

**INVESTIGATING THE ROLE OF INTERCELLULAR  
COMMUNICATION ON SPATIAL DIFFERENTIATION THROUGH  
AGENT-BASED MODELING**

A Dissertation

Presented to

The Academic Faculty

by

Chad Michael Glen

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in Biomedical Engineering in the

Wallace H. Coulter Department of Biomedical Engineering

Georgia Institute of Technology

Emory University

August 2018

**COPYRIGHT © 2018 BY CHAD M. GLEN**

**INVESTIGATING THE ROLE OF INTERCELLULAR  
COMMUNICATION ON SPATIAL DIFFERENTIATION THROUGH  
AGENT-BASED MODELING**

Approved by:

Dr. Melissa L. Kemp, Advisor

Department of Biomedical Engineering

*Georgia Institute of Technology*

Dr. Todd C. McDevitt, Advisor

Department of Bioengineering &

Therapeutic Sciences

*University of California*

Dr. Krishnendu Roy

Department of Biomedical Engineering

*Georgia Institute of Technology*

Dr. Brani Vidakovic

Department of Biomedical Engineering

*Georgia Institute of Technology*

Dr. Eberhard O. Voit.

Department of Biomedical Engineering

*Georgia Institute of Technology*

Dr. Michael Levin

Department of Regenerative &

Developmental Biology

*Tufts University*

Date Approved: July 25, 2018

To caffeine and my sanity

## **ACKNOWLEDGEMENTS**

I would like to start by acknowledging that I am not particularly adept at personal acknowledgements and that this will accordingly be brief and to the point. I would like to thank my advisors, Melissa Kemp and Todd McDevitt. They have both provided significant guidance along the way and helped me to achieve significantly more than I would have thought possible. Also, I can not understate my appreciation for their willingness to accept my sometimes ill-timed humor. I would like to thank members of the Kemp lab, both past and present, for creating an enjoyable environment to work in. I would like to thank two specific friends, Liana Chan and Chantal Trepanier, who have always had an open ear for me over the years. I am not sure I have ever been entirely sane but I am positive that I would have been entirely insane if they had not been around to talk to over the last five years. I would like to thank my parents for always supporting and believing in me. They have both been amazing role models for me and taught me how to think independently, which has been invaluable throughout my life. Finally, I need to thank Alyssa Duck. She has had a profound effect on me, my life, and my appreciation of cats. We have been through a lot together over the last 3 years and I expect there will be many more years to come.



## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS .....</b>	<b>IV</b>
<b>LIST OF TABLES .....</b>	<b>VII</b>
<b>LIST OF FIGURES .....</b>	<b>VIII</b>
<b>LIST OF SYMBOLS AND ABBREVIATIONS .....</b>	<b>XI</b>
<b>SUMMARY .....</b>	<b>XII</b>
<b>CHAPTER 1 INTRODUCTION.....</b>	<b>1</b>
<b>1.1. Research Objectives and Specific Aims .....</b>	<b>1</b>
<b>CHAPTER 2 BACKGROUND .....</b>	<b>4</b>
<b>2.1. Embryonic Stem Cells .....</b>	<b>4</b>
<b>2.2. Morphogenesis.....</b>	<b>5</b>
<b>2.2.1 Mathematical and Computational Approaches to Morphogenesis.....</b>	<b>9</b>
<b>2.2.2 Agent-Based Models for Morphogenesis .....</b>	<b>17</b>
<b>2.3. Gap Junctions and Intercellular Communication .....</b>	<b>46</b>
<b>2.4. Connexins Structure and Function .....</b>	<b>46</b>
<b>2.5. Connexins and Differentiation.....</b>	<b>48</b>
<b>2.6. Connexins and Cell Cycle.....</b>	<b>50</b>
<b>CHAPTER 3 CHARACTERIZATION OF MULTISCALE INTERCELLULAR COMMUNICATION DURING EARLY NEURAL DIFFERENTIATION.....</b>	<b>51</b>
<b>3.1. Introduction.....</b>	<b>51</b>
<b>3.2. Methods.....</b>	<b>53</b>
<b>3.2.1 Experimental Methods .....</b>	<b>53</b>
<b>3.2.2 Computational Methods.....</b>	<b>57</b>
<b>3.3. Results .....</b>	<b>64</b>
<b>3.3.1 Spatially organized differentiation occurs during retinoic acid treatment</b>	<b>64</b>
<b>3.3.2 Differentiation causes a transient redistribution of Cx43 and increased permeability for communication between neighboring cells .....</b>	<b>67</b>
<b>3.3.3 Intercellular transport as a function of cell cycle .....</b>	<b>70</b>
<b>3.3.4 Cell-cycle dynamics generate intercellular network heterogeneity and system complexity .....</b>	<b>74</b>

3.3.5 Intercellular diffusion dynamics generate spatiotemporal trajectories of differentiation .....	77
3.4. Discussion.....	80
<b>CHAPTER 4 THE EFFECT OF INTERCELLULAR PERTURBATION ON DIFFERENTIATION POTENTIAL .....</b>	<b>84</b>
4.1. Introduction.....	84
4.2. Methods.....	85
4.2.1 Experimental Methods .....	85
4.2.2 Computational Methods.....	87
4.3. mESC Results .....	87
4.3.1 Selected Perturbations.....	87
4.3.2 Perturbations to the D3 mESC Cell Line .....	88
4.3.3 Perturbations to the G-Olig2 RW4 mESC Cell Line.....	93
4.4. hiPSC Results .....	98
4.4.1 Connexin43 Knockdown .....	98
4.5. Discussion.....	101
<b>CHAPTER 5 MODULAR COMMUNICATION IN AGENT-BASED MODELS</b>	<b>105</b>
5.1. Multiscale Communication Agent-Based Model Generator.....	106
5.1.1 Network Creation.....	107
5.1.2 Model Parameters .....	110
5.1.3 State-Change Mechanism & Parameters.....	114
5.1.4 Compile & Run .....	120
5.2. Application Example of MsCAMgen .....	122
5.3. Discussion.....	123
<b>CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS.....</b>	<b>126</b>
6.1. Conclusions.....	126
6.2. Future Directions .....	129
6.2.1 Agent-Based Bioelectricity .....	129
6.2.2 Computational Improvements .....	130
6.2.3 hiPSC Cx43-KD Mixed Populations .....	133
<b>REFERENCES.....</b>	<b>135</b>

## LIST OF TABLES

Table 1: Cell types present in nervous system and their corresponding connexins and interactions.....	49
Table 2: Cell type specific parameters used in ABM of intercellular communication.....	58

## LIST OF FIGURES

Figure 1: Overview of the two most popular theories for pattern formation during morphogenesis - reaction-diffusion (RD) systems and positional information (PI) - as well as common features of agent-based models (ABMs) for morphogenesis .....	13
Figure 2: The three most common physical representations of cells in ABMs.....	16
Figure 3: Classification used to categorize three types of agent-based models for morphogenesis. ....	19
Figure 4: Comparison of three crypt model implementations and agent descriptions. ....	25
Figure 5: Overview of the analysis of spatial features using a combination of network analysis and dimensional reduction techniques. ....	36
Figure 6: Quantifying the loss of fluorescence in the cells adjacent to the photobleached cell during gap-FRAP. ....	56
Figure 7: Comparison of cell cycle time and the respective DTnorm transform.....	59
Figure 8: CellProfiler pipeline used to convert experimentally imaged colonies to digital networks synonymous with model outputs.....	62
Figure 9: The seven selected metrics used to quantify the spatial patterning within colonies. ....	63
Figure 10: Loss of Oct4 and maintenance of Sox2 expression during retinoic acid-induced differentiation.....	65
Figure 11: Sox2 and Oct4 concomitant loss through LIF withdrawal in the presence of serum containing media .....	66
Figure 12: Consistent Cx43 stain throughout membrane in mitotic cells.....	67
Figure 13: Cx43 signal (green) increases during retinoic acid-induced differentiation, with compartmentalization of transitioning cells between 24 and 48 hours. ....	69
Figure 14: Analysis of intercellular transport rates as a function of cell-cycle state.....	71
Figure 15: Defining the asynchronous distribution of recovery constants by cell cycle state. ....	73

Figure 16: Comparison of recovery constants collected over the first 48 hours of retinoic acid (RA) exposure that were calculated from GAP-FRAP results.....	75
Figure 17: Computational analysis of cell-cycle modulation on intercellular communication.....	76
Figure 18: Quantification of spatial patterning during RA differentiation. ....	79
Figure 19: Experimental colonies of D3 mESC cell line at 24 and 48 hour time points of retinoic acid treatment.....	90
Figure 20: Simulation colonies at 24 and 48 hour time points of retinoic acid treatment for the Vehicle control (n=75), SQ-treatment (n=30), and $\beta$ -GA-treatments (n=30) converted into latent space.....	91
Figure 21: Perturbation to the intercellular network of a multicellular D3 ESC population affects RA-accelerated differentiation in a temporal manner. ....	92
Figure 22: Experimental colonies of RW4 mESC cell line at 24 and 48 hour time points of retinoic acid treatment.....	94
Figure 23: Simulation colonies at 19 and 43 hour time points of retinoic acid treatment for the Vehicle control (n=75), SQ-treatment (n=30), and BGA-treatments (n=30) converted into latent space.....	95
Figure 24: Perturbation to the intercellular network of a multicellular RW4 ESC population affects neural differentiation in a temporal manner. ....	96
Figure 25: RW4 colony patterning and morphology after 24 and 48 hours of retinoic acid exposure, respectively, with adenylyl cyclase or gap junction inhibition compared to the vehicle control.....	97
Figure 26: Dual-SMAD inhibition in the hiPSC line LBC-GJA1 with and without doxycycline for days 4 through 6.....	99
Figure 27: Average Oct4 and Pax6 expression over 6 days of dual-SMAD inhibition in the hiPSC line LBC2-GJA1 with (Cx43-KD) and without (WT) doxycycline.....	100

Figure 28: Cx43 expression during dual-SMAD inhibition in the hiPSC line LBC-GJA1 with (Cx43-KD) and without (WT) doxycycline for days 0, 2 and 6.....	101
Figure 29: A schematic diagram of our proposed mechanism for the influence of AC and GJ inhibition on differentiation potential.....	102
Figure 30: Overview of the primary features and forms of communication of MsCAMgen across different timescales. ....	106
Figure 31: The first frame “Network Creation” and initial window that appears upon launching MsCAMgen.....	108
Figure 32: The four possible cell transition hierarchies.....	109
Figure 33: Using an experimental image as a template for the morphology of the cell network in MsCAMgen. ....	110
Figure 34: The second frame "Model Parameters" of MsCAMgen. ....	111
Figure 35: The three types of colony growth dynamics available in MsCAMgen. ....	112
Figure 36: Defining non-transportable intracellular species and their respective differential equations. ....	114
Figure 37: The third frame “State-Change Mechanism” of MsCAMgen.....	116
Figure 38: Setting up a direct interaction mechanism in MsCAMgen. ....	117
Figure 39: The fourth frame "State-Change Parameters” for Direct Interaction in MsCAMgen.....	118
Figure 40: Setting up a phenotypic expression mechanism in MsCAMgen.....	119
Figure 41: The fourth frame "State-Change Parameters” for Phenotypic Expression in MsCAMgen.....	120
Figure 42: The fifth frame "Compile & Run" for MsCAMgen. ....	122
Figure 43: Simulated BMP4 activity, indicated by pSMAD1, as a function of cell density using a MsCAMgen generated model.....	123

## LIST OF SYMBOLS AND ABBREVIATIONS

ABM	Agent-based model
AC	Adenylyl Cyclase
Async	Asynchronous (population)
$\beta$ -GA	$\beta$ -glycyrrhetic acid
cAMP	Cyclic adenosine monophosphate
CP	Cortical plate
CPM	Cellular Potts Model
Cx	Connexin
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
ECM	Extracellular matrix
ENS	Enteric nervous system
Epi	Epiblast
ESC	Embryonic stem cell
Fn	Fibronectin
GFP	Green fluorescent protein
GJC	Gap junction communication
GT	Genital tubercle
hiPSC	Human induced pluripotent stem cell
HMI	Hoechst-mediated identification
IZ	Intermediate zone
LIF	Leukemia inhibitory factor
MZ	Marginal zone
NDS	Normal donkey serum
NOC	Nocodazole
PCA	Principal component analysis
PDE	Partial differential equation
ODE	Ordinary differential equation
PI	Positional information
PrE	Primitive endoderm
RA	Retinoic acid
RD	Reaction-diffusion
SQ	SQ 22536

## SUMMARY

The initiation of heterogeneity within a population of phenotypically identical progenitors is a critical event for the onset of morphogenesis and differentiation patterning. Information flow between adjacent cells informs cell fate decisions and can occur by a number of mechanisms. Gap junction communication within multicellular systems produces complex networks of intercellular connectivity that result in heterogeneous distributions of intracellular signaling molecules. In this work, we investigated the emergent systems-level behavior of the intercellular network within embryonic stem cell (ESC) populations and the corresponding spatial organization during early neural differentiation.

An agent-based computational model of ESC collective behavior was designed to prompt the state change of individual cells through intracellular accumulation of molecular drivers of differentiation throughout a colony. Transfer rates in the intercellular network were defined from experimentally-determined parameters, with cell cycle modulation of gap junctions stimulating dynamic connectivity. The model yielded complex, dynamic transport networks for delivery of differentiation cues between neighboring cells, reproducing the distribution and variety of observed morphogenic trajectories that result during retinoic acid-induced mouse ESC differentiation. Furthermore, the model correctly predicted the delayed differentiation and preserved spatial features of the morphogenic trajectory that occurs in response to perturbation to intercellular communication. These findings suggest an integral role of gap junction communication in the temporal coordination of emergent patterning during early differentiation and neural commitment of pluripotent stem cells.

The relationship between intercellular communication and neural differentiation was further interrogated through the CRISPRi-mediated knockdown of connexin43 (Cx43), the predominant gap junction protein in pluripotent cells. The selective removal of Cx43 during the differentiation of human induced pluripotent cells (hiPSCs) reiterated the



role of intercellular communication in the temporal control of differentiation by delaying neural commitment. Furthermore, it was discovered that in both mESC and hiPSC the loss of pluripotency was concomitant with transient, localized increases in Cx43 expression. These results suggest that the coordination of differentiation events within a multicellular population involves the dynamic regulation of gap junction communication.

While this work considered intercellular communication in isolation, there is significant cross-talk between the various modes of cellular communication present within multicellular systems. The diffusion and binding of extracellular molecules can modulate production of intracellular molecules that are able to diffuse through gap junction channels [1]. Even more complicated, some intracellular molecules can be both secreted extracellularly and diffuse intercellularly, while modulating the secretion and production of other extracellular and intercellularly-diffusing molecules [2-4]. Furthermore, molecules that diffuse intercellularly can have a significant impact on gene expression [5], cell cycle progression [6], and even apoptosis [7]. As such, there is a need for tools that are able to integrate multiple modes of communication and decipher their interactions in a spatial context.

To facilitate future studies of emergence in multicellular systems, a multiscale communication agent-based model generator (MsCAMgen) was developed in Python. MsCAMgen provides a framework for modeling various spatial aspects of a multicellular network without requiring explicit programming by the user. Each model is capable of accounting for cell division and growth, state changes between different cell types, extracellular diffusion of molecules that are secreted and consumed by cells, intercellular communication of small molecules between neighboring cells, and intracellular gene/protein networks. The ability to quickly add and remove these features at the discretion of the user makes MsCAMgen an ideal platform for investigating emergence in biological systems. Furthermore, the ease of simulating diverse morphological structures

that can include and integrate each of these processes distinguishes MsCAMgen as a uniquely suited tool for optimizing the design of engineered living systems.

In summary, cell cycle state was coupled with intercellular transport rates and incorporated into an agent-based model capable of replicating and predicting spatiotemporal trajectories of differentiation. This mode of intercellular communication as a driver of differentiation was validated with small molecule inhibitors as well as CRISPR interference targeting of connexins. In addition, we developed a graphical user interface for generating agent-based models containing various modes of communication. This platform provides a tool to interrogate the spatial aspects of information transfer within multicellular systems and facilitates accelerated design of engineered living systems such as organoids by enabling the analysis of multiscale communication within cell populations of any morphology or organization.

# CHAPTER 1 INTRODUCTION

The spatial organization of heterogeneous cells within multicellular systems, such as tissues and organs, is a primary determinant in deriving their respective functionality [8]. Understanding the processes and mechanisms that contribute to the spatial arrangement of cells within such systems is a vital step towards the creation of *in vitro* replica. Although developmental biology has focused intensely for decades on understanding the role of reaction-diffusion morphogens in driving cell fate decisions, one often overlooked mechanism when considering pattern formation is the intercellular communication between cells. The ability to directly couple the cytosolic compartment between two cells is a fast and effective way to exchange information and allow coordination amongst local cells. At a global scale, this local coordination can enable the propagation of signals across a population of cells at a significantly faster time scale than the secretion and diffusion of extracellular signaling molecules. As such, elucidating the dynamics of intercellular communication would provide significant insight into how signaling progresses across a multicellular system. However, these processes not only need to be understood as separate entities but also as interactive units during the development of multicellular systems. Computational modeling provides an elegant solution for handling the complexity of this task. In particular, agent-based models are adept at the integration of various processes and the study of their dynamics. The main objectives of this research were to 1) identify the role of intercellular communication in differentiation patterning and 2) facilitate advances in the development of engineered living systems. These objectives were addressed by the following three aims:

## 1.1. Research Objectives and Specific Aims

**Aim 1: Investigate the influence of cell cycle state on gap junction communication and differentiation patterning in pluripotent colonies.** This aim sought to model intercellular communication within a multicellular population using cell cycle to regulate gap junction

permeability. *The working hypothesis was that intercellular transport governed by the natural asynchrony of the cell cycle could generate intracellular gradients capable of directing fate decisions.* Intercellular transport rates were quantified at the single-cell level and the distribution of calculated transport rates was linked to cell cycle state. The garnered single-cell transport information was incorporated into an agent-based model and the resultant intracellular gradients used to instigate differentiation. Spatial patterns were quantified and compared between simulations and pluripotent colonies differentiated *in vitro*.

**Aim 2: Evaluate the effect of intercellular perturbation on differentiation potential and the emergence of differentiation patterns.** In this aim, the intercellular model from Aim 1 was used in tandem with perturbation experiments to assess the role of intercellular communication in pattern formation during differentiation. *The working hypothesis was that perturbations to the intercellular network would alter the spatiotemporal trajectory of differentiation within a pluripotent colony.* Small molecules were used to modulate (i) the production of an intracellular molecule and (ii) gap junction connectivity over the course of differentiation. Furthermore, the effect of knocking down the predominantly expressed connexin in pluripotent cells was assessed. The inhibited intercellular network was found to have a strong temporal delay but minimal influence on spatial patterning during early differentiation events.

**Aim 3: Develop a graphical interface for constructing diverse agent-based models applicable to other biological systems.** This aim sought to produce an easy-to-use application for generating biologically related agent-based models with a myriad of features that was not limited by the programming knowledge of the user. *The working hypothesis was that the modular nature of agent-based models would enable the selection and removal of features from a graphical interface.* An application was developed in Python that allows a user to designate the cellular organization and morphology of a cell population. The user can specify whether extracellular and/or intercellular transport is

occurring within the system and how each of those factors affect state change. Furthermore, an intracellular system of non-transportable molecules can be included that is solved independently for each cell. The user-defined model can be compiled and run within the application, with real-time updates on the morphology and organization of cell types within the population.

## CHAPTER 2      BACKGROUND

### 2.1.    Embryonic Stem Cells

Embryonic stem cells (ESCs) are derived from cells originating in the inner cell mass of a blastocyst during development [9, 10]. These pluripotent cells are unique in their ability to differentiate into the three germ lineages and thus any somatic cell of the fully developed organism. Furthermore, ESCs are capable of undergoing self-organization and forming organ-like structures when differentiated under specific conditions. Therefore, ESCs also provide an invaluable tool for studying the mechanisms that regulate pattern formation during morphogenesis [11, 12]. The ability to recapitulate these naturally arising cellular processes *in vitro* is essential in order to design more intricate, multicellular systems. As such, ESCs have potential for two types of therapeutics: (i) as a replacement source of a specific cell type, and (ii) as a system for deriving tissue-engineered constructs. The stem cell field, in past and present, has predominantly focused on the establishment of differentiation protocols that can efficiently convert pluripotent cells into nearly homogenous populations of a desired cell type, but has generally ignored the trajectory of differentiation and the spatial patterns that transpire during the transition of phenotype. While current methodologies provide a framework for achieving efficient differentiation, usually towards a single phenotype, they do not explore how initial fate decisions affect future pattern emergence.

## 2.2. Morphogenesis

Morphogenesis is the complex chain of biological processes with which cellular populations self-organize, in a reproducible manner, into predetermined structures or patterns. It involves a multitude of mechanisms and systems and is governed by signal transduction across various spatial and temporal scales. In addition to evident outward patterning events, such as the formation of stripes on a tiger's back or the regular spacing of hairs or feathers, morphogenesis encompasses all molecular processes that convert a fertilized egg cell into a blastula, then into an embryo with germ layers that have their unique roles, and ultimately into a functional organism. It has been known for some while that the numerous simultaneous events during this journey, such as the development of fingers out of a limb bud or the organization of neurons into functional networks in the brain, involve fundamental processes of cell migration and differentiation, but it is extraordinarily difficult to ascertain and characterize the molecular, mechanistic underpinnings guiding these processes and allowing the often complicated structures to form.

Notwithstanding these challenges, the fact that an extremely complex organism evolves out of a single cell or a seemingly homogeneous group of cells is very intriguing, and it is hardly surprising that the biological and chemical study of morphogenesis eventually coalesced with mathematical—and later computational—approaches that attempted to distill the essence of pattern formation out of the overall complex developmental process. While the first mathematical approaches relied on simple diffusion gradients and biochemical reactions, the emergence of unprecedented computer power and its wide accessibility increasingly permitted more complicated and realistic simulation studies, which have culminated by now in sophisticated agent-based models (ABMs). These models are uniquely qualified for spatially and functionally representing the complexity of a system that is the collective result to a multiplicity of well-timed, fine-tuned cues. This review summarizes the development of morphogenetic models from

relatively simple reaction-diffusion models to today's complex ABMs and places particular emphasis on proliferation, migration, and differentiation as the main mechanisms for pattern formation.

The history of morphogenetic observations and investigations goes back a long time, but theory-based explanations were not proposed until the 20<sup>th</sup> Century. A landmark was D'Arcy Thompson's work *On Growth and Form* [13], in which he described similarities between mechanical and physical systems and the shapes of biological organisms. Due to severe limitations with respect to both theoretical analysis and experimental validation, his observations and calculations were purely hypothetical, as he freely acknowledged. Nevertheless, they marked the beginning of an illustrious scientific development. A decade later, Alan Turing proposed in his treatise *The Chemical Basis of Morphogenesis* [14] a mechanistic explanation that dominated the field for several decades. The core concept of this theoretical explanation was the by now widely accepted *reaction-diffusion* (RD) mechanism, in which, under the right conditions, a two-molecule reaction system is capable of producing periodic patterning through diffusion instability. Specifically, a fast-diffusing *global inhibitor* interacts with a slow-diffusing *local activator*, and their functional coupling can be shown to exhibit non-linear reaction dynamics that can generate repetitive patterns, such as spots or stripes. For instance, the inhibitor prevents features, such as hair follicles from forming too close to each other, an important and wide-spread effect sometimes called *lateral inhibition* [15].

The RD patterns produced by the inhibitor and activator gradients can be considered chemical pre-patterns that act as templates for future differentiation. Thus, the apparent initial homogeneity of an egg or cell cluster morphs into spatially distinct profiles of "invisible" high and low concentration regions, which later guide the implementation of cellular fate decisions and the emergence of visible shapes and forms. Importantly for the field of computational morphogenesis, Turing's RD mechanism demonstrated that it is feasible to represent morphogenetic patterns using a simple, biochemically plausible



system governed by simple mathematical rules. Alas, while the theory was conceptually convincing, it did not gain significant traction in the field for almost five decades.

In another landmark publication, two decades after Turing's proposal, Lewis Wolpert introduced the conceptual framework of *positional information* (PI) as a mechanism of pattern formation during morphogenesis [16]. This theoretical framework was inspired by old observations during the morphogenesis of sea urchins, which Hans Driesch had made as early as 1891 [17]. According to the tenets of PI, a cell is able to determine its assigned fate from its position relative to other parts of the organism [18, 19]. The position, in turn, is characterized by the concentration of a morphogen. Thus, a cell senses its positional value by interpreting a morphogen concentration and makes a fate decision based on this local information. It is assumed that the cell interprets the position based on its genetic make-up and its developmental history, but a central claim of PI is that there is no pre-pattern in the embryo. A reasonable biological implementation of positional information could be a morphogen source leading to a spatial morphogen gradient that gradually decreases with the distance from the source, thereby providing positional information. The concept of PI is quite intuitive and was quickly accepted in the field, partly because experimental findings corroborated its existence [20]. For instance, numerous experiments, especially in the field of limb development, regeneration and transplantation, clearly suggested that cells indeed possess characteristic information regarding their position, which may be acquired through dedicated regulatory programs involving genes such as *sonic hedgehog*, *hunchback* and *Hox* (e.g., [21]).

In spite of the intuitive appeal of PI and intense research over several decades, it remains to be unclear even today how the necessary gradient is established, how the cell senses it, and how a cell correctly interprets it. Diffusion comes to mind, but diffusion processes are not particularly reliable, precise or robust toward external perturbations, and one must wonder how interactions between morphogens and their environment would be realized in terms of effective molecular events [22]. Furthermore, while sensing of a

morphogen by a cell is easy to imagine in principle, the cells along a gradient would have to be able to distinguish very subtle concentration differentials [23]. As Wolpert [19] himself recently stated: “There is no good evidence for the quantitative aspects of any of the proposed gradients and details of how they are set up.” Thus, to summarize PI, many experiments have convincingly suggested that positional information exists, but it is not clear how it is implemented in living organisms. Finally, one should note that PI and RD are not mutually exclusive but in fact complementary (*e.g.*, [24]).

The advent of the digital computer, and its eventual widespread availability, drastically accelerated research on morphogenesis. Computational modeling permitted additional complexity to be considered with minimal effort and effectively decreased the limitations caused by mathematical tractability. In particular, the RD mechanism experienced burgeoning interest [25-27], and the Turing paradigm was successfully put to the test with sophisticated differential equation modeling and improved experimental techniques. Ultimately, research in recent years has substantiated the role of RD mechanisms alongside PI mechanisms during development [28-31], despite their slow start. In addition to the two prominent morphogen-based mechanisms, newer approaches to understanding morphogenesis with computational means have also taken mechanical, electrical, and environmental cues into account [22, 32-35]. However, similar to the treatment of chemical mechanisms, both mechanical and electrical cues have typically been investigated independently. This is a clear shortcoming, as these factors often act simultaneously and presumably synergistically. Thus, in moving forward, it seems wise to pay attention to the multiscale nature of morphogenesis and to the integration of the diverse signals that control it.

## 2.2.1 Mathematical and Computational Approaches to Morphogenesis

### 2.2.1.1 Reaction-Diffusion Strategies

The original reaction-diffusion system proposed by Turing was comprised of two partial differential equations (PDEs). In generalized format, these equations read

$$\frac{\partial A}{\partial t} = D_A \nabla^2 A + F(A, B),$$

$$\frac{\partial B}{\partial t} = D_B \nabla^2 B + G(A, B),$$

where the  $\nabla$ -terms represent diffusion and  $F$  and  $G$  are the nonlinear functions representing the reaction kinetics. For this system to achieve diffusion driven instability, the nonlinearity of the reactions is obligatory, and the diffusivities between the two species must be different. Various particular reaction formats have been analyzed as pattern generators using the RD paradigm. The first examples satisfying these requirements were proposed several decades ago, for instance, by Gierer and Meinhardt [36] and by Schnakenberg [37]. At the time of these studies, computers did not have the bandwidth or speed for realistic simulation studies, which mandated mathematical solution by hand, which quickly became very laborious, given the sensitivity of the RD mechanism to parameter values and initial conditions. An important “trick” toward solving the equations was the transformation, or scaling, of the equations into a non-dimensional format. The resulting dimensionless equations maintain continuity between diffusion and reaction terms of both species, but retain the feature of being quite general with respect to parameter values and also simplify the visualization of the admissible parameter space. Murray provided solutions for this set of non-dimensional reaction-diffusion systems and describes a methodology to isolate the *Turing space*, which consists of the range of parameter values that successfully generate patterns [38].

Structurally, the RD approach has not changed much since its inception: It is still mostly driven by two interacting molecules and uses PDEs, inspired by continuum mechanics, on a predefined 2D grid. A few cases have extended the physical space to 3D-

domains to represent morphological features more realistically, while others have resorted to growing domains that can account for expanding cell populations [39-42]. A notable variation of the generic model is the inclusion of additional species and also of immobile environmental factors. This extension broadens the Turing space of the corresponding two-molecule system while simultaneously removing the requirement of differential diffusivities [43]. The potential of immobile factors for enhancing RD pattern formation had been suggested previously [44, 45], but was particularly highlighted by Macron *et al.* [43] who introduced a sophisticated automated mathematical analysis that directly identified these network topologies along with their respective parameter constraints.

#### 2.2.1.2 Positional Information

In contrast to the RD mechanism, there is no universal mathematical representation of positional information. The reason is that PI is a phenomenological concept describing the capacity of a cell to interpret its location from its immediate environment [18]. The original theory focused solely on the interpretation of morphogen signals for differentiation by defining threshold values that separate the concentration ranges associated with specified fates. The premise of PI, that cells can ascertain their location within developing tissue, has been fundamental to the current understanding of morphogenetic patterns and confirmed in experiments, as discussed before. However, no formalism has been established and generally accepted that can be used to verify PI, and PI has therefore mostly been applied as a qualitative descriptor [46, 47]. Some developing systems seem to suggest that the time spent by a cell in a certain position is critical [48]. Specifically, if a signal is restricted to the front end of a growing domain, the duration of exposure to that signal can be inferred as PI. Taken together, PI has not been strictly tethered to morphogen gradients but instead has encompassed a range of mechanisms for designating fate decisions via cell positions.

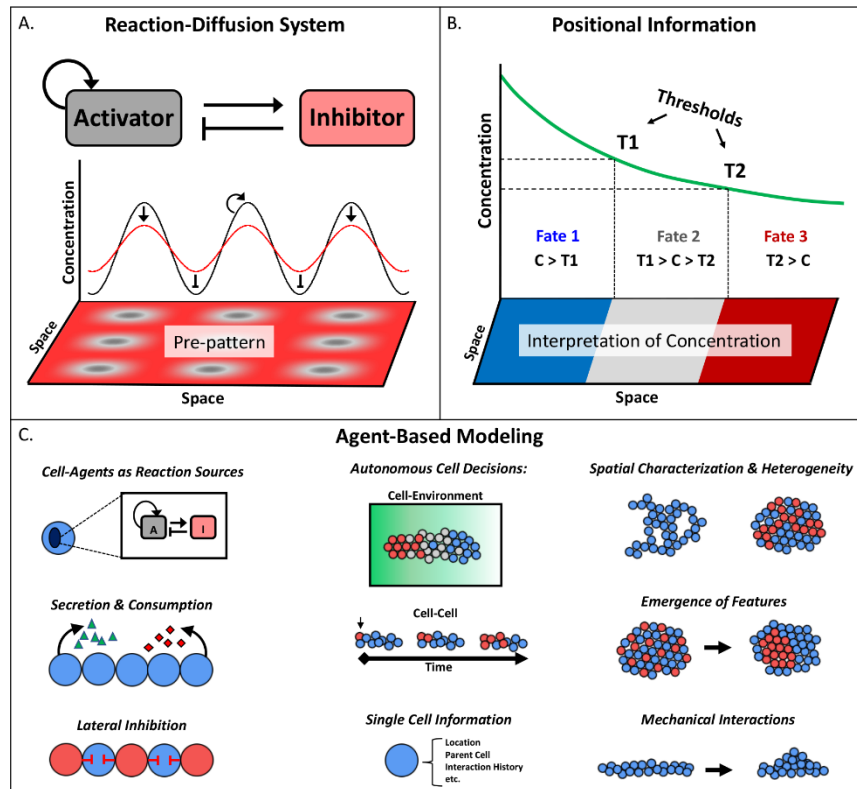
For instance, there has been a recurring theme of local cell-cell interactions contributing to the acquisition of PI. These interactions could be mediators of PI with varying degrees of influence. At one end of the spectrum, cell-cell interactions were used as noise modulators, providing a spatial averaging effect that enhanced detection of morphogen signaling [49, 50]. At the other end, patterning was derived from a localized, secreted factor affecting cell-cell interactions, without evoking the use of global positional information from a morphogen [51]. To what degree cell-cell interactions play an absolutely necessary role in the interpretation of PI is unclear, but the fact that such interactions appear to play a role in various systems suggests that computational approaches are likely to benefit from an implementation at a cellular resolution.

#### 2.2.1.3 Agent-Based Models

RD models employ relatively simple PDEs, which are sufficient for generic analyses of morphogen gradients and their interactions. However, as soon as the target of morphogenetic modeling is a more realistic, complex space, such as a cell, a blastula, or a developing limb bud, the use of PDEs becomes cumbersome. In fact, it seems almost impossible for analytic PDE models to account for genuinely heterogeneous milieus, in which cells or organisms develop.

An alternative that emerged for such purposes over the past three or four decades is the use of *agent-based models* (ABMs). The main ingredients of ABMs are autonomous *agents* that represent some type of entity, such as a molecule, cell or person. An agent can, in theory, be anything; generically, it is a definable, active or responsive element. The dynamics of an agent may consist of a variety of actions, which must adhere to *rules* that govern the behavior of the agent. All agents independently make rule-based decisions as they interact with other agents or their simulated environment. Usually these decisions are probability-based as described in later sections below. ABMs are extremely flexible in their execution and provide an exceptionally powerful framework for the creation of multiscale

models. In particular, they permit the integration of processes across time scales. By their nature, ABMs are particularly adept at capturing spatial phenomena. For example, in a RD system modeled with PDEs, an identical set of equations is solved at each point of an equally spaced grid. Since the reactions are usually cell-mediated, it is implied that each grid point is a cell. This rigidity makes it a difficult task to define a system of equations that account for heterogeneity in a cell distribution, heterogeneous environmental features, or the inclusion of more than one phenotypic state. In contrast, biologically based ABMs easily represent the system at the resolution of the cell. The properties and history of each cell can be monitored, multiple cell states are readily implemented, and movement does not have to be constrained to a grid. This flexibility of ABMs permits unlimited options for exploring the dynamics of a complex system. Of importance here is that the ABM framework is very well equipped to simulate morphogenetic events in molecular and cellular detail, and to investigate the role of PI in developmental cell fate decisions. The following sections first review generic features of ABMs and then focus on various properties and capabilities that pertain specifically to morphogenesis.



