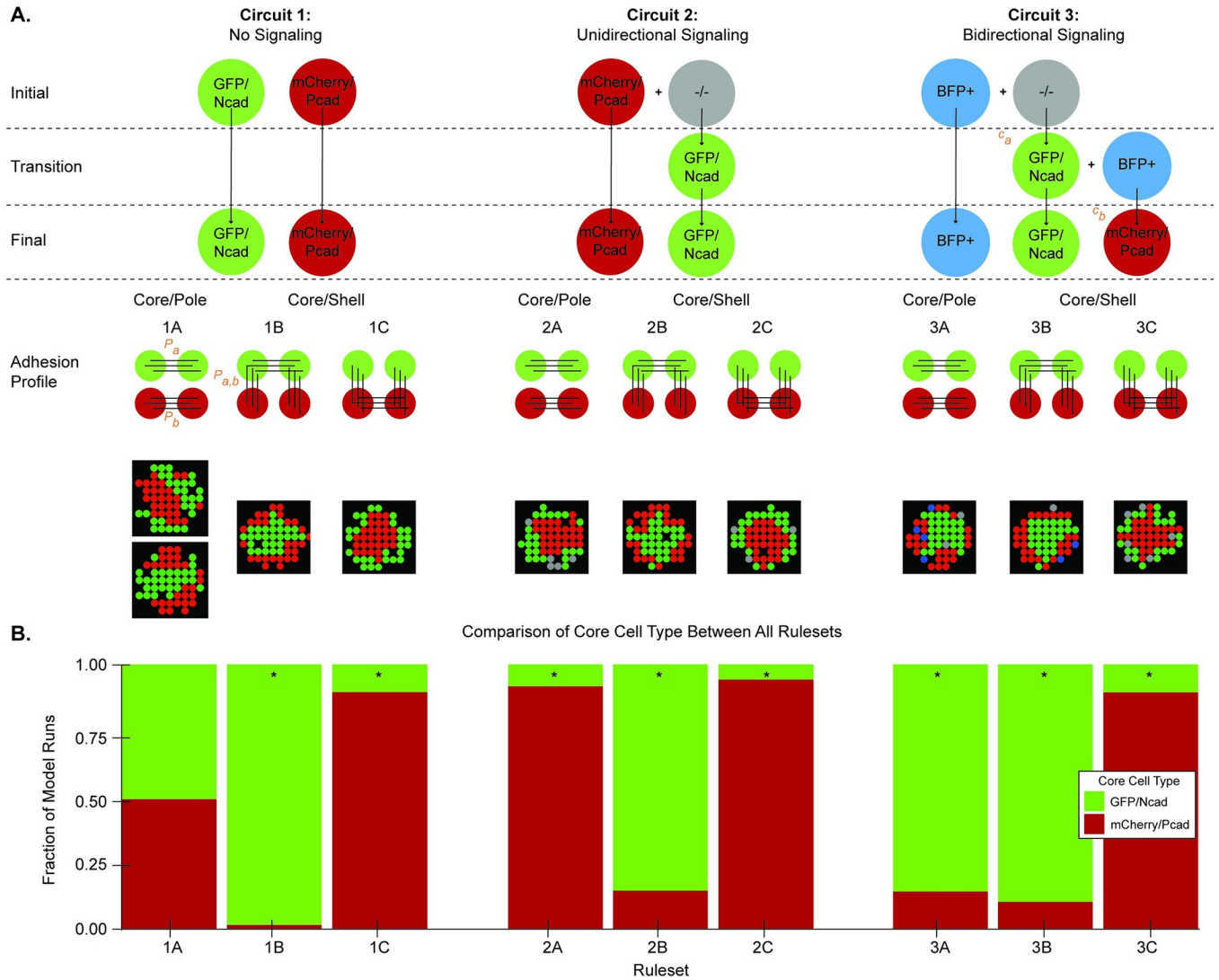


Fig 6. Dynamic cell-cell signaling has differential effects on core/pole and core/shell spheroids. (A) A schematic of three different ABM signaling circuits and three adhesion profiles is shown. Circuit 1 contains constitutive cadherin expression and no juxtacrine signaling. Circuit 2 includes unidirectional signaling. Circuit 3 includes bidirectional signaling (as described in Fig 1A). In Circuit 3, induction constants for GFP/Ncad and mCherry/Pcad phenotypes are parametrized separately (c_a and c_b , respectively). In each adhesion profile, the homotypic adhesion strengths of GFP/Ncad and mCherry/Pcad phenotypes are also parametrized separately (P_a and P_b , respectively). All depicted adhesion profiles have a homotypic adhesion strength of 1.0. Adhesion profile A gives rise to core/pole spheroids, whereas adhesion profiles B and C give rise to core/shell spheroids. The combination of circuit type and adhesion profile leads to nine rulesets (1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, and 3C). (B) The model was run 100 times, each for 100 timesteps, for each of the nine rulesets. Each ruleset was seeded with 200 cells at a 1:1 ratio of the seeding cell types. The fraction of trials with either GFP/Ncad or mCherry/Pcad spheroid core was compared between rulesets using a Chi-Squared test with $df = 8$; significant differences between observed and expected fraction (0.50) of core cell types were found across all nine rulesets with $p < 0.001$. The expected fraction was set to 0.5 because the null hypothesis states that signaling has no effect on which cell type forms the spheroid core. * indicates groups with residuals > 1 .

<https://doi.org/10.1371/journal.pcbi.1010701.g006>

Dynamic cell-cell signaling and homotypic adhesion strength impact pattern formation

To test the hypothesis that the dynamic expression of different cadherin isoforms impacts spheroid patterning, three different signaling circuits were explored (Fig 6A). In circuit 1, the cadherins were constitutively expressed, as opposed to expression via induction by juxtacrine signaling. In circuit 2, only unidirectional signaling was present such that seeded mCherry/

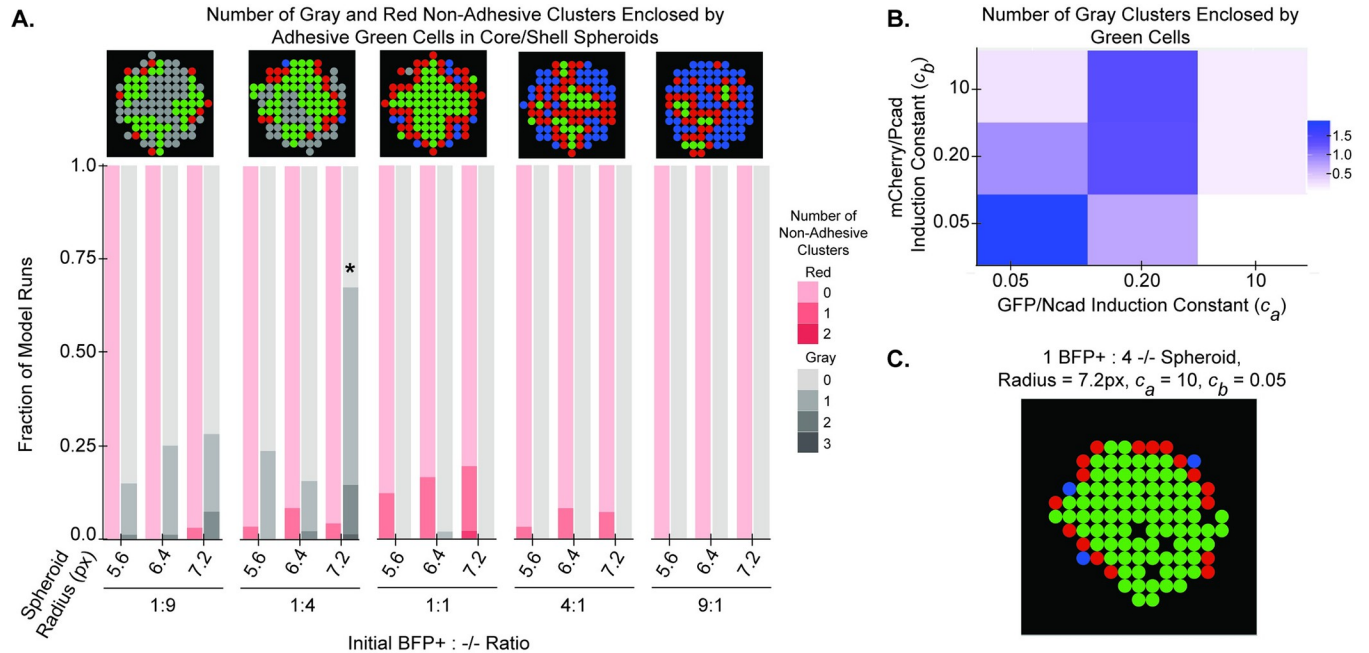


Fig 7. Increasing the relative fraction of seeded -/- cells in Ruleset 3B leads to instances where non-adhesive cell types are surrounded by adhesive cell types. (A) The number of gray and red non-adhesive cell clusters with > 75% enclosure by green adhesive cells was averaged for 100 model runs, each run for 200 timesteps, for spheroids with radii 5.6, 6.4, and 7.2 pixels (px) and varying initial BFP+/- ratios (1:9, 1:4, 1:1, 4:1, 9:1). All groups were compared with a Chi-Squared test with $df = 14$, resulting in a $p < 0.001$; * indicates groups with greater than 10% contribution to the chi statistic. (B) c_a and c_b were simultaneously varied over values of 0.05, 0.2, and 10 in spheroids with radius 7.2 px and initial BFP+/- ratio of 1:4. Total number of gray clusters with > 75% enclosure by green cells was quantified for 100 model runs, each run for 200 timesteps, for each model condition. (C) A representative output of a simulated spheroid with radius 7.2 px, initial BFP+/- seeding ratio of 1:4, and $c_a = 10$ and $c_b = 0.05$.