**Figure 1: Overview of the two most popular theories for pattern formation during morphogenesis - reaction-diffusion (RD) systems and positional information (PI) - as well as common features of agent-based models (ABMs) for morphogenesis**

**(A) Archetypal Turing-RD system with an activator and inhibitor generating repetitive patterns from differential diffusivities and non-linear reaction terms.** The RD system depends on the concept of a chemical pre-pattern developing in advance of cell-fate decision, and emphasizes the ability to induce pattern formation from an allegedly homogeneous initial state. **(B) Wolpert's PI theory proposes an interpretation step based on concentration thresholds that alleviates the need for a morphogenetic pattern to match the chemical pre-pattern.** In PI, a cell is capable of multiple fate decisions from a single molecular gradient by discerning subtle variances in concentration along the gradient. **(C) AB modeling provides a framework capable of implementing features from both theories.** Cell-agents can act as the sources of activators and inhibitors, permit localized reactions, and make autonomous decisions in response to

**their local environment. In addition to the generation of static patterns, AB modeling allows for the investigation of dynamic, spatio-temporal patterning.**

#### *2.2.1.3.1 Generic Features of Agent-Based Models*

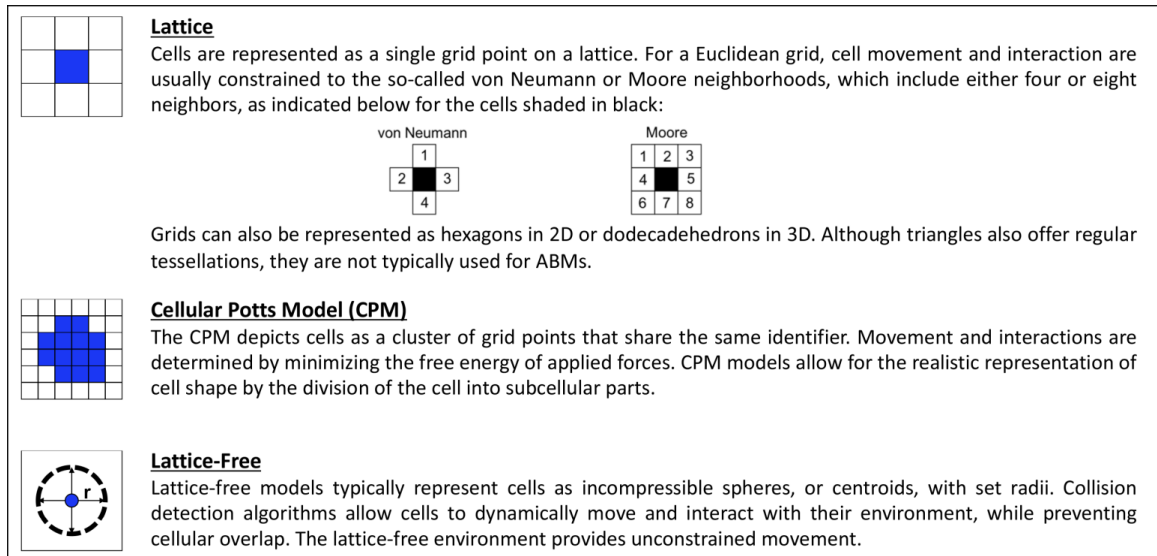
The dynamic system represented by an ABM is a collection of constituent parts and rules. For example, in an immunological model, the agents might be various T-cells, B-cells, and cytokines. In general, the designation of agents is mostly a question of whether the selected parts are capable of replicating the behavior and/or structures of the examined system. This question is directly related to the dominant size scale of the system, which in turn tends to be tied to the time scale of the system. Agents can be portrayed as computational entities in a variety of ways, and the choice of representations is important as it can affect interactions with other agents and movements throughout the environment. Typical agent representations for cell-based models are illustrated in Figure 2. In the simplest models, agents are simply grid points on a two-dimensional (2D) or three-dimensional (3D) lattice that have defined properties. Movement occurs only to neighboring grid points. The somewhat more complex Cellular Potts Model (CPM) allows agents to consist of often irregularly shaped clusters of grid points that share the same or similar properties. The movement of such a cluster is governed by forces and energy. Some of the newer, more flexible models do not rely on lattices anymore. A cell is often represented as a compressible sphere or a *centroid*, which corresponds to the geometric center of a shape, and a sophisticated algorithm determines movements while avoiding the overlapping of agents.

The rules governing the dynamics of the agents can vary in their degree of abstraction, both across and within models, but within a single model there is usually a focus on either mechanistic or phenomenological descriptions. For instance, a rule may dictate what happens if two agents encounter each other. Such mechanistic rules are



established based on empirical data or mathematical representations. An example of the latter might be a random walk model to describe the motion of a molecule by diffusion. In contrast, phenomenological rules are utilized either to mimic an observed behavior or to condense a cluster of mechanistic events into a single representative behavior. For example, cell migration involves a cascade of intracellular signaling, cytoskeleton remodeling, and numerous adhesion forces, but is typically modeled as a direct movement response [52]. ABMs that predominantly consist of mechanistic as opposed to phenomenological rules are deemed “bottom-up” as opposed to “top-down” models, respectively. In either case, additional rules can be defined in a manner that the system is able to learn and evolve as a function of past behavior. Thus, agents can become “smarter” in their actions and may exhibit new behaviors as a result of past behaviors and experience.

The agents of an ABM typically move in an artificial environment that represents the biological space of interest in a simplified manner. The simulated environment of an ABM can be as simple or as complex as the research questions demand or the modeler chooses. In the simplest cases, agents move and interact on a 2D grid space within a plane. They may also diffuse through a 3D grid representing a fluid or viscous solution. The environment may be static and passive, in which case the agents interact only with each other, or it can dynamic, in which case the environment can have an effect on the actions of the agents and/or the agents can change the environment. An example is a predator-prey system, where the environment could be constant or provide the prey with a feed source that dynamically grows and is being grazed up.



**Figure 2: The three most common physical representations of cells in ABMs. Lattice models are generally the least complex, given the constraints to movement and interaction, while both CPM and Lattice-Free models possess varying degrees of complexity depending on the features included in the model. Miniature depictions of the respective images will be placed beside each subsection discussing morphogenetic ABMs in later sections to designate the**

Any additional complexity in model design requires extra rules, definitions, and caveats. For instance, in an interactive environment, constant updates are necessary to resolve interactions with agents. Movements and interactions in a 3D system require more attention to the physical properties of the system. Clearly, the computational cost rises significantly for 3D simulations, especially if the entire space needs to be assessed at every iteration of the simulation. As a partial remedy, it is often beneficial to design an ABM in a phased manner: One might begin with a representation of only the most essential features, and later add complexity step by step, as soon as the dynamics and the repertoire of possible behaviors of the current model becomes clear.

A hallmark feature of ABMs is their potential to (re)produce emergent behavior. Expressed differently, the collective decisions of independent agents within a multi-agent

simulation may result in realistic complexity and behaviors that are not predictable from the governing rules. In effect, the system in such a case exhibits a synergistic response that emerges from the cumulative interactions of its various components. This emergent behavior is not always intuitive or even explainable *a priori*, especially when the rules are intricate or adaptive. In these cases, simulations are often the only means for identifying and characterizing emergent phenomena within the system [53].

Other benefits of ABMs derive from the realistic nature of autonomous agents and their flexibility, which imposes very few limitations. The autonomy is very important for the definition of rules, because autonomous agents control their own behavior and react to local factors without needing to know the global environment. At the same time, the autonomous agent representation permits heterogeneity, both with respect to agents and rules. The flexibility of ABMs is immense: agent behavior, types of agents, agent interactions, agent adaptation, variability in agent scales, stochasticity, and environmental factors are all unlimited in scope and easily manipulated through corresponding settings of parameter values.

### **2.2.2 Agent-Based Models for Morphogenesis**

One hallmark of morphogenesis is the coexistence of different types of cells, such as stem cells and differentiated cells. ABMs for morphogenesis use different agents to represent such cells, or various other phenotypes, and study how they organize into the targeted patterns. ABM models for morphogenesis fall coarsely into three methodologies for detailing interactions that make use of the three agent representations depicted in Figure 2, namely: they may be points on a rigid 2D or 3D lattice, clusters according to a Cellular Potts Model (CPM), or entities that operate without a lattice. The lattice-based representation is the simplest and computationally cheapest; not surprisingly, it is also the most limited in terms of mechanistic features and details of cell movement. The CPM is the most flexible for representing irregular cell shapes and movements, and it is also the

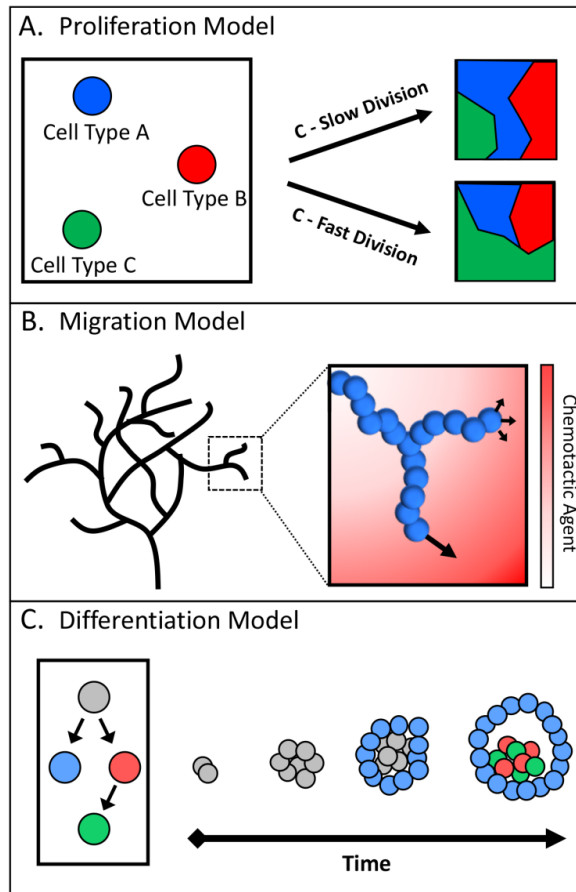
most computationally intensive model, as it requires forces to be converted to energy. The lattice-free models are well-suited for generating 3D organizations and emulating deformations, but require more up-front effort for defining realistic details of collision and movement events. As so often is the case in modeling, the “best” representation depends on the questions to be asked, the context of the model, and the availability of data.

The specific models to be discussed in this review have been organized into three main categories that denote the predominant mechanism for generating the targeted morphology or type of anticipated pattern in each model: Proliferation, migration, and differentiation (Figure 3). As a caveat to this categorization one must note that there is certainly overlap and that models frequently employ two or all three mechanisms.

Proliferation models place their main focus on *differential division rates* and on *volume exclusion* to produce patterns (Figure 3a). Furthermore, the patterns are frequently defined by the organization of cell types relative to each other, *i.e.*, the clustering of Cell Type X surrounded by Cell Type Y. To generate discernible patterns, these models typically require at least two distinct cell types and/or a method for identifying unique populations. The discussion of proliferation in the next section is divided into subsections based on biological mechanisms that can affect morphological patterning.

Migration models can be classified by the inclusion of *directed cellular movement* and their emphasis on the gain of *cell polarity*. A further criterion is the *desired morphological pattern* of these models in comparison to other patterns. Specifically, the shape and growth of the entire network or system is often more important than the relative location of cell types within the network (Figure 3b). For example, when assessing the formation of the vascular network, one might use the degree of branching and interconnected vessels as a metric, rather than the arrangement of cells within each vessel. The discussion of migration models is therefore split into two subsections representing extracellular factors that can induce migration.

Differentiation models consider systems that are derived from a single cell type but *produce patterns or morphologies* that contain two or more cell phenotypes. In all cases but one, the system is initially homogeneous with respect to phenotype and gradually gains heterogeneity from differentiation events (Figure 3c). The exception to this strategy is an initialization with heterogeneity, which subsequently requires differentiation into multiple lineages to maintain the desired morphology. Therefore, patterns from this class are described by the relative organization of cell types, which all originate from a stem cell. The subsections highlight two differentiation events that are integral for development: the loss of pluripotency and the gain of heterogeneity in tissue.



**Figure 3: Classification used to categorize three types of agent-based models for morphogenesis. Each model class describes the primary mechanism that induces pattern formation. (A) Proliferation models depend on the differential division of cells, either**

**between cell types with variable growth rates, between cell generations, or as an alternative to cell death. The illustration depicts volume exclusion of Cell Types A and B when Cell Type C has a fast division rate. (B) Migration models concentrate on directional movement, both in a purely migrational sense and in terms of polarized growth. The patterning of migration models is typically centered on overall morphology, such as the branched network of a vascular network. (C) The focus of differentiation models is on patterning that is reliant on fate change, in most cases following a hierarchy of lineage commitments. The interactions between the various cell types are major determinants for the resultant behavior.**

If an ABM is designed in a phased manner, the model complexity increases with each sequential phase. As a case in point, the classification into proliferation, migration, and differentiation models itself follows a sequential increase in complexity due to the way cell-cell interactions are involved in these events and how they are encoded. Proliferation models place their main attention on individual cell behavior (division, apoptosis, *etc.*). Therefore, the most complicated rules that are to be defined only affect the agent performing the rule. Such rules governing individual agent behavior are the easiest to implement because they are self-contained and straightforward to interpret. Migration models account for two additional aspects, namely the sensing of environmental factors and appropriate responses to such factors. As a consequence, rules are required that describe how an environmental signal is perceived and converted into movement, and how this movement is coordinated with neighboring cells. Cell-cell and cell-environment interactions both involve repeated activity updates and, for migration, a balance between adhesion and motility. Differentiation models typically combine the features of proliferation and migration models and additionally introduce differentiation mechanisms which are usually regulated by environmental factors and/or more complex cell-cell interactions. Thus, the level of detail increases from one section to the next.

### 2.2.2.1 ABMs Focused on Proliferation

The key feature of ABMs in this category is their capability to specify unique growth rates for each cell type within a simulation. In a developmental context, this variability in proliferation can occur through numerous mechanisms, but the most archetypal instances involve inherent differences in division times between cell types. As a case in point, consider the transition from a pluripotent stem cell to a multipotent progenitor: the genetic distance between the two cell types is relatively short but the difference between their cell cycle lengths is significant [54]. This fluidity in growth rate as a function of phenotype creates a vast potential for unique spatial organizations that form dynamically over time, even if one only accounts for cell division and the original spatial orientation of each cell type. ABMs provide an optimal mechanism for investigating the effects of proliferation and cellular densities or arrangements during morphological processes.

#### 2.2.2.1.1 *Generational Patterning*

The first class of proliferation ABMs makes heavy use of *cellular invasion waves*, which arguably provide the best example for illustrating the capacity of proliferation-based patterning. The phrase ‘cellular invasion’ undoubtedly brings to mind tumorigenesis, but cellular invasion is also an integral part of normal, physiological processes, such as wound healing and morphogenesis [55]. Invasion waves contain two actions, namely cell migration (the invasion), which is followed by proliferation of those cells (the wave). Of particular importance here is the ‘wave’ that occurs after the ‘invasion’.

As a specific example, the development of the enteric nervous system (ENS) involves the migration of enteric neural crest cells to the foregut where they proceed as an ‘invading wave’ that colonizes the entire gastrointestinal tract [56-58]. Initially, it seems intuitive that the invading wave of cells will maintain an equal composition of progeny from the original population as it grows. However, clonal dominance during ENS

colonization has been observed [59, 60]. Experimentally, a single GFP-labelled ENC cell was added to a population of (~8,000) unlabeled ENC cells and fused with gut tissue before the start of colonization [60]. After colonization, large variability was observed in the number of GFP-positive cells contributing to the ENS across experiments, and in one case one-third of the entire ENS was composed of GFP-positive cells. To characterize this appearance of clonal dominance in a population of phenotypically identical cells, a cell invasion ABM was devised that was capable of tracking cell lineage and generation [60]. The agents were placed on a 2D grid with equal growth rates and with equal ability to move stochastically (in the form of a random walk), but only if grid space was available. It is easy to imagine that a cell with a faster division time than others within the initial population would have a competitive advantage and ultimately dominate. However, in this case all cells were assigned exactly the same division time. Intriguingly, tracking cell generations within this model revealed that 50% of the final population could be attributed to the progeny of only a few cells ('superstars') from the initial population. In other words, despite the identical division time for each agent, a few cells were able to monopolize the available space and create a spatially distinct distribution of their progeny. Analysis of the model mechanics yielded the following: (1) location within the initial population affects the probability of a cell becoming a superstar but it is not the sole contributing factor; and (2) stochastic competition via volume exclusion is a sufficient mechanism for instigating clonal dominance. Specifically, an accumulation of stochastic movement and daughter cell placement events allows the progeny of a single cell to preclude the growth of other cells by preventing access to grid space. This competitive advantage is reflected in many proliferation models that demonstrate domination through volume exclusion [61-66].

#### *2.2.2.1.2 Models of Balanced Growth and Death: Apoptosis and Homeostasis*

Apoptotic events are integral and necessary for shaping tissue during development; two excellent examples are the formation of limb buds and the blastocyst [67, 68].



Combined with proliferation, the balance between cell-specific apoptosis and proliferation allows organisms to maintain morphological homeostasis while at the same time enabling the formation of heterogeneously organized structures.

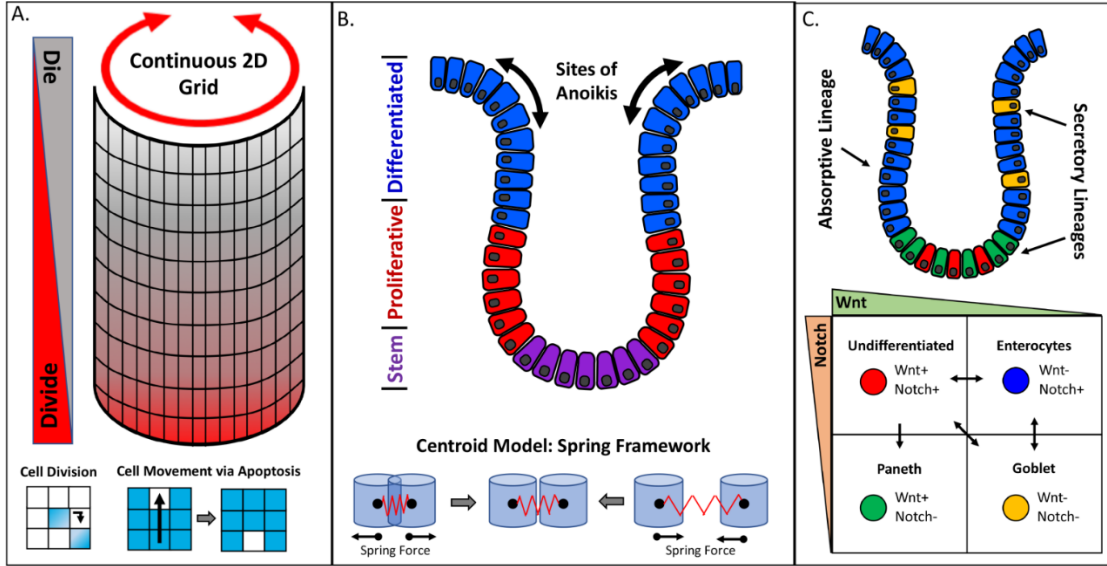
A popular model system that is highly dependent on apoptotic signals is the formation of small cavities, called *acini*, in the mammary gland during epithelial morphogenesis of this gland. Several ABMs have investigated acini formation [69-71], but the most recent model stands out because it was implemented on a three-dimensional grid [72]. 2D model representations of actual 3D systems are informative, but a higher level of abstraction is required for implementing the projection, which often makes model results more difficult to interpret. A good example for this situation is the comparison of proliferation rates: One can easily show that a 2D model requires a slower rate of proliferation than a 3D model to create the same cross-sectional structure. As a thought experiment demonstrating this difference, imagine either a circle or a sphere full of cells. In the former, the maximum number of cells is proportional to  $\text{Circle-radius}^2/\text{Cell-radius}^2$ , while for the latter it is  $\text{Sphere-radius}^3/\text{Cell-radius}^3$ . As a consequence of the different powers, the same cell doubling time allows the circle to grow faster than the sphere.

The recent 3D acini ABM [72] begins with a single cell type that expands into the external basement membrane with a designated proliferation potential. As the acinus grows, cells that are not adjacent to the basement membrane receive a signal that triggers apoptosis with a set probability. As cells die over the course of a simulation, the space they occupied becomes part of the acinus lumen. Of particular note from a modeling point of view is that a minimalistic rule set is sufficient to capture the dynamics of normal and aberrant acini morphologies. This rule set slightly modulates the balance between proliferation and apoptosis. It is also important to mention that this delicate growth-death balance only leads to the correct production and maintenance of patterning if two requirements are met: (1) The signals must be relatively equal in magnitude at the

population level; and (2) at least one of the signals must be effective in a spatially distinct manner.

The colon crypt is another system that establishes a dynamic steady state between cellular growth and death. The bottom of the crypt is inhabited by a small number of adult stem cells, identifiable by the *Lgr5* marker gene [73], that give rise to a large population of proliferative cells. These proliferative cells both self-renew and produce terminally differentiated cells that make up the majority of the crypt. The patterning capability, which is based on high turnover rates of cells within the intestine, is probably one reason that intestinal and colon crypts have been a popular target system for ABMs. Furthermore, monoclonal conversion, which is synonymous with clonal dominance, is commonly found in crypt systems [74].

Several of the following sections cover a variety of crypt models that illustrate how the same phenomenon may be approached with different types of representations governing the mechanics within the system. Figure 4 contrasts the representations of some of these models. The first crypt ABM was designed on a simple 2D lattice, with cell growth and death defined probabilistically as functions of two pre-set gradients [75]. The two gradients in this model (termed *Divide* and *Die*) are not associated with specific molecules but instead used to provide positional information. The probability of a cell transitioning from a quiescent cell to a proliferating cell and then to a terminal cell increases as a cell moves down the *Divide* gradient. The *Die* gradient runs in the opposite direction, with the highest probability of cell death occurring at the top of the crypt. Cell movement up the column is defined as a function of cell death, with cells moving upwards to fill any unoccupied space. Similar to the mammary gland acini model, discussed before, the growth and death signals are balanced to maintain crypt size but are also spatially distinct to preserve the cellular organization.



**Figure 4: Comparison of three crypt model implementations and agent descriptions. (A)** Cells are defined on a continuous 2D lattice such that a cell moving off the right edge of the grid reappears on the left edge. Divide and Die gradients are used to describe the behavior of different cell states in the crypt while movement up the crypt is the result of apoptosis. **(B)** A centroid model may be employed to investigate the role of crypt geometry on the location of anoikis within a crypt. The proliferative state of a cell is defined by pre-specified regions along the crypt and cell-death is solely dependent on the occurrence of anoikis. **(C)** Centroid model that includes differentiation and dedifferentiation between the four main phenotypes present in an intestinal crypt. Wnt signaling is defined by position within the crypt while Notch signaling is determined by the phenotype of the cell and its neighbors.

#### 2.2.2.1.3 Models Accounting for External Proliferation Cues

Numerous diffusible factors in the extracellular environment can influence the proliferation of a cell. These factors range from something as simple as the availability of nutrients to growth-specific proteins that are aptly termed *growth factors*. Indeed, many growth factors have been classified as morphogens due to their ability to promote proliferation and modulate differentiation potential during development [76, 77]. Many examples of ABMs rely on such morphogens.

The development of the genital tubercle (GT) provides a clear example for the critical role of morphogens. GT development is identical in males and females until stage E15.5 of embryogenesis when the male GT is exposed to androgens. The androgen signal promotes proliferation of both mesenchymal and endodermal cell types in the GT, resulting in sexual dimorphism. An ABM of GT morphogenesis was implemented in the software CompuCell3D using a cellular Potts model that captured adequate movement and interaction dynamics [26, 78]. A simple signaling network was integrated into the model to account for the three principal morphogens that direct GT development: SHH, FGF10, and androgens. SHH and FGF10 induce growth in mesenchymal and endodermal cells, respectively, while androgens enhance sensitivity of urethral plate endoderm and preputial mesenchyme to FGF10. In the model, SHH was secreted by endodermal cells at a constant rate, thus stimulating the secretion of FGF10 in mesenchymal cells in a concentration-dependent manner. Androgens were assumed to permeate the system at a set time point (E15.5), causing an instantaneous change to the FGF10 sensitivity in all affected cells. The model was initialized with an idealized geometry of the tubercle at E13.5, containing five distinct cell types. Simulations using this model were able to recapitulate the sexual dimorphism that occurs during GT development by modulating proliferation in response to androgen signaling.

During puberty, the mammary gland undergoes extensive proliferation and ductal morphogenesis [79]. Previous studies had shown that exposure to ionizing radiation before or during puberty significantly increases the risk of developing breast cancer in women by causing an increase in the mammary stem cell population [80]. To test the mechanisms eliciting the morphogenetic changes observed after puberty, an *in vitro* and an *in silico* ABM model were defined [62]. The *in vitro* model identified TGF $\beta$  as a major activator of growth during ductal morphogenesis while the *in silico* model addressed the mechanism by which radiation could modulate the TGF $\beta$ -induced growth to produce a larger stem cell population. The ABM was contained on a 2D grid and included three cell types: bi-potent

progenitor cells, terminal basal cells, and terminal luminal cells. The progenitors had the option to divide symmetrically or asymmetrically, producing two stem cells or one stem cell and one basal or luminal cell, respectively. Furthermore, a dedifferentiation mechanism allowed terminal cells to convert back to bi-potent progenitors. The model parameters associated with the probabilities that these events occurred were fit using experimentally determined distributions of the cell types as the objective metric. These experimental distributions were obtained from the *in vitro* model using TGF $\beta$  to stimulate growth, with and without irradiation. By comparing the calculated best-fit parameters between these conditions, the likely mechanism precipitating the increased stem cell population was identified as enhanced self-renewal in response to TGF $\beta$ .

Interestingly, while differentiation was an important component of this model, it played a less vital role than the proliferation rates for yielding the desired patterning. Specifically, the majority of daughter cells actually originated from a parent of the same phenotype rather than via differentiation of a progenitor cell. This situation arose because differentiation in the model was coupled to cell division, with the consequence that the rate of differentiation could never surpass the rate of cell growth. Furthermore, because growth could only occur where space was available, the growth of progenitor cells was quickly hindered by other cell types through competition due to volume exclusion.

#### 2.2.2.1.4 *Models of Growth Directed through Mechanical Forces*

Mechanical forces are involved in nearly every aspect of morphogenesis, albeit with varying degrees of influence. At the macroscopic scale, mechanical forces can shape and organize tissue in response to the cumulative effect of numerous forces at the microscopic scale [81]. In particular, cell adhesion molecules allow cells to bind together and to their environment, and these adhesion forces affect how the population expands outwards as cells divide.

Up to this point, the ABMs discussed have been constrained to lattices and, in many cases, division was only allowed to occur when space was available. While lattice structures are useful and computationally efficient, they offer limited spatial resolution and lack the ability to describe intercellular forces. Furthermore, the division constraints are contextual and require scrutiny: a constraint could represent contact inhibition or quiescence, but it also frequently creates an artificial scenario where mitosis is restricted to the edge of a cell cluster. Centroid models overcome these limitations by allowing lattice-free movement. A centroid model represents each cell as a central point that is connected to neighboring cells by stiff springs. If a cell comes within the confines of another cell, whether by movement or division, a collision occurs. The collision produces a force that moves both cells to prevent overlap and, if necessary, propagates this movement through the population.

An effective demonstration of a centroid model is again a colon crypt model [82]. This particular ABM sought to probe the mechanisms behind crypt *anoikis*, *i.e.*, programmed cell death in response to detachment from the basement membrane. In this ABM, cellular agents grow and move freely along a 3D rendering of the crypt membrane, experiencing cell-cell collisions and an attachment force to the membrane. In the biological system, anoikis typically occurs at the top of the crypt, and the authors hypothesized that the crypt geometry and attachment forces were the crucial factors for simulating the process. Intriguingly, by modulating the attachment forces, the localization of anoikis events could be replicated, and the system was able to self-regulate the rate of anoikis, thereby maintaining homeostasis. Results like these highlight the advantage of lattice-free models for investigating behavior that emerges as a function of interaction forces, especially within complex geometries.

When pondering mechanical forces during proliferation, one might also consider the effects of external resistance. It is known that the extracellular matrix (ECM) can act as a scaffold for growth; a pertinent example is the basement membrane of the second crypt

model. The ECM can also act as a boundary or an anchoring point [83]. The stiffness and elasticity of ECM are highly variable, and these properties influence the response of cells and their organization during growth.

An example is adipose tissue, which consists of adipocytes clustered into distinct lobules by ECM, with variable morphologies. Using a lattice-free ABM platform, the self-organization of adipose tissue was investigated with adipose cells modeled as growing spheres and ECM fibers as short lines that could cross-link together. Each ECM fiber had a defined, constant unit strength, with larger strands of cross-linked fibers able to exert more force on neighboring adipose cells through the fiber network. Additionally, the adipose cells were able to exert pressure on the ECM network to prevent compression as the available space for growth decreased. To explore the dynamics of lobule formation, the linking-unlinking frequency ( $\nu_d$ ) of the fibers was investigated, which revealed that adjustments of this quantity could result in three distinctive morphologies. Since modulation of  $\nu_d$  was representative of the degree of ECM restructuring, the three morphologies were regarded as consecutive ‘phases’ that developed over time.

#### 2.2.2.2 ABMs Focused on Migration

Migration events are commonplace throughout development and pivotal for the morphogenesis of numerous tissues [84]. The forces behind cell motility are predominantly mechanical, but the signals that trigger and direct migration can be mechanical, chemical, electrical, or all three. As mentioned previously, cells *in vivo* are experiencing a multitude of forces from the environment and neighboring cells. Furthermore, expansion of a cell population can lead to passive movement of the cells in response to physical interaction effects (collisions, boundaries, *etc.*). In contrast to passive movement, migration involves the active generation of forces by a cell to induce movement, usually in response to an external stimulus. Concurrently, cell polarity accords directionality by establishing a leading or front edge during movement. The ABMs discussed in this section emphasize

two key mechanisms whereby heterogeneous environmental factors can direct and coordinate cell movement.

#### *2.2.2.2.1 Extracellular Matrix Influence on Migration*

Cell migration during development exhibits a collective form of organization whereby cell populations are able to traverse long distances as coordinated groups. The emergence of this behavior requires one or more environmental cues that guide the migration of each cell. One source of directional information for migrating cells is the variable composition of adhesion and signaling proteins within the ECM [85]. The organization of ECM components is dynamically modulated by local cells that can secrete and degrade ECM macromolecules [86]. The resultant remodeling of the ECM can bias migration [87] and even reinforce specific migration paths [88].

The gastrulation in amphibians provides a representative example of collective migration that is dependent on an ECM component: fibronectin (Fn) is essential for mesendoderm cells to migrate as a sheet-like cluster across the inner surface of the blastocoel [89, 90]. As the cells progress, they bind to the ECM and exert traction forces via integrin-Fn interactions, which restructure the ECM in the wake of the first migratory cells [91, 92]. In addition, mesendoderm cells exhibit ‘shingling’ behavior: cells overlap each other and maintain cell-cell adhesion, mediated by cadherins, during movement.

An ABM, reminiscent of a standard CPM, was developed to investigate the capability of these interactions to guide coordinated migration [93]. Specifically, each cell is represented by a 3x3 grid on a 2D lattice of Fn, where the center represents the cell body and its Moore neighborhood the cell’s “edges.” The edges of adjacent cells are allowed to overlap and produce an adherence force that mimics shingling behavior. In the model, the Fn matrix is initialized with randomly distributed concentrations along the x-axis but with a mild gradient along the y-axis. The restructuring of the matrix is modeled by a decrease in Fn concentration each time a cell moves through it. To determine the magnitude and



direction of movement, the Fn gradient is calculated along each edge, with larger forces being generated by larger Fn differentials. A similar force is calculated for each pixel overlapped by another cell, and the net force vector is computed as the sum of both forces. This customized framework is able to achieve cellular movement along the Fn gradient but cannot capture the observed, coordinated sheet-like clustering. In particular, the balance between integrin and cadherin turns out to be disproportionate along the edge of the cell colony. To remedy the situation, the authors introduced an intracellular feedback network based on Wnt/B-catenin signaling, whereby integrin binding triggers cadherin production, which in turn equalizes the integrin and cadherin forces and allows a sheet-like migration to occur.

A second representative example in this category is the development of the neocortex, which involves an interesting pattern of migration from the intermediate zone (IZ) towards the marginal zone (MZ) [94-96]. The neocortex consists of six distinct layers of pyramidal neurons. The migrating cells form a growing cortical plate (CP) between the IZ and MZ, one layer at a time. Cortical layer VI is the first layer of cells to form and is positioned closest to the migration source (IZ), with each sequential layer migrating over previous layers in an inside-out manner and increasing cortical thickness [88]. Reelin is an ECM glycoprotein that is essential for the proper formation of the neocortex [97-99]. In fact, the cortical layers in a Reelin-null mutation mouse (*reeler* mouse) are inverted, with layer VI becoming the most superficial layer [100]. However, the exact function of Reelin and its effect on migration through the CP have been debated [94]. To explore the different hypotheses regarding the role of Reelin during cortical development, a set of migration ABMs was generated with various rule sets to reflect each proposed mechanism [101]. It is unnecessary to describe each individual rule combination but, in general, the following held for these models: each layer of cells was introduced to the system independently and uniquely colored, one row at time; furthermore, cell movement and the conversion to an immotile state were defined as functions of Reelin. Extensive model testing revealed that

it is possible to isolate a rule set that very closely mimics the biological system. This inference of rules was accomplished by comparing the model output (*i.e.*, the colored cell distributions of CP) against the known behavior in *reeler* mutants and in a Reelin-dependent mutant, *Dab1* [102]. Other ABMs involving ECM-mediated migration include [103-105].

#### 2.2.2.2.2 *Models Accounting for Chemotactic Cues*

A widespread mechanism driving morphogenesis is a branching process. Indeed, this process can be found across multiple organ systems, including the lungs, kidneys, and vasculature [106-109]. In each of these tissues, the branched morphology is achieved through the recurrence of three events affecting the bud or vessel: formation, extension, and splitting. The branching morphogenesis in each of these organ systems is sensitive to the spatial distribution of a key morphogen, namely, FGF10, GDNF, and VEGF for lungs, kidneys, and vasculature, respectively. Multiple mechanisms have been proposed for initiating the spatial distribution of each morphogen, including ligand-receptor-based Turing mechanisms for the lung and kidney systems [31, 110]. In these cases, the morphogen induces expression of its receptor which correspondingly increases sensitivity to the morphogen. This receptor-ligand cooperativity creates regions of high receptor density with enhanced morphogen activity that are interpreted as locations for branching events. A more common approach for modeling branching events is the consideration of a morphogen as both an inducer of proliferation and as a chemoattractant. High concentrations of the chemoattractant polarize the tip of a growing bud or vessel and cause proliferation in the respective direction. As the network expands, the vessel/bud cells consume the morphogen and establish a gradient that directs future growth and branching events.

A Cellular Potts model of a ureteric bud in the kidney was compiled to determine the relative influence of mechano-chemical factors on the observed branching morphology

[111]. Specifically, this model assesses the factors affecting a single splitting event within the kidney. Model analysis revealed that the morphology is most strongly influenced by two model parameters representing the strength of chemotaxis and the proliferation rate. In fact, if chemotaxis and proliferation are perfectly balanced, the physiological morphology is achieved, whereas skewing the ratio leads to pathological morphologies.

A different role of branching morphogenesis can be found at the network level. A good example is a recently published 3D model of vasculogenesis that delves deep into the mechanical aspects of these branching processes [112]. In contrast to what one might expect, the morphogen (VEGF) is not implemented in the model as a direct activator of proliferation. Rather, VEGF acts exclusively as a chemoattractant that stimulates migration of vessel tip cells. Nonetheless, proliferation is important in this ABM. It is stimulated through mechanical stretch forces generated by the ‘pull’ force of chemotactically migrating cells, which has indeed been experimentally observed in vascular endothelial cells. The mechanism by which the stretch affects proliferation in these experiments consisted of upregulating the VEGF receptor. The ABM for this system uses a lattice-free, centroid description of two cell types: tip cells and vessel cells. Both cells types are subject to multiple local forces that govern their behavior. Tip cells experience: a chemotactic force along a VEGF gradient; a persistent force that results in the tendency of cells to continue moving in the same direction; an environmental drag force; and interaction forces from neighboring cells. The vessel cells are exposed to: interactions forces; environmental drag forces; and an angular persistence force that stabilizes and corrects possible buckling caused by cell division. Furthermore, mechanical stretch and compression of vessel cells are used to regulate proliferation and sprouting events, respectively. The complex rule set of this ABM is unique compared to other models of branching morphogenesis, which typically assume chemical dominance in extension and splitting events. The topics of angiogenesis and vasculogenesis have been extensively researched and reviewed previously [113], with copious AB models for both phenomena [109, 114-123].

### 2.2.2.3 ABMs Focused on Differentiation

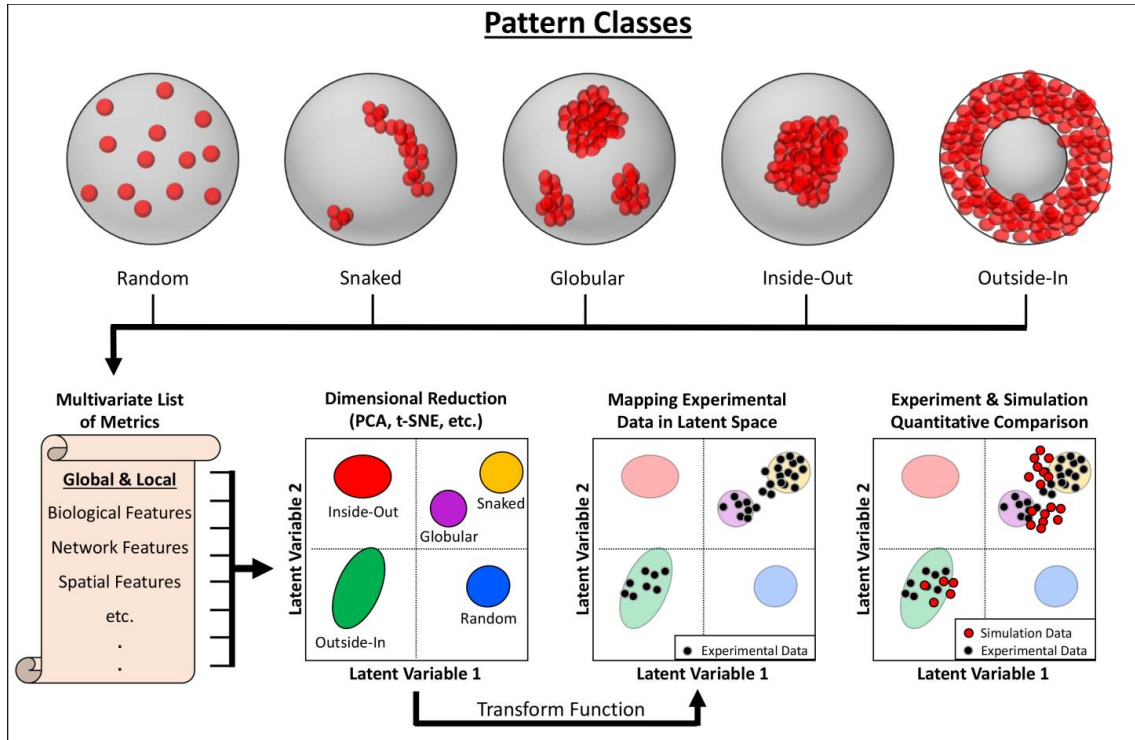
Morphogenesis and differentiation are highly interdependent. All organs and tissues are composed of heterogeneous assemblies of cells, and the acquisition of phenotypic heterogeneity through differentiation often occurs concurrently with the gain of organization. Thus, while differentiation is not as directly involved in shaping tissue as proliferation and migration, it is clearly essential for morphogenetic events. The following describes representative ABMs that focus is on: (1) Initial differentiation from the pluripotent state; and (2) differentiation within tissues.

#### 2.2.2.3.1 *Gain of Organization during Loss of Pluripotency*

The loss of pluripotency at the beginning of development prefaces nearly all instances of morphogenesis. As development proceeds, pluripotent cells differentiate into the three germ layers and eventually every somatic cell type. These initial fate decisions are crucial for embryogenesis and require coordination between the processes that maintain pluripotency or designate cell fate [124]. The transcription factors governing the pluripotent state can be reduced to a ‘core’ network consisting of Oct4, Sox2, and Nanog [125-128]. These key transcription factors dynamically regulate their own expression and have essential roles in directing self-renewal and fate specification [129-132]. The self-renewal aspect is of particular importance for the study of pluripotent cells *in vitro*. Pluripotent cells are only transiently present *in vivo* and they quickly acquire germ layer fates. Therefore, sustaining expression of the key transcription factors associated with self-renewal is necessary for the culturing of pluripotent cells *in vitro*. Typically, this task is accomplished by the addition of exogenous factors such as leukemia inhibitory factor (LIF), which indirectly activates Sox2 and Nanog expression in mouse embryonic stem cells [133, 134]. Since the discovery of means for maintaining pluripotent cells *in vitro* [9, 10, 135, 136], the opposite has become very alluring as well: harnessing the specific organ and tissue forming potential of these cells.

Aggregates of pluripotent embryonic stem cells (ESCs) can serve as powerful *in vitro* platforms for studying morphogenesis and early differentiation events, both experimentally and with ABMs. In a landmark study, White and colleagues found that spontaneous differentiation of these aggregates produces transitional patterns as pluripotency is gradually lost [137, 138]. Specifically, spontaneous differentiation was instigated by culturing the cell aggregates in the absence of LIF. The transition out of the pluripotent state was monitored by collecting representative images of the pluripotency-regulating transcription factor Oct4 expression within the aggregates over the course of differentiation.

To investigate the emergence of patterns, a 3D ABM utilizing the centroid schema was constructed. The ABM simulations considered two cell types (Oct4+ and Oct4-), with the initial population consisting entirely of Oct4+ cells. Three mechanisms for triggering state change were explored: stochastic processes, juxtacrine signaling, and paracrine signaling. In total, seven unique rule sets for governing differentiation were formulated and compared. In general, for both juxtacrine and paracrine signaling the Oct4+ cells were modeled as the source of inhibitory factors while Oct4- cells acted as the source of differentiation activators. To compare the experimental and simulation results, network metrics were extracted from a training set of computationally generated pattern classes. Principal component analysis (PCA) was applied to the multivariate set of metrics calculated from the training set networks and mapped onto a dimensionally-reduced latent space. By applying the same PCA transform to metrics calculated from the experimental and simulation data, spatial patterns of Oct4 expression could be evaluated as a function of proximity to each pattern class within latent space. This strategy facilitated a direct and quantitative comparison between the simulation results and the experimental data, thus identifying the state-change mechanism best able to describe the observed patterning.



**Figure 5: Overview of the analysis of spatial features using a combination of network analysis and dimensional reduction techniques. A set of metrics is calculated or extracted from a series of pattern classes that depict typical cell organizations within the system; here, using the defined pattern classes from [137, 138]. The selected metrics should be equally represented in the simulated and experimental systems. Dimensional reduction techniques allow the multivariate data to be condensed to a few axes, ideally separating the defined pattern classes into distinct regions of the latent space. A transform function trained on the pattern class data can then be applied to metrics calculated from experimental and modeling results, mapping both into latent space and compared to the locations of the pattern classes.**

As with other studies, this model investigation faced a generic challenge of AB modeling, namely the unbiased quantification of results. Here, model validation was approached by assessing population-based and phenotype-specific metrics that were able to quantify spatial features rather than relying on visual comparison. In fact, the specific methodology of this study may serve as a generic means of validating ABMs by

comparison of spatial characteristics, which is particularly important in morphogenetic studies. For example, the same network-based analysis was used to examine the evolution of spatial patterning during cichlid gastrulation *in vivo* [138]. Figure 5 provides an overview of this approach for comparing spatial patterning between simulation and experimental systems.

ABMs have been used to elucidate another aspect of early differentiation *in vivo*, namely the loss of pluripotency and the concomitant morphogenesis of a blastocyst [139]. This process of blastocyst formation is remarkably robust and entirely self-contained, receiving no external maternal information. Rather than attempt to delineate every signal and interaction, the authors of this ABM decided to define overarching rules in a top-down approach. The question was simple: What is a minimalist set of rules that can adequately capture the complexity of blastocyst formation. Extensive exploration of various ABM implementations suggested that four rules, derived from four main regulatory events, were sufficient to recreate the structuring of the blastocyst. The first rule targets polarity at E3.0, thereby establishing the inner cell mass (ICM) with trophectoderm cells along the periphery. The second rule mimics FGF signaling at E3.5 and creates a salt-and-pepper distribution of epiblast (Epi) and primitive endoderm (PrE) cells. The third rule causes lineage segregation of the Epi and PrE cells through differential adhesion. Finally, the fourth rule causes apoptosis of any PrE cells that were not segregated properly from the Epi cells. The model was defined as lattice-free and implemented in 2D, where the success rate for achieving correct blastocyst formation using these four rules was 79%. In a slight modification, the model was also validated in 3D with minor changes to account for differences in the number of nearest neighbors.

Several biological insights related to the FGF/ERK pathway emerged from this work. First, the specification of Epi and PrE cells in a salt-and-pepper pattern resembles a Turing-like mechanism. Second, FGF4 is secreted by ICM and Epi cells and inhibits Nanog. Third, Nanog and Gata6 mutually inhibit each other, which effectively creates a

positive feedback loop outside the steady state. The consequence is a local intracellular amplification mechanism which, in conjunction with a global inhibitor, produces a spot pattern. Furthermore, the initiation of FGF/ERK signaling appears to be invariant to the number of cells within the embryo and is instead dependent on the time since fertilization. As experimental validation, scaling experiments were conducted whereby mouse embryos at the 8-cell stage, before trophectoderm specification, were merged into 16-cell and 24-cell embryos. Blastocyst formation proceeded normally in these double- and triple-embryos, albeit with approximately 2- and 3-fold more final cells, respectively. Of important note, the ratio of PrE to Epi cells was maintained within the ICM despite the scaling, indicating that the patterning mechanism does not rely on—nor is sensitive to—the quantity of cells within the ICM. Experimental evidence of the time-dependence was provided by the response to transient inhibition of ERK: the PrE/Epi ratio decreased in proportion to the duration of inhibition. The model predicts that the time-dependence of these initial fate decisions is due to a necessary accumulation of FGF, or a similar signaling component, from the onset of fertilization to the point of activation.

#### 2.2.2.3.2 *Multi-phenotypic Tissue Models*

As a representative ABM in this category, we return again to crypts, but focus here on intestinal crypts, which share the same physiological structure with the colon crypt but contain an additional cell type, called Paneth cells [140]. In the previous colon crypt models, cell types were not explicitly defined beyond their proliferation capacity. Here, the individual cell types are considered along with their type-specific interactions; they include: undifferentiated (stem-like) cells, secretory-progenitors, secretory-Goblet cells, secretory-Paneth cells, and enterocyte progenitors. The model has some similarities with the second crypt model discussed before in that it is a centroid model where each cell experiences adherence forces to the basement membrane. In addition, the model here includes active migration, extracellular signaling, intercellular signaling, and



differentiation. Migration is implemented as a constant upward movement for all cell types, except for Paneth cells that migrate downwards. The extracellular signaling molecule, Wnt, is considered to be a function of local crypt curvature. As a consequence, Wnt exhibits a constant gradient, with the highest concentrations at the bottom of the crypt. Intercellular signaling is defined to be a function of Notch, where undifferentiated- and enterocyte progenitor cells produce the receptor, and secretory lineage cells produce the ligand. In addition, Notch signaling acts as an inhibitory signal for secretory cell differentiation in neighboring cells. Therefore, the secretory cells inhibit the differentiation of nearby cells into secretory cells in a process termed lateral inhibition [15]. The various lineage progressions between cell types, including dedifferentiation, are summarized in Figure 4c. The complex model is capable of replicating numerous biological phenomena reported in the literature, such as recuperation of the undifferentiated cell population after ablation, and predicts that a similar recovery is possible for each functional cell type within the crypt. This prediction reflects the primary focus of the model to emphasize the role of cell-environment interactions in establishing the functional phenotype of a cell. Specifically, it stresses that fate decisions are fluid and progenitors are capable of interconverting or dedifferentiating if exposed to the right set of cues.

Many morphogenetic events include populations with diverse sets of transitioning phenotypes and cell-type specific interactions. As the number of interactions expand, it becomes more difficult to intuit the role of any single interaction within the system. The capability of ABMs to capture the emergence of system-level features in response to the addition or removal of single interactions is a potent tool for probing developmental events. Indeed, other differentiation-focused ABMs have explored a myriad of developmental processes, such as somite formation [141], establishment of the germline in *C. elegans* [142, 143], and others [144-149].

#### 2.2.2.4 Manipulation of Morphogenesis Based on ABMs

While morphogenesis has been studied for over a century, a modern goal of morphogenesis research has become the manipulation of developmental mechanisms for purposes of targeted tissue engineering. Specifically, the short-term goal is to achieve functioning organ systems by replicating environmental conditions that regulate the targeted morphogenetic events *in vivo*. This modern line of research fundamentally asks questions regarding the number and character of conditions that are sufficient to emulate the morphogenesis of any given tissue. ABM provides a unique platform for mimicking realistic single-cell behavior at the tissue level in response to spatially and temporally diverse signals. Indeed, AB modelers hope that it is possible to derive the necessary conditions and interactions by iteratively simulating organogenesis from its inception under slightly altered rules and conditions. An excellent illustration of this pursuit is the modeling work of Setty *et al.* on early pancreatic organogenesis [150]. Their ABM consists of three main components: a reactive system engine for running the model, a front-end animation, and a GUI for mathematical analysis. The reactive system engine permits user interactions during runtime, such as pausing a simulation or adding or removing stimuli. In parallel, the front-end provides a 3D animation of the model during runtime, and the GUI analyzes the system. This implementation permits real-time analysis and manipulation of the simulated system. The obvious advantage of this approach is its ability to gauge the impact of specific conditions as they emerge.

The agents in this ABM consist themselves of three interacting components: the cell itself, a nucleus, and a membrane. The cell component makes growth and fate decisions, using information from the nucleus and membrane. The nucleus contains a set of genes that may be expressed or silent as a function of the environment and the cell state. The membrane responds to environmental factors, such as the binding or releasing of extracellular molecules, which triggers cell movement. The various features of the model are defined in a *statechart*, which designates the cell types and their independent rules for

cell-cell and cell-environment interactions. With these methodological settings, the model is able to recapitulate 2D histological features of the developing pancreas with high fidelity and generates results that are visually similar to 3D histology samples. While quantifiable validation is limited, the model is capable of creating tissue-scale morphological features that depend solely on single-cell decisions in response to environmental cues. Thus, with appropriate caution, this type of model can be a useful tool for determining and fine-tuning conditions that are necessary for deriving complex tissue structures *in vitro*.

#### 2.2.2.5 Limitations of ABMs for Morphogenesis

Like all modeling approaches, ABMs have clear strengths, but also germane weaknesses. The description of ABMs in the previous sections has demonstrated that a particular strength of ABMs is the relative ease with which specific hypotheses can be explored and tested. Thus, a typical question for an ABM is: Is this hypothesized mechanism sufficient to generate an observed pattern in time and space? While such a question can often be answered, a seemingly similar question is comparatively very difficult to assess with ABMs: Is this hypothesized mechanism actually driving the biological phenomenon and/or are there other mechanisms that are operating in parallel? More generally, “forward simulations,” which mimic “what-if” scenarios, are natural for ABMs, while it is difficult to address “inverse” questions, which attempt to determine feature representations from high-level data. For instance, it is difficult to infer the mathematical format of a mechanism, such as a growth or migration process. It is similarly difficult to answer questions such as “how many steady states can this system have?” or to characterize such steady states, especially if they are unstable. Along the same lines, it is sometimes difficult to answer questions like “is it possible for this system to exhibit a particular behavior?” Nonetheless, it is not entirely infeasible to characterize ABMs with rigorous mathematical analysis. For instance, a very thorough analysis of CPMs demonstrated some shortcomings with the depiction of cells with subcellular parts [151].

While a significant portion of the analysis was dependent on the infrastructure of CPMs, the approach itself highlights the potential for applying mathematical and model theory from other fields to answer some of these questions on an individual model basis.

A second complex of challenges pertains to fitting parameters. In particular, the designation of rules does not always lend itself to parameters that can be directly measured. In some cases, parameter values may be inferred from experimental data, if they are independent of other parameters, but when multiple parameters are interdependent, they have to be fit simultaneously. However, ABMs are not naturally amenable to parameter estimation algorithms, and the typical steepest-descent methods or genetic algorithms face issues with ABMs. Two factors that further complicate the situation are the stochastic nature of ABMs and the phenomenon of emergence, which can lead to high parameter sensitivities, that is, they can lead to large fluctuations in system behaviors in response to small changes in parameter values. As a consequence, typical parameter estimation techniques would require very large numbers of iterations for each parameter set, which would incur significant computational costs. Addressing these difficulties, new methods for parameter estimation in ABMs have begun to appear [106, 152].

Specifically with respect to morphogenesis, the usual output from an ABM is a set of agent-objects that contain spatial and state information at a cellular resolution. However, the patterns and morphological structures that are being targeted are formed at the cell population level. It is quite easy to qualify similarities between experimental observations and simulation results by visual comparison, but some form of quantification is necessary for model validation and the ability to make definitive claims. The derivation of such quantifiable metrics, especially those that characterize spatial features, is a non-trivial task. Spatial metrics need to be able to classify the same types of often irregular features, and denote them with similar representative values, for both experiments and simulations. Due to difficulties associated with this task, the use of spatial metrics for validation has been relatively rare, which is surprising for models targeting morphogenetic events. Instead, it

has been more common to attempt model validation through easily quantified population-based metrics. While population metrics do provide some connection between the simulations and the true system, they are not optimal for models that are designed to explain or predict pattern formation.

Two methods that are used most frequently for evaluating pattern formation are image analysis and network analysis. In the former, the simulated agents are converted into images that match the experimentally obtained images, whereas in the latter, the experimentally obtained images are converted into digital networks that have the same format as the agents. In both scenarios, representative features are extracted from data or calculated. For instance, cell locations are determined by identifying nuclei in the experimental images. If the critical features cannot easily be identified, customized algorithms are needed that allow pattern classification without the need of explicitly defined metrics [153, 154]. The acini model, for example, demonstrates the use of image analysis for model validation, with images acquired at multiple z-planes for 3D comparisons [72]. An excellent example of validation per network analysis can be found in the ESC aggregate model [138]. Finally, the combination of a dimensionality reduction technique, such as PCA, with either of these analysis methods can be an effective way to visualize dynamic changes in patterning in addition to quantification (see Figure 5). Here, we apply this methodology of network analysis in conjunction with PCA for exploring spatial pattern progression during early neural differentiation.

The functionality of ABMs in morphogenetic research has greatly improved recently through technological improvements, with one example being GPU-parallelization [155-157]. An evident advantage of parallelization is the ability to simulate much larger cell populations without compromising computational time. Thus, more realistic tissue scale models can be produced even on regular desktop computers. An added benefit is that models simulating fewer cells can run faster, thereby potentially eliminating some of the issues regarding parameter estimation. However, parallelization requires

certain constraints that can either limit functionality entirely or limit certain functionality based on the user's programming knowledge. At this point, ABMs seem to be gaining the capacity to model tissue-level morphogenetic events but given the paucity of specific examples so far it is hard to predict how the various interactions will translate into a parallel infrastructure.

#### 2.2.2.6 Summary and Application to Current Work

Morphogenesis is a paradigm for the benefits of merging traditional, reductionist biology with some of the newer concepts of experimental and computational systems biology. Clearly, understanding morphogenesis requires the very detailed elucidation of individual processes, but it also depends critically on solid knowledge of the dynamic interactions among these processes. ABMs are unique in their ability to investigate and integrate combinations of processes and their respective dynamics. Here, the process of intercellular communication is explored through experimental and computational means by integrating single-cell, experimentally-derived parameters into an agent-based model of differentiation at the population-scale.

ABMs are useful even for the exploration of poorly studied systems that lack sufficient data. They can be developed initially with rather coarse, top-down, behavioral-driven rule sets that ultimately generate hypotheses regarding those processes that are most influential for producing a given morphology or pattern. These hypotheses can in turn guide the design of laboratory experiments that increase the likelihood of identifying key events or pathways within the system. Even if model predictions are not entirely correct, the insights gained can be used to adjust, refine or alter the model or a previously posed hypothesis, for instance, by accounting for mechanistic features suggested by experimental data. In this alternation of experimental and computational methods, one side informs and fertilizes the other and, iteratively, both often improve.

As with any computational modeling strategy, it is important to note that clear goals and questions are needed up front, because they will determine the implementation details for the ABM. As a case in point, some early ABMs in biology lacked a unique objective and emphasized the replication of a biological phenomenon rather than the discovery of new characteristics or behaviors of a system. An early mesendoderm migration model, for instance, was framed around the functional incorporation of various elements into a model. These rather vague objectives can in retrospect be attributed to the fact that ABMs were relatively new in the field and that it was necessary to gain experience with exploring specific model features and the role of synergism among processes and rules. Since these early days of ABMs in biology, the field has substantially matured, and most modern ABMs are crisply focused on specific questions surrounding a particular application, as they should be.

One of the primary advantages of designing models for a specific application is that it becomes possible to simplify extraneous features. For example, the ABM developed here primarily considers intercellular communication as the mechanism of differentiation in pluripotent cells. While various factors contribute to cellular fate decisions, the inclusion of these factors would add complexity without providing additional information about the mechanism of interest.

While the history of ABMs is quite short, trends suggest that these models are quickly becoming mainstream tools in biology. Not long ago, they simply tried to replicate what biologists were observing. In their next phase of development, they began to explain the roles of hypothesized mechanisms and their interactions. The field is now at the threshold of using ABMs for predictions of scenarios that had never been tested in the laboratory. Such predictions will not only help with the formulation of novel, testable hypotheses, but may become a foundation for manipulating developing systems in a targeted manner, which will be fundamental to tissue engineering and regenerative

medicine and possibly the creation of “tissue factories” that permit the production of pure, valuable organic compounds.

### **2.3. Gap Junctions and Intercellular Communication**

A gap junction is a channel that connects the cytosol of two neighboring cells, allowing transport of ions and small molecules below 1 kDa in size [158]. Gap junctions are comprised of two connexons located in the plasma membrane of adjacent cells, with each connexon being composed of a hexamer of connexin proteins. The permeability of a single gap junction channel to specific molecules is dependent on the connexin subunits that form the gap junction [158, 159]. With 21 known isoforms in the human connexin family, the permutations of different connexin combinations results in high variability in gap junction transport potential [160]. The overall rate of intercellular transport is the cumulative diffusion across every open gap junction channel formed between two cells. Whether a gap junction is in an open or closed conformation is a function of various intrinsic factors, such as membrane potential [161], pH [162], calcium concentration [163], and the phosphorylation state of connexin subunits[164]. Each of these factors is of particular importance as ions and secondary messengers freely diffuse through gap junctions, allowing signals to propagate and creating a dynamic intercellular network.

### **2.4. Connexins Structure and Function**

All connexins contain four transmembrane domains connected by two extracellular loops with both N- and C- termini located in the cytosol. The connexin isoforms differ mainly by the length of the cytoplasmic loop and carboxy-terminal domain. The C-terminus is particularly interesting as it is involved in trafficking, localization, and gap junction turnover rates[160]. Furthermore, it contains the phosphorylation and ubiquitination regions affecting open and closed configurations, and degradation of gap junction channels, respectively. The C-terminus of connexin43 (Cx43) is 146 amino acids long and has been extensively studied in response to being the most ubiquitously expressed,



as well as involved in cell signaling[165] and proliferation[166]. Separate domains of the Cx-43 C-terminus have been implicated as necessary for correct assembly, transport to the membrane from the Golgi apparatus, and degradation of Cx43-composed gap junctions[160, 167]. In contrast, Cx26 has a small 10 amino acid C-tail with minimal post-translational modification potential. As such, these two connexins require different environmental conditions to affect their transport, with the short C-terminus of Cx26 not having as many domains to be affected compared to the C-terminus of Cx43. Therefore, between the large variability in gap junction configuration and connexin-specific sensitivities to environmental factors, the gap junction-mediated network is highly adaptable.

Additionally, connexons can form functioning channels independent of their ability to form gap junctions. When an adjacent cell is not present a connexon can form an open hemichannel to the extracellular environment [163, 168, 169]. These hemichannels result in an outward flux of cytosolic molecules, particularly ATP and glutamate [170], and an influx of larger membrane-impermeable metabolites [168]. Therefore, a cell that is adjacent to only a few cells has a higher probability of interacting with the extracellular space since it has more undocked connexons that can act as hemichannels. The phenomenon of opening hemichannels has predominantly been studied in astrocytes[163, 170] and lens epithelial cells[171, 172], but isolating connexon-mediated transport is challenging under physiological conditions where multiple channels can contribute. For a colony of cells this could result in spatial differences in the mechanism of transport for cells along the perimeter of the colony compared to cells centrally located. Therefore, if this occurs, cells along the periphery can interact with the extracellular space to a larger degree than centrally-located cells, with the potential for releasing or accumulating molecules that can stimulate differentiation.

## **2.5. Connexins and Differentiation**

The differentiation of a cell is defined by significant phenotypic changes, including morphologic and proteomic [173]. Most cell phenotypes have a corresponding set of connexin isoforms that are typically expressed [174], providing a mechanism for selective communication between cell types that are proximally located, with distinct transport rates for each unique channel composition. For example, the brain's neural network is comprised of three main cell types: astrocytes, oligodendrocytes, and neurons, all of which arise from a common neural progenitor stem cell lineage[175]. An astrocyte can form a gap junction channel with any other astrocyte, whereas oligodendrocytes do not form channels with other oligodendrocytes but only astrocytes. Neurons form electrical synapses composed of gap junction channels between other neurons but are not directly connected to astrocytes or oligodendrocytes. However, astrocytes can release glutamate through connexon hemichannels which creates a stimulatory response in the neuron network. The ability of these three neural cell lineages to selectively interact is a function of the different connexin profile that each phenotype expresses. The complexity of interaction in the neural system, and the derivation of three terminal phenotypes from a common progenitor make it an appealing lineage for studying intercellular dynamics. Table 2 outlines the specific connexins expressed by each cell type and the cell-to-cell interactions that occur.

**Table 1: Cell types present in nervous system and their corresponding connexins and interactions.**

Cell Type	Expressed Connexins	Cell Interactions
Astrocyte (A) [176]	Cx26, Cx30, Cx43	Intercellular Communication: A/O, A/A Glutamate Release
Neuron (N) [170, 177, 178]	Cx36, Cx45	Electrical Synapse: N/N Intercellular Communication: A/N * Glutamate Activation
Oligodendrocyte (O) [178]	Cx29, Cx32, Cx47	Intercellular Communication: A/O

\* Cells decouple as neurons mature

Regulation of cell-to-cell communication consequently means that connexins can modulate differentiation potential [165, 179-184] and can lead to disease states [184-189]. Mutations in Cx32 cause one variation of Charcot-Marie-Tooth (CMT) disease, where the Cx32 distribution in Schwann cells is functionally necessary for proper myelination to occur [185]. Directed knockouts of genes during murine development revealed that Cx26 is embryonically lethal [186], Cx32 generated susceptibility to liver tumors [187], Cx37 lead to female infertility [188], and Cx43 produced lethal cardiac abnormality after birth [189]. The effect of atypical connexin function during development led to controlled *in vitro* studies with specific modifications to intercellular communication. In an embryonic stem cell model system, non-specific gap junction inhibitors and siRNA targeting Cx43 were employed to establish the involvement of intercellular communication in maintaining pluripotency [190]. In both cases the cells lost their ability to maintain pluripotency and

transitioned to a more differentiated morphology. Interestingly, if the inhibitors were removed within 24-hours the cells regained their pluripotent state.

## **2.6. Connexins and Cell Cycle**

During the cell cycle, a highly conserved pathway involving the interaction of numerous proteins, there is a significant loss in gap junction connectivity [191]. Gap junctions containing Cx43 become internalized, decreasing intercellular communication to nearly negligible levels in cells that rely on this connexin. The isolation from intercellular diffusion from neighboring cells allows the cell to maintain tighter control over specific molecules necessary for cell cycle progression. Cyclic AMP (cAMP), for example, is a secondary messenger that is both able to pass through gap junctions and can inhibit cell proliferation at high concentrations [6, 192, 193]. Specifically, the expression of G1 cyclin, a necessary protein for transitioning from G1 to S phase, is inhibited by high cAMP levels. Only gap junctions and connexons composed of Cx26 are able to remain open during the cell cycle[6]. This results in redistribution of cAMP levels throughout the population of cells, preventing cell division. In comparison, Cx43 and Cx50 have more direct effect on cell proliferation. Rather than acting as channels for a separate molecule, these two connexins promotes the ubiquitination of another protein, Skp2 [166, 183]. Focusing on Cx43, in the presence of high levels of Cx43 there is an increased degradation rate of Skp2, preventing the role of Skp2 in ubiquitinating the cyclin-dependent kinase inhibitor p27. Thus, Cx43 causes a direct increase in p27 when the Skp2 gene is present which inhibits the transition from G1 to S phase of the cell cycle. The effect of Cx43 on cell cycle also partially explains why Cx43 accumulates, without forming functional channels, in the plasma membrane of cells undergoing cell division: it would prevent the ability of Cx43 to promote ubiquitination of Skp2 and also allow immediate channel formation upon completion of the cell cycle [191].

## CHAPTER 3 CHARACTERIZATION OF MULTISCALE INTERCELLULAR COMMUNICATION DURING EARLY NEURAL DIFFERENTIATION

### 3.1. Introduction

During embryogenesis, pluripotent cells migrate and differentiate to form complex multicellular structures in a reliable and reproducible manner. An incomplete understanding of the dynamic signaling mechanisms that affect differentiation and morphogenic patterning limits faithful and accurate replication of emergent behavior *in vitro*. To create more sophisticated engineered living systems (ELS), it is necessary to elucidate the collective impact of the numerous processes that shape multicellular constructs during normal development. Embryonic stem cells (ESCs) are an excellent model system for mimicking aspects of embryonic morphogenesis and investigating the various modes of communication amongst pluripotent populations [194]. The process of secretion, diffusion, and uptake of molecules is a well-established mechanism of biochemical communication across tissues, with the formation of extracellular morphogen gradients providing positional information that instructs cell fate decisions during differentiation, both *in vitro* and *in vivo* [195-197]. However, emerging evidence in recent years suggests that direct cell-cell communication plays an equally significant role in pattern formation during morphogenesis [32, 33, 198-200]. Ascertaining the role of intercellular communication as a regulator of differentiation is crucial for deciphering the diversity of spatial cues present during developmental processes and for the future derivation of more complex ELS.

Gap junction communication (GJC) provides direct channels that facilitate intercellular diffusion of small molecules (<1 kDa) between the cytosol of adjacent cells. Gap junctions assemble from hemichannels of connexin proteins present in the plasma membrane of adjacent cells and the connexin composition of each channel dictates the

permeability of specific metabolites [201]. Furthermore, the transcription and translation of connexin isotypes is regulated by cellular phenotype, allowing cells to exercise considerable dynamic control over intercellular connectivity during differentiation and tissue development [175]. The collective GJC across a population of cells produces an intercellular network of cells with fluid connectivity. The versatility of GJ-connectivity creates vast potential for the development of intracellular gradients of small molecules - such as cAMP, ATP, and serotonin - that influence many downstream metabolic and transcriptional processes governing cell-fate decisions [5, 6, 192, 202, 203]. Unfortunately, accurately interpreting molecular gradients within a network of differentiating ESCs is challenging due to the close-packed density of epithelial cells and development of gradients across various length scales. While some sensors are capable of discerning concentration gradients of small molecules, many rely on FRET-based detections and have noted limitations [204]. Specifically, bleed through of the FRET-donor can skew measurements and an inherently low signal-to-noise ratio severely limits the sensitivity of these sensors. Furthermore, while several techniques exist for characterizing GJ transport [205, 206], they typically offer limited capability to quantify fluctuations in connectivity at a single-cell resolution simultaneously with the transport behavior at the population level. The difficulty of quantifying the influence of individual cells on the intercellular network is compounded when considering connectivity that can both modulate and be modulated by dynamical differentiation processes occurring throughout the population. For such instances, computational modeling offers an attractive approach, in combination with single-cell transport data, to investigate the dynamics of multicellular GJ communication and its relationship with differentiation.