<https://doi.org/10.1371/journal.pcbi.1010701.g007>

Pcad sender cells induced the GFP/Ncad phenotype in seeded -/- receiver cells. In circuit 3, the full bidirectional signaling circuit was modeled, with seeded BFP+ sender cells inducing GFP/Ncad expression in seeded -/- receiver cells and newly formed GFP/Ncad cells inducing mCherry/Pcad expression in BFP+ cells. In this circuit, the induction constants for GFP/Ncad and mCherry/Pcad phenotypes were parametrized separately (c_a and c_b , respectively, instead of c_i). In the simulations discussed for circuit 3 in this section, $c_a = c_b = c_i$, but c_a and c_b are varied in later Results sections. For each circuit, three homotypic adhesion strength profiles were tested, in which the homotypic adhesion strengths of GFP/Ncad and mCherry/Pcad phenotypes were parametrized separately (P_a and P_b , respectively, instead of P_i). In profile A, GFP/Ncad and mCherry/Pcad cell types exhibited maximal homotypic adhesion and minimal heterotypic adhesion ($P_a = P_b = 1.0$, $P_{a,b} = 0.0$), giving rise to core/pole spheroids, in which one adhesive cell type assembled in a core and the other assembled in smaller peripheral clusters. In profile B, GFP/Ncad cells exhibited maximal homotypic adhesion, mCherry/Pcad cells exhibited minimal homotypic adhesion, and heterotypic adhesion was maximal ($P_a = 1.0$, $P_b = 0.0$, $P_{a,b} = 1.0$). In profile C, GFP/Ncad cells exhibited minimal homotypic adhesion, mCherry/Pcad cells exhibited maximal homotypic adhesion, and heterotypic adhesion was maximal ($P_a = 0.0$, $P_b = 1.0$, $P_{a,b} = 1.0$). Both profiles B and C gave rise to core/shell spheroids, in which the more adhesive cell type formed a core surrounded by a complete shell of the less adhesive cells (Fig 6A). Combinations of each signaling circuit and adhesion profile yielded 9 rulesets. For all rulesets, simulations were seeded with a 1:1 ratio of initial cell types for a fixed total of 200 cells.

The effects of uni- and bi-directional signaling between cells were explored by comparing the core cell types in spheroids for circuits 1–3 (Fig 6B). Adhesion profile A tended to produce

core/pole spheroids for which the core cell type varied based on the signaling circuit. Ruleset 1A gave rise to either GFP/Ncad or mCherry/Pcad cores, whereas in Rulesets 2A and 3A the adhesive phenotype that was first present in the circuit formed the spheroid core. Thus, the adhesive cell type that is present longer in the spheroid exhibited a higher likelihood of condensing into a central core. The same trends for core cell type were not observed for adhesion profiles B and C, which tended to produce core/shell patterns. In core/shell spheroids, the more adhesive phenotype formed the spheroid core in > 90% of model simulations regardless of the signaling circuit (Fig 6B). Thus, for profiles B and C, the signaling dynamics produced outcomes that were consistent with the differential adhesion hypothesis in 200-cell spheroids seeded with a 1:1 ratio.

Varying cell seeding in Ruleset 3B leads to patterns that violate the differential adhesion hypothesis

To test whether it is possible to generate spheroids that violate the differential adhesion hypothesis, the 2D ABM was used to predict patterns of cell subpopulations when the seeding ratios and total cell numbers in Ruleset 3B were varied. Ruleset 3B was chosen because it is based on the fundamental premise of the differential adhesion hypothesis, in which two phenotypes with differing homotypic adhesion strengths self-organize. As the relative proportion of seeded $-/-$ cells and spheroid radius increased, the likelihood that adhesive cells would enclose non-adhesive cells increased (Fig 7A). In the 1:1 ratio group, regardless of spheroid radius, a small number of simulated spheroids (~12–20%) contained one group of non-adhesive red cells enveloped by adhesive green cells, while virtually no spheroids contained non-adhesive $-/-$, or gray, cells enveloped by adhesive green cells. For the 1:4 and 1:9 ratios, the fraction of spheroids with groups of non-adhesive red cells enclosed by adhesive green cells decreased significantly, and the overall fraction of spheroids with groups of non-adhesive gray cells enclosed by adhesive green cells increased significantly, with the latter effect becoming more apparent as the spheroid radius increased. This trend is partially explained by the presence of fewer seeded BFP+ cells, which leads to fewer $-/-$ cells transitioning to the GFP/Ncad phenotype and mCherry/Pcad phenotypes. Due to the increased tendency of smaller green clusters to spontaneously arise at the periphery of the spheroid, non-adhesive cells surrounding each individual green cluster appeared to be trapped in the center of the spheroid, giving rise to situations where non-adhesive cells appeared enclosed by adhesive cells. This interpretation is reinforced by the fact that when the induction constant of the GFP/Ncad phenotype (c_a) was increased, a single, larger green core formed, and the spheroid patterns resembled those observed in the 1:1 group (Fig 7B and 7C). When the induction constant of the GFP/Ncad phenotype (c_a) was larger, fewer BFP+ sender cells were required to induce the GFP/Ncad phenotype in $-/-$ receiver cells, and a larger central GFP/Ncad core formed more quickly in the spheroid, instead of the slower formation of GFP/Ncad clusters and their subsequent enclosure of non-adhesive cells.

Homotypic adhesion strengths impact heterogeneity in spheroid patterns

Systematically covarying the homotypic adhesion strengths of GFP/Ncad and mCherry/Pcad cells (P_a and P_b , respectively) and setting $P_{a,b} = 0$ in the 2D ABM for signaling circuit 3 (bidirectional signaling) for spheroids seeded with ~350 cells produced a range of different spheroid patterns. We chose to increase the spheroid size for these models because distinct patterns emerged more clearly in larger spheroids. To display the results of these simulations, heatmaps were created to represent the numbers and average areas of green or red clusters in the spheroids (Fig 8A). Recall that the extremes of these maps describe core/shell ($P_a = 1.0$ and $P_b = 0.0$ or $P_a = 0.0$ and $P_b = 1.0$) and core/pole patterns ($P_a = P_b = 1.0$). With $P_a > P_b$, there were

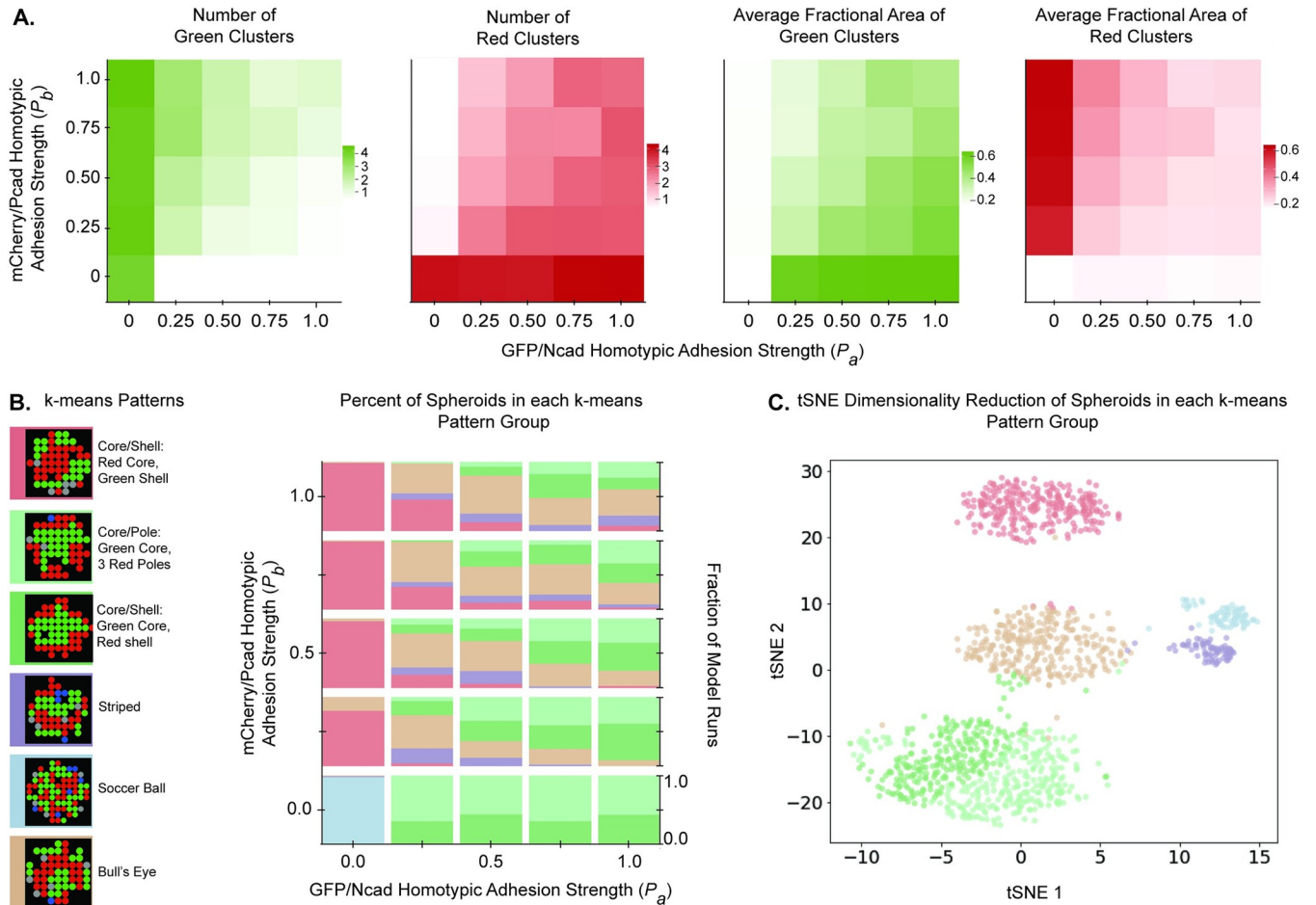


Fig 8. Covarying homotypic adhesion strengths of activated cell types in Ruleset 3A leads to a variety of different patterns. (A) Homotypic adhesion strengths of the first and second activated phenotypes (P_a , P_b) were simultaneously varied with values of 0.0, 0.25, 0.50, 0.75, and 1.0 for signaling circuit 3 (bidirectional signaling circuit) in the 2D ABM. For each parameter combination, the simulations were run for spheroid radius 6.4 px at a 1:1 BFP+/- ratio and were run 100 times, each for 300 timesteps. The average values of the indicated metrics were plotted in heat maps. (B) Results from this parameter space exploration were clustered using a *k*-means algorithm into six groups, with representative ABM outputs shown for each group and color coded as the key for this graph. (C) A t-distributed stochastic neighbor embedding (tSNE) algorithm was used to show where the *k*-means clusters projected in two dimensions. The tSNE algorithm was run with the perplexity parameter = 100.