Here, intercellular transport rates were quantified from single cells within ESC colonies, the cell cycle state was identified as a significant modulator of these rates, and this knowledge was used to construct a computational model of intercellular transport in a multicellular system. This agent-based model, regulated by cell cycle and considering

growth, division, and differentiation, generated a complex, dynamic network topology of communication that was capable of predicting spatiotemporal perturbations of Oct4 expression during early neural commitment. Spatial patterns were quantified through dimension reduction techniques using derived network metrics to directly and quantitatively compare experimental results and simulation data; this approach enabled the development of accurate computational models for investigating communication within multicellular systems. This work highlights the importance of asynchronous cell division in establishing molecular gradients across tissue-scale systems and provides a framework for investigating the spatial evolution of differentiation within multicellular systems.

## **3.2. Methods**

### **3.2.1 Experimental Methods**

#### **3.2.1.1 Cell Culture**

The murine embryonic stem cell line (D3) was cultured at 37 °C in monolayer on 100-mm tissue culture plates coated with 0.1% gelatin (Millipore EmbryoMax) in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% fetal bovine serum (FBS) (Atlanta Biologicals, Atlanta, GA), 2 mM L- glutamine (Lonza), 100 U/ml penicillin, 100 µg/ml streptomycin, and 0.25 µg/ml amphotericin (MP Biomedicals), 1x MEM nonessential amino acid solution (Corning), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), and 10<sup>3</sup> U/ml leukemia inhibitory factor (LIF) (EMD Millipore). Cells were passaged every 2–3 days, using 0.05% trypsin (Corning) to dissociate cells, centrifuged at 200 rcf for 5 minutes, and plated at a density of 20000 cells/cm<sup>2</sup>.

#### **3.2.1.2 Differentiation Protocol**

Cells were plated at a density of 10000 cells/cm<sup>2</sup> on Ibidi µ-slides coated with a 0.1% gelatin, 0.05% fibronectin solution. Cells were allowed to grow under regular cell culture conditions described earlier for 24 hours before being transitioned to N2B27

medium containing 1  $\mu$ M retinoic acid (Sigma-Aldrich). The N2B27 medium was composed of 50% DMEM/F12 (Thermo Fisher Scientific) and 50% Neurobasal Media (Thermo Fisher Scientific), supplemented with 0.5% N-2(100x) (Thermo Fisher Scientific) and 1% B-27(50x) (Thermo Fisher Scientific). Every 24 hours, the medium was replaced with fresh N2B27 medium containing 1  $\mu$ M retinoic acid.

#### 3.2.1.3 Immunofluorescence

For staining of Oct4 and Sox2, cells were fixed in 4% paraformaldehyde (Thermo Fisher Scientific) for 5 minutes and then permeabilized with 0.1% Triton X100 for 15 minutes. Samples were blocked with 2% normal donkey serum (NDS) (Sigma-Aldrich) for 1 hour and then incubated overnight at 4°C with the following primary antibodies in 2% NDS: goat polyclonal Oct-3/4 (Santa Cruz, 1:200) and rabbit polyclonal Sox-2 (Thermo Fisher Scientific, 1:200). After washing, samples were incubated in a 2% NDS secondary antibody solution of donkey anti-goat Alexa Fluor 568 (Thermo Fisher Scientific, 1:200) and donkey anti-rabbit Alexa Fluor 647 (Thermo Fisher Scientific, 1:200) for 45 minutes before counterstaining with Hoechst 33342 (1:1000) in DI water. All images were collected on a PerkinElmer UltraVIEW VoX spinning-disk confocal microscope with a sCMOS camera at 20X magnification.

For Cx43 and Oct4 staining, the same procedure was utilized but with the following changes: The duration of fixation was increased to 15 minutes, the primary antibodies were goat polyclonal Oct-3/4 (Santa Cruz, 1:200) and rabbit polyclonal Cx43 (Sigma, 1:400), and the images were collected at 60x magnification.

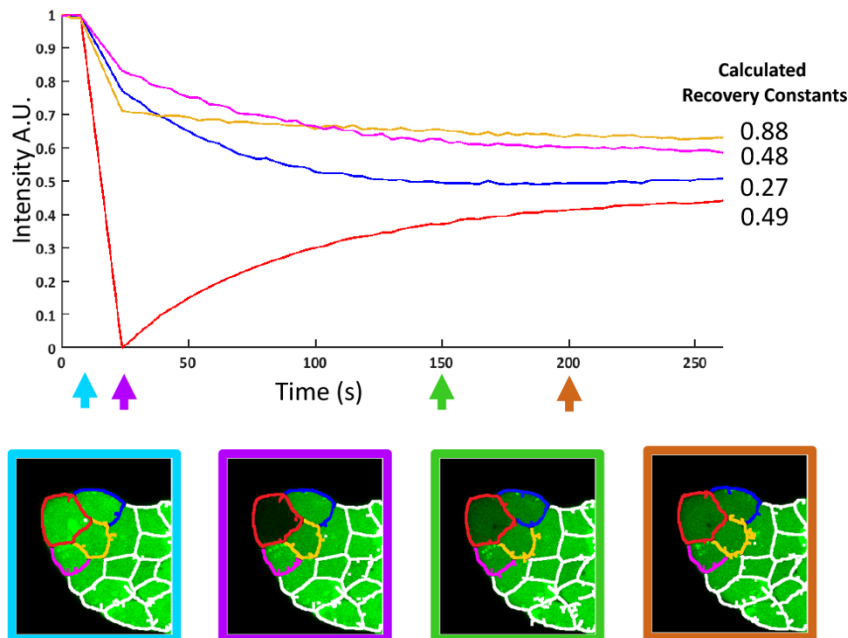
#### 3.2.1.4 GAP-FRAP

To form colonies, mouse (m)ESCs were plated at a density of 10000 cells/cm<sup>2</sup> on Ibidi  $\mu$ -slides, coated with a 0.1% gelatin, 0.05% fibronectin solution, for 24 hours. A 1 mM calcein-AM (Thermo Fisher Scientific) stock solution in DMSO was diluted into 100  $\mu$ L DMEM media and vortexed. The diluted calcein-AM solution was added to each well



of the Ibidi  $\mu$ -slide to a final concentration of 1  $\mu$ M for 40 minutes at 37°C. After 40 minutes, Alexa Fluor 647–conjugated wheat germ agglutinin (WGA) (Thermo Fisher Scientific) was added to each well to a final concentration of 10  $\mu$ g/mL, for 5 minutes at 37°C. Each well was rinsed once before phenol red–free DMEM (Corning) medium was added for imaging. Images were acquired with a Zeiss NLO 710 confocal microscope, using built in software for photobleaching individual cells. Bleaching was initiated after acquiring three images, then maximum laser intensity was pulsed for 30 iterations before measuring fluorescence intensity in the photobleached cell every 3.8 seconds. For identifying mitotic cells, Hoechst 33342 was added with the WGA for 5 minutes at a 1:1000 dilution. When quantifying nocodazole-treated cells, each well was treated with 100 ng/mL of nocodazole (Sigma-Aldrich) for 4 hours before replacing with new DMEM medium containing 1  $\mu$ M calcein-AM and incubating for 40 minutes at 37°C, followed by 5 minutes with WGA, then changing to phenol red–free DMEM for imaging.

From each GAP-FRAP experiment, the measured fluorescence intensity value in the photobleached area was compiled as a function of time, setting  $t=0$  as the time point with the minimum intensity value after bleaching. The intensity values were normalized to the maximum intensity after bleaching and fitted to an exponential function of the form  $I = 1 - \exp(-\frac{t}{R_c})$  where  $R_c$  is the recovery constant. The recovery constant was then normalized to the maximum fluorescence intensity before bleaching. In a few cases, it was also possible to quantify the recovery lost from neighboring donor cells (Figure 6).



**Figure 6: Quantifying the loss of fluorescence in the cells adjacent to the photobleached cell during gap-FRAP. In the circumstance where the change in intensity of primary neighbors is sufficient to measure, the relative transport of those cells can be quantified in a similar manner to typical FRAP experiments. Interpretation of these transport rates is colony specific because non-photobleached cells are experiencing both a loss of fluorescence into the photobleached cell and a gain of fluorescence from other adjacent cells.**

However, the context of the intercellular transport is different for the neighboring cells compared to the photobleached cell. Specifically, the photobleached cell only experiences a net influx of unbleached calcein whereas the neighboring cells have both a loss of calcein to the photobleached cell and an influx from secondary neighbors. As such, measurements from secondary cells are not included in any of the distributions for characterizing transport rates in Figure 13.

### 3.2.1.5 Flow Cytometry

For quantifying differentiation progression, cells were grown for 24 hours on 100 mm tissue culture plates coated with 0.1% gelatin. Each plate was rinsed with PBS then

differentiated using our previously stated differentiation protocol, with samples taken at 0, 24, 48, and 72 hours from the start of the protocol. Each sample was dissociated using 0.05% trypsin, centrifuged for 5 minutes at 200 rcf, and fixed using 4% paraformaldehyde for 10 minutes. Cells were permeabilized by centrifuging at 200 rcf with 5% Triton X100 for 5 minutes. To prevent non-specific binding, cells were incubated for an hour in 10% NDS before being treated with primary antibodies Oct-3/4 (Santa Cruz, 1:200) and Sox-2 (Thermo Fisher Scientific, 1:200) in 10% NDS for 1 hour. Cells were washed 3 times then incubated for 30 minutes in 10% NDS containing the following secondary antibodies: Donkey anti-goat Alexa Fluor 488 (Thermo Fisher Scientific, 1:200) and Donkey anti-rabbit Alexa Fluor 647 (Thermo Fisher Scientific, 1:200).

For cell cycle analysis, the same protocol was used but with incubation in Hoechst 33342 (1:1000) for 15 minutes instead of primary and secondary antibody staining. All flow cytometry measurements were collected on a BD LSR Fortessa flow cytometer.

### **3.2.2 Computational Methods**

#### **3.2.2.1 Agent-Based Model**

An agent-based model was developed for simulation of pluripotent colonies by refining a python-based custom platform first described in White et al [137]. Here, each cellular agent with a set of properties of cell radius, spatial coordinates, division time, cell type, and intracellular concentration. Cell radius was set at 6.5  $\mu\text{m}$  based on previous data collected in our lab for D3 cells, with cells modeled as rigid bodies. Spatial coordinates were initially set using a digitized set of coordinates from an experimental colony, assuming 100% pluripotent cells. Division time is randomly distributed between 0 and 18 hours for each cell and is defined as the time since the cell last divided, with pluripotent cells dividing every 18 hours and differentiated cells dividing every 51 hours. As the colony grows it is restricted to a 2D-plane, with a collision detection algorithm optimizing cellular movement to prevent overlap of cells in the model. The asymmetrical growth of the colony

is accomplished by preferential movement of cells away from the central mass of the colony when resolving overlap during collision detection. The initial intracellular concentration of each cell is randomly selected between 0.6 and 0.65, with a nascent permeability, production rate, and degradation rate specific to each cell type (Pluripotent, Differentiated). Table 3 presents parameter values for each cell type.

**Table 2: Cell type specific parameters used in ABM of intercellular communication**

<b>Cell Type</b>	<b>Pluripotent</b>	<b>Differentiated</b>
<b>Cell Radius (r)</b>	6.5	6.5
<b>Cell Cycle Length (CCL)</b>	18 hours	51 hours
<b>Nascent Permeability (PM<sub>n</sub>)</b>	0.45	0.85
<b>Production Constant (vp)</b>	4.89E-6	6.0E-6
<b>K50</b>	0.03	0.04
<b>Degradation Constant (Kd)</b>	1.3E-7	1.3E-7

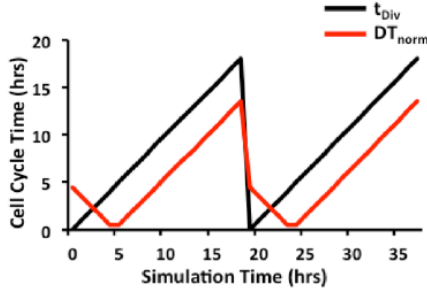
The intercellular network is determined at each time step by calculating the intercellular distance between every cell, and if the cell membranes are within 2 microns ( $2 \times \text{radius} + 2 \text{ um}$ ) then they are “connected”. Since the resistance to diffusion is significantly higher at the gap junction interface, we consider the intracellular compartment to have instantaneous mixing after every diffusion time step. The flux between each cell follows Fick’s first law, accounting for the individual permeability of the connected cells (function of cell cycle and cell type, Equation 1-6).

$t_{\text{Div}}$ : Current division time, between 0 and CCL for current cell

$\text{PM}_{\text{max}}$ : Maximum permeability, the maximum flux between any two cells if all gap junctions were open

$$[1] \quad DT_{\text{norm}} = \text{Abs} \left( t_{\text{Div}} - \left( \frac{CCL}{4} \right) \right)$$

$DT_{\text{norm}}$  is a transform function, shifting the minimum value forward by  $\frac{1}{4}$  of the cell-cycle length (Figure 7).



**Figure 7: Comparison of cell cycle time and the respective DTnorm transform.**

This shift decreases the steep transition in cell-cycle time that is seen in  $t_{\text{div}}$  when a cell starts a new division cycle (@18 hours simulation time), while maintaining individual regions of minimum and maximum values per period.

$$[2] \quad \text{Cell Cycle Effect (CCE)} = \frac{1}{1 + \left( \frac{DT_{\text{norm}}}{CCL} \right)^2} + \frac{0.69 t_{\text{Div}}^6}{t_{\text{Div}}^6 + CCL^6}$$

The cell-cycle effect function is defined such that maximum transport would occur during the first few hours of a new division cycle when a cell would be in G1-phase, followed by a decline throughout S-phase to reach a minimum transport value near the end of a division cycle when a cell would be undergoing mitosis.

$$[3] \quad PM_{\text{n,eff}} = PM_{\text{max}} (CCE_{\text{cell1}} * PM_{\text{n,cell1}})(CCE_{\text{cell2}} * PM_{\text{n,cell2}})$$

The effective permeability between any two cells is equal to the maximum permeability adjusted by the product of the nascent permeability of the adjacent cells. The

nascent permeability is defined by cell type (differentiated or undifferentiated) and regulated by the Cell cycle effect function.

$$[4] \quad Prod = \frac{vp}{1+C_t/K_{50}}$$

$$[5] \quad Deg = -k_d C_t$$

$$[6] \quad Flux = PM_{n,eff} \Delta C + Prod + Deg$$

Production, degradation, and flux in and out of a cell are calculated every 3 seconds, or 1200 times per differentiation time step, with a continuously updating intracellular concentration. For the production function, we use an inhibitory Hill function with an n-value of 1 to represent decreased production when intracellular concentrations are high. A simple degradation function is used with the degradation constant the same for both cell types. Diffusive flux between adjacent cells is calculated as the product of the effective permeability and the concentration gradient between each neighboring cell. The final concentration at the end of every hour is used to determine the probability of differentiation, as per Equation 7.

$$[7] \quad P(diff) = \frac{tC^n}{tC^n + K_{th}^n}$$

Threshold Concentration ( $C_{th}$ ): 0.74

tC: Threshold count, number of timesteps a cell is over  $C_{th}$

*ie. IF  $C > C_{th}$ :  $tC = tC + 1$  ELSE:  $tC = tC - 1$*

$K_{th} = 2$

$n = 4$

A stochastic differentiation term is incorporated to represent external factors affecting differentiation that are not mechanistically captured by our intercellular diffusion module. Since the main external factor was the addition of retinoic acid, which was replaced every

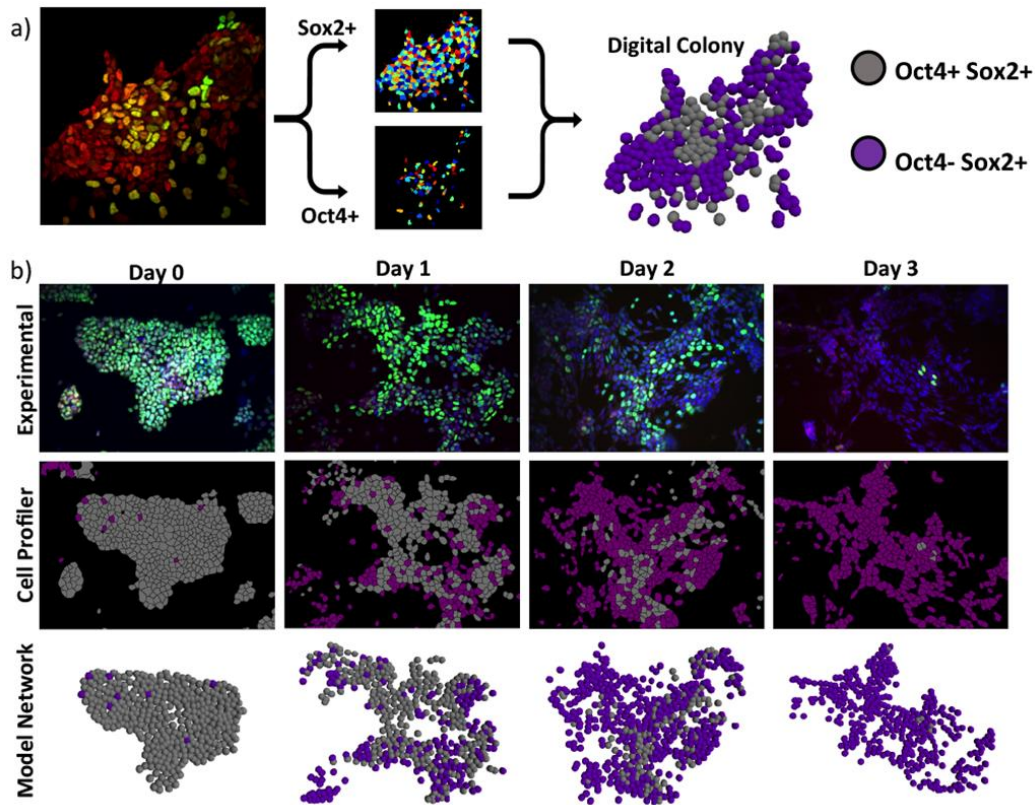
24 hours, the stochastic term is made oscillatory to correspond with RA-addition. Specifically, differentiation potential increases for 12 hours after RA-addition then decreases back to the base differentiation potential, with this repeating every 24 hours (Eq 8).

$$[8] \quad P(stochastic) = 0.002 + Abs[0.001 * \sin\left(\frac{t_{sim}}{8}\right)]$$

where  $t_{sim}$  is the simulation time.

### 3.2.2.2 Image Analysis

To convert our experimental images into digital networks, we used a CellProfiler (<http://cellprofiler.org/>) pipeline (Figure 8). The pipeline uses the blue channel (Hoechst) to identify nuclei within the image, and then uses propagation to define cellular boundaries. Both red and green channels, representing Sox2 and Oct4 respectively, first use a local MCT threshold to separate high expressers from low expressers and background noise. After thresholding, the red and green channels are converted to masks and applied to the cell objects identified using the blue channel. This data is exported to a CVS file and imported into Python, where it is represented in an analogous format as our simulation data.

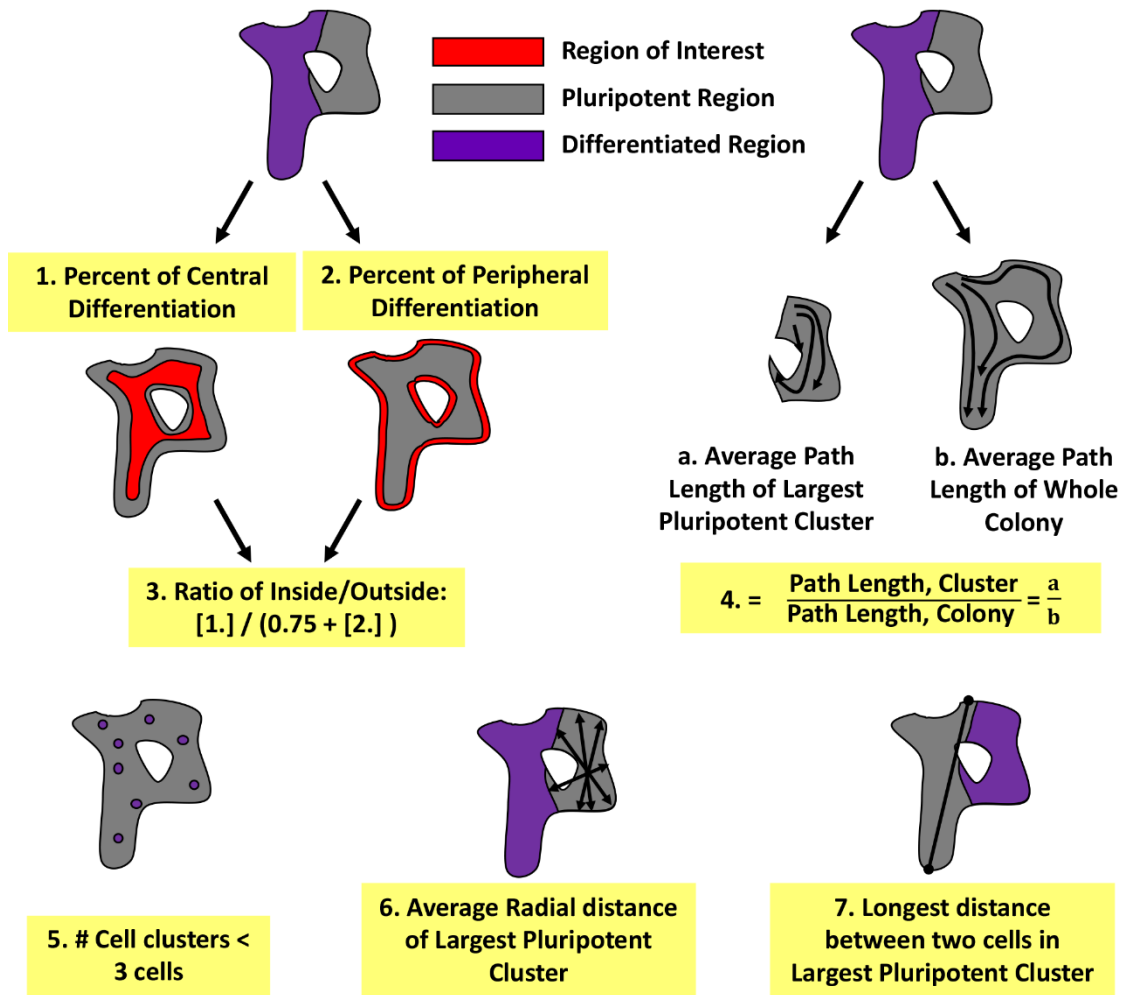


**Figure 8: CellProfiler pipeline used to convert experimentally imaged colonies to digital networks synonymous with model outputs. (A) Conversion of a single image using Sox2 and Oct4 staining into a digital colony. (B) Experimental images, with digitally enhanced Oct4 stain for ease of comparison, before conversion to CellProfiler representations and respective digital networks.**

### 3.2.2.3 Latent Space Analysis

We computationally generated 2D pattern classes defined for embryoid body differentiation. Each pattern class was produced using a previously digitized experimental network, resulting in 120 unique colony structures. To better distinguish spatially localized clusters, the differentiation status of a cell was also a function of the differentiation status of neighboring cells. Specifically, if the percent of differentiated neighbors is greater than the average number of differentiated neighbors for every cell in the population, it is classified as differentiated. From each network, seven metrics were extracted as represented in Figure 9.





**Figure 9: The seven selected metrics used to quantify the spatial patterning within colonies. Each metric is represented by a single numerical value and every sample (experimental & computational) is depicted by the same seven metrics. Colony-based metrics: (1,2,3,5), Cluster-based metrics: (4,6,7).**

After collecting these metrics for every pattern class generated in our training set (7 metrics x 120 networks), the data is mean centered, scaled to unit variance, then transformed using Principal Component Analysis (PCA). The same metrics were extracted from both simulation and experimental data, scaled/normalized in an identical fashion to our training set, and transformed into latent variable space using the previously trained

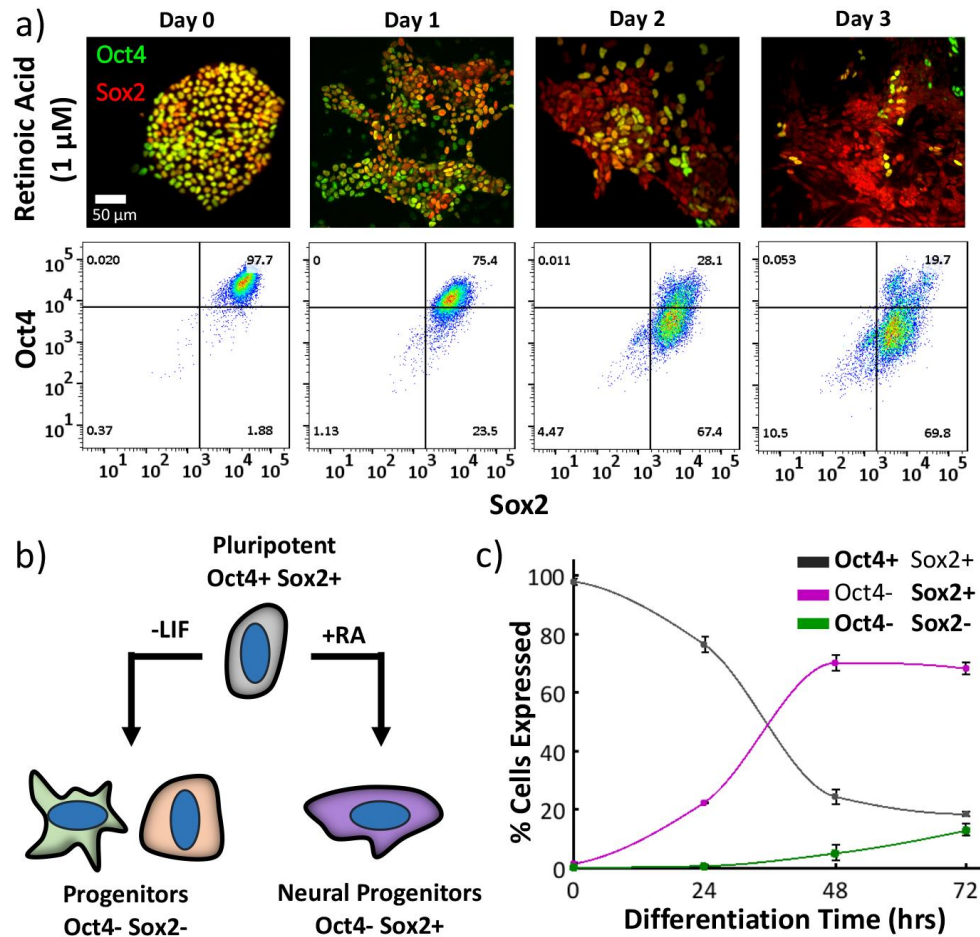
PCA transform. The average trajectories for simulations were calculated by averaging the data points at each specific time.

### **3.3. Results**

#### **3.3.1 Spatially organized differentiation occurs during retinoic acid treatment**

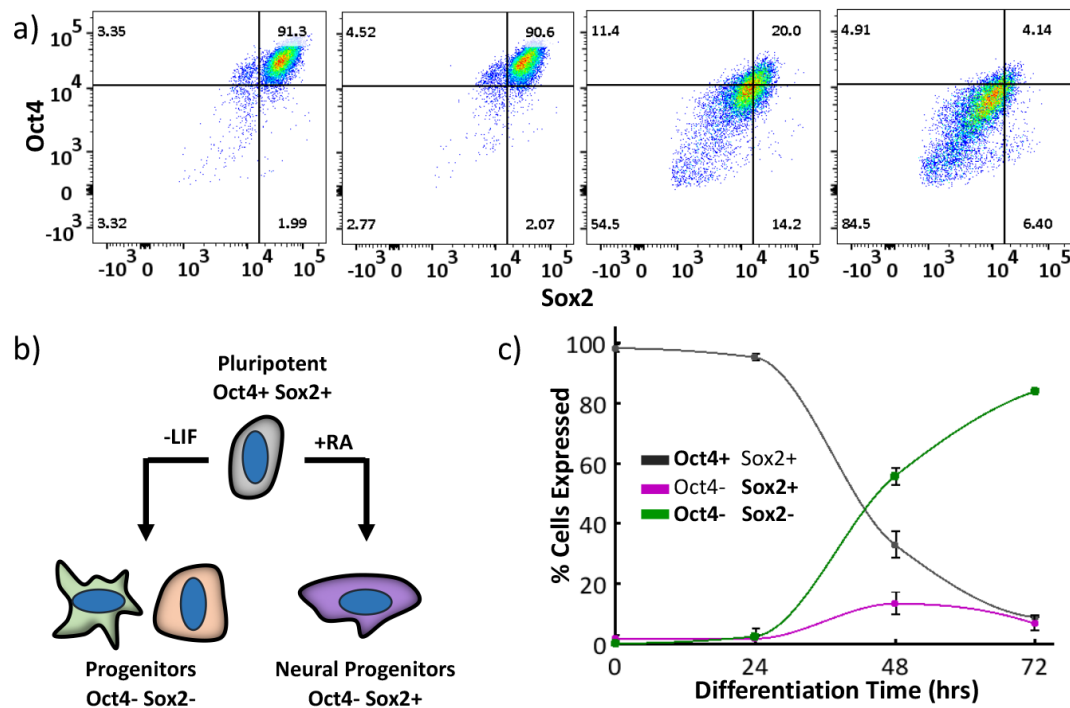
To evaluate spatial patterning during neural differentiation, monolayers of murine ESC colonies were treated for 72 hours with retinoic acid (RA, 1  $\mu$ M), a potent morphogenic promoter of neural progenitor cells that emerge during hindbrain development [207]. A 2D monolayer system was selected to enable a more comprehensive link between the patterning produced and intercellular communication between neighboring cells. In particular, ESCs are known to secrete paracrine factors that can accumulate within the confines of an aggregate and affect differentiation in a spatial manner [124]. The 2D system minimizes this potential for environmental heterogeneity and consequently allows heterogeneity caused by intercellular mechanisms to be studied in a less confounded context. The progression of differentiation was assessed based on expression of the two pluripotency factors, Oct4 and Sox2. The Oct4+Sox2+ state has been well characterized to reflect the pluripotent state of ESCs, whereas Sox2 expression alone (in the absence of Oct4) is associated with a neural ectoderm fate, indicating that the Oct4-Sox2+ phenotype is the initial neural progenitor state after loss of pluripotency [129, 131, 132].

As differentiation proceeded, a transient, patterned loss of Oct4 was observed while Sox2 expression was maintained, as expected for neural progenitors (Figure 10a). The majority of the cell population (70 $\pm$ 4%) assumed an Oct4-Sox2+ neural progenitor state within 48 hours of RA treatment. In contrast, spontaneous differentiation induced by LIF-withdrawal from serum-containing media resulted in a decrease in both Oct4 and Sox2 expression, associated with primitive endoderm and mesoderm specification (Figure 10b,c, Figure 11) [208].



**Figure 10: Loss of Oct4 and maintenance of Sox2 expression during retinoic acid–induced differentiation.** Oct4 expression begins to decline after 24-hour exposure to retinoic acid (1  $\mu$ M), with the main transition to an Oct4<sup>-</sup> state occurring between 24 and 48 hours of RA treatment. After 72 hours, ~ 70% of the population retain an Oct4<sup>-</sup>Sox2<sup>+</sup> phenotype (A,C).

A schematic depicting the change in gene expression in ESC populations towards the predominant progenitor state during two differentiation protocols: LIF-withdrawal in serum-containing media and retinoic acid addition (B). Quantification of flow cytometry data in (A) demonstrates the maintenance of the Oct4<sup>-</sup>Sox2<sup>+</sup> state between 48 and 72 hours of RA treatment (C), n=9 (3 biological replicates, 3 technical replicates for each biological replicate).

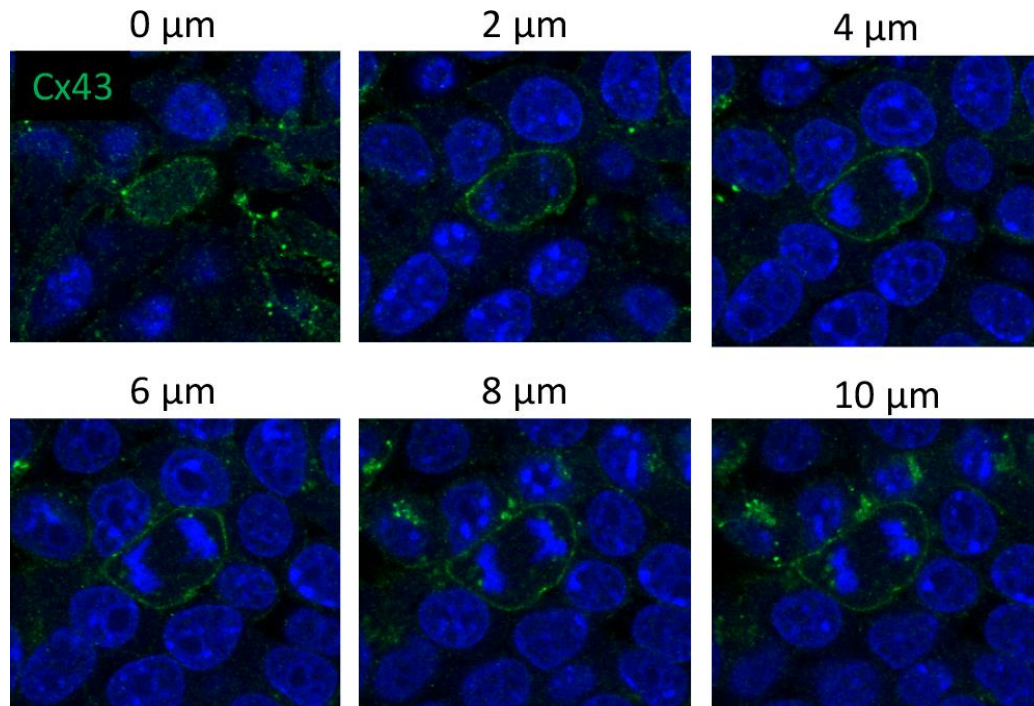


**Figure 11: Sox2 and Oct4 concomitant loss through LIF withdrawal in the presence of serum containing media characterized by flow cytometry. (B) Schematic representing the change in gene expression towards the predominant progenitor state in ESC populations during two differentiation protocols, LIF-withdrawal in serum containing media and retinoic acid addition. (C) Quantification of flow cytometry data in (A), with a large population of Oct4<sup>-</sup> Sox2<sup>-</sup> progenitors accumulating between 24 and 72 hours in addition to a smaller population of Oct4<sup>-</sup> Sox2<sup>+</sup> cells, n=9 (3 biological replicates, 3 technical replicates for each biological replicate).**

RA-accelerated differentiation was used hereafter in this study due to the selective commitment of ESCs to the neural lineage. The selectivity provides a robust system for studying the emergence of spatial organization during the transition of a cell population from a pluripotent to a largely homogenous neural progenitor state.

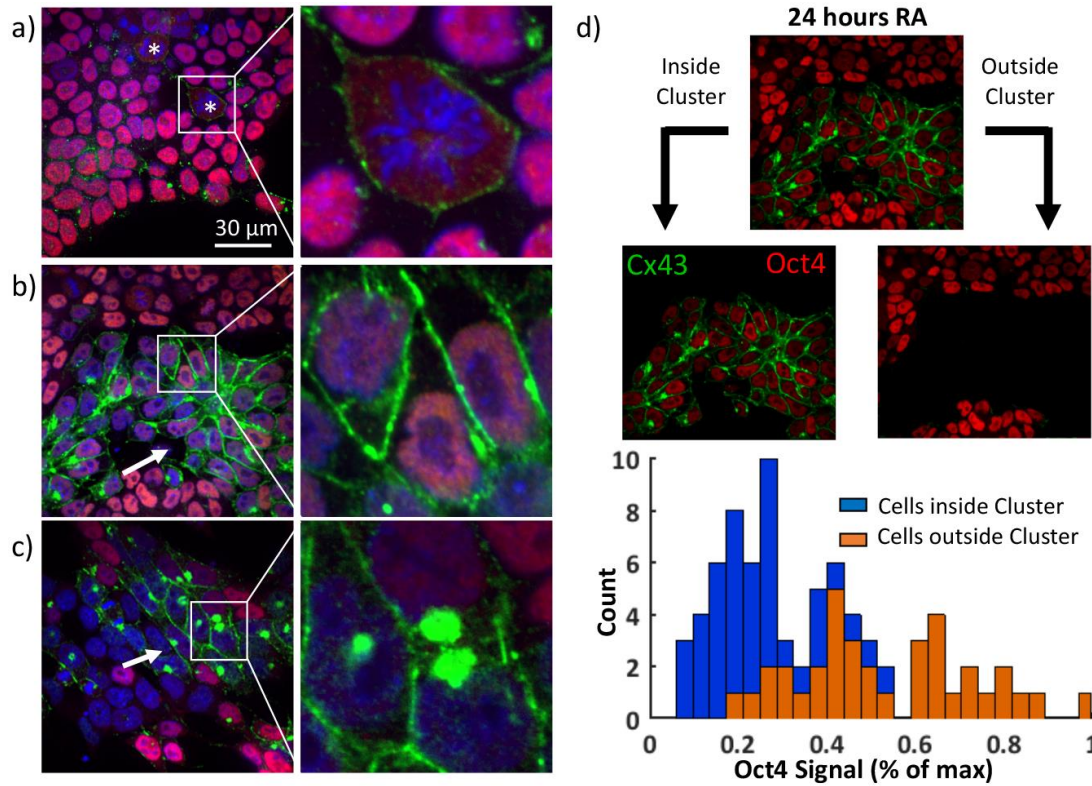
### 3.3.2 Differentiation causes a transient redistribution of Cx43 and increased permeability for communication between neighboring cells

The gap junctional intercellular network has been implicated as a regulator of spatial patterning during the development of neural networks [199, 209]. As such, we examined whether RA treatment initiated spatial remodeling of the intercellular network by assessing the membrane localization of connexin43 (Cx43), the primary connexin isotype expressed by ESCs [210]. Cx43 expression in untreated colonies (0 hours) was concentrated at the cell-cell interfaces, usually as a series of punctate spots (Figure 13a). In contrast, mitotic cells typically had diffuse Cx43 spread across their membrane that was not limited to cellular interfaces (Figure 12). This diffuse Cx43 ‘ring’ pattern has previously been found to coincide with mitosis-specific phosphorylation of Cx43 that causes decreased plaque formation [191].



**Figure 12: Consistent Cx43 stain throughout membrane in mitotic cells. A confocal z-stack shows a diffuse ‘ring’ structure at each focal plane in the membrane of mitotic cells while non-mitotic cells display a more punctate staining profile.**

After 24 hours, distinct clusters of cells with enhanced Cx43 staining between cells were observed (Figure 13b). Each Cx43-enhanced cluster also showed diminished Oct4 expression (Figure 13d), implying that the loss of pluripotency is concomitant with an increase in intercellular communication. After 24 hours, distinct clusters of cells with enhanced Cx43 staining between cells were observed (Figure 13b). Each Cx43-enhanced cluster also showed diminished Oct4 expression, implying that the loss of pluripotency is concomitant with an increase in intercellular communication. A similar increase in Cx43 signal was apparent in clusters of differentiated cells after 48 hours, but with fewer cells per cluster compared to the 24-hour time point, indicating that the change was transient (Figure 13c). Furthermore, the clusters at 48 hours also displayed an accumulation of Cx43 in the cytoplasm. Interestingly, a similar accumulation of Cx43 is present in the Golgi apparatus of proliferative neural progenitor cells that is temporarily lost during differentiation [211]. Thus, ESCs exhibit a transient increase in Cx43 localization to their membranes as they undergo differentiation towards a neural progenitor state in response to RA-treatment.



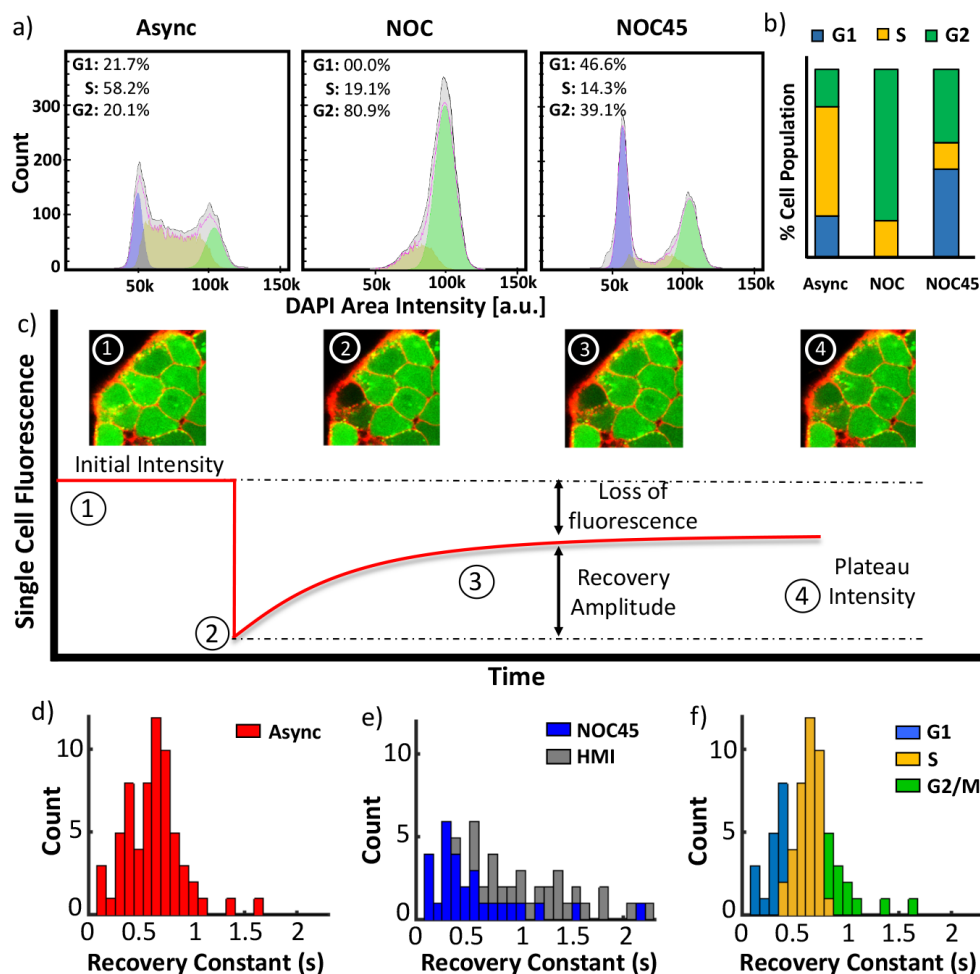
**Figure 13: Cx43 signal (green) increases during retinoic acid-induced differentiation, with compartmentalization of transitioning cells between 24 and 48 hours. Mitotic cells are prevalent in pluripotent colonies (A) and show a diffuse ‘ring’ of Cx43 in the membrane, designated by an asterisk, that is typical of Cx43 not forming GJ plaques. After 24 hours of RA treatment (B), Cx43 noticeably increased between clusters of cells with low Oct4 expression (red), characterized in (D). At 48 hours of treatment (C), cells that have low (but non-zero) Oct4 expression have large expression of Cx43 in the cytoplasm. Previous studies have linked an accumulation of Cx43 expression in the cytoplasm to localization in the Golgi apparatus, specifically in proliferative neural progenitor cells. The average Oct4 intensity was calculated for cells that were inside and outside the clusters displaying enhanced Cx43 at 24 hours (D). Cells within the Cx43 enhanced clusters at 24 hours exhibited a significant lower Oct4 expression compared to cells outside of the cluster with low Cx43 signal.**



### **3.3.3 Intercellular transport as a function of cell cycle**

Given the capability of the cell cycle to modulate GJ connectivity and the rapid cycling time of pluripotent cells, we hypothesized that the cell cycle was a source of dynamic heterogeneity within the pluripotent-intercellular network. To characterize the degree of influence, intercellular transport rates were quantified at each stage of the cell cycle. The various cell cycle stages were distinguished by treatment with nocodazole to cause cell cycle arrest at G2/M-phase followed by the removal of nocodazole and subsequent recovery [212]. Sampling before (Async), immediately after (NOC), and 45 minutes after nocodazole (NOC45) treatment yielded three conditions: an asynchronous population, a synchronized population of G2/M-phase cells, and a recovered population containing both G1- and G2/M-phase cells (Figure 14a,b). In lieu of the NOC condition, Hoechst-mediated identification (HMI) of M-phase cells in an asynchronous population was adopted for quantification purposes. Hoechst treatment enabled the distinct DNA configurations of M-phase cells to be visually distinguished and provided data that was more representative of communication in the naturally asynchronous multicellular environment. The GAP-FRAP technique was used to establish distributions of intercellular transport rates from the three conditions: Async, HMI, and NOC45. Briefly, GAP-FRAP monitors the diffusive transport of an intracellular fluorophore (calcein) into a single photobleached cell from neighboring interconnected cells (Figure 14c)[213]. Recovery constants were mathematically derived using the perturbation-relaxation equation described in [213] and have an inverse relationship with the functional transport rate of a cell.

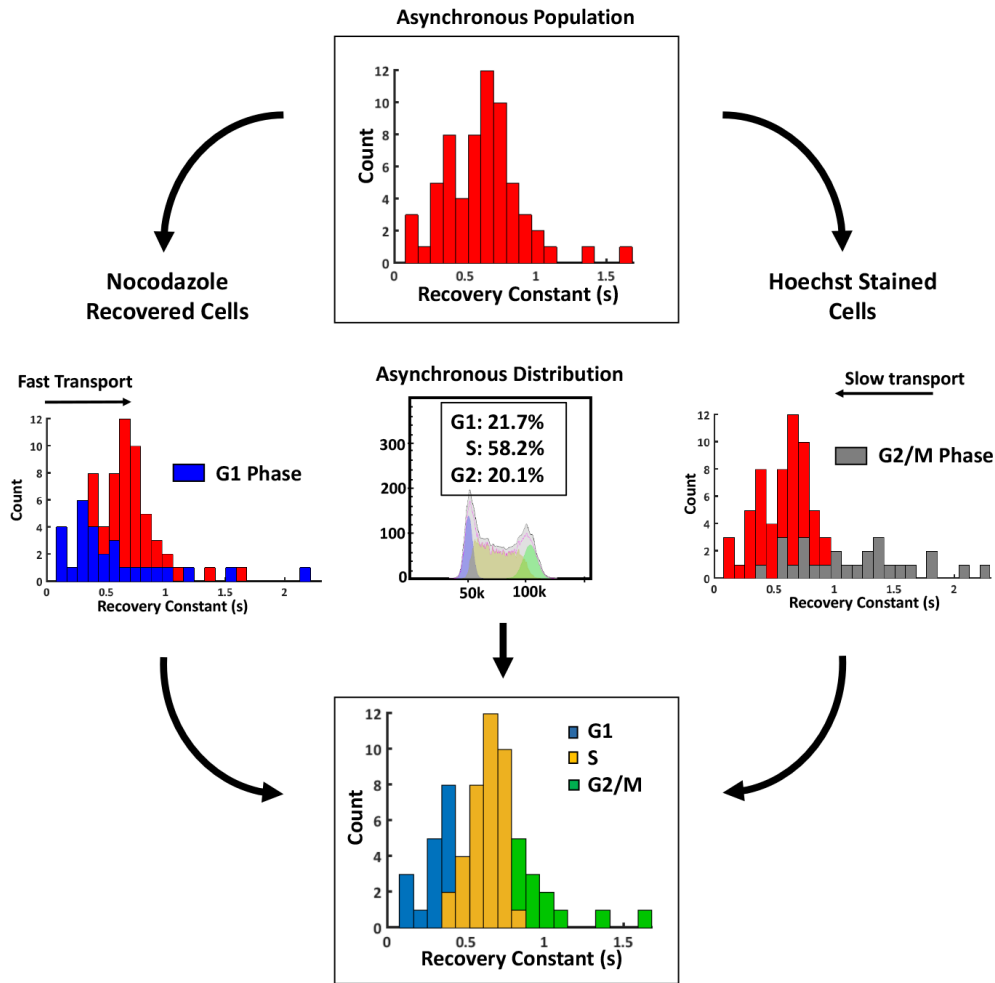




**Figure 14: Analysis of intercellular transport rates as a function of cell-cycle state. (A)** Asynchronous cell population (Async) shifted to G2-phase after nocodazole treatment (NOC), and after 45 minutes of recovery (NOC45) the population shifted to G1-phase. **(B)** The population distribution averages from (A) were calculated for the Async, NOC, and NOC45 treatment conditions. **(C)** An illustration of the GAP-FRAP technique for quantification of relative diffusion rates between adjacent cells. **(D)** A histogram of recovery time constants collected using gap-FRAP in the Async population, where high and low recovery constants represent slow and fast transport rates, respectively. **(E)** The distributions of recovery constants for Hoechst-identified mitotic cells and NOC45 were shifted to the right and left of the Async population, respectively, indicating slower and

**faster transport in these populations. (F) A projection of each cell cycle state onto the Async distribution using information from (A) and (E), as described in Figure 15.**

The Async population yielded a unimodal distribution of recovery rates, predominantly representing S-phase cells (~60%)(Figure 14a,d). The HMI population had significantly larger recovery constants on average (~50%,  $p < 1e-5$ ) than that of the Async population, indicating slower transport for mitotic cells which is consistent with the literature [191]. In contrast, the NOC45 population primarily exhibited low recovery constants and fast intercellular transport. Since the NOC45 population consists mostly of G1- and G2/M-phase cells (Figure 14b) and analysis of the HMI population showed mitotic cells with slow transport, it was deduced that G1-phase was associated with the observed fast transport. As a secondary validation, morphological analysis of the NOC45 tested cells showed that the rounded morphology typical of M-phase cells accompanied each case of slow transport in the distribution. Therefore, by establishing that G1-phase and G2/M-phase produce fast and slow transport rates, respectively, the asynchronous distribution could be divided according to the proportion of each cell cycle state (Figure 14b,f,Figure 15).



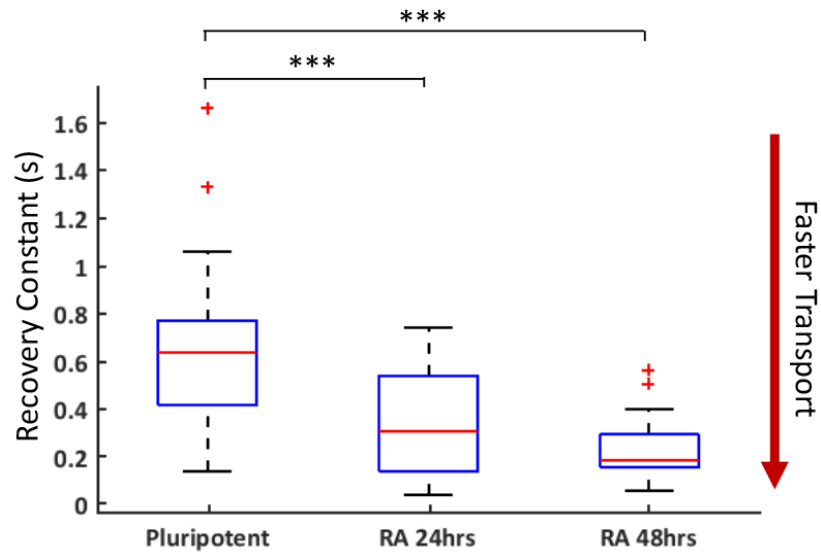
**Figure 15: Defining the asynchronous distribution of recovery constants by cell cycle state.** Through identification of mitotic cells using Hoechst-staining, it was discovered that the recovery constants of G2/M phase cells cluster along the right tail of the asynchronous distribution. Similarly, performing GAP-FRAP on a nocodazole-recovered (NOC45) population yielded a cluster along the left tail of the asynchronous population and a few cells with slower transport. Since the NOC45 population consists predominantly of G1- and G2/M-phase cells (see Fig 14a) and G2/M-phase were shown to have decreased transport rates, we can associate the cluster along the left tail with G1-cells. Using the distribution of cells in each cell cycle state within an asynchronous population, the recovery constants can be divided by: 21.7% of the fastest transporting cells were in G1-phase, 20.1% of the slowest transporting cells were in G2/M-phase, and the remaining 58.2% were in S-phase.

### **3.3.4 Cell-cycle dynamics generate intercellular network heterogeneity and system complexity**

To further understand the collective effect of cell cycle heterogeneity on spatial patterning, we implemented intercellular diffusion as a conduit for initiating differentiation into an agent-based computational model in which individual cells divide, differentiate, and modulate their connectivity to neighboring cells according to state-specific rules [137]. The overall permeability between two cells was defined as the product of the individual base permeability ( $PM_n$ ) of each cell, reflecting the underlying biological mechanism of gap junction channel formation from neighboring connexon hemichannels [167]. To assess phenotypic differences in permeability, GAP-FRAP was performed on randomly selected cells within the differentiating populations after 24 and 48 hours of RA exposure. In agreement with the increased Cx43 signal detected after RA-treatment (Figure 14), there was a significant increase in the intercellular transport at both time points (Figure 16). As such, a higher base permeability was assigned to differentiated cells in the model (Figure 17a). The base permeability of each cell type was modulated by cell cycle state, with G1-phase having the highest permeability (100%  $PM_n$ ) and mitotic cells having the lowest (30%  $PM_n$ ). The degree of modulation was defined from the asynchronous distribution of recovery constants (Fig 3d) and cell cycle length previously established [137].

Differentiation was initiated via two means: a) through a stochastic process, or b) when a small (<1 kDa), diffusible molecule accumulated within a cell past a specified threshold. The small molecule was assumed to be produced by both cell types and to be capable of intercellular transport via concentration gradients. A higher production rate of the small molecule was assigned to differentiated cells to reflect a possible neuro-inductive positive feedback in their local environment. The variability in transport potential was

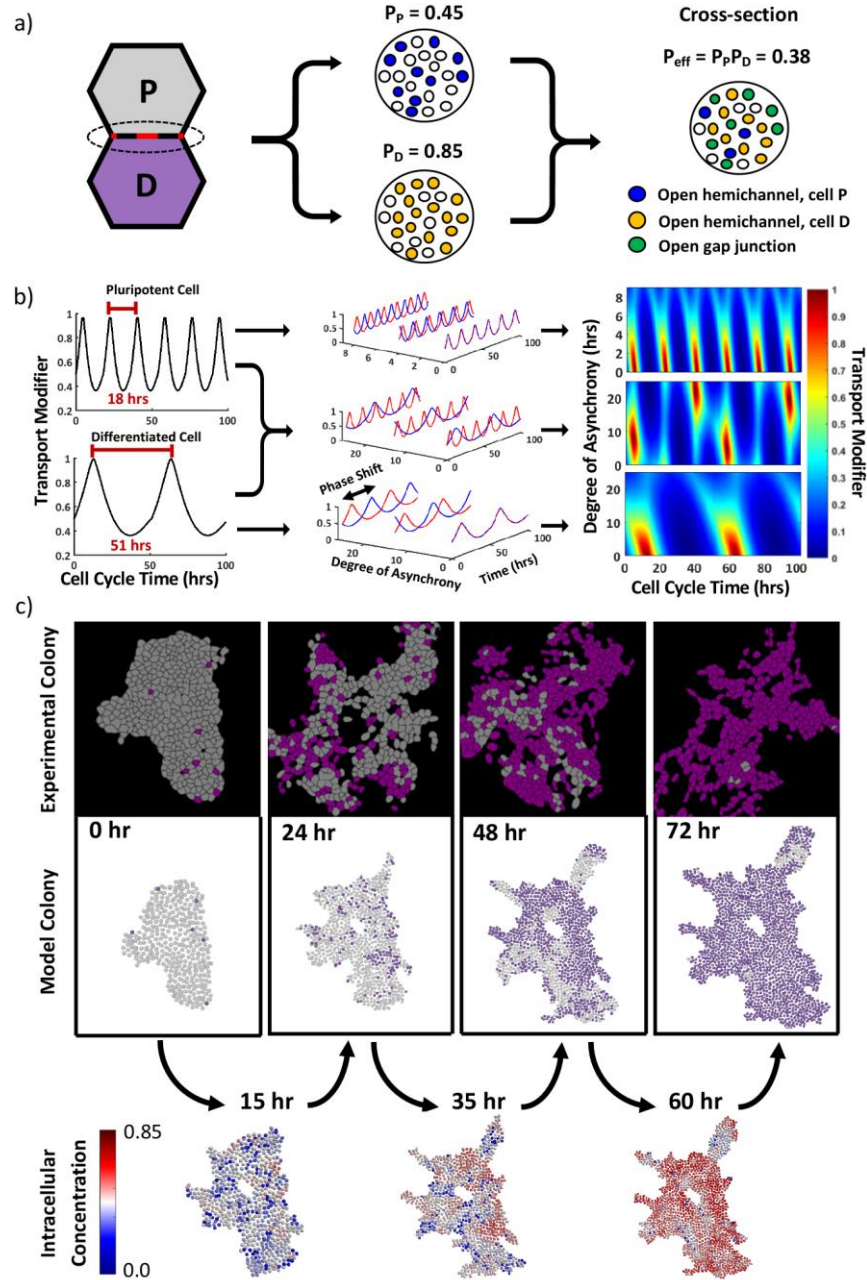
extensive with these differentiation and intercellular transport mechanisms implemented in the model.



**Figure 16: Comparison of recovery constants collected over the first 48 hours of retinoic acid (RA) exposure that were calculated from GAP-FRAP results. The measured cells were randomly selected within the populations at each time point. At both 24 and 48 hours of RA-treatment, there was a statistically significant ( $p < 1e-5$ ) increase in the rate of intercellular transport compared to the pluripotent population, denoted by the decrease in the recovery constant. By 48 hours, the recovery constants showed less overall variance in comparison to the pluripotent and 24 hour RA populations. Box plots show the median, the 25<sup>th</sup> and 75<sup>th</sup> percentiles, Tukey whiskers (median  $\pm$  1.5 times interquartile range), and outliers (+).**

Unique intercellular transport profiles were produced as a function of time from cells dynamically interacting with adjacent cells at various stages of the cell cycle and differentiation states (Figure 17b). In contrast, a nearly uniform intracellular distribution within the population occurred when cell cycle was synchronized in the model, followed by rapid differentiation with no pattern formation. While this result is exaggerated by the model constraints (i.e. intercellular communication as the primary differentiation mechanism), it emphasizes the role of the cell cycle in producing intercellular

heterogeneity within cell populations. Indeed, a cell cycle gating mechanism is sufficient for generating a dynamic intercellular network that is capable of pattern formation.



**Figure 17: Computational analysis of cell-cycle modulation on intercellular communication.**

(a) Each cell type (P, pluripotent; d, differentiated) has a base permeability, representing the average percent of gap junction hemichannels open at a cell-cell interface, with the total percent of open channels being the product of the two base permeabilities. (b) A function

for cell-cycle modulation over time was defined based on the transport trends noted in Fig 3, with a convolution of all possible transport profiles that two cells could experience over time for P-P, P-D, and D-D. (c) The intercellular transport model was implemented within an agent-based model and compared to digitized experimental colonies over time.

### **3.3.5 Intercellular diffusion dynamics generate spatiotemporal trajectories of differentiation**

The structure of the intercellular gap junction network rapidly established a non-uniform distribution of the small diffusing molecule in the model, starting from a relatively homogenous population of pluripotent cells randomly distributed throughout different phases of the cell cycle in the same ratio as an asynchronous population. Differentiation began in regions of the cellular network that accumulated high levels of the small diffusing molecule. The differentiated cells further modulated the intercellular network by having longer division times, thus longer G1- and S-phases, and instigated future differentiation patterns. Visually, the differentiation patterns produced by the model appeared similar to those observed in an experimental data set of time-lapse digitized images of Oct4 immunofluorescence (Figure 17c). To quantitatively compare similarities in the spatial patterning, representative metrics were extracted from both experimental and model data for dimension reduction by principal component analysis (PCA).

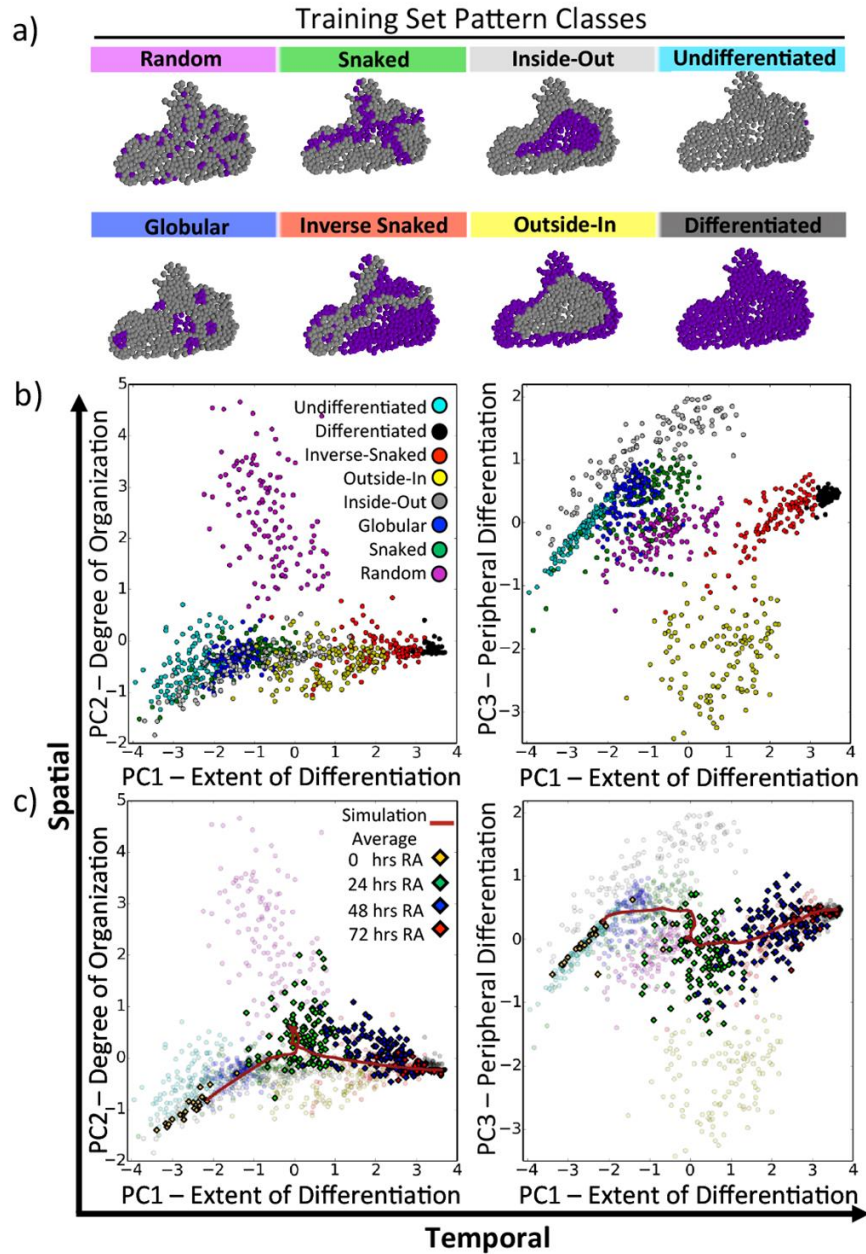
A training set of 2D patterns classes that occur during ESC differentiation were computationally produced for pattern classification and to define descriptive metrics (Figure 18a) [137]. An original set of 15 potential network metrics, where each metric represented a physical or spatial characteristic of the colony, were calculated from the 960 computationally defined patterns (8 patterns x 120 colony structures). PCA was performed on the set of network metrics, condensing the multivariate characterization of each pattern class into latent variables and conferring a unique ability to visualize similarities and relationships between the patterns. Seven of the metrics were found to be the most capable of separating the various pattern classes into discrete clusters when plotting the principal

component (PC) loadings (Figure 18b, Figure 9). Furthermore, the three PCs derived from the seven selected metrics each represent a particular property of the differentiation within a colony. Specifically, PC1 represents the extent of differentiation, with each pattern effectively being separated by the percent of differentiated cells in the colony. PC2 reflects the relative spatial organization, ranging from the Random pattern class to highly organized, single cluster patterns, such as Inside-Out and Outside-In. PC3 illustrates the locale of differentiation, or whether Oct4<sup>+</sup> cells tend to cluster on the periphery of a colony or in central regions with high cell density. The first three components accounted for ~88% of data variance (PC1: 59%, PC2: 15%, PC3: 14%) and provided a quantitative means of describing the spatiotemporal properties of differentiating cell populations.

The previously mentioned time-lapse images and simulation data of differentiating cell populations were converted to latent space by applying the trained-PCA transform, which is an effective means for analyzing dynamic morphogenic trajectories (Figure 18c) as we have previously reported [137, 138]. Our experimental data exhibited a temporal trajectory as it transitioned from an undifferentiated to a differentiated fate along PC1. Although there was large variability in patterning at 24 hours along PC2, with colonies spread across multiple pattern classes, the simulation data indicated that this phenomena represented a transitional state in the spatial organization of differentiation. Specifically, each simulation showed a slow accumulation of spatially heterogeneous differentiation within the population followed by a fast differentiation event, occurring ~24 hours, that forms clusters of differentiated cells. Differentiation preferentially occurred along the edges of colonies, as indicated by the data and trajectory approaching the Outside-In pattern class (Figure 18c, PC3). In addition to differentiation initially advancing along the outside of a colony, loss of Oct4 also tended to propagate between differentiated cells. The manifestation of this behavior became apparent at 48 hours where nearly every differentiated cell was connected as a single cluster. For example, in Figure 18c the edges of the experimental colony are almost entirely differentiated and strands of differentiated



cells form through the center to connect the edges. A similar pattern occurs within the intercellular diffusion model where paths of high concentration form and propagate differentiation between already differentiated cells. Thus, our model was able to accurately recapitulate the trajectory of differentiation using the intercellular network to inform cell fate decisions in a spatially organized fashion.



**Figure 18: Quantification of spatial patterning during RA differentiation. (a) The computationally generated pattern class structures used to train the principle component**

analysis, derived from [137], were applied to 120 experimental colony structures. (b) The seven selected metrics were calculated from each of the training set pattern classes (8 classes x 120 colony structures) and transformed into latent variable space through principal component analysis. PC1 represents extent of differentiation (temporal), and PC2 and PC3 represent organization/stochasticity and spatial locale, respectively (spatial characteristics). (c) The same metrics were calculated from experimental images of 0- (n=24), 24- (n=113), 48- (n=139), and 72-hour (n=22) RA-treated colonies and transformed into latent variable space. The average simulation trajectory was capable of capturing the spatiotemporal trajectory of the experimental data. At 24 hours there is a steep transition along both spatial axes, indicating that there is a gain in random differentiation and that it propagates along the edges of colonies. By 48 hours, the majority of differentiated cells are connected within a single, asymmetrical cluster.

### 3.4. Discussion

In this work, we characterize the spatial organization that occurs during early neural differentiation of pluripotent mESC colonies. Differentiation initiates a redistribution of Cx43 within the population and increased connectivity between differentiating cells. In addition to phenotypic changes, intercellular heterogeneity is modulated by cell cycle state. The pro-differentiation molecular gradients driving intercellular transport, as a function of cell cycle and phenotype, result in dynamic Oct4 patterning in our computational model that recapitulates the spatiotemporal trajectory of differentiation during early neural commitment.