<https://doi.org/10.1371/journal.pcbi.1010701.g008>

typically fewer and larger green clusters than red clusters, but the average number and area of green clusters were not as low as those observed in the core/pole pattern with a red core. Conversely, with $P_b > P_a$, there were fewer and larger red clusters than green clusters, but the average number and area of red clusters were not as low as those observed in the core/pole pattern with a green core. Deviations from core/pole and core/shell patterns could occur when P_a and P_b have non-extreme values.

To identify the different spheroid patterns formed, we performed *k*-means clustering, an unsupervised machine learning technique, to identify six groups within the data (Fig 8B). Each of the six groups corresponded to a distinct spheroid pattern: a core/pole pattern with a green core and three red poles, a core/shell pattern with green core and red shell, a core/shell pattern with red core and green shell, a striped pattern of alternating green and red linear clusters, a soccer ball pattern of circular green and red clusters, and a bull's eye pattern with radially alternating green and red shells. As an alternative way to visualize the *k*-means clustering results, the data were projected in two dimensions using t-distributed stochastic neighbor embedding

(tSNE), with pattern identities retained from k -means clustering (Fig 8C). These results also visually support the binning of spheroids into more than just two categories.

Each combination of homotypic adhesion strengths led to simulation outputs that exhibited varying frequencies of the six patterns identified by the clustering algorithm described above. When $P_a = P_b = 0.0$, a soccer ball pattern was generated more than 90% of the time. In alignment with the differential adhesion hypothesis, when $P_b > P_a$ and $P_a = 0.0$, patterns with a red core were generated almost 100% of the time. Conversely, when $P_a > P_b$ and $P_a = 0.0$, patterns with a green core were generated almost 100% of the time. The distribution of core/shell, core/pole, bull's eye, and striped patterns generated was relatively similar for all conditions where the homotypic adhesion strengths of both cell types are equal (e.g., $[P_a, P_b] = [0.25, 0.25]$, $[0.5, 0.5]$, $[0.75, 0.75]$, and $[1.0, 1.0]$). Interestingly, when $P_a = P_b = 1.0$, patterns with a green core and surrounding red cells formed, but a bull's eye pattern emerged with equal frequency. In larger spheroids, the possibility for there to be alternating clusters of cell types was higher.

When the GFP/Ncad phenotype had relatively weak homotypic adhesion (e.g., $[P_a, P_b] = [0.25, 0.25]$ or $[0.25, 0.75]$), striped and bull's eye patterns emerged more frequently than in conditions when the GFP/Ncad phenotype had stronger homotypic adhesion (e.g., $[P_a, P_b] = [0.75, 0.25]$ and $[0.75, 0.75]$). When the GFP/Ncad phenotype had relatively weak homotypic adhesion, the frequency of heterotypic interactions between cells was greater, allowing for a higher frequency of patterns characterized by longer borders between heterotypic cells.

In conditions where differential adhesion was present (e.g., $[P_a, P_b] = [0.25, 0.75]$ and $[0.75, 0.25]$), the most frequent patterns involved the more adhesive cells being surrounded by the less adhesive cells. With $P_a = 0.75$ and $P_b = 0.25$, the most frequent patterns were the core/pole and core/shell spheroids with a green core and bull's eye spheroids with a green core. With $P_b = 0.75$ and $P_a = 0.25$, the most frequent patterns were the core/pole spheroids with a red core and the bull's eye spheroids with a red core. This trend suggests that when differential adhesion is present, but the adhesive strength of the second induced phenotype is not maximal, hybrids of the core/shell pattern arise.

Cadherin induction constants impact core/pole and core/shell patterns

Varying the induction constants of GFP/Ncad and mCherry/Pcad cells (c_a and c_b , respectively) was also predicted to impact the final numbers of green and red clusters and their average areas (Fig 9A). For Ruleset 3A (bidirectional signaling core/pole ruleset) and when $c_a = 0.05$ and $c_b = 10$, one green cluster with a relatively small area was formed, and one red cluster with a larger area was formed. When c_b was only slightly larger than c_a , multiple small green clusters were formed, and one or two red clusters were formed, suggesting the possibility that different patterns may occur for different parameter combinations. Seven groups were identified by k -means clustering (Fig 9B). The first two patterns included a green core and three or four red poles. The third pattern comprised spheroids with a green core and red shell. The fourth pattern exhibited an inversion of the first two, with a red core and two or more green poles. The fifth pattern included one red core and one green pole. The sixth included a red core and single green and gray peripheral cells. Finally, the seventh pattern included striped or soccer ball patterns. As in Fig 9A, when $c_a \ll c_b$, a red core formed with one or multiple green poles. However, when c_a was sufficiently high, core/pole structures with a green core and multiple red poles formed more frequently. To produce an alternative visualization, tSNE was used to reduce the dimensionality of the system (Fig 9C). tSNE results further supported the existence of multiple subtypes of spheroids. Modifying the cadherin induction constants affects the overall spheroid pattern because various combinations of c_a and c_b directly impact the final ratio of BFP+, -/-, GFP/Ncad, and mCherry/Pcad cells in the spheroid (S1A Fig).

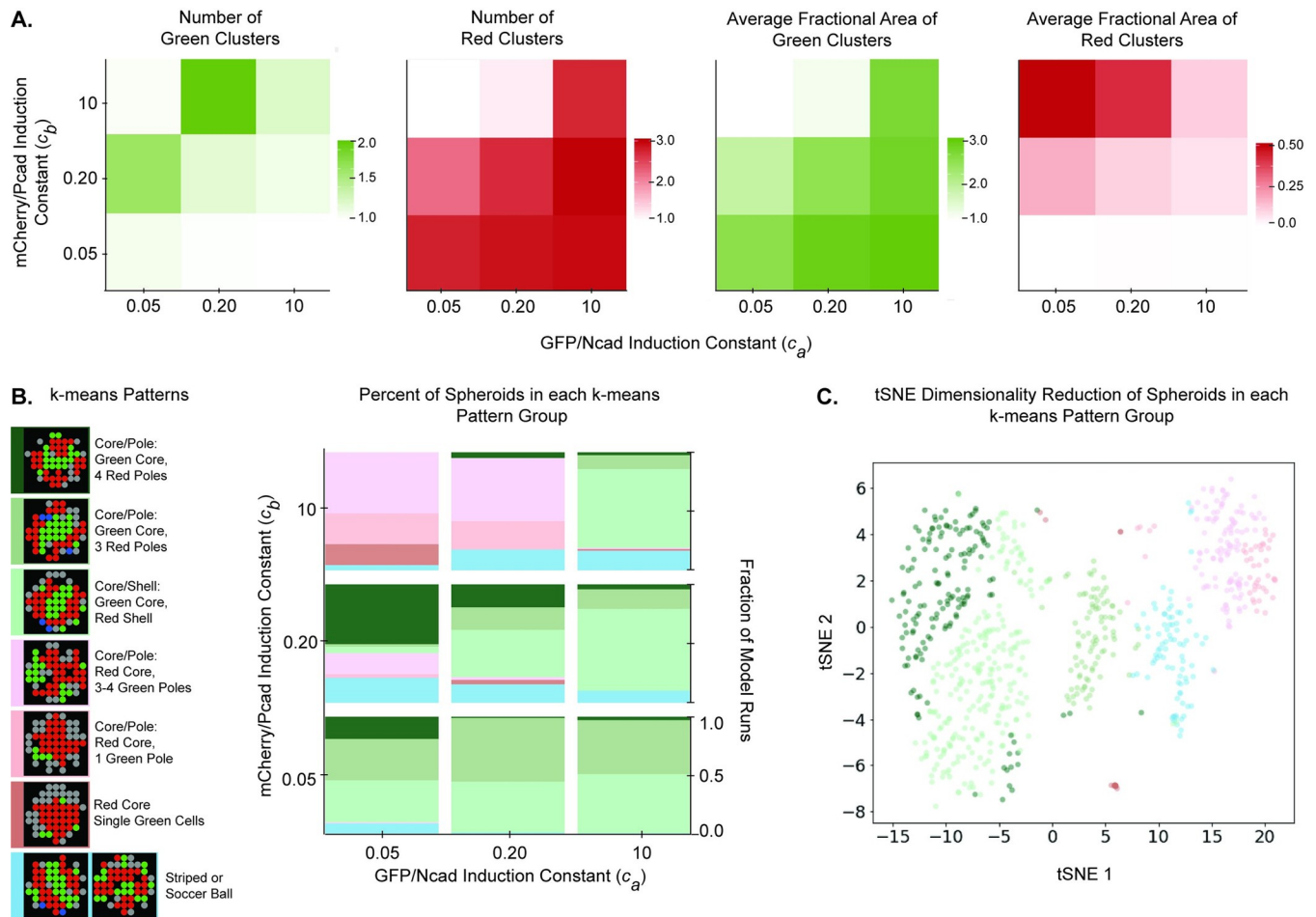


Fig 9. Cadherin induction constants (c_i) control the number of poles in core/pole spheroids. (A) The cadherin induction constants (c_a, c_b) were simultaneously varied across values of 0.05, 0.2, and 10 for Ruleset 3A (bidirectional signaling core/pole ruleset) in the 2D ABM. For each parameter combination, simulations were seeded with 200 cells at a 1:1 BFP+/- ratio and run 100 times, each for 100 timesteps. The average from 100 model runs of the following metrics were plotted for each parameter combination: number of green clusters, number of red clusters, average fractional area of green clusters, and average area of red clusters. (B) Results from this parameter space exploration were clustered using a k -means algorithm into seven groups, with representative ABM outputs shown for each group and color coded as the key for this graph. (C) t-distributed stochastic neighbor embedding (tSNE) algorithm was used to project the data into two dimensions with colors for spheroid cluster types retained from the k -means clustering analysis in panel (B). The tSNE algorithm was run with perplexity = 100.

<https://doi.org/10.1371/journal.pcbi.1010701.g009>

The parameters c_a and c_b were also varied for Ruleset 3B (bidirectional signaling core/shell ruleset). The ratio of the shell's thickness to the core's radius progressively increased as c_b became larger than c_a (S1B and S1C Fig), which is consistent with more mCherry/Pcad cells at the final time point. Because this parameter describes sensitivity of a cell to its neighbors, and the number of neighbors in 3D is larger than in 2D, results from these 2D model calculations may not translate directly to a 3D model.