We find that the intercellular network gains complexity as neural commitment occurs. Most cell types express a unique profile of connexin isotypes and, as differentiation occurs, both the composition and proportion of Cxs within the cell can change. Furthermore, Cx isotypes each have distinct permeabilities for diffusing molecules and capacity for forming channels with other isotypes. The result is an intercellular network

capable of phenotypic-compartmentalization, opposing gradients of small molecules, and even unidirectional transport [175, 201]. Also, most Cxs are susceptible to post-translational modifications that affect GJ formation, degradation, and permeability [214]. As such, intercellular transport can be modulated during differentiation both as a function of transcription and Cx activity. The rapid and sequential gain and loss of Cx43 membrane localization between 24 and 48 hours of RA treatment suggests that the increased connectivity in Oct4-diminished clusters was predominantly due to a redistribution rather than a change in expression of Cx43. We hypothesize that enhanced assembly of Cx43 via cAMP/PKA activation, which has been linked with RA treatment, produced the redistribution [215]. In terms of the cytoplasmic accumulation of Cx43, the locale and staining pattern agrees with previous studies demonstrating localization in the Golgi apparatus [211]. Since Cx43 connexons are formed in the Golgi network before being trafficked to the membrane [211, 214], an accumulation in the Golgi network would agree with enhanced assembly. Despite the decreased number of cells at 48 hours displaying enhanced Cx43 signal, a faster intercellular transport rate was maintained compared to pluripotent cells (Figure 13). Therefore, it is possible that either Oct4- cells transcribe a separate Cx isotype or that the prolonged cell cycle associated with differentiation masks the decreased Cx43 expression relative to the Oct4-diminished cells. Specifically, the extended G1/S-phase relative to M-phase associated with differentiation would also cause the mean transport rate to be faster.

We developed a multicellular model that describes cell fate transitions coupled with intercellular communication that is regulated by the dynamics of connexins as a function of the cell cycle. Although previous computational models of intercellular communication have included the biophysics of gap junction transport, they have primarily focused on small molecule diffusion through individual channels [216, 217]. One notable multicellular model that focused on ion transfer relied upon membrane potential to regulate intercellular connectivity [218]. While capable of examining the development of intracellular gradients

at a multicellular level, the bioelectric model is constrained to a static system. A primary advantage of our selected framework is its ability to investigate the impact of dynamic intercellular topologies on a multicellular system that includes growth, division, and differentiation. The dynamic network facilitates the study of evolving spatial organization and the resultant effect on the formation of intracellular gradients.

Interestingly, a recent paper proposed an alternative, hidden Markov model framework that explored the role of stochasticity during early neural differentiation, but without consideration of spatial organization and cell-cell communication [219]. While the mechanistic framework is different from the work presented here, they share several similarities. In particular, both models consider differentiation to be an autonomous process with discrete state changes that are probabilistic. The probability of a state transition in our model, however, is dependent on the accumulation of an intracellular metabolite within a cell rather than the purely stochastic progression through microstates. The description of microstates typically considers gene and protein expression patterns, but in the present work the metabolic profile of a cell also impacts susceptibility to differentiation signals [220, 221]. The cellular communication integrated within our model via gap junction-mediated transport leverages both stochastic state transitions (consistent with Stumpf et al.) and the higher-order collective behavior of the multicellular system, suggested by our observations of pattern propagation within pluripotent colonies. Therefore, our model both complements this prior modeling study while also advancing the description of spatial organization effects on the emergence and propagation of micro- and macrostates during early neural differentiation.

The superposition in latent variable space between experimentally derived results and agent-based simulations is a powerful feature of the analysis pipeline presented here. The extraction of network metrics is particularly necessary for the accurate quantification of systems with variable and asymmetrical structures, such as during unconstrained monolayer growth (Figure 17c). Since each colony structure can have a unique

configuration, the spatial characteristics used to distinguish pattern classes need to be independent from the underlying morphology. Network metrics accomplish this by evaluating relationships between neighboring nodes/cells and colony-wide. Furthermore, the transformation of each multivariate set of metrics to latent space enables visual and numerical comparison between individual colonies. The training set of artificially generated colony patterns provides the ability to directly map regions of latent space to specific spatial organizations (e.g. the “random” cloud versus a highly organized “outside-in” or “inside-out” cloud). Accordingly, transformed sets of colony metrics from both experimental and computational sources can be delineated according to their proximity to each mapped pattern class in latent space. Computational models therefore can be designed to replicate morphogenic trajectories in a quantifiable manner and elucidate the factors that influence specific pattern formations.

In summary, a computational model for investigating multiscale intercellular communication was produced that demonstrates intercellular heterogeneity can be generated by dynamic fluctuations in GJ permeability via the cell cycle. The resultant dynamic intercellular network is capable of producing similar spatial patterning to that observed during early neural differentiation and provides a means for predicting the influence of intercellular perturbation on spatial patterning.. Ultimately, identifying the effect of molecular regulators of differentiation at a multicellular scale is an initial step towards replicating morphogenic events *in vitro* and successfully directing the emergence of desired engineered multicellular structures.

## CHAPTER 4      THE EFFECT OF INTERCELLULAR PERTURBATION ON DIFFERENTIATION POTENTIAL

### 4.1.    Introduction

Intercellular communication has the capacity to form a rich variety of intracellular gradients, but it is unknown to what extent these gradients affect pattern formation relative to other forms of communication. The deletion of both Cx43 and Cx45 in mouse ESCs was reported to disrupt the formation of primitive endoderm within differentiating aggregates [198]. In a more extreme example, the inhibition of GJ channels caused decapitated planarian flatworms to regenerate with head and brain morphologies that were similar to *other* flatworm species [33], emphasizing the potential of intercellular communication to alter morphological features. With that in mind, exploring the capacity of intercellular communication to modulate differentiation could lead to significant control over patterning events. By understanding the influence of a given cell on neighboring cells, it could be possible to direct differentiation in a spatially localized manner. Such a feat would have significant implications for the derivation of defined multicellular systems

In this Chapter, the intercellular communication of pluripotent cells was perturbed during neural differentiation using various methods. A small molecule was used to inhibit adenylyl cyclase and prevent cAMP production. While this perturbation has effects beyond influencing differentiation or intercellular transport, there was strong evidence for cAMP involvement during neural differentiation. A different small molecule was used to globally inhibit gap junctions. The non-specific inhibition of gap junctions was a necessary perturbation because if intercellular communication was affecting differentiation then the complete prevention of communication should have a significant impact, with the caveat being that adaptation to the inhibitor could occur. Lastly, Cx43 was knocked-down using a CRISPRi system. The advantage of this approach was the specificity of the target

compared to the small molecule perturbations. Computational modeling was used in conjunction with the small molecule perturbations to study intercellular communication as a differentiation mechanism, which allowed intercellular communication to be studied in isolation from other types of communication. Overall, this work emphasizes the role of intercellular communication during differentiation. We report the previously unrecognized capability of intercellular communication to delay differentiation.

## **4.2. Methods**

### **4.2.1 Experimental Methods**

#### **4.2.1.1 Mouse ESC Perturbation Experiments**

For both cell lines, D3 and RW4, the same plating and differentiation protocols were used as described in the Chapter 3 Methods but with minor adjustments. The concentration of retinoic acid was decreased to 0.3  $\mu$ M and either  $\beta$ -glycyrrhetic acid (Sigma-Aldrich) or SQ22536 (R&D Systems) was added to a final concentration of 50  $\mu$ M or 100  $\mu$ M, respectively. The DMSO concentration for each condition and the vehicle control was set at 0.1%. The media was changed every 24 hours with the inhibitors added immediately prior to ensure proper concentrations and activity. The same immunofluorescence and image analysis methods were used here as described in Chapter 3 Methods.

#### **4.2.1.2 Human Induced Pluripotent Stem Cell Culture**

A human induced pluripotent stem cell (hiPSC) line, LBC2-GJA1, was derived from the WTC11 cell line, courtesy of Ashley Libby and David Joy at UCSF. The LBC2-GJA1 line has constitutive expression of GFP-Lamin B and a CRISPRi system targeting the GJA1 (Cx43) gene. hiPSCs were cultured in 6 well plates coated with growth-factor reduced Matrigel® (Corning) in mTeSR™1 (STEMCELL Technologies) media and passaged

every 3-4 days upon reaching ~70% confluency. For passaging, cells were washed with PBS followed by disassociation with Accutase (EMD Milipore) for 5 minutes. The Accutase-cell mixture was diluted 1:5 in PBS and centrifuged for 3 minutes at 200g. The cells were seeded at a 10000 cell/cm<sup>2</sup> density in mTeSR™1 with 10 µM of the Rho-associated protein kinase inhibitor (Ri), Y-27632 (Selleck Chemicals). The media was changed every 24 hours with mTeSR™1 not containing Ri.

#### 4.2.1.3 Immunofluorescence

For staining of Oct4 and Pax6, cells were fixed in 4% paraformaldehyde (Thermo Fisher Scientific) for 15 minutes and then permeabilized and blocked with 0.3% Triton X100 in 5% normal donkey serum (NDS) for 60 minutes. Samples were incubated overnight at 4°C with the following primary antibodies in 5% NDS: goat polyclonal Oct-3/4 (Santa Cruz, 1:200) and mouse monoclonal Pax6 (DSHB, 1:20). After washing, samples were incubated in a 5% NDS secondary antibody solution of donkey anti-goat Alexa Fluor 568 (Thermo Fisher Scientific, 1:200) and donkey anti-mouse Alexa Fluor 647 (Thermo Fisher Scientific, 1:200) for 45 minutes before counterstaining with Hoechst 33342 (1:1000) in DI water for 10 minutes. All images were collected on a PerkinElmer UltraVIEW VoX spinning-disk confocal microscope with a sCMOS camera at 10X magnification.

For Cx43 and Oct4 staining, the same procedure was utilized but with the following changes: The primary antibodies were goat polyclonal Oct-3/4 (Santa Cruz, 1:200) and rabbit polyclonal Cx43 (Sigma, 1:400), and the images were collected at 20x magnification.

#### 4.2.1.4 hiPSC Neural Differentiation

Dual-SMAD inhibition was utilized to differentiate the hiPSCs towards the neural lineage. The LBC2-GJA1 cells were seeded at a 10000 cell/cm<sup>2</sup> density in mTeSR™1 with 10 µM Ri on ibidi 8-well µ-slides coated with growth-factor reduced Matrigel®. The media was changed every 24 hours with mTeSR™1 not containing Ri. After 3 days, the media was



supplemented with 10  $\mu$ M SB431542 (Sigma-Aldrich) and 2  $\mu$ M dorsomorphin (Sigma-Aldrich). Differentiation proceeded for 6 days following the addition of the two inhibitors. To induce knockdown of the GJA1 gene, the media was supplemented with 2  $\mu$ M doxycycline (Sigma-Aldrich) throughout the entire protocol, from seeding to fixation.

## **4.2.2 Computational Methods**

### **4.2.2.1 Image Analysis**

For all cell lines, mESC and hiPSC, the same image analysis was used as described in Chapter 3. However, for the hiPSC cell lines the red channel represented Pax6 rather than Sox2.

## **4.3. mESC Results**

### **4.3.1 Selected Perturbations**

Nocodazole synchronization was originally proposed as a fundamental perturbation to test the predictive capability of the intercellular differentiation module. However, this proved to be technically unfeasible since cell cycle asynchrony is regained rapidly after short-term nocodazole treatment and long-term nocodazole treatment decreases the expression of pluripotency markers at a similar time scale as RA treatment [222]. Instead, two other perturbations were selected: (i) inhibition of the production of an abundant small molecule that could be represented by the generic diffusible molecule in the computational model, and (ii) global inhibition of gap junction channels.

#### **4.3.1.1 Adenylyl Cyclase Inhibition**

Cyclic AMP (cAMP) is a canonical secondary messenger that has been directly associated with creating spatially diverse intracellular distributions as a function of intercellular communication [6]. It was hypothesized that the distribution of cAMP plays a role in the spatial aspects of RA-accelerated differentiation since RA treatment can activate

PKA, a cAMP dependent kinase [215]. While cAMP is involved in multiple cellular processes, the availability of reagents to perturb intracellular cAMP levels makes it an attractive molecular target for investigating GJ-mediated communication and testing the validity of our model. Adenylyl cyclase (AC) was inhibited with SQ22536 (SQ) to prevent the conversion of ATP to cAMP, and consequently decrease the intracellular cAMP concentration [223].

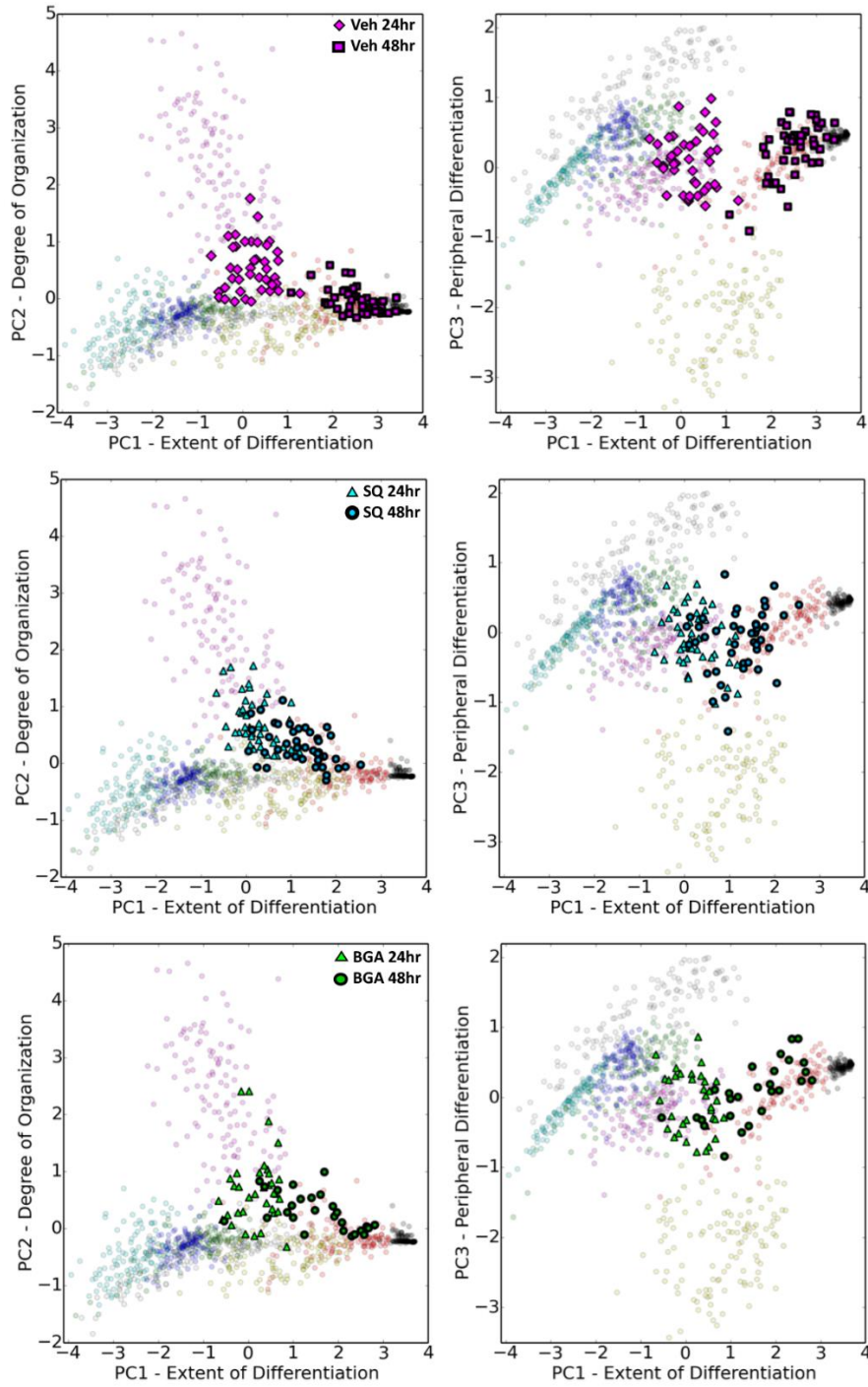
#### 4.3.1.2 Gap Junction Inhibition

A few molecules with different mechanisms of action have commonly been used to globally inhibit gap junction channels within cell populations. Some molecules, such as heptanol, cause GJ channels to uncouple by inducing changes in the structure of the lipid membrane. The insecticide lindane initiates an acute internalization of GJ channels followed by a longer delayed-onset inhibition, both mediated by oxidative stress. In contrast, 12-O-tetradecanoylphorbol-13-acetate (TPA) and  $\beta$ -glycyrrhetic acid ( $\beta$ -GA) modulate the phosphorylation state of connexins and affect the localization and assembly of gap junctions. The small molecule  $\beta$ -GA was deemed the least harsh of these chemical treatments and was selected as the means for GJ inhibition.

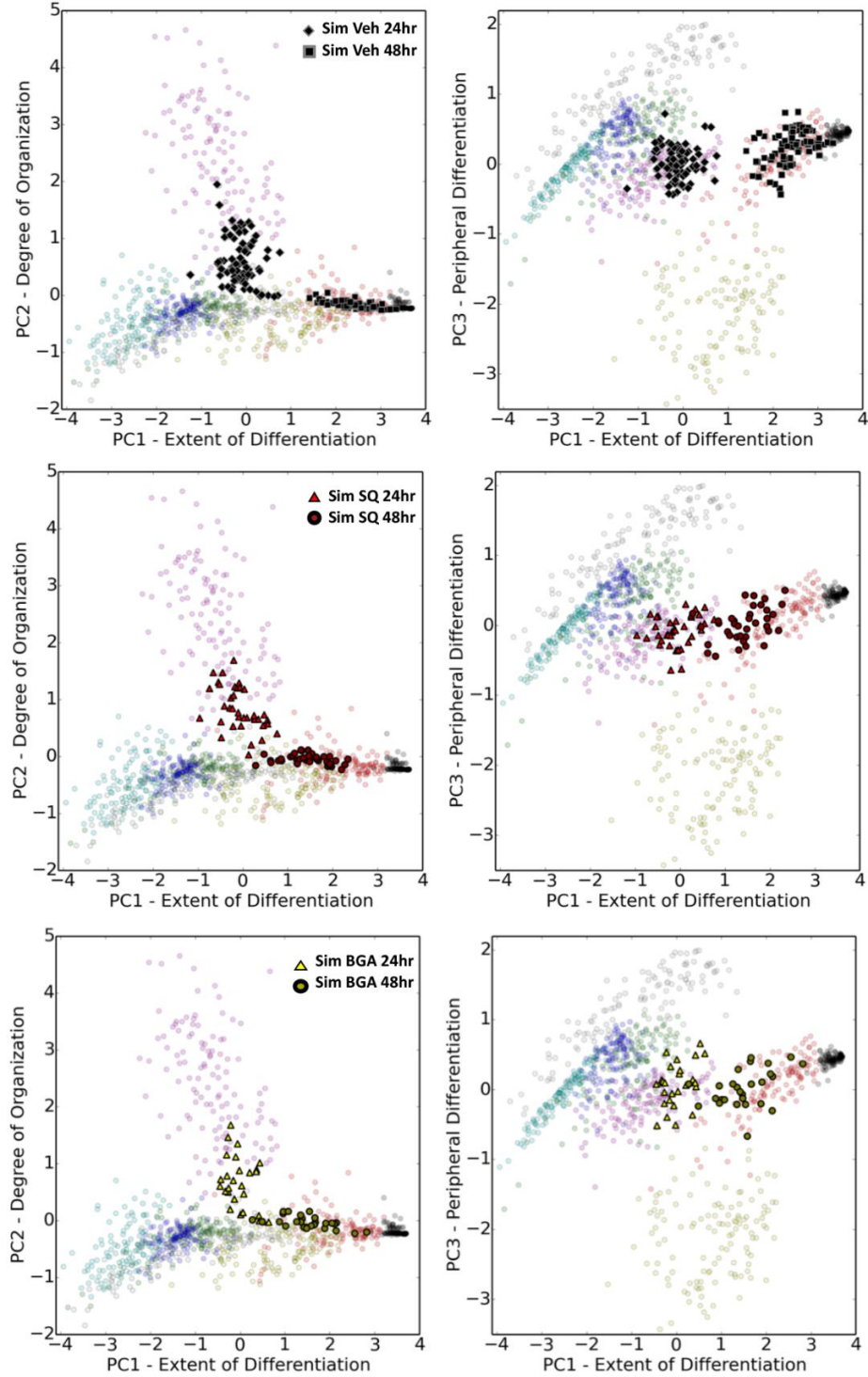
#### **4.3.2 Perturbations to the D3 mESC Cell Line**

The initial perturbation studies were conducted with the D3 mESC cell line that the model parameters were derived from. Therefore, computational simulations were used iteratively with experiments to gain insight both about the biological system and the model. Specifically, due to the abstraction of differentiation potential into a single intercellular molecule, the degree of experimental perturbation does not correspond exactly with the degree of perturbation within the model. A simple example would be considering two intercellular molecules that equally affect differentiation: if only a single molecule is

inhibited, the other molecule still exerts its influence. Therefore, full inhibition of the molecule in the biological system only maps to 50% inhibition in the model. As such, the model was originally able to provide coarse predictions of expected behavior followed by more detailed information from refining the degree of modulation to fit the experimental data. To implement cAMP inhibition, the production constant (defined in Section 3.2.2.1) was decreased by 14% in differentiated cells. The reason it was only decreased in the differentiated cells was also in response to the abstraction. The base production constant for pluripotent cells was set such that the steady state concentration within a cell would be below the differentiation threshold by a specific margin. This steady state concentration can be interpreted as representing the “naïve” pluripotent state and it becomes “primed” as it approaches the differentiation threshold. Therefore, in this abstraction, decreasing the production rate in pluripotent cells would be equivalent to an increase in pluripotency, which seemed unlikely in the biological system. To simulate GJ inhibition, the nascent permeability (defined in Section 3.2.2.1) was decreased by 9% for both cell types. The coarse prediction for inhibition of cAMP production was that SQ treatment will decrease the rate of differentiation and, depending on the extent of its influence, an increased residence time in the ‘Random’ pattern class. Simulations of GJ inhibition by  $\beta$ -GA yielded a similar delay in the differentiation rate and minimal variation in the spatial organization or locale of differentiation, comparable to the results from inhibiting cAMP production. The data clouds for experimental and fit simulations are depicted in Figure 20 with the average values shown in Figure 21.

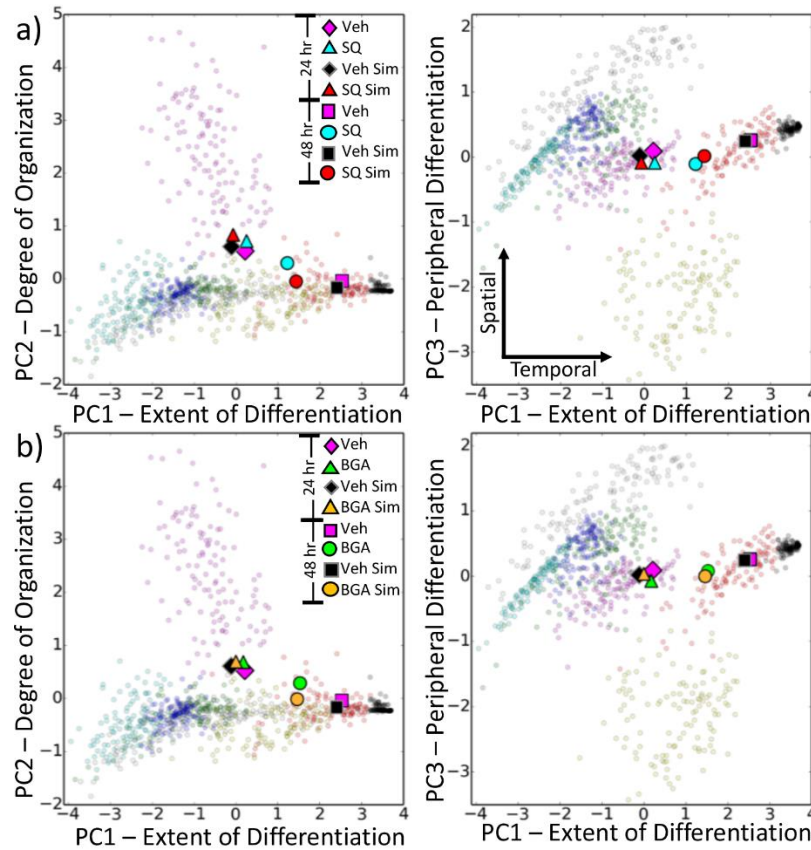


**Figure 19: Experimental colonies of D3 mESC cell line at 24 and 48 hour time points of retinoic acid treatment Averages of these values are illustrated in Figure 21.**



**Figure 20: Simulation colonies at 24 and 48 hour time points of retinoic acid treatment for the Vehicle control (n=75), SQ-treatment (n=30), and  $\beta$ -GA-treatments (n=30 converted into latent space. Averages of these values are illustrated in Figure 21.**

The stochasticity and peripheral differentiation in SQ-treated colonies increased slightly at 24 hours compared to the vehicle DMSO control. Furthermore, SQ-treated colonies showed significantly delayed differentiation between 24 and 48 hours, retaining a large proportion of pluripotent cells. The simulation of SQ-treatment in the model resulted in similar dynamic changes and comparable distributions of spatial patterns at 24 and 48 hours relative to the experimental data (Figure 22a).



**Figure 21: Perturbation to the intercellular network of a multicellular D3 ESC population affects RA-accelerated differentiation in a temporal manner. At 24 hours, neither adenylyl cyclase (AC) inhibition (a, SQ-treatment) nor gap junction (GJ) inhibition (b,  $\beta$ -GA-treatment) induced a significant change in spatial or temporal characteristics of differentiation compared to the vehicle control. By 48 hours, a temporal shift along PC1 is observed for both treatments (a,b), depicting a decrease in the rate of differentiation. The intercellular model accurately predicted these dynamics, represented by average values for**

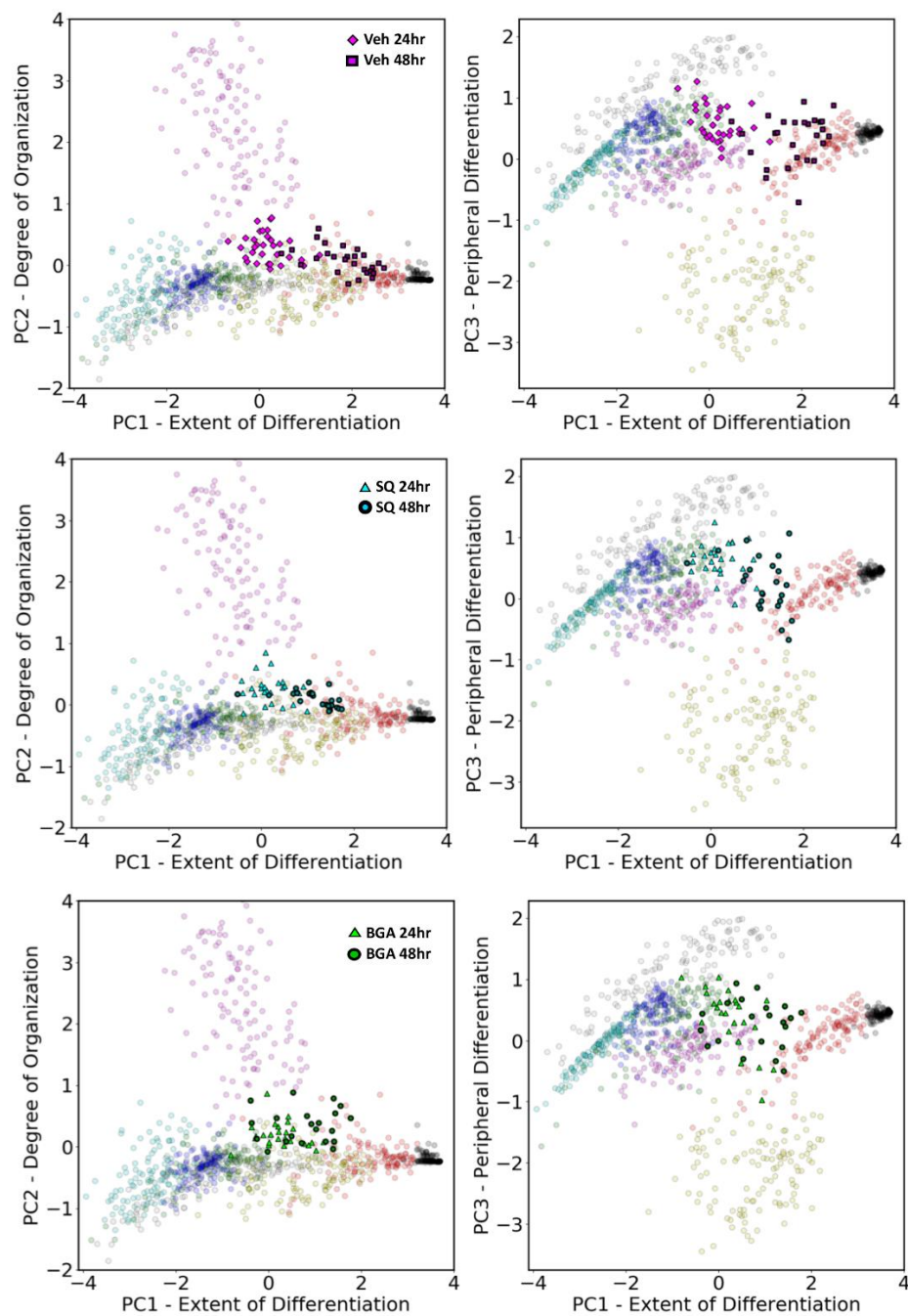
**SQ-treatment (Exp: n=87, Sim: n=30) and  $\beta$ -GA-treatment (Exp: n=62, Sim: n=30)  
compared to the vehicle control (Exp: n=87, Sim: n=75) (a,b).**

To test the degree of influence from the intercellular network, we investigated the effect of gap junction inhibition during differentiation. The experimental perturbation to GJ-channels by  $\beta$ -GA resembled the outcome predicted by the simulations, with differentiation delayed at 48 hours in a manner similar to, but less potent than, AC-inhibition (Figure 21b).

Collectively, these results lead to several conclusions: i) delaying differentiation by SQ-treatment implies that cAMP has a role in stimulating neural differentiation during RA-treatment; ii) the ability of the model to capture the temporal change indicates that this role is at least partially mediated by the intercellular network; and iii) the minor spatial changes upon perturbation reflect that the topology of the gap junction network remains the same despite the slower differentiation rate.

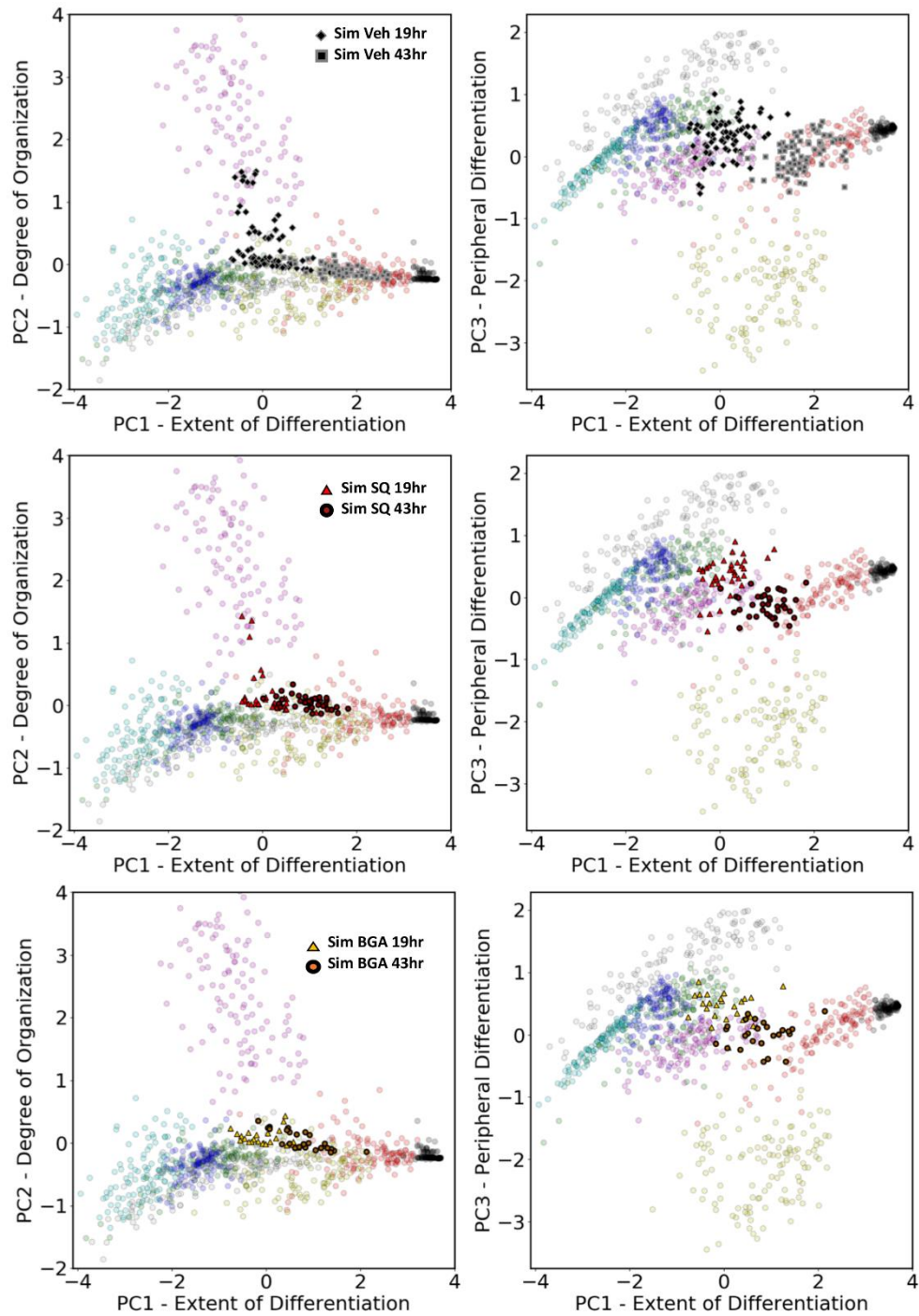
#### **4.3.3 Perturbations to the G-Olig2 RW4 mESC Cell Line**

As further validation, both perturbations were applied to ATCC G-Olig2 cells, derived from the RW4 ESC line (Figures 22-24). In both conditions, the temporal delay was recapitulated, with  $\beta$ -GA having a slightly more significant effect than the SQ-treatment. Interestingly, the intercellular model was able to accurately predict each condition from the 19- and 43-hour time points of the D3 simulations, indicating that the RW4 cells are either inherently more pluripotent than D3 cells or that they are less sensitive to RA-treatment.