Adjusting ABM parameters enables customization of multicell patterns in spheroids

To demonstrate the utility of the model for developing design principles for spheroid engineering, we used the 2D ABM to design multicellular spheroids with desired patterns. To control the overall spheroid configuration, the homotypic adhesion strengths of induced cell types should be controlled (Fig 10A). The ABM predicted that, to produce a core/pole pattern, both adhesion molecules should have homotypic adhesion strengths > 0.75 . In order to produce a

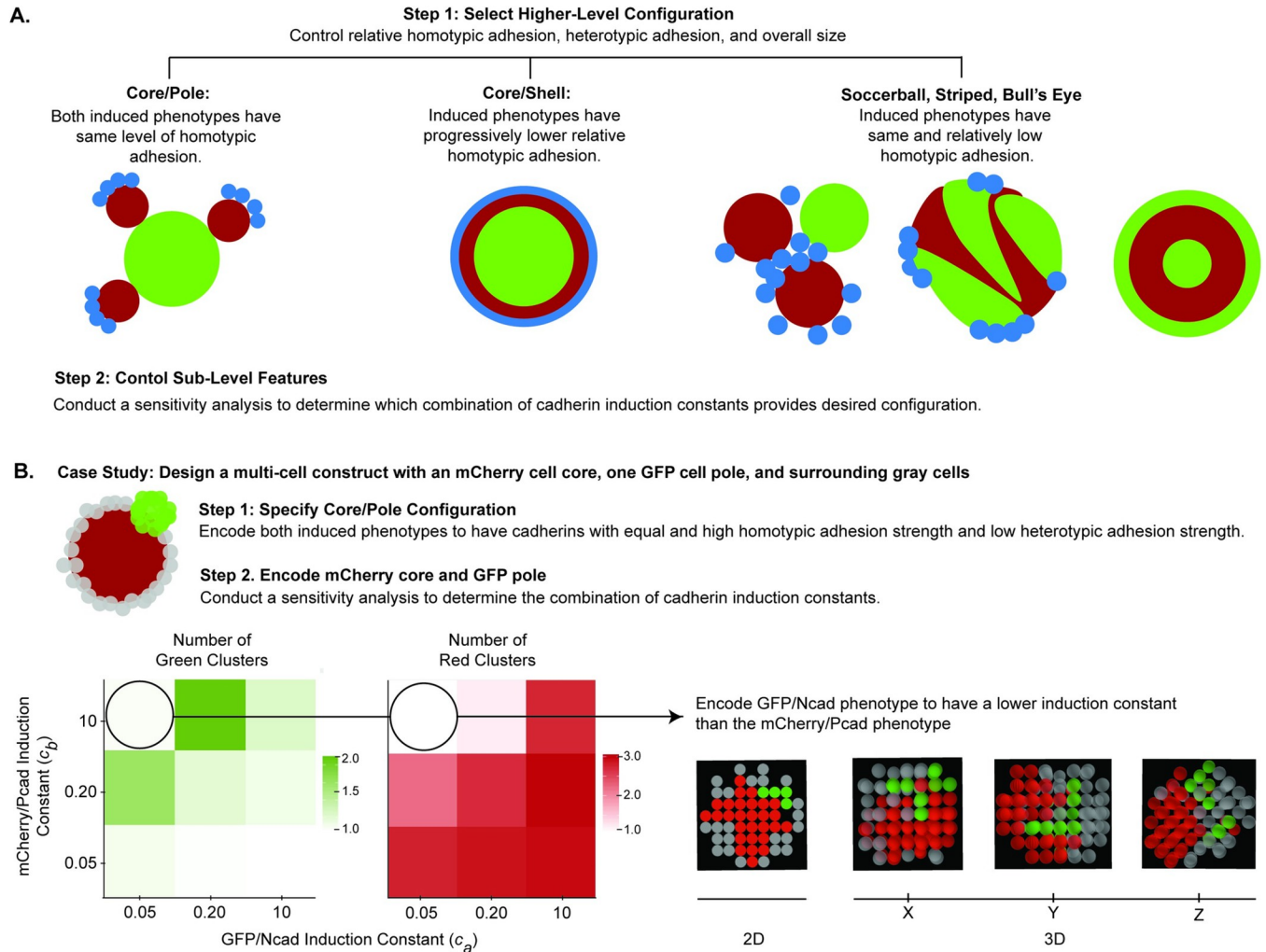


Fig 10. ABMs may be used to design synthetic signaling circuits encoding customizable core/pole, core/shell, soccer ball, striped, and bull's eye spheroid patterns. (A) For a desired spheroid pattern (e.g., core/pole, core/shell, soccer ball, striped, or bull's eye), homotypic and heterotypic adhesion strengths of cell types can be modulated. Features of the core/pole and core/shell configurations, such as the number of poles and ratio of shell thickness to core radius, may be further controlled by modifying the induction constants of cadherin phenotypes in the system. (B) An outline of the process to create a structure with mCherry/Pcad core, GFP/Ncad pole, and surrounding single cells is shown.

<https://doi.org/10.1371/journal.pcbi.1010701.g010>

core/shell pattern, one induced phenotype must have homotypic adhesion > 0.75 , while the other induced phenotype must have a homotypic adhesion strength < 0.25 . To produce soccer ball, striped, or bull's eye patterns, both activated phenotypes must have homotypic adhesion strengths in the range 0.25–0.75.

The schematics in Fig 10B demonstrate how bivariate parameter explorations of the 2D ABM can guide the forward design of a user-specified spheroid with an mCherry/Pcad core, one GFP/Ncad pole, and surrounding $-/-$ cells. Based on the results of the model, the user should set P_a and P_b to be maximal and equal to achieve a core/pole configuration and engineer the system such that $c_a \ll c_b$ to achieve an mCherry core and one GFP/Ncad pole.

Discussion

We created an ABM of bidirectional cell-cell interactions that drive patterning of multicellular spheroids inspired by the signaling circuits engineered by Toda et al. [2]. The ABM provides

an alternative to time-consuming cell culture experimentation by predicting patterns that can arise from a multicell signaling system as a result of adjusting initial conditions, signaling interactions, and adhesion strengths alone or in combination. We used the model to explore the impact of dynamic cell-cell signaling on adhesion-driven pattern formation, and we observed the emergence of new patterns that were inconsistent with the classical differential adhesion hypothesis. Moreover, we used machine learning to map combinations of homotypic adhesion strengths to unique spheroid patterns. Additionally, we found that modulating the induction constants for the GFP/Ncad and mCherry/Pcad phenotypes controls the number of poles in core/pole patterns and the ratio of shell thickness to core radius in core/shell patterns. Our study supports the use of computational models to identify parameters that will yield customized spheroid patterns. Because these parameters have physical meaning, they may be used as design constraints for multicell, synthetic circuits. For example, homotypic and heterotypic adhesion strengths may be controlled by selecting specific cadherin isoforms [6]. Similarly, promoter regions for synthetic transgenes may be selected according to the intended phenotype induction constants [55].

Previous studies have leveraged computational models to explore cell patterning in heterogeneous systems [8–10,43–45,47–49,56,57]. For example, a prior Cellular Potts computational model of the differential adhesion hypothesis [8,9] demonstrated how cell subpopulations with varying levels of cohesiveness self-assemble and how environmental variables such as temperature, shear force, and pressure affect these self-assembly processes. These types of models have been expanded to develop more complex systems such as CompuCell3D, a multi-scale modeling framework that recapitulates morphogenesis in a wide array of tissues and species and incorporates cell differentiation in response to morphogens [10]. Taylor et al. [11] developed an ABM to explore boundary formation and border sharpening between cell subtypes in the Eph-ephrin signaling system, which is critical in embryogenesis. Their model predicts a range of configurations that arise between Eph- and ephrin-expressing cells, and as a function of varying the homotypic and heterotypic adhesion strengths between cell types. Although these published models have thoroughly explored pattern formation in heterogeneous systems with static cell phenotypes, they have not considered how patterns may be affected by the modulation of differential adhesion via known, programmable cell-cell signaling interactions. Our model differs from prior efforts because it simulates phenotypic transitions controlled by a synthetic gene circuit, which allowed us to explore how the phenotype induction probabilities, timing, and sequencing of dynamic cell-cell differential adhesion drives pattern formation.

Our study also contributes to the literature on robustness of pattern formation in multicell systems. Here, we define robustness as lack of substantial change in spheroid patterns for specified changes in model parameters, or for no change in parameters at all, consistent with prior definitions [58]. Our model identified regions in parameter space where patterns (or distributions of patterns) were predicted to be robustly formed and other regions where spheroid patterns were sensitive to changes in parameters. Prior work has modeled multicell pattern emergence using Turing models [58–60], which make predictions based on considerations of sharp spatial gradients of chemical morphogens that arise due to competition between reaction and diffusion [61]. Turing systems are sensitive to small changes in parameters, and prior studies have noted that integrating biological processes such as cell-cell signaling and differential adhesion may increase robustness of model predictions [58–60]. It is possible that incorporating the morphogen reaction-diffusion processes considered by Turing models in our ABM, and into the design of synthetic circuits for cell culture experiments, could increase robustness of pattern formation.

This work also contributes to recent efforts to elucidate design principles for multicell patterns using a combination of computational and synthetic approaches [62–64]. Depending on

the desired pattern (e.g., core/shell, core/pole), different combinations of adhesion molecules or methods to induce their expression may be selected. For example, quantitative studies of homotypic and heterotypic adhesion strengths of cadherins, such as those by Duguay et al. [6] and Foty & Steinberg [7], may guide the selection of adhesion molecules for synthetic circuits. To control details such as the number of poles or ratio of shell-to-core radius, induction constants for cadherin expression may be adjusted. Similar parameters can be found in other types of models. For example, Matsuda et al. [55] designed a synthetic experimental system to study signal propagation in heterogeneous cell populations. Their ODE-based model determined conditions that optimize signal propagation and included a “promoter coefficient” similar to our c_i parameter. The synthetic feature related to this model parameter accounts for the copy number of the promoter and likelihood of epigenetic modification based on insertion region.

Aspects of the basic modeling approach utilized here will also be useful for the study of endogenous cell signaling circuits that regulate cell-cell adhesion. An application of relevance in cancer biology is epithelial-mesenchymal transition (EMT), a cell process that aberrantly occurs in numerous carcinomas (e.g. breast cancer, pancreatic ductal adenocarcinoma), occurs early in metastasis [65], and promotes resistance to chemotherapy and targeted therapeutics [23,24,66]. During EMT, epithelial-derived cancer cells lose polarity and expression of cell adhesion molecules to become more mesenchymal, motile, and invasive. This process is highly regulated by signaling and transcriptional networks that can be initiated by growth factors, cell-matrix interactions, and tumor microenvironmental factors (e.g., low oxygen tension). ABMs capable of incorporating signaling dynamics that lead to EMT and the resultant effects on differential cell-cell adhesion could be used to understand the morphogenesis of primary tumors and the processes that lead tumors of different compositions to shed metastases [28,54]. Thus, elements of our modeling framework could help expand upon prior computational models of how differential adhesion impacts tumor metastasis and cell migration [67].

There are limitations of both our computational model and the *in vitro* system. The computational model does not explicitly simulate the effects of cell proliferation nor the fact that induced cadherin expression could decrease after loss of contact between sender and receiver cells. The ABM also assumes that cadherin expression is uniform among cells, but realistic heterogeneity of expression within a population of even clonal cells could impact pattern formation. These issues could be addressed in future model implementations. For example, distributions of cadherin expression among cells could be quantified by flow cytometry or single-cell RNA sequencing, and this data could inform individual simulated cell behaviors in the ABM. Experimental limitations, including an inability to count the individual cells within spheroids, also limited the degree to which we could quantitatively compare model results with experimental data. Furthermore, neither our model nor the experimental spheroids contain diffusible morphogen cues, which are instructive in the patterning of natural (non-synthetic) multicell biological systems [41]. Future extensions of the ABM could incorporate diffusible morphogen cues, such as those needed for inducing cell differentiation [44,45]. An additional model limitation is that signaling dynamics were encoded as simple time delays for cadherin expression after cell contacts were made. In the future, it will be worthwhile to incorporate more realistic models of intracellular signaling, as has been done in recently published multiscale ABMs that integrate information across intracellular and intercellular scales [68–71]. The attendant need to solve differential equations describing signaling reaction kinetics in each cell within the ABM would greatly increase the number of model parameters and the computational resources needed to simulate the system. These are common issues encountered in multiscale, mechanistic biological models [42,67,72], and tradeoffs must be carefully weighed in designing such models. Encouragingly, recent studies have presented novel

methods for reducing the computational burden of complex mechanistic models by using neural networks and hybrid continuum-based modeling approaches [73–76].

An emerging trend in multiscale computational modeling of biological systems is ABM integration with machine learning and other data science approaches to exploit high throughput experimental and simulated data. We demonstrated the utility of interpreting ABM results with machine learning through our use of *k*-means clustering, which mapped cell-specific ABM parameters (homotypic adhesion strengths and cadherin induction constants) to tissue-level spheroid patterns. Previous studies have similarly used ABMs with machine learning methods to study and predict multicellular interactions [46,77]. Future iterations of our ABM could combine data-driven modeling and machine learning to further inform rule development [73,74], model calibration [78], and parameter exploration [77] in order to create an even more comprehensive and predictive model.

Materials and methods

Cell lines and propagation

L929 murine adipocytes engineered with a bidirectional cell-cell signaling circuit based on the synthetic notch (synNotch) receptor system [2] were generously provided by Dr. Wendell Lim (University of California, San Francisco). Cells were maintained in DMEM (ThermoFisher) supplemented with 10% fetal bovine serum (VWR), 1 mM L-glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin. Cell lines were used within 25 passages and maintained in a 37°C, 5% CO₂ ThermoForma i160 incubator.

Spheroid culture

A total of 60 or 200 cells (BFP+/- ratios = 1:9, 1:4, 1:1, 4:1, 9:1) were seeded per well in Ultra Low Attachment 96-well microplates (Corning #7007, Corning, N.Y). Cells aggregated into spheroids at the bottom of the well. 12 spheroids were plated per condition. For 200-cell spheroids, 50 µL of 4 cells/µL cell suspension and 50 µL of DMEM were loaded in each well. For 60-cell spheroids, 50 µL of 1.2 cells/µL cell suspension and 50 µL of DMEM were loaded in each well.

Spheroids were imaged 50 hr after plating using a Zeiss AxioObserver Z1 widefield microscope at 10× magnification. Six to seven representative spheroids were selected for imaging for each condition, by selecting wells without residual plastic particles (from plate manufacture) and wells in which all three cell colors were easily detected due to spheroid orientation. For each spheroid, a phase contrast image and three fluorescence images were taken: BFP (DAPI filter set), GFP (AF488 filter set), and mCherry (AF546 filter set). Zen Zeiss software was used to merge phase, BFP, GFP, and mCherry channels.

ABM design and implementation

2D and 3D on-lattice ABMs were developed to simulate collections of cells containing the synthetic signaling circuit represented in Fig 1. An ABM platform was chosen because this modeling technique allows for prediction of spheroid-level patterns based on individual cell-cell interactions. The 2D and 3D models were developed using NetLogo software [79].

The ABMs treat cells as individual agents in a spheroid. The 2D ABM simulates cells interacting in a plane through the central axis of a spheroid, whereas the 3D ABM simulates all cells in a spheroid. Each grid space in the on-lattice model represents a 10 µm × 10 µm area capable of accommodating one cell. The same cell size and one-cell-per-site assumptions are necessary limitations of on-lattice models. Each model timestep represents 30 min, such that 100

timesteps represent the 50-hr period over which spheroids were studied *in vitro*. To determine the impact of different timesteps, we tested a range including 15, 30, and 60 min, and we updated the total runtime and time-dependent parameters (a , b , Δa , Δb , t) accordingly (but without refitting the model). All timestep values led to fractional red and green areas being within the standard deviation of the *in vitro* measurements (S2A–S2C Fig). The deviation for the blue area for the 15 min timestep seen in S2A Fig is primarily due to the constraint that fractional areas must sum to one and the bump in red area seen for 15 min, which is within the error of the *in vitro* measurements. Based on those observations, and noting that the 30-min timestep yielded spheroids that were most qualitatively similar to *in vitro* results (S2D Fig), we concluded that our results were not extremely sensitive to the timestep value and selected 30 min. Unless otherwise noted, simulations were run for 100 timesteps because this value corresponds to the physical time over which spheroids were observed *in vitro*. Additionally, simulations compared with *in vitro* data were seeded with the twice the number of cell agents as the number of cells seeded experimentally. This was done to account for the effects of cell proliferation over this time period. Specifically, based on published data in Toda et al. [2] and our own observations of their cells, we assumed that cells doubled once in 50 hr. Seeded model agents represent BFP+ sender cells and -/- receiver cells, which can become mCherry/Pcad and GFP/Ncad cells, respectively. State transitions were encoded by setting Boolean values of private variables. Intercellular interactions were encoded as six rules, characterized by ten parameters, that execute on each timestep (Fig 1C, Table 2).

Fig 1C describes the rules of the ABM. At the initial timestep, the “setup” method randomly places BFP+ and -/- cells within a user-specified radius and according to a user-specified ratio and percent saturation (percent of cell-occupied grid spaces). On every subsequent timestep of the model, cells move to the open grid space (of eight neighboring spaces in 2D and 26 neighboring spaces in 3D) closest to the center of the simulation space to represent the gravitational and adhesive forces that tend to pull them together at the bottom of a low-attachment well. If there are no open spaces available on any of the neighboring eight grid spaces surrounding a cell, that cell will not move.

Clusters of cells containing fewer than a threshold number of cells (μ) identify the cell closest to the spheroid center as a cell that moves to a randomly selected space on one of its neighboring eight grid spaces (or 26 grid spaces in 3D) that is unoccupied. If there are no open spaces surrounding the leader cell, the cluster will not move. Like cells in a contiguous cluster follow the leader as a connected unit, displacing cells in previously occupied grid spaces as needed to avoid two cells occupying the same grid-space simultaneously. Displaced cells move randomly to the positions vacated by the cluster that moved in response to a leader cell move. As the model is on-lattice, movement is discrete (i.e., between grid spaces), and each grid space is occupied by only one cell at a time.

After a time delay (a), -/- receiver cells can convert to GFP/Ncad cells depending on the number of neighboring BFP+ sender cells. Similarly, after another time delay (b), BFP+ sender cells can express mCherry/Pcad depending on the number of neighboring sender GFP/Ncad cells. Neighboring cells are defined as cells occupying the eight surrounding grid spaces in 2D or the 26 surrounding grid spaces in 3D. The probability of phenotypic induction ($P_{induction}$) for receiver cells in contact with sender cells was determined as

$$P_{induction,2D} = 1 - e^{-c_i \left(\frac{n}{8}\right)} \text{ or } P_{induction,3D} = 1 - e^{-c_i \left(\frac{n}{26}\right)}$$

where c_i is a tunable parameter that controls the probability of cadherin expression induction, and n is the number of neighbors with the activating phenotype. If the cadherin phenotype is

induced, cells change colors and express Ncad or Pcad after specified times (Δa and Δb , respectively). In this baseline ruleset, both cadherin phenotypes have the same induction constant, c_i .

Cadherin-expressing cells attach to other cadherin-expressing cells based on homotypic and heterotypic adhesion probabilities (P_i and $P_{a,b}$, respectively). These parameters dictate the probability that a cadherin-expressing cell will form a link with a neighboring like or unlike cadherin-expressing cell within a single simulation timestep, or tick. Additionally, when a cell expressing one type of cadherin is wedged between two cells expressing another type of cadherin, the cells switch places based on the homotypic adhesion probability such that the cells with the two like cadherins occupy adjacent grid spaces. Similarly, if a cell expressing one type of cadherin is separated from a cell expressing a different cadherin by a non-cadherin expressing cell, cells switch places based on the probability $P_{a,b}$ such that cells expressing the two different cadherins occupy adjacent spaces. This type of position switching behavior has been experimentally observed [4]. In the baseline ruleset, both cadherin phenotypes have the same homotypic adhesion strength, P_i . Additionally, to provide the opportunity for cells to leave one cluster and join other clusters, 50% of the cadherin interactions across a grouping of connected cells were randomly broken at each timestep, and cadherin interactions between groups of cells could only be reformed after a specific number of ticks specified by the parameter t .