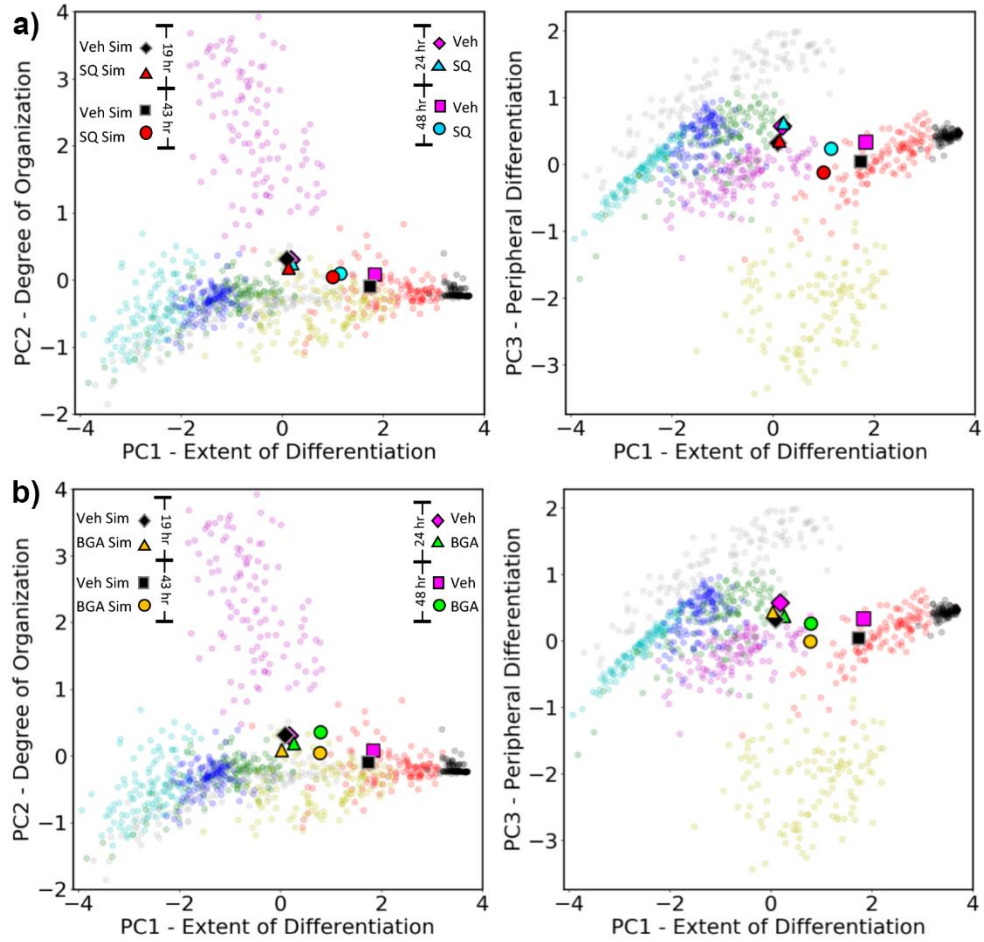


**Figure 22: Experimental colonies of RW4 mESC cell line at 24 and 48 hour time points of retinoic acid treatment Averages of these values are illustrated in Figure 24.**





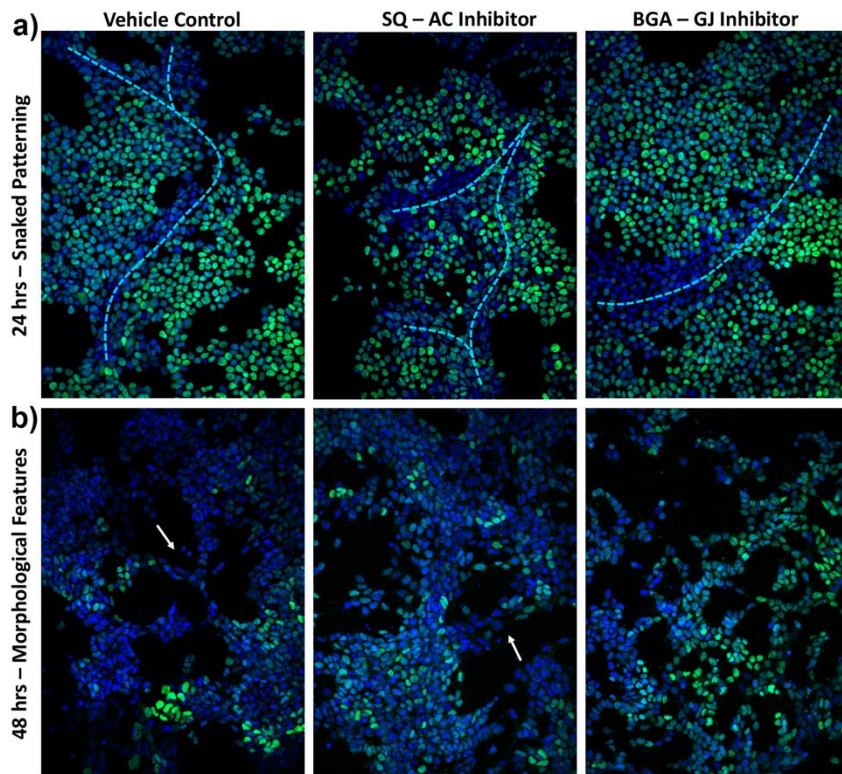
**Figure 23: Simulation colonies at 19 and 43 hour time points of retinoic acid treatment for the Vehicle control (n=75), SQ-treatment (n=30), and BGA-treatments (n=30 converted into latent space. Averages of these values are illustrated in Figure 24.**



**Figure 24: Perturbation to the intercellular network of a multicellular RW4 ESC population affects neural differentiation in a temporal manner. Similarly to Fig 21, neither adenylyl cyclase (AC) inhibition (a, SQ-treatment) nor gap junction (GJ) inhibition (b,  $\beta$ -GA-treatment) induced a significant change in spatial or temporal characteristics of differentiation compared to the vehicle control at 24 hours. By 48 hours, a temporal shift along PC1 was observed for both treatments (a,b), but with a more significant decrease in the differentiation rate for GJ-inhibition. Furthermore, GJ-treatment at 48 hours produced an increase in stochasticity compared to the vehicle and SQ-treatment groups. In comparison to Fig 22, the RW4 tended to differentiate in a more “inside-out” manner than the noted ‘outside-in’ differentiation of the D3 ESC line. Interestingly, the intercellular model suggests that RW4 cells differentiate slower than D3 cells in response to RA, as**

indicated by the accurate prediction of both treatments at the simulated 19 and 43 hour time points. (Veh: n=59, SQ: n=49, BGA: n=46)

A few other notable differences between the cell lines became apparent in our analysis. For instance, the RW4 colonies appeared to be less specific for initiating differentiation along colony edges. However, the propagation of differentiation between clusters remained a prominent spatial feature, as demonstrated by the frequent emergence of the ‘Snaked’ pattern class (Figure 25a). Also, GJ inhibition by  $\beta$ -GA induced a pronounced morphological change in RW4 colonies at the 48 hour time point (Figure 25b). Specifically, a web-like pattern consisting of strands that were 1-2 cells thick formed and was interpreted in latent space as an increase in stochasticity.



**Figure 25: RW4 colony patterning and morphology after 24 and 48 hours of retinoic acid exposure, respectively, with adenylyl cyclase or gap junction inhibition compared to the vehicle control. At 24 hours, obvious examples of the snaked pattern class emerge**

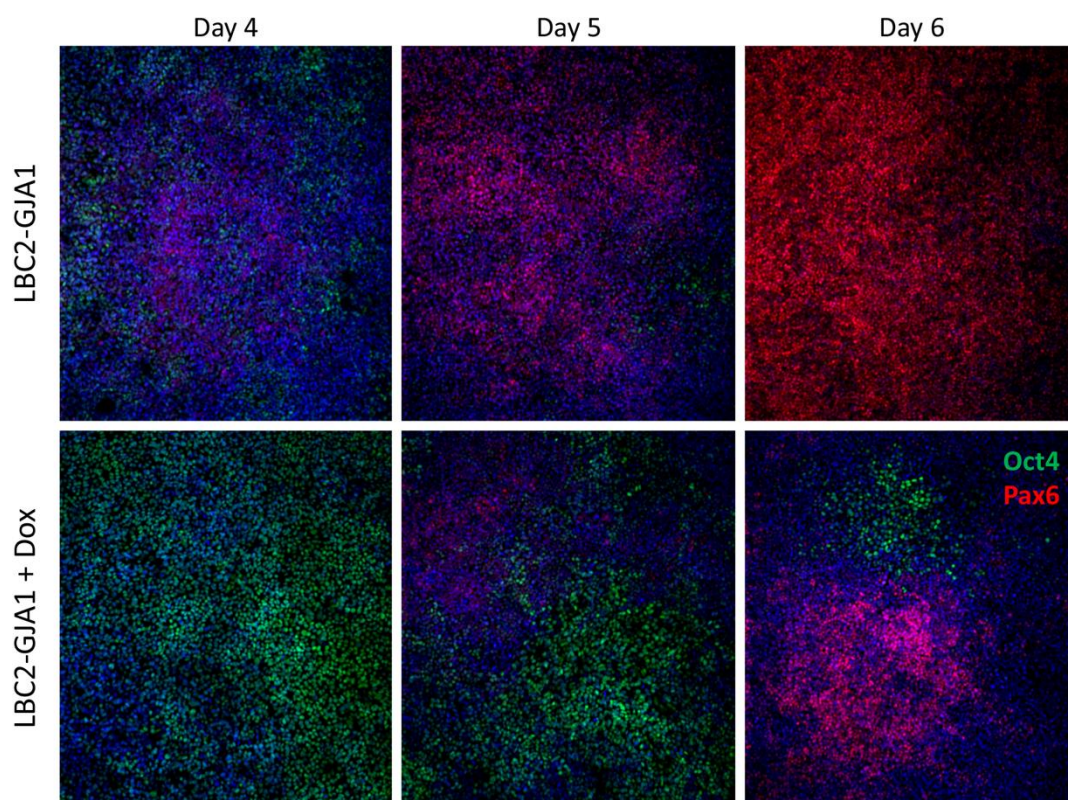
(demonstrated by dotted lines). (b) At 48 hours, both the Vehicle and SQ colonies tend to form larger cell clusters with a few smaller extensions between clusters (indicated by white arrows). In contrast, after GJ inhibition the majority of cells in the population form a network of branching extensions and develop numerous circular pockets. Oct4: Green, Blue: Hoechst.

## 4.4. hiPSC Results

### 4.4.1 Connexin43 Knockdown

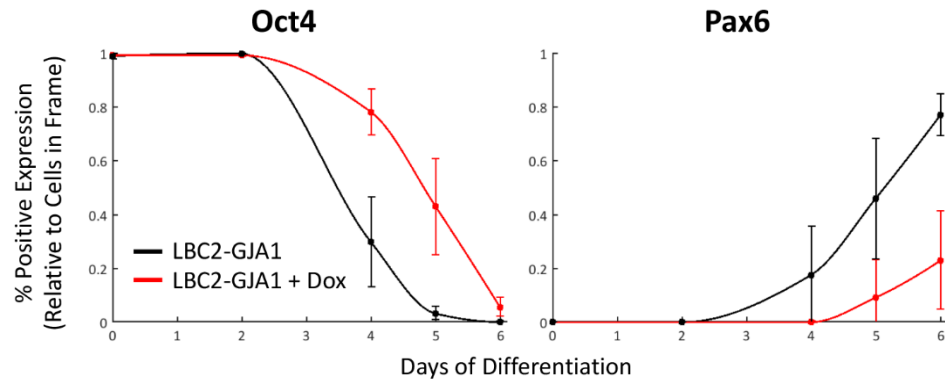
To selectively perturb the intercellular network, Cx43 was knocked down in an hiPSC cell line prior to neural differentiation through a doxycycline-induced CRISPRi system [224]. For ease of reading, the LBC2-GJA1 and LBC2-GJA1+Dox conditions will be referred to as WT and Cx43-KD, respectively. Differentiation was initiated by dual-SMAD inhibition after the cell culture had reached 80-90% confluency, which increases the efficiency of neural differentiation [225]. As such, the collected images are not of individual colonies but rather a number of individual frames collected from one large population. Oct4 and Pax6 were selected as markers of the differentiation status in this set of experiments. We observed divergence of expression in these markers between the WT and Cx43-KD cells, with the Cx43-KD maintaining a high level of Oct4 cells compared to the WT at Days 4 and 5 (Figure 26). The WT showed significant and increasing expression levels of Pax6 that peaked at Day 6, whereas the Cx43-KD had minimal Pax6 expression until Day 6 with regions of Oct4+ cells still present.





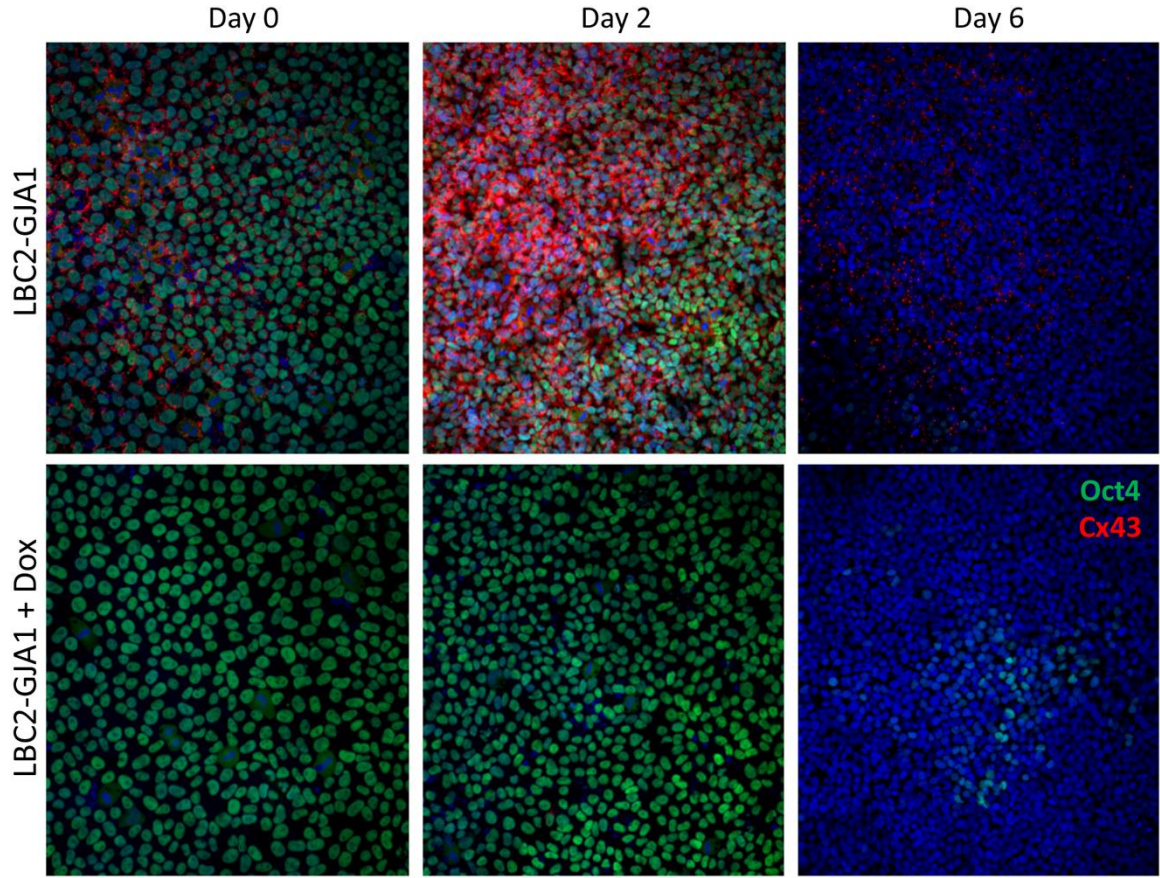
**Figure 26: Dual-SMAD inhibition in the hiPSC line LBC-GJA1 with and without doxycycline for days 4 through 6.**

To quantify this observation, the images were analyzed in CellProfiler as described in 3.2.2.2 and the average number of Oct4 and Pax6 positive cells within each image was calculated. Comparing the means for each day confirmed the image observations, with the Cx43-KD exhibiting a statistically significant delay in differentiation compared to the WT (Figure 27).



**Figure 27: Average Oct4 and Pax6 expression over 6 days of dual-SMAD inhibition in the hiPSC line LBC2-GJA1 with (Cx43-KD) and without (WT) doxycycline.**

In parallel with these experiments, the expression of Cx43 within the populations was assessed to validate the knockdown over the course of differentiation (Figure 28). At every time point examined, the Dox treatment very effectively knocked down the expression of Cx43 to nearly negligible levels (<2%). Interestingly, at Day 2, the WT cells exhibited a similar increase in expression of Cx43 as was observed in mESCs in Figure 13. Also in agreement with the mESC Cx43 expression data was that the increased Cx43 appeared to be occurring in cells with diminished Oct4 and was eventually lost. In fact, by Day 6 there is an apparent decrease in Cx43 in the WT population.

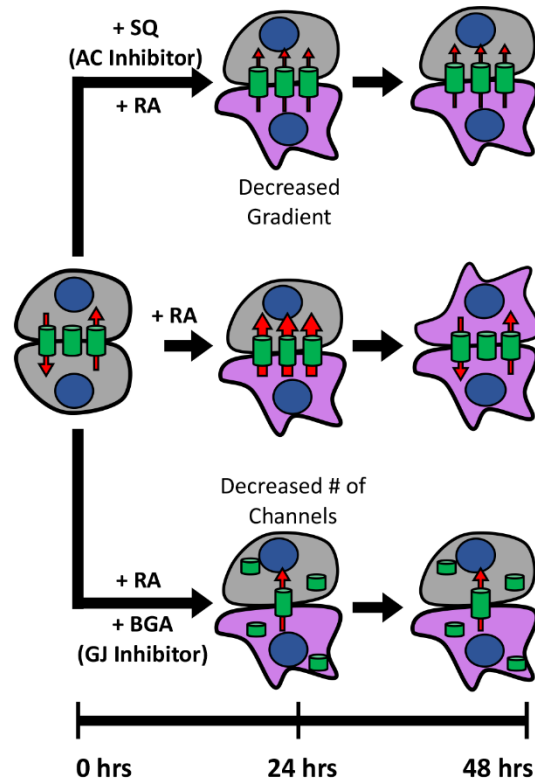


**Figure 28: Cx43 expression during dual-SMAD inhibition in the hiPSC line LBC-GJA1 with (Cx43-KD) and without (WT) doxycycline for days 0, 2 and 6.**

#### 4.5. Discussion

While the ability to direct differentiation is undeniably an essential step in the goal of generating ELS, it could be argued that inhibiting the propagation of *undesirable* differentiation plays an equally important role. In this work, we find that perturbing intercellular communication by disruption of GJ formation, inhibiting levels of a critical molecule transported via GJCs, or directly inhibiting a specific connexin protein can delay the propagation of differentiation. Furthermore, this temporal modulation was consistent across three different cell lines, two different species, and three different mechanisms of perturbation. There are several potential mechanisms that could be behind this modulation.

The most obvious mechanism involves the transfer of an intracellular species, possibly cAMP, that acts as a pro-differentiation cue. The similarity in response to the two small molecule inhibitors can be attributed to the decreased intracellular accumulation of this pro-differentiation molecule caused by limiting the intercellular transport, either by decreasing the number of active channels ( $\beta$ -GA) or by decreasing the concentration gradient between cells (SQ) (Figure 29).



**Figure 29: A schematic diagram of our proposed mechanism for the influence of AC and GJ inhibition on differentiation potential. Specifically, both AC and GJ inhibition are suggested to decrease the intercellular flux between cells but via separate mechanisms: modulating the concentration gradient and the number of open channels for AC and GJ inhibition, respectively.**

Other interactions exist between cAMP and Cx43 that could also explain the similarity. For instance, the PKA/cAMP pathway is directly involved in the enhanced assembly of Cx43. In this case, cAMP inhibition would indirectly cause GJ inhibition. Hence, the



mechanism would still be reliant on the diffusion and spatial gradient of a pro-differentiation molecule, but cAMP would predominantly affect the transport rather than the differentiation. Another possibility is that the loss of GJ connectivity induces a morphological rearrangement, similar to that observed in the RW4 cells after  $\beta$ -GA treatment (Figure 25). By changing the cell-cell arrangement, sensitivity to extracellular signaling cues can change in response to receptor localization, akin to the TGF $\beta$  receptor lateralization in densely packed hiPSC colonies [226]. However, morphological differences were not observed in the colony morphology of the Cx43-KD cells or to any significant extent in the GJ-inhibited D3 cells. Therefore, changes in the morphological arrangement are unlikely to be the primary cause of the delayed differentiation, but they may have been a secondary factor in the RW4 cell line.

The repeated phenomenon of enhanced Cx43 expression during early differentiation of mESCs and hiPSCs suggests that increased intercellular communication is a typical feature as pluripotency is lost. We had originally proposed that the increased Cx43 in mESCs was due to enhanced assembly, with the hypothesis that retinoic acid was activating the cAMP/PKA pathway. However, since neither of the SMAD inhibitors are known to cause PKA activation, an alternative mechanism must exist. The concomitance with the loss of Oct4 in both scenarios, despite different treatments and cell lines from different species, suggests that the change is genetic. Although, it is unknown whether this transient transcription of Cx43 is a ubiquitous feature of early differentiation or if it is specific to early neural commitment. In either case, these results indicate that early differentiation events involve dynamic regulation of gap junction communication. A later stage of neural differentiation has also been linked to dynamic modulation of intercellular communication [211]. Interestingly, transport rates were found to transiently decrease in the first day of neural progenitor differentiation then increase rapidly in the following days of differentiation. Therefore, the regulation of intercellular communication during

differentiation appears to be a general mechanism of coordination not limited to pluripotent cells.

## **CHAPTER 5      MODULAR COMMUNICATION IN AGENT-BASED MODELS**

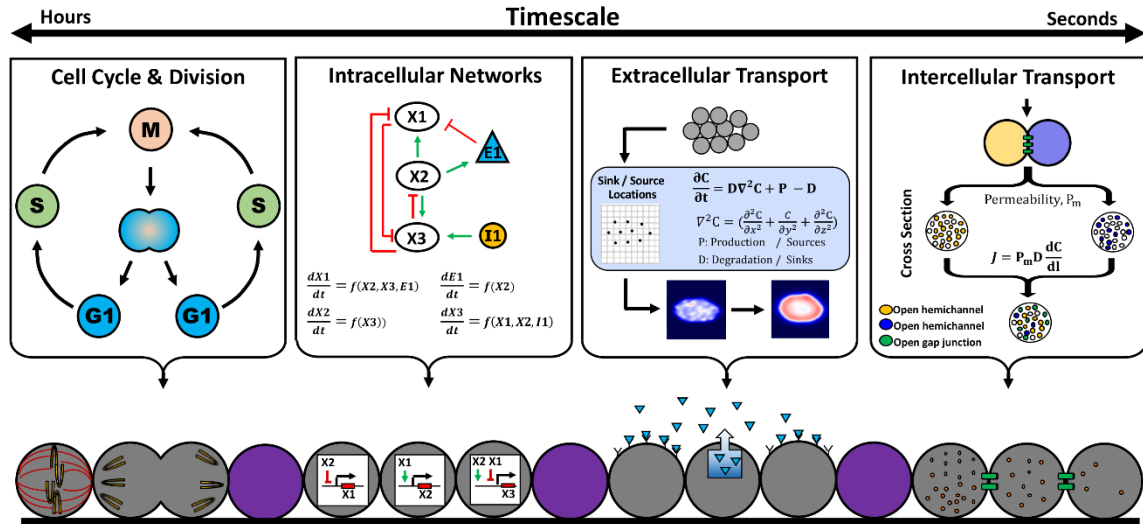
Within an evolving multicellular system, there are numerous opportunities for spatial patterns to emerge across a variety of scales. From macro to micro scales, spatial features can occur in the overall morphology of the cell population, the organization of cells within the population, gradients of extracellular molecules, and even intracellular gradients between neighboring cells. Furthermore, heterogeneity at any one of these scales can influence the development of spatial variability at the other scales. Consequently, it can be difficult to ascertain the role of these factors, both individually and in tandem, in the initiation and progression of spatial patterning within a multicellular system. Here, we present a graphical user interface (GUI) for producing agent-based models that are capable of exploring these various facets of pattern formation. Agent-based models are unique in their ability to integrate combinations of heterogeneous processes and investigate their respective dynamics, especially in the context of spatial phenomena. However, while ABMs are an ideal tool for studying spatial characteristics, the required programming knowledge and initial time investment can be daunting. Other interfaces for developing agent-based models exist but the user must explicitly code the agents, their properties, their interactions, and the environment. For example, NetLogo [227] and CompuCell3D [26] both provide an interface for the creation of agent-based models but the possible complexity of the generated models in these interfaces is constrained by the users knowledge of the respective programming languages.

The GUI application presented here avoids these requirements by allowing the user to add or remove hard-coded features considered essential processes for studying pattern formation. While this may initially seem to limit flexibility, the ABMs produced can be modified to include any additional processes, as the user sees fit. The primary advantage of this infrastructure is the ability to quickly iterate through different combinations of

processes, cell organizations, and morphologies. Various designs for engineering living systems (ELS) can be investigated and optimized *in silico* to accelerate the production of *in vitro* systems. Alternatively, current ELS can be modeled and improved upon by testing variations in the designs or conditions. Ultimately, we provide a tool that facilitates the study of multiple modes of cellular communication in a spatial manner that can be applied to a wide variety of biological questions.

### 5.1. Multiscale Communication Agent-Based Model Generator

A multiscale communication agent-based model generator (MsCAMgen) was developed in Python 2.7 to provide a framework for modeling various spatial aspects of a multicellular network without requiring explicit programming by the user. Each model is capable of accounting for cell division and growth, state changes between different cell types, extracellular diffusion of molecules that are secreted and consumed by cells, intercellular communication of small molecules between neighboring cells, and intracellular gene/protein networks (Figure 30). For simplification, the small molecules that are transported between cells are called “intercellular molecules” to differentiate them from the non-transportable species in the intracellular networks.



**Figure 30: Overview of the primary features and forms of communication of MsCAMgen across different timescales. Each cell type can be assigned a cell cycle length, allowing**

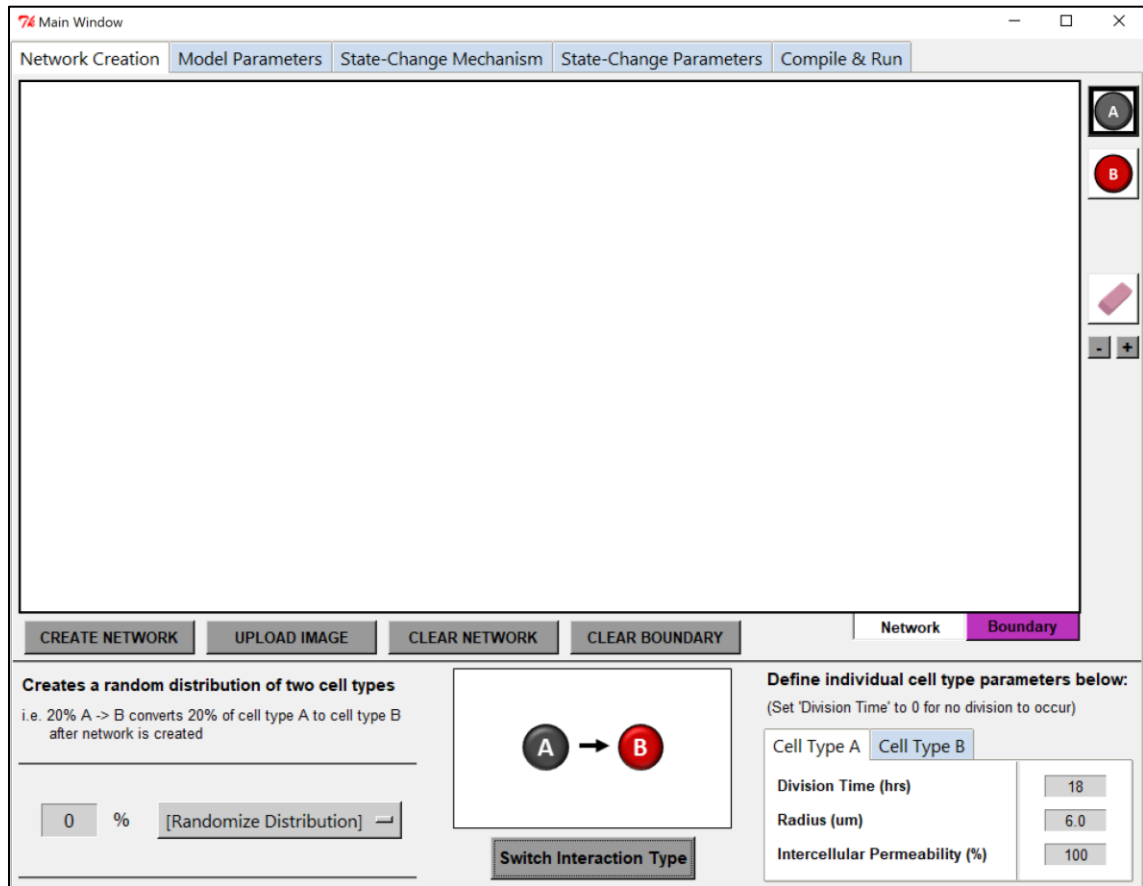
**growth and division of cells over the course of multiple hours. Intracellular gene/protein networks interact with extracellular and intercellular molecules, updating every hour. Extracellular and intercellular transport are both updated every second, with intercellular molecules typically diffusing significantly faster.**

The application is open-source and installation involves installing Python from the official website [www.python.org](http://www.python.org), followed by simple instructions to download additional modules. Upon installation, the user runs the `start_GUI.py` script (by double-clicking) and is ready to design a model. The application has 5 tabs/frames that the user proceeds through from left to right: Network Creation, Model Parameters, State-Change Mechanism, State-Change Parameters, and Compile & Run. The following sections will detail the various features available in each tab.

### **5.1.1 Network Creation**

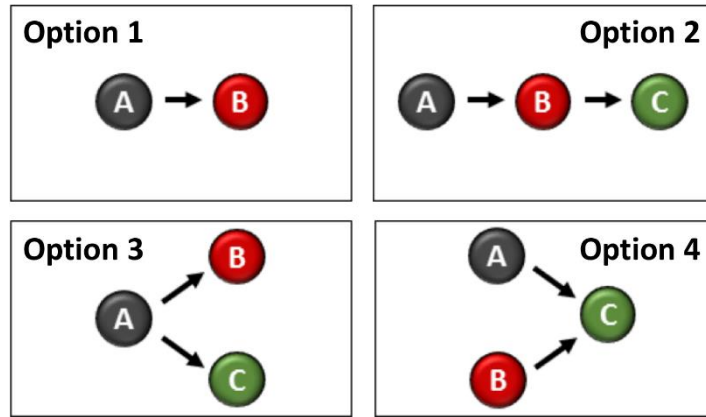
The first frame focuses on the macro aspects of the model: the hierarchy of cell types, the initial morphology of the colony, and the organization of cells within that colony (Figure 31). At the top of the frame are the five tabs a user will progress through, left to right. Moving forward, there will be a linear progression through the choices that a user has and a suggested order for making those decisions, however any step could be ignored along the way and the model would instead use the default settings. These default values are not suggested to be used for actual model analyses, but they allow a user to quickly test features and learn how MsCAMgen works.

Within the panel in Figure 31, the large white space in the center is an area for drawing a cell colony or inputting an image, where on the far right side one can select which cell type to draw. On the bottom right, the user specifies properties for each cell type. Once the network is drawn, the user clicks the “Create Network” button for the drawn region to be converted to a cloud of cells. On the right, the purple Boundary button lets the user specify a growth boundary that prevents cellular movement.



**Figure 31: The first frame “Network Creation” and initial window that appears upon launching MscAMgen.**

The first consideration is the hierarchy of cell types because it influences multiple input options during the model design. Four options are included, as shown in Figure 32. Briefly, there is a simple transition between two cell types, a linear progression between three cell types, a multipotent cell type capable of divergent fates, or two cell types capable of transitioning to the same state. The first three options are relatively common hierarchal schemes but the fourth is more context specific and can represent different biological phenomena. For instance, the two initial cells could represent a mutant and wildtype of the same cell type or the final state could represent an apoptotic state of the two initial cell types.

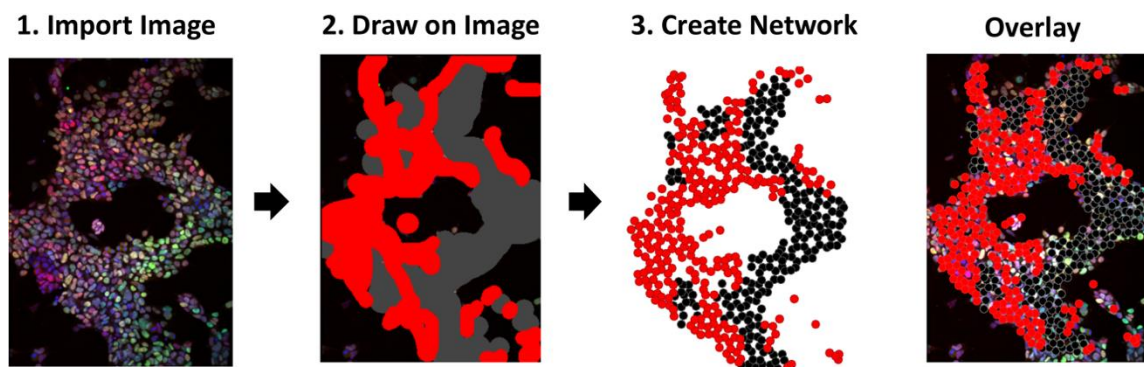


**Figure 32: The four possible cell transition hierarchies. Option 1 is a simple transition between two cell types. Option 2 depicts a linear progression between three cell types, where A transitions to B followed by B transitioning to C. Option 3 represents a multipotent cell (A) that can transition into two different states, B and C. Option 4 covers special cases such as mutant cells (B is a mutant of A) or apoptosis (C is an apoptotic cell).**

After selecting the state change options, the user draws the desired shape of their system on the large white canvas. On the top right side of the frame are buttons that depict each cell type, with the black outline indicating which button is selected. Once satisfied with the morphology and cellular organization, clicking the button labelled “Create Network” will convert the filled in area into a cluster of cells. To change the relative size of each cell type within the network, the user can change the value for the cell radii in the “Cell Type Parameters” table on the bottom right of the frame. If a random mixture of cells is required, on the bottom left is the option to create a random distribution of two cell types once the network is created. Furthermore, an image can be uploaded using the ‘Upload Image’ button. The image is overlaid on the white canvas and can be drawn over in the same way as the canvas. In this way, it is possible to coarsely mimic the morphology and organization of an experimental system in a generated model, as illustrated by Figure 33.