***In vitro* and *in silico* quantification of multicell patterns**

A variety of characteristics that describe the spatial patterns of cells in the *in vitro* spheroids and *in silico* spheroids were quantified (Fig 1B). A MATLAB pipeline was developed to quantify the following characteristics for each color/channel from images of *in vitro* spheroids: fractional area, total area, and distance between cluster centroid and spheroid center. In order to do this, thresholds were applied to images to isolate purely green, red, or blue pixels. Then, pixel areas and centroids of contiguous cluster for each color were computed based on the MATLAB method 'RegionProps,' which computes these metrics using a connected components algorithm.

The same features were quantified in the 2D ABM simulations. Similarly, the MATLAB connected components 'RegionProps' algorithm was used to quantify the total fractional area, average area of homotypic cell clusters, and average distance between the centroid of homotypic clusters and the spheroid center for each color channel. A cluster was defined as a group of adjacent cells containing more than two cells. Additionally, the number of non-clustered cells in each color channel and total cell count of each color channel was quantified. Finally, a contiguous area ratio, which represents the fraction of the border of a spheroid core that is surrounded by cells expressing the non-core forming cadherin, was also computed. In the 3D ABM, the total fractional volume of each color channel was quantified. This metric was computed as the cell count of a certain color channel divided by the total number of cells in the spheroid.

Parameterization of the ABM using experimental data

Table 2 summarizes the ABM parameter values and how they were determined. Parameters a , b , Δa , and Δb were estimated according to data published in Toda et al. [2], which indicated that for 200-cell spheroids seeded with an initial 1:1 BFP+/-/- ratio, GFP/Ncad cells appeared after 13 hr and mCherry/Pcad cells appeared adjacent to GFP/Ncad cells after 21 hr [2]. Values for a , Δa , b , and Δb were determined such that the model matched these results at the corresponding time points of 26 ticks and 42 ticks, since one model tick (or timestep) represents 30 min. The remaining five model parameters (c_i , P_i , $P_{a,b}$, μ , and t) could not be determined from experimental data. As a first approximation, these parameters were manually adjusted to

baseline values so that the 2D ABM yielded spheroids that were qualitatively similar to images of the Ncad/Pcad signaling circuit published in Toda et al. [2]. Specifically, values for μ and t were manually set to encourage continual movement of homotypic clusters in the spheroid over the 100-tick period and prevent the model from reaching a static state composed of smaller homotypic clusters at early timepoints.

Univariate parameter sensitivity analysis

A univariate sensitivity analysis was performed to identify which of the five parameters that could not be experimentally determined (c_i , P_i , $P_{a,b}$, μ , and t) most control model outputs. Each model parameter was perturbed, one at a time, by a factor of 0.9 or 1.1, and a sensitivity coefficient (S) was calculated as

$$S = \frac{y_b - y_p}{|x_b - x_p|} * \frac{x_p}{y_p}$$

where x_b and x_p represent baseline and perturbed parameter values, respectively, and y_b and y_p represent the baseline and perturbed values of the model output, respectively. Sensitivity coefficients were computed for the following model outputs: number of green clusters, average fractional area of green clusters, number of red clusters, and average fractional area of red clusters. The mean sensitivity for each model output was computed from 100 model runs and reported for the 0.9 and 1.1 perturbation of each of the 10 parameters.

Computational tuning with an evolutionary algorithm

An evolutionary algorithm (EA) was used to tune parameters that could not be determined from published data. Due to the high complexity and computational expense of parameter tuning, only the top two sensitive parameters from the univariate sensitivity analysis (c_i and P_i) were selected for tuning. The EA finds a parameter combination that minimizes an error function through a natural selection process wherein optimal parameter combinations in parent generations are selected for cross-over to create offspring parameter combinations in subsequent generations. The error, E , that the EA minimized was calculated as

$$E = (e^{\sum_j (|y_j - Y_j|)} - 1)^2,$$

where y_j represents the value of a calculated model output, and Y_j represents the experimental value for that output. The following model outputs (j) were used in the error equation: total fractional green area, total fractional red area, total fractional blue area, average fractional green cluster area, average fractional red cluster area, number of green clusters, and number of red regions. The exponential form of the error equation was used to encourage EA convergence. The EA was run for a population size of 20 (i.e., 20 randomly generated sets of initial guesses for c_i and P_i), with 10 parallel ABM runs to compute an average error and 100 generations for each set of initial guesses. At each generation, a cross-over rate of 50% and mutation rate of 80% was used. This process (20 sets of initial guesses \times 10 ABM runs \times 100 generations) was repeated 30 times, and the combination of c_i and P_i producing the overall lowest error was chosen to model the Ncad/Pcad signaling circuit developed by Toda et al. [2]. This approach was used to account for stochasticity of the ABM and EA and to reduce the likelihood of the EA becoming stuck in local error minima. We developed our own Python script for the EA and used the NL4Py Python library [80] to enable communication between Python and NetLogo. To improve run time, our EA implementation ran the ABM in parallel across multiple computing units.

Comparison of ABM predictions with experimental data

The 2D and 3D ABMs were validated against experimental data by simulating different BFP+/- seeding ratios (1:9, 1:4, 1:1, 4:1, 9:1) for 60 and 200 total cells, replicating the conditions used for *in vitro* experiments. Simulations for each set of conditions were run 100 times, and the average results from these runs were analyzed. The locations of each simulated cell were saved at the end of each model run, and metrics were quantified as described above. Fractional areas/volumes of each color were compared between *in vitro* results and ABM predictions.

Rule set implementations

Different rulesets were explored using the 2D ABM by varying the order of signaling events, the activated homotypic and heterotypic adhesion strengths, and the cadherin induction constants. In these rulesets (described in Figs 6–10), the GFP/Ncad phenotype and mCherry/Pcad phenotype can take on different homotypic adhesion strengths and induction constants (e.g., $[c_a, P_a]$ were assigned to GFP/Ncad cells, and $[c_b, P_b]$ were assigned to mCherry/Pcad cells). Each ruleset and its associated observations are explained in detail in *Results*.

k-means clustering and tSNE dimensionality reduction

k-means clustering of the parameter space exploration results was conducted using the scikit-learn library in Python [81]. The optimal number of clusters for each clustering problem was determined by plotting the within-cluster standard deviation for *k*-means clustering using two to fifteen clusters. This analysis was used to determine the minimal number of clusters that reduced within-group standard deviation. tSNE was implemented using the same Python library.

Statistics

Statistical comparisons for continuous data were made by a one-way Kruskal-Wallis test followed by Dunn's post-hoc analysis. This test was specifically applied when comparing fractional areas for *in vitro* spheroids and in predicted ABM spheroids (Fig 2). Statistical comparisons for qualitative metrics (e.g., core cell type) and discrete metrics (e.g., number of clusters) were made using a Chi-Squared test with $p < 0.05$ considered significant for all tests.

Supporting information

S1 Fig. Cadherin induction constants (c_i) control the ratios of cell types formed in the system and the shell thickness to core radius in core/shell spheroids. (A) The average fractions of BFP+, -/-, GFP/Ncad, and mCherry/Pcad cells in spheroids were calculated for each parameter combination. (B) The cadherin induction constants (c_a, c_b) were simultaneously varied across values of 0.05, 0.3, 1, 5, and 10 for Ruleset 3B (bidirectional signaling core/shell ruleset) in the 2D ABM. For each parameter combination, simulations were seeded with 200 cells at a 1:1 BFP+/- ratio and run 100 times, each for 100 timesteps. (C) Representative images of spheroids from extremes of 9E are shown. (TIF)

S2 Fig. Model results are not highly sensitive to time step. Comparisons were made of (A) blue, (B) green, and (C) red fractional areas between *in vitro* (i.v.) experiments and 2D ABM predictions, where the ABM was run for three different timestep values to represent experiments seeded with 200 cells at 1:1 ratio of BFP+ and -/- cells. ABM-predicted fractional areas

for the 2D model conditions were calculated as an average of 100 model runs, each run for 200, 100, and 50 total timesteps for the 15-, 30-, and 60-min time steps, respectively. Fractional areas from *in vitro* images were calculated from a sample of 10 images. Error bars indicate standard deviation. **(D)** Representative images of ABM predictions for each timestep are compared against a representative experimental image. (TIF)

Acknowledgments

The authors are grateful to Dr. Wendell Lim (University of California San Diego) for providing engineered L929 cell lines and to Dr. Satoshi Toda (Kanazawa University) for helpful technical discussions.

Author Contributions

Conceptualization: Nikita Sivakumar, Shayn M. Peirce, Matthew J. Lazzara.

Data curation: Nikita Sivakumar.

Formal analysis: Nikita Sivakumar.

Funding acquisition: Nikita Sivakumar, Shayn M. Peirce, Matthew J. Lazzara.

Investigation: Nikita Sivakumar.

Methodology: Nikita Sivakumar, Helen V. Warner, Shayn M. Peirce.

Project administration: Shayn M. Peirce, Matthew J. Lazzara.

Resources: Shayn M. Peirce, Matthew J. Lazzara.

Software: Nikita Sivakumar, Helen V. Warner, Shayn M. Peirce.

Supervision: Shayn M. Peirce, Matthew J. Lazzara.

Validation: Nikita Sivakumar.

Visualization: Nikita Sivakumar, Shayn M. Peirce, Matthew J. Lazzara.

Writing – original draft: Nikita Sivakumar, Shayn M. Peirce, Matthew J. Lazzara.

Writing – review & editing: Nikita Sivakumar, Shayn M. Peirce, Matthew J. Lazzara.

References

1. Battle E, Wilkinson DG. Molecular mechanisms of cell segregation and boundary formation in development and tumorigenesis. *Cold Spring Harb Perspect Biol.* 2012; 4(1):a008227. <https://doi.org/10.1101/cshperspect.a008227> PMID: 22214769
2. Toda S, Blauch LR, Tang SKY, Morsut L, Lim WA. Programming self-organizing multicellular structures with synthetic cell-cell signaling. *Science.* 2018; 361(6398):156–62. <https://doi.org/10.1126/science.aat0271> PMID: 29853554
3. Steinberg MS. Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *J Exp Zool.* 1970; 173(4):395–433. <https://doi.org/10.1002/jez.1401730406> PMID: 5429514
4. Steinberg MS. Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science.* 1963; 141(3579):401–8. <https://doi.org/10.1126/science.141.3579.401> PMID: 13983728
5. Steinberg MS, Takeichi M. Experimental specification of cell sorting, tissue spreading, and specific spatial patterning by quantitative differences in cadherin expression. *Proc Natl Acad Sci U S A.* 1994; 91(1):206–9. <https://doi.org/10.1073/pnas.91.1.206> PMID: 8278366

6. Duguay D, Foty RA, Steinberg MS. Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. *Dev Biol.* 2003; 253(2):309–23. [https://doi.org/10.1016/s0012-1606\(02\)00016-7](https://doi.org/10.1016/s0012-1606(02)00016-7) PMID: 12645933
7. Foty RA, Steinberg MS. The differential adhesion hypothesis: a direct evaluation. *Dev Biol.* 2005; 278(1):255–63. <https://doi.org/10.1016/j.ydbio.2004.11.012> PMID: 15649477
8. Glazier JA, Graner F. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics.* 1993; 47(3):2128–54. <https://doi.org/10.1103/physreve.47.2128> PMID: 9960234
9. Zhang Y, Thomas GL, Swat M, Shirinifard A, Glazier JA. Computer simulations of cell sorting due to differential adhesion. *PLoS One.* 2011; 6(10):e24999. <https://doi.org/10.1371/journal.pone.0024999> PMID: 22028771
10. Cickovski TM, Huang C, Chaturvedi R, Glimm T, Hentschel HG, Alber MS, et al. A framework for three-dimensional simulation of morphogenesis. *IEEE/ACM Trans Comput Biol Bioinform.* 2005; 2(4):273–88. <https://doi.org/10.1109/TCBB.2005.46> PMID: 17044166
11. Taylor HB, Khuong A, Wu Z, Xu Q, Morley R, Gregory L, et al. Cell segregation and border sharpening by Eph receptor-ephrin-mediated heterotypic repulsion. *J R Soc Interface.* 2017; 14(132). <https://doi.org/10.1098/rsif.2017.0338> PMID: 28747399
12. Steinberg MS. Differential adhesion in morphogenesis: a modern view. *Curr Opin Genet Dev.* 2007; 17(4):281–6. <https://doi.org/10.1016/j.gde.2007.05.002> PMID: 17624758
13. Akitaya T, Bronner-Fraser M. Expression of cell adhesion molecules during initiation and cessation of neural crest cell migration. *Dev Dyn.* 1992; 194(1):12–20. <https://doi.org/10.1002/aja.1001940103> PMID: 1421517
14. Bronner-Fraser M, Sternberg PW. Pattern formation and developmental mechanisms. The cell biological basis of inductive signaling. *Curr Opin Genet Dev.* 2000; 10(4):347–9. [https://doi.org/10.1016/s0959-437x\(00\)00094-0](https://doi.org/10.1016/s0959-437x(00)00094-0) PMID: 11023301
15. Bronner-Fraser M, Wolf JJ, Murray BA. Effects of antibodies against N-cadherin and N-CAM on the cranial neural crest and neural tube. *Dev Biol.* 1992; 153(2):291–301. [https://doi.org/10.1016/0012-1606\(92\)90114-v](https://doi.org/10.1016/0012-1606(92)90114-v) PMID: 1397686
16. Halbleib JM, Nelson WJ. Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev.* 2006; 20(23):3199–214. <https://doi.org/10.1101/gad.1486806> PMID: 17158740
17. Takeichi M. The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. *Development.* 1988;639–55. <https://doi.org/10.1242/dev.102.4.639> PMID: 3048970
18. Nose A, Nagafuchi A, Takeichi M. Expressed recombinant cadherins mediate cell sorting in model systems. *Cell.* 1988; 54(7):993–1001. [https://doi.org/10.1016/0092-8674\(88\)90114-6](https://doi.org/10.1016/0092-8674(88)90114-6) PMID: 3416359
19. Nose A, Tsuji K, Takeichi M. Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell.* 1990; 61(1):147–55. [https://doi.org/10.1016/0092-8674\(90\)90222-z](https://doi.org/10.1016/0092-8674(90)90222-z) PMID: 2317870
20. Grant CM, Kyprianou N. Epithelial mesenchymal transition (EMT) in prostate growth and tumor progression. *Transl Androl Urol.* 2013; 2(3):202–11. <https://doi.org/10.3978/j.issn.2223-4683.2013.09.04> PMID: 25346895
21. Pastushenko I, Blanpain C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol.* 2019; 29(3):212–26. <https://doi.org/10.1016/j.tcb.2018.12.001> PMID: 30594349
22. Gloushankova NA, Zhitnyak IY, Rubtsova SN. Role of Epithelial-Mesenchymal Transition in Tumor Progression. *Biochemistry (Mosc).* 2018; 83(12):1469–76. <https://doi.org/10.1134/S0006297918120052> PMID: 30878022
23. Hutchins EJ, Bronner ME. A Spectrum of Cell States During the Epithelial-to-Mesenchymal Transition. *Methods Mol Biol.* 2021; 2179:3–6. https://doi.org/10.1007/978-1-0716-0779-4_1 PMID: 32939707
24. Aclouque H, Adams MS, Fishwick K, Bronner-Fraser M, Nieto MA. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *J Clin Invest.* 2009; 119(6):1438–49. <https://doi.org/10.1172/JCI38019> PMID: 19487820
25. Bocci F, Gearhart-Serna L, Boareto M, Ribeiro M, Ben-Jacob E, Devi GR, et al. Toward understanding cancer stem cell heterogeneity in the tumor microenvironment. *Proc Natl Acad Sci U S A.* 2019; 116(1):148–57. <https://doi.org/10.1073/pnas.1815345116> PMID: 30587589
26. Ouled Dhaou M, Kossai M, Morel AP, Devouassoux-Shisheboran M, Puisieux A, Penault-Llorca F, et al. Zeb1 expression by tumor or stromal cells is associated with spatial distribution patterns of CD8+ tumor-infiltrating lymphocytes: a hypothesis-generating study on 113 triple negative breast cancers. *Am J Cancer Res.* 2020; 10(10):3370–81. PMID: 33163276
27. Gomez EW, Chen QK, Gjorevski N, Nelson CM. Tissue geometry patterns epithelial-mesenchymal transition via intercellular mechanotransduction. *J Cell Biochem.* 2010; 110(1):44–51. <https://doi.org/10.1002/jcb.22545> PMID: 20336666

28. Ramis-Conde I, Drasdo D, Anderson AR, Chaplain MA. Modeling the influence of the E-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach. *Biophys J*. 2008; 95(1):155–65. <https://doi.org/10.1529/biophysj.107.114678> PMID: 18339758
29. Astin JW, Batson J, Kadir S, Charlet J, Persad RA, Gillatt D, et al. Competition amongst Eph receptors regulates contact inhibition of locomotion and invasiveness in prostate cancer cells. *Nat Cell Biol*. 2010; 12(12):1194–204. <https://doi.org/10.1038/ncb2122> PMID: 21076414
30. Mateo-Lozano S, Bazzocco S, Rodrigues P, Mazzolini R, Andretta E, Dopeso H, et al. Loss of the EPH receptor B6 contributes to colorectal cancer metastasis. *Sci Rep*. 2017; 7:43702. <https://doi.org/10.1038/srep43702> PMID: 28262839
31. Makrilia N, Kollias A, Manolopoulos L, Syrigos K. Cell adhesion molecules: role and clinical significance in cancer. *Cancer Invest*. 2009; 27(10):1023–37. <https://doi.org/10.3109/07357900902769749> PMID: 19909018
32. Stanta G, Bonin S. Overview on Clinical Relevance of Intra-Tumor Heterogeneity. *Front Med (Lausanne)*. 2018; 5:85. <https://doi.org/10.3389/fmed.2018.00085> PMID: 29682505
33. Sun XX, Yu Q. Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment. *Acta Pharmacol Sin*. 2015; 36(10):1219–27. <https://doi.org/10.1038/aps.2015.92> PMID: 26388155
34. Dietz S, Harms A, Endris V, Eichhorn F, Kriegsmann M, Longuespee R, et al. Spatial distribution of EGFR and KRAS mutation frequencies correlates with histological growth patterns of lung adenocarcinomas. *Int J Cancer*. 2017; 141(9):1841–8. <https://doi.org/10.1002/ijc.30881> PMID: 28699162
35. Johnson MB, March AR, Morsut L. Engineering multicellular systems: using synthetic biology to control tissue self-organization. *Curr Opin Biomed Eng*. 2017; 4:163–73. <https://doi.org/10.1016/j.cobme.2017.10.008> PMID: 29308442
36. Carvalho A, Menendez DB, Senthivel VR, Zimmermann T, Diambra L, Isalan M. Genetically encoded sender-receiver system in 3D mammalian cell culture. *ACS Synth Biol*. 2014; 3(5):264–72. <https://doi.org/10.1021/sb400053b> PMID: 24313393
37. Morsut L, Roybal KT, Xiong X, Gordley RM, Coyle SM, Thomson M, et al. Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell*. 2016; 164(4):780–91. <https://doi.org/10.1016/j.cell.2016.01.012> PMID: 26830878
38. Toda S, Brunger JM, Lim WA. Synthetic development: learning to program multicellular self-organization. *Current Opinion in Systems Biology*. 2019; 14:41–9.
39. Guye P, Ebrahimkhani MR, Kipniss N, Velazquez JJ, Schoenfeld E, Kiani S, et al. Genetically engineering self-organization of human pluripotent stem cells into a liver bud-like tissue using Gata6. *Nat Commun*. 2016; 7:10243. <https://doi.org/10.1038/ncomms10243> PMID: 26732624
40. Zhang S, Yan L, Altman M, Lassle M, Nugent H, Frankel F, et al. Biological surface engineering: a simple system for cell pattern formation. *Biomaterials*. 1999; 20(13):1213–20. [https://doi.org/10.1016/S0142-9612\(99\)00014-9](https://doi.org/10.1016/S0142-9612(99)00014-9) PMID: 10395390
41. Toda S, McKeithan WL, Hakkinen TJ, Lopez P, Klein OD, Lim WA. Engineering synthetic morphogen systems that can program multicellular patterning. *Science*. 2020; 370(6514):327–31. <https://doi.org/10.1126/science.abc0033> PMID: 33060357
42. Walpole J, Papin JA, Peirce SM. Multiscale computational models of complex biological systems. *Annu Rev Biomed Eng*. 2013; 15:137–54. <https://doi.org/10.1146/annurev-bioeng-071811-150104> PMID: 23642247
43. Shvartsman SY, Muratov CB, Lauffenburger DA. Modeling and computational analysis of EGF receptor-mediated cell communication in *Drosophila* oogenesis. *Development*. 2002; 129(11):2577–89. <https://doi.org/10.1242/dev.129.11.2577> PMID: 12015287
44. Barkai N, Shilo BZ. Modeling pattern formation: counting to two in the *Drosophila* egg. *Curr Biol*. 2002; 12(14):R493–5. [https://doi.org/10.1016/S0960-9822\(02\)00967-3](https://doi.org/10.1016/S0960-9822(02)00967-3) PMID: 12176350
45. Ben-Zvi D, Barkai N. Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proc Natl Acad Sci U S A*. 2010; 107(15):6924–9. <https://doi.org/10.1073/pnas.0912734107> PMID: 20356830
46. Ozik J, Collier N, Heiland R, An G, Macklin P. Learning-accelerated discovery of immune-tumour interactions. *Mol Syst Des Eng*. 2019; 4(4):747–60. <https://doi.org/10.1039/c9me00036d> PMID: 31497314
47. Robertson SH, Smith CK, Langhans AL, McLinden SE, Oberhardt MA, Jakab KR, et al. Multiscale computational analysis of *Xenopus laevis* morphogenesis reveals key insights of systems-level behavior. *BMC Syst Biol*. 2007; 1:46. <https://doi.org/10.1186/1752-0509-1-46> PMID: 17953751
48. Thorne BC, Bailey AM, Peirce SM. Combining experiments with multi-cell agent-based modeling to study biological tissue patterning. *Brief Bioinform*. 2007; 8(4):245–57. <https://doi.org/10.1093/bib/bbm024> PMID: 17584763

49. Thorne BC, Bailey AM, DeSimone DW, Peirce SM. Agent-based modeling of multicell morphogenic processes during development. *Birth Defects Res C Embryo Today*. 2007; 81(4):344–53. <https://doi.org/10.1002/bdrc.20106> PMID: 18228259
50. Walpole J, Chappell JC, Cluceru JG, Mac Gabhann F, Bautch VL, Peirce SM. Agent-based model of angiogenesis simulates capillary sprout initiation in multicellular networks. *Integr Biol (Camb)*. 2015; 7(9):987–97. <https://doi.org/10.1039/c5ib00024f> PMID: 26158406
51. Longo D, Peirce SM, Skalak TC, Davidson L, Marsden M, Dzamba B, et al. Multicellular computer simulation of morphogenesis: blastocoel roof thinning and matrix assembly in *Xenopus laevis*. *Dev Biol*. 2004; 271(1):210–22. <https://doi.org/10.1016/j.ydbio.2004.03.021> PMID: 15196962
52. Izaguirre JA, Chaturvedi R, Huang C, Cickovski T, Coffland J, Thomas G, et al. CompuCell, a multi-model framework for simulation of morphogenesis. *Bioinformatics*. 2004; 20(7):1129–37. <https://doi.org/10.1093/bioinformatics/bth050> PMID: 14764549
53. Zeng W, Thomas GL, Glazier JA. Non-Turing stripes and spots: a novel mechanism for biological cell clustering. *Physica A: Statistical Mechanics and its Applications*. 2004; 341:482–94.
54. Ramis-Conde I, Chaplain MA, Anderson AR, Drasdo D. Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Phys Biol*. 2009; 6(1):016008. <https://doi.org/10.1088/1478-3975/6/1/016008> PMID: 19321920
55. Matsuda M, Koga M, Nishida E, Ebisuya M. Synthetic signal propagation through direct cell-cell interaction. *Sci Signal*. 2012; 5(220):ra31. <https://doi.org/10.1126/scisignal.2002764> PMID: 22510469
56. Tan RZ, Chiam KH. Computational modeling reveals that a combination of chemotaxis and differential adhesion leads to robust cell sorting during tissue patterning. *PLoS One*. 2014; 9(10):e109286. <https://doi.org/10.1371/journal.pone.0109286> PMID: 25302949
57. Letort G, Montagud A, Stoll G, Heiland R, Barillot E, Macklin P, et al. PhysiBoSS: a multi-scale agent-based modelling framework integrating physical dimension and cell signalling. *Bioinformatics*. 2019; 35(7):1188–96. <https://doi.org/10.1093/bioinformatics/bty766> PMID: 30169736
58. Vittadello ST, Leyshon T, Schnoerr D, Stumpf MPH. Turing pattern design principles and their robustness. *Philos Trans A Math Phys Eng Sci*. 2021; 379(2213):20200272. <https://doi.org/10.1098/rsta.2020.0272> PMID: 34743598
59. Maini PK, Woolley TE, Baker RE, Gaffney EA, Lee SS. Turing's model for biological pattern formation and the robustness problem. *Interface Focus*. 2012; 2(4):487–96. <https://doi.org/10.1098/rsfs.2011.0113> PMID: 23919129
60. Scholes NS, Schnoerr D, Isalan M, Stumpf MPH. A Comprehensive Network Atlas Reveals That Turing Patterns Are Common but Not Robust. *Cell Syst*. 2019; 9(3):243–57 e4. <https://doi.org/10.1016/j.cels.2019.07.007> PMID: 31542413
61. Turing AM. The chemical basis of morphogenesis. 1953. *Bull Math Biol*. 1990; 52(1–2):153–97; discussion 19–52. <https://doi.org/10.1007/BF02459572> PMID: 2185858
62. Norfleet DA, Park E, Kemp ML. Computational modeling of organoid development. *Current Opinion in Biomedical Engineering*. 2020; 13:113–8.
63. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R. A synthetic multicellular system for programmed pattern formation. *Nature*. 2005; 434(7037):1130–4. <https://doi.org/10.1038/nature03461> PMID: 15858574
64. Karig D, Martini KM, Lu T, DeLateur NA, Goldenfeld N, Weiss R. Stochastic Turing patterns in a synthetic bacterial population. *Proc Natl Acad Sci U S A*. 2018; 115(26):6572–7. <https://doi.org/10.1073/pnas.1720770115> PMID: 29891706
65. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, et al. EMT and dissemination precede pancreatic tumor formation. *Cell*. 2012; 148(1–2):349–61. <https://doi.org/10.1016/j.cell.2011.11.025> PMID: 22265420
66. Buonato JM, Lazzara MJ. ERK1/2 blockade prevents epithelial-mesenchymal transition in lung cancer cells and promotes their sensitivity to EGFR inhibition. *Cancer Res*. 2014; 74(1):309–19. <https://doi.org/10.1158/0008-5472.CAN-12-4721> PMID: 24108744
67. Warner VH, Sivakumar N., Peirce M. S., Lazzara J. M. Multiscale computational models of cancer. *Current Opinion in Biomedical Engineering*. 2019;11.
68. Cilfone NA, Perry CR, Kirschner DE, Linderman JJ. Multi-scale modeling predicts a balance of tumor necrosis factor-alpha and interleukin-10 controls the granuloma environment during *Mycobacterium tuberculosis* infection. *PLoS One*. 2013; 8(7):e68680.
69. Pienaar E, Matern WM, Linderman JJ, Bader JS, Kirschner DE. Multiscale Model of *Mycobacterium tuberculosis* Infection Maps Metabolite and Gene Perturbations to Granuloma Sterilization Predictions. *Infect Immun*. 2016; 84(5):1650–69. <https://doi.org/10.1128/IAI.01438-15> PMID: 26975995

70. Warsinske HC, Wheaton AK, Kim KK, Linderman JJ, Moore BB, Kirschner DE. Computational Modeling Predicts Simultaneous Targeting of Fibroblasts and Epithelial Cells Is Necessary for Treatment of Pulmonary Fibrosis. *Front Pharmacol*. 2016; 7:183. <https://doi.org/10.3389/fphar.2016.00183> PMID: 27445819
71. Rikard SM, Athey TL, Nelson AR, Christiansen SLM, Lee JJ, Holmes JW, et al. Multiscale Coupling of an Agent-Based Model of Tissue Fibrosis and a Logic-Based Model of Intracellular Signaling. *Front Physiol*. 2019; 10:1481. <https://doi.org/10.3389/fphys.2019.01481> PMID: 31920691
72. Myers SHL P. J., Lazzara M. J. Mechanistic and data-driven models of cell signaling: Tools for fundamental discovery and rational design of therapy. *Current Opinion in Systems Biology*. 2021;28.
73. Cess CG, Finley SD. Data-driven analysis of a mechanistic model of CAR T cell signaling predicts effects of cell-to-cell heterogeneity. *J Theor Biol*. 2020; 489:110125. <https://doi.org/10.1016/j.jtbi.2019.110125> PMID: 31866395
74. Cess CG, Finley SD. Multi-scale modeling of macrophage-T cell interactions within the tumor microenvironment. *PLoS Comput Biol*. 2020; 16(12):e1008519. <https://doi.org/10.1371/journal.pcbi.1008519> PMID: 33362239
75. Lee JJ, Talman L, Peirce SM, Holmes JW. Spatial scaling in multiscale models: methods for coupling agent-based and finite-element models of wound healing. *Biomech Model Mechanobiol*. 2019; 18(5):1297–309. <https://doi.org/10.1007/s10237-019-01145-1> PMID: 30968216
76. Virgilio KM, Martin KS, Peirce SM, Blemker SS. Agent-based model illustrates the role of the microenvironment in regeneration in healthy and mdx skeletal muscle. *J Appl Physiol (1985)*. 2018; 125(5):1424–39. <https://doi.org/10.1152/jappphysiol.00379.2018> PMID: 30070607
77. Macklin P. Key challenges facing data-driven multicellular systems biology. *Gigascience*. 2019; 8(10). <https://doi.org/10.1093/gigascience/giz127> PMID: 31648301
78. Macklin P, Edgerton ME, Thompson AM, Cristini V. Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): from microscopic measurements to macroscopic predictions of clinical progression. *J Theor Biol*. 2012; 301:122–40. <https://doi.org/10.1016/j.jtbi.2012.02.002> PMID: 22342935
79. Wilensky U. 1999. Netlogo <http://cclnorthwesternedu/netlogo/>. Center for Connected Learning and Computer-Based Modeling: Northwestern University, Evanston, IL.
80. Gunaratne C, Garibay I. NL4Py: Agent-Based Modeling in Python with Parallelizable NetLogo Workspaces. 2018.
81. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2021; 12(85):2825–30.