Lastly, by clicking on the purple “Boundary” button, a growth boundary can be specified around the cell network that cells cannot pass beyond. An experimental

equivalent of this type of growth boundary is the use of micropatterning, which has been used for a number of differentiation studies [226, 228, 229]. To account for the limited growth space, each cell agent has a compression value calculated from the degree of overlap with other cells that prevents cell division if the local cell density is too high. An example of the growth boundary is shown in the next section when the different growth dynamics are discussed (Figure 35).



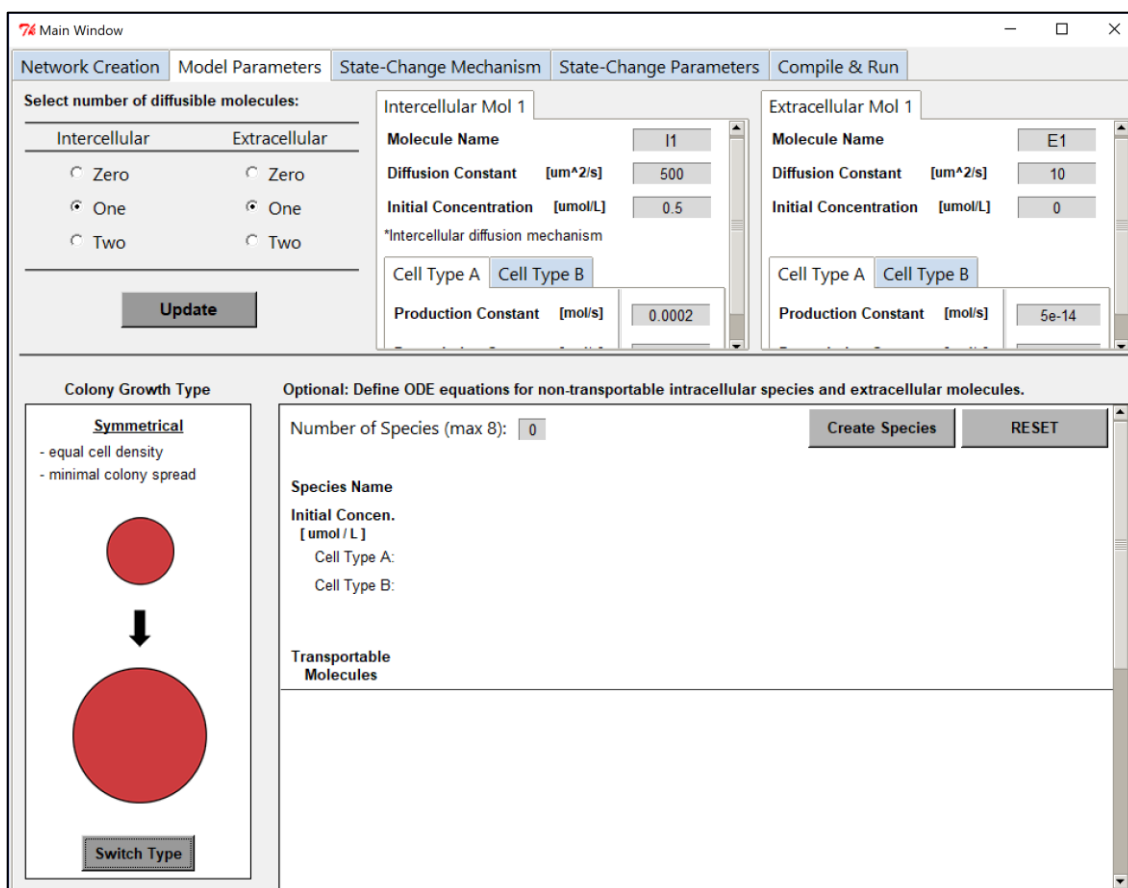
**Figure 33: Using an experimental image as a template for the morphology of the cell network in MsCAMgen. An uploaded image can be drawn over, using the different colors to designate cell type, and converted into a network of cells with similar morphology and cellular organization.**

### 5.1.2 Model Parameters

The second tab/frame of MsCAMgen mostly consists of selecting the transportable and non-transportable molecules within the system, but also contains options for selecting the type of growth dynamics (bottom left of Figure 34). Along the top, the number of extracellular and intercellular molecules can be specified with their associated parameters, such as their diffusion constants and production/degradation rates for each cell type. At the bottom left, the growth dynamics can be selected and on the bottom right, a system of ODEs can be inputted for non-transportable intracellular species.

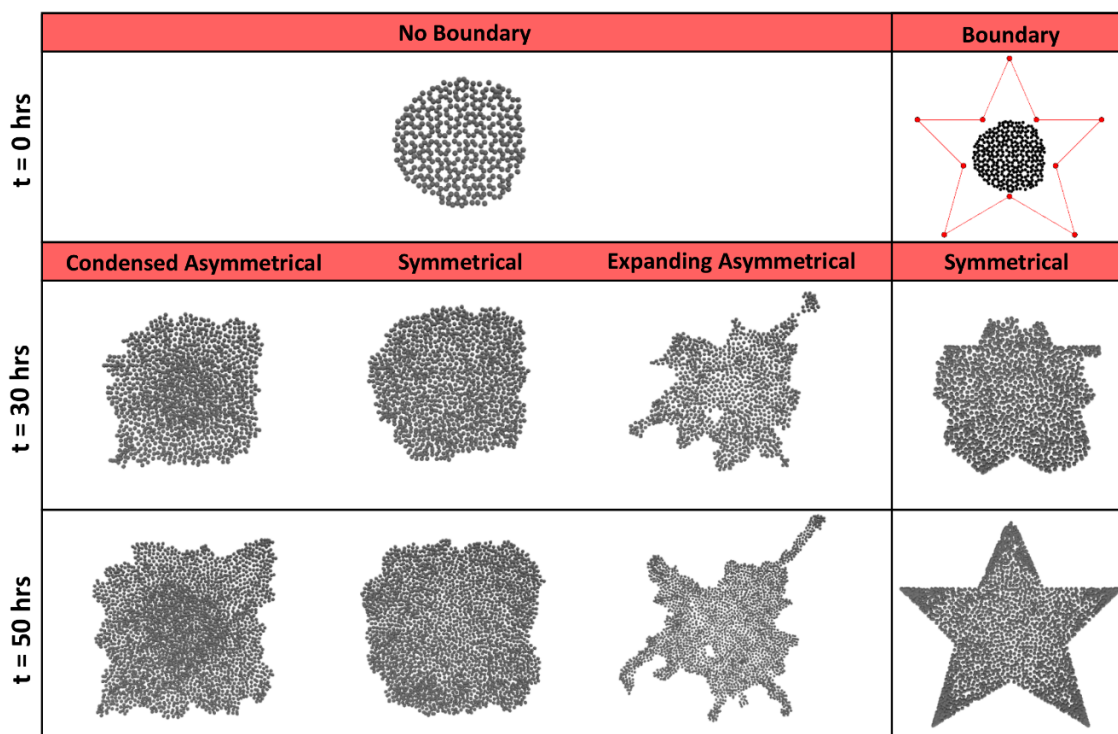
As illustrated in Figure 35, there are three colony growth choices: *symmetrical*, *condensed asymmetrical*, and *expanding asymmetrical* growth.





**Figure 34: The second frame "Model Parameters" of MsCAMgen.**

The symmetrical growth type produces colonies with equal cell densities and uniform expansion along each edge. Typically, a rather circular colony will result from the symmetrical growth type over time. The condensed asymmetrical growth type exhibits a high-density region at the center of the colony while cells along the edges spread outwards in an asymmetrical fashion. A similar phenomenon occurs in pluripotent colonies where cells pack closely together and become less dense approaching the edges [228]. The expanding asymmetrical type depicts motile cells that create low-density and highly asymmetrical morphologies. Expanding asymmetrical growth was implemented for the intercellular model, described in Chapter 3, to mimic the cellular extensions out of the mESC colonies that occurred during neural differentiation.



**Figure 35: The three types of colony growth dynamics available in MsCAMgen.**

**Symmetrical growth has equal density and minimal spread of the colony edges. Condensed asymmetrical maintains a highly dense center of cells while the edges expand outwards. Expanding asymmetrical has low density throughout the colony and develops asymmetrical extensions from the colony. On the right is an example of using a star-shaped growth boundary, with cells unable to move out of the specified region.**

Along the top of the second frame (Figure 34), the user can select and parameterize up to two extracellular and/or intercellular molecules. Once the number of each type of species is decided, clicking the “Update” button will add or remove those molecules. The parameters for each species include the diffusion constant, a production rate, a degradation rate, and an initial concentration either within the cell or in the extracellular space. The chosen molecules will become options when assigning the state-change mechanism in the next section.

Additionally, up to eight non-transportable intracellular species can be defined as a set of ordinary differential equations (Figure 36). First, the user can choose the number of intracellular molecules, followed by clicking the “Create Species” button. A number of entry boxes will appear for naming the respective species and assigning initial concentrations of each of those species within the different cell types. Clicking the “Create Equations” button instantiates another set of entry boxes for defining the respective differential equations. The equations can include any transportable or non-transportable species, represented by their associated shorthand  $X_n$ ,  $E_n$ , or  $I_n$  for the intracellular, extracellular and intercellular species, respectively. Constants can be typed directly into the equations or they can be defined using the shorthand  $C_n$ , where  $n$  is an integer between 0 and 99. If using the shorthand constants, then once the equations are finalized, the “Set Constants” button can be clicked. An entry box for each constant that was defined in the equations will appear and values can be assigned. Later, when the model is compiled, all of the equations and initial values will be extracted, parsed into their respective variables, and compiled in the correct syntax within the model.

Optional: Define ODE equations for non-transportable intracellular species and extracellular molecules.

Number of Species (max 8):  Create Species RESET

	X1	X2	X3	X4	X5
Species Name	TF1	TF2	TF3	TF4	TF5
Initial Concn. [umol / L]					
Cell Type A:	100	100	100	100	100
Cell Type B:	100	100	100	100	100
Cell Type C:	100	100	100	100	100

Transportable Molecules:   Create Equations

**ODE Equations**  
 Write ODE equations using  $X_n$ ,  $I_n$ , and  $E_n$  abbreviations as shown above for each species.  
 Define coefficients/constants using  $C_n$  where  $n$  is an integer 0-99.  
 $P$  and  $D$  terms are production and degradation rates defined above.  
 \*\* for exponents, \* for multiplication, / for division

$d(E1)/dt =$    
 $d(TF1)/dt =$    
 $d(TF2)/dt =$    
 $d(TF3)/dt =$    
 $d(TF4)/dt =$    
 $d(TF5)/dt =$

Set Constants

Insert Constant Values (time scale in hours).  
 ODE equations will import after compiling model.

C1	C2	C3	C4	C5	C6	C7	C8
<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>
C9	<input type="text" value="0"/>						

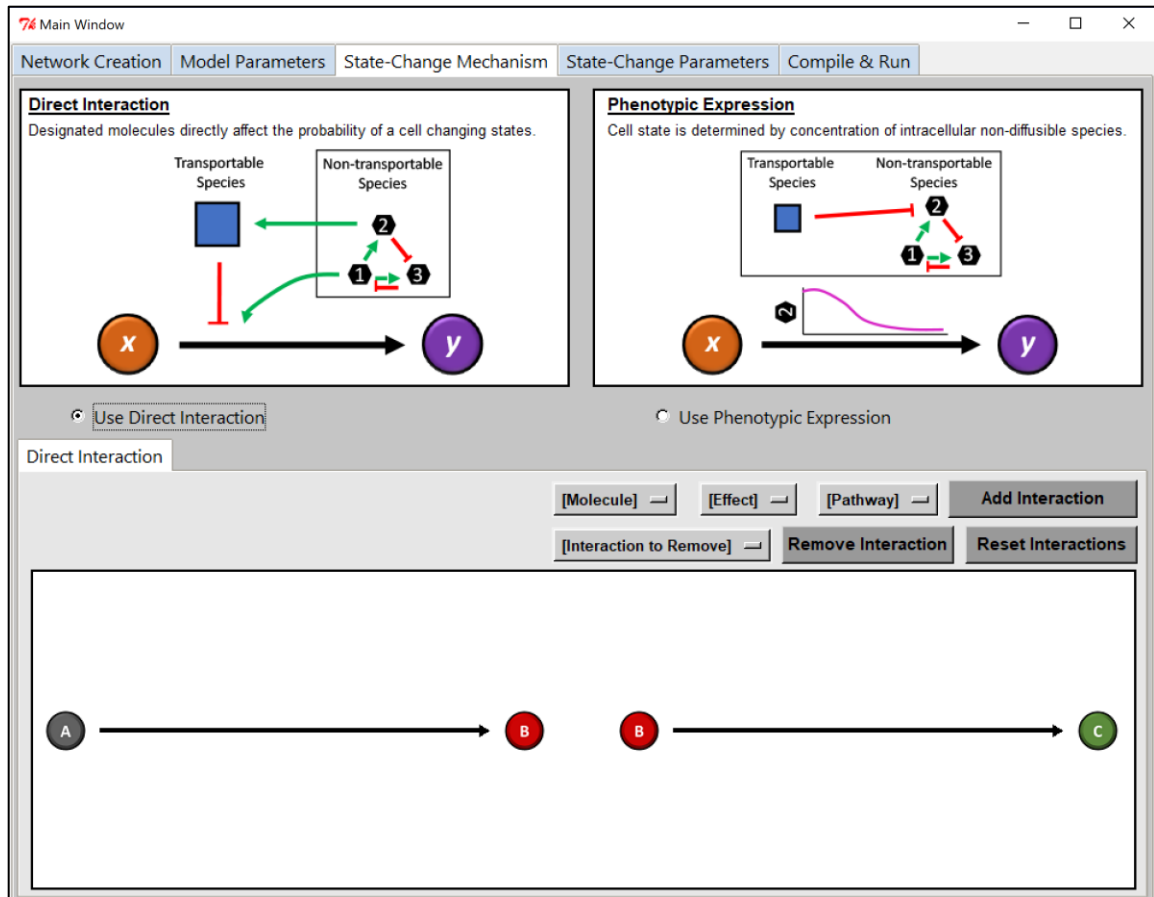
**Figure 36: Defining non-transportable intracellular species and their respective differential equations. Starting from the top, the user selects how many species and clicks “Create Species”, followed by giving names and initial concentrations to those species before clicking “Create Equations”. A series of entry boxes will appear for each intracellular and extracellular species where the differential equations can be typed. The equations can include any transportable or non-transportable species and constant values can be defined and assigned values by clicking “Set Constants”.**

### 5.1.3 State-Change Mechanism & Parameters

The third frame centers on describing the cell interactions with the various molecules/species in the system and their influence on state transitions (Figure 37). Along the top are descriptions of the two mechanisms, Direct Interaction and Phenotypic Expression. Selecting either option will change the “State-Change Parameters” frame to reflect the different mechanism. At the bottom, interactions can be assigned by selecting

the molecule from the drop-down list, the effect that molecule has, and the pathway it affects.

There are two possible mechanisms for inducing state change: direct interaction or phenotypic expression. The direct interaction method allows for molecules to directly affect the probability of a state change occurring. Any defined molecule can behave as an activator or an inhibitor in a concentration-dependent manner and the state change probability for any particular cell is the cumulative effect of every interacting molecule the cell is exposed to. In contrast, the phenotypic expression method assumes that each cell type can be represented by a unique steady state of the non-transportable intracellular species. For example, if the pluripotency gene network was considered then cell type A and B could represent Oct4+ and Oct4- cells, respectively. Since the fourth frame “State-Change Parameters” is dependent upon the selected mechanism, this section will be split into two subsections for each mechanism that will also discuss the respective parameter options.

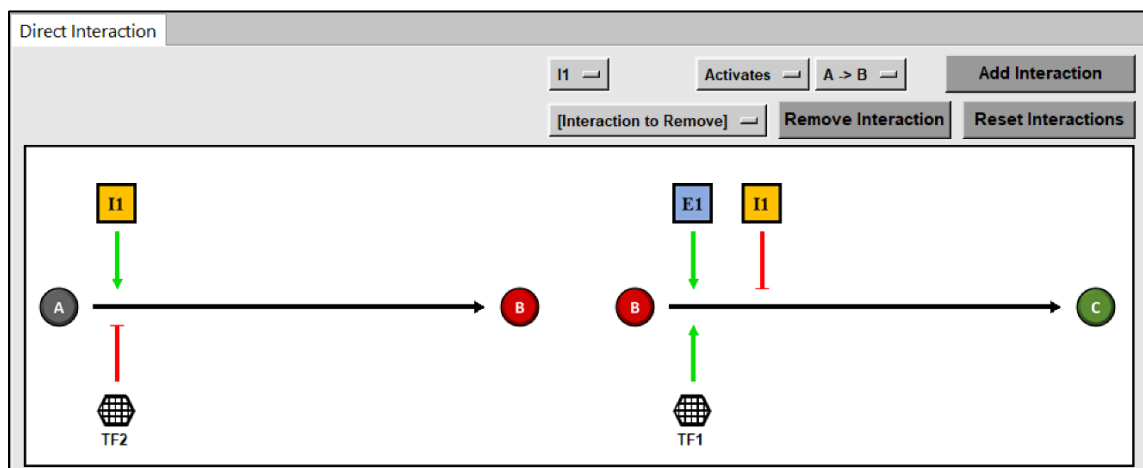


**Figure 37: The third frame “State-Change Mechanism” of MsCAMgen.**

#### 5.1.3.1 Direct Interaction Mechanism

For direct interaction, at the bottom of the third frame there are three drop-down menus labelled “[Molecule]”, “[Effect]”, and “[Pathway]”. Using these three drop-down menus, the user can specify which molecules can activate or inhibit the probability of one cell type converting to another. The molecule list will include the given or default names of each non-transportable species as well as the shorthand notations for the extra- and intercellular molecules,  $E_n$  and  $I_n$  respectively. An example is shown in Figure 38, “I1 Activates A -> B”. When the “Add Interaction” button is clicked a representation of I1 with a green activation arrow appears above the A to B transition. The same procedure can be repeated for every other interaction that is occurring in the system. Furthermore, all

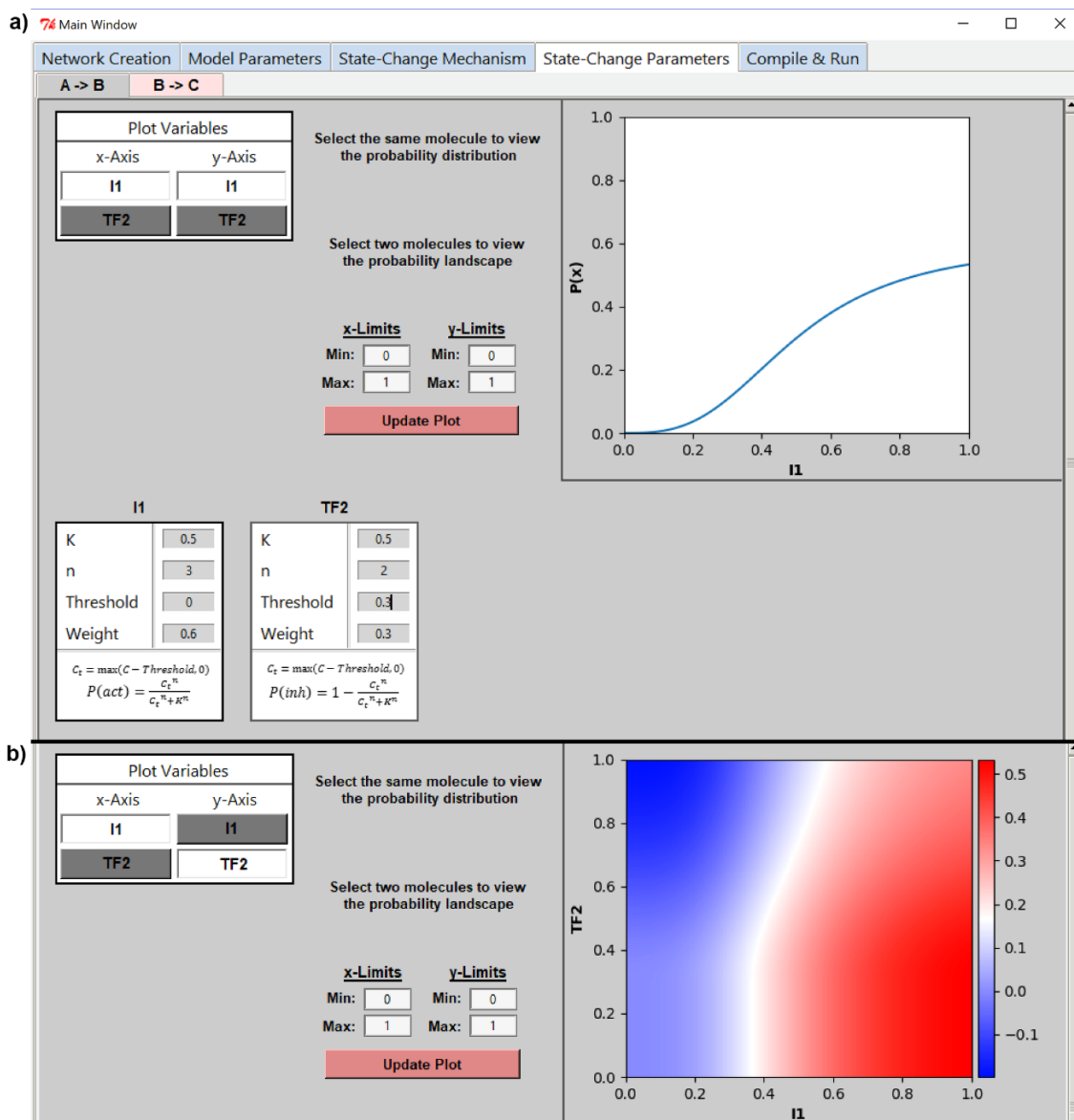
interactions can be reset or removed one at a time using the “Interaction to Remove” drop-down list.



**Figure 38: Setting up a direct interaction mechanism in MsCAMgen. Transportable molecules are represented by squares while the non-transportable species are hexagons. The Molecule, Effect, and Pathway drop-down lists at the top right are selected as I1, Activates, and A -> B, respectively.**

Moving on to the fourth frame “State-Change Parameters”, there is a tab for each pathway i.e. A -> B (Figure 39a). Each tab has the same layout, with the top left having a list of all molecules that affect the pathway. At the bottom of the frame, for each tab, each molecule has a set of parameters that define its probability function. A positive and negative Hill function were used to represent the probability function for activation and inhibition, respectively. A threshold can also be set which shifts the origin of the Hill function to that value, effectively making any concentrations below the threshold have no influence on the transition. The Weight parameter is a multiplier that changes the magnitude of the probability function. As a visualization, if the same molecule is selected for the x- and y-axes in the Plot Variables box followed by clicking the “Update Plot” button then the probability function is plotted on the top right. To change the scale or only look at a specific concentration range, the x- and y-limits can be modified for the plot. In

addition to visualizing a single probability function, if two different molecules are selected in the Plot Variables box then a probability landscape is generated (Figure 39b). Specifically, a heatmap of the cumulative probability from the two selected molecules within the concentration ranges set in x- and y-limits.

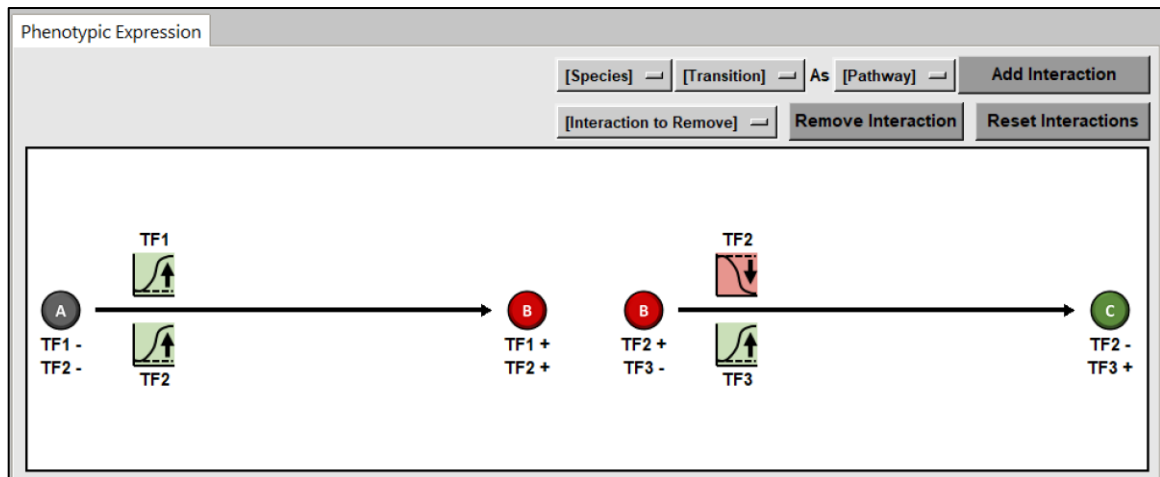


**Figure 39: The fourth frame "State-Change Parameters" for Direct Interaction in MscAMgen. (a) The top right of the frame shows a plot of the probability function for I1 with the associated parameter values in the box on the bottom left. (b) The probability landscape using the parameter values in (a) for I1 vs TF2.**



### 5.1.3.2 Phenotypic Expression Mechanism

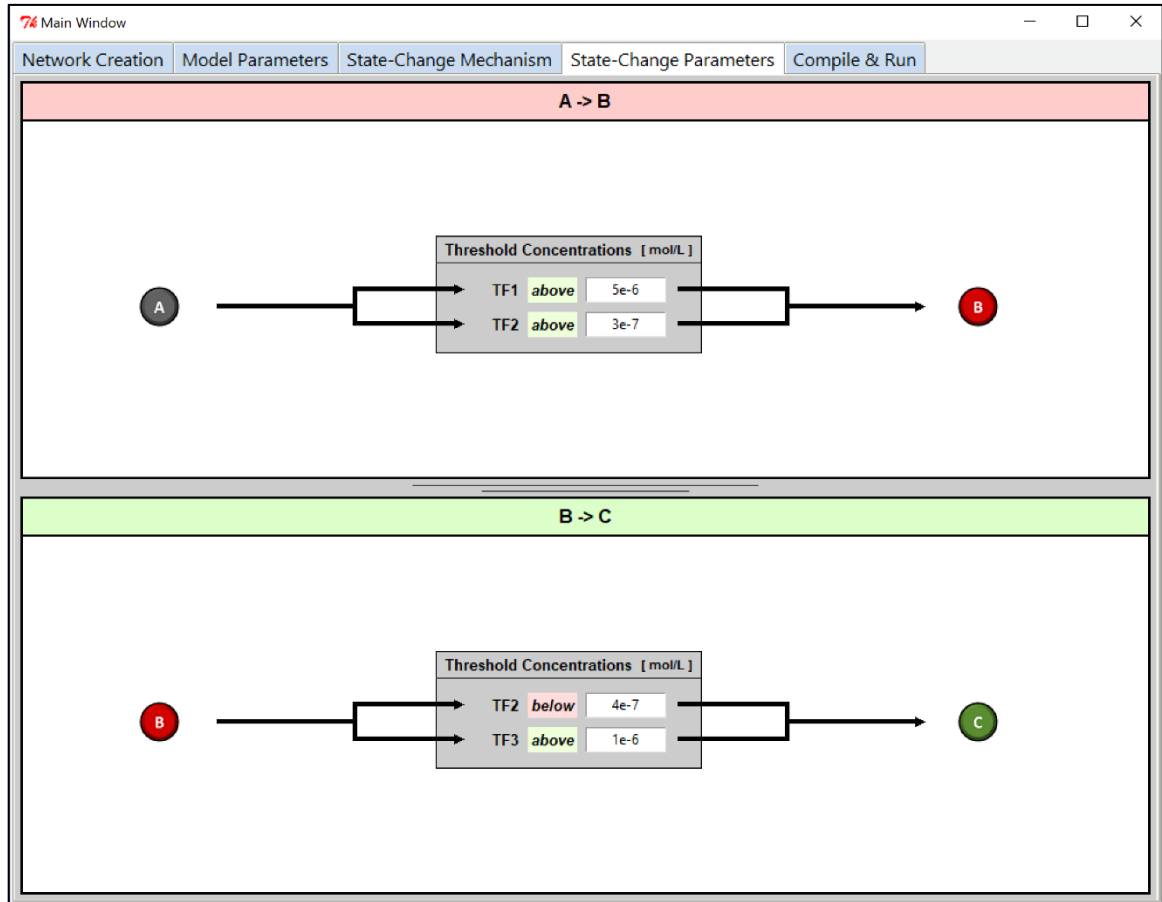
The layout for phenotypic expression is nearly identical to direct interaction, with the main difference being that instead of an “[Effect]” drop-down there is a “[Transition]” drop-down to represent whether the concentration of the species increases or decreases during the transition. The same procedure as described for Direct Interaction is used to Add, Remove, and Reset the interactions.



**Figure 40: Setting up a phenotypic expression mechanism in MsCAMgen. Non-transportable species are depicted by miniature plots that either increase (green) or decrease (red). Additionally, the species are labelled underneath each cell type with a plus or minus sign to illustrate whether it is starting at a higher or lower concentration than the cell type it is transitioning towards.**

The only State-Change Parameters required for the phenotypic expression mechanism are threshold concentration values (Figure 41). At every time point (hour increments by default), the intracellular concentrations are updated according to their differential equations and compared to the thresholds. If all assigned threshold conditions are met then there is a 20% chance for the cell to change states. A probability was implemented to minimize the effect of fluctuations initiating a state change, but the

assigned 20% was arbitrarily selected during the development of MsCAMgen. A future version will allow the user to choose this probability value for each transition pathway.



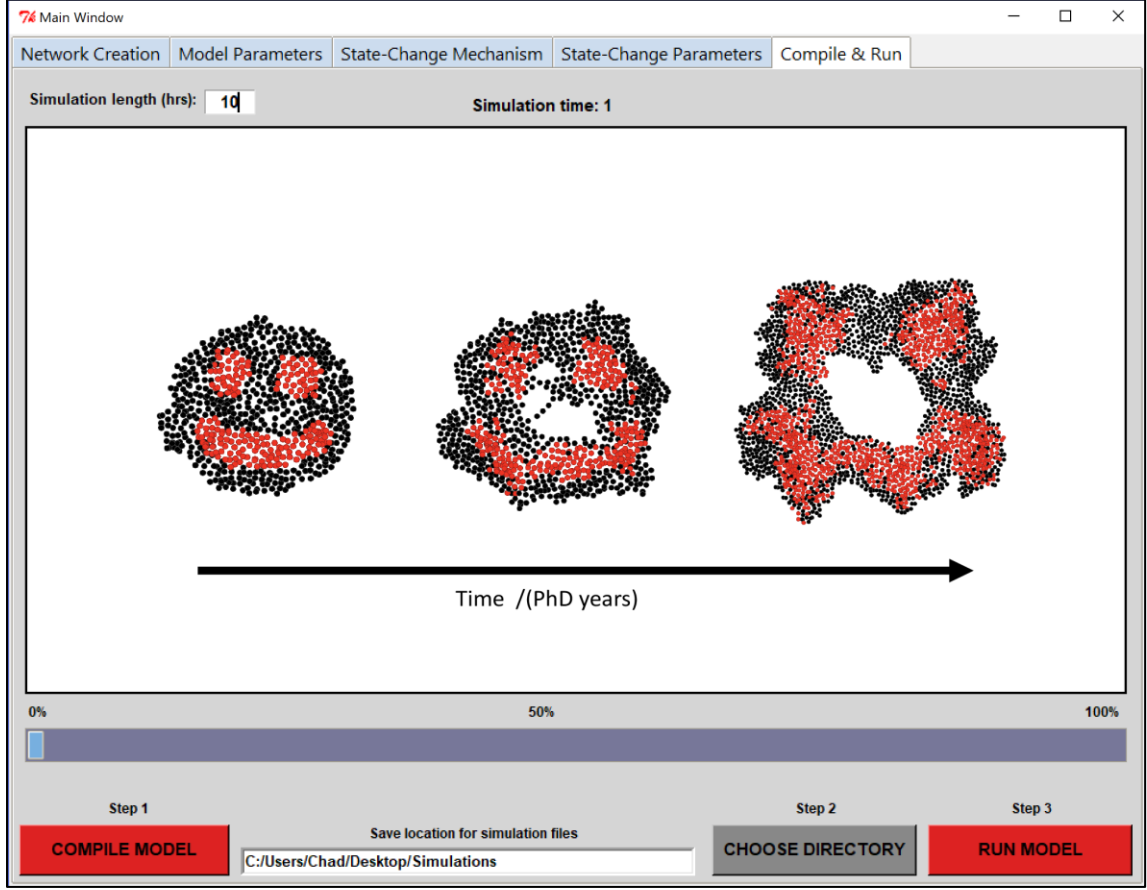
**Figure 41: The fourth frame "State-Change Parameters" for Phenotypic Expression in MsCAMgen. For a transition to occur, a cell must simultaneously meet the threshold requirements for each species. If the requirements are met then the cell has a set 20% probability of changing state.**

#### 5.1.4 Compile & Run

In the final frame, all of the previously input information is extracted from other frames and compiled (Figure 42). Specifically, template model scripts have insertion points where different parameters and/or functions are written. Therefore, each model is generated in real-time in a minimalistic form that does not contain numerous conditional statements.

The advantage of this tactic is that readability of the model code is enhanced and variables can be dynamically written. The “Compile Model” button formulates a new agent-based model in the same folder as the `start_GUI.py` script using the parameters from the other frames. Once compiled, the user can select or type in a directory for saving simulation files and then click “Run Model” to begin. During runtime of the simulation, the central region of the frame is updated to show the evolution of the morphology and organization of the colony.

Once started, a new folder is created in the provided directory that contains a copy of the model code. As the model runs, files containing the simulation data for each hour time point are added to the folder. The “network.pickle” files contain all the information about the cell-agents, while other saved files are named after the extracellular molecules and contain the respective concentrations in the extracellular space, represented by a discrete grid. All of this information can be extracted and either independently analyzed or converted into images using scripts provided with the generated model.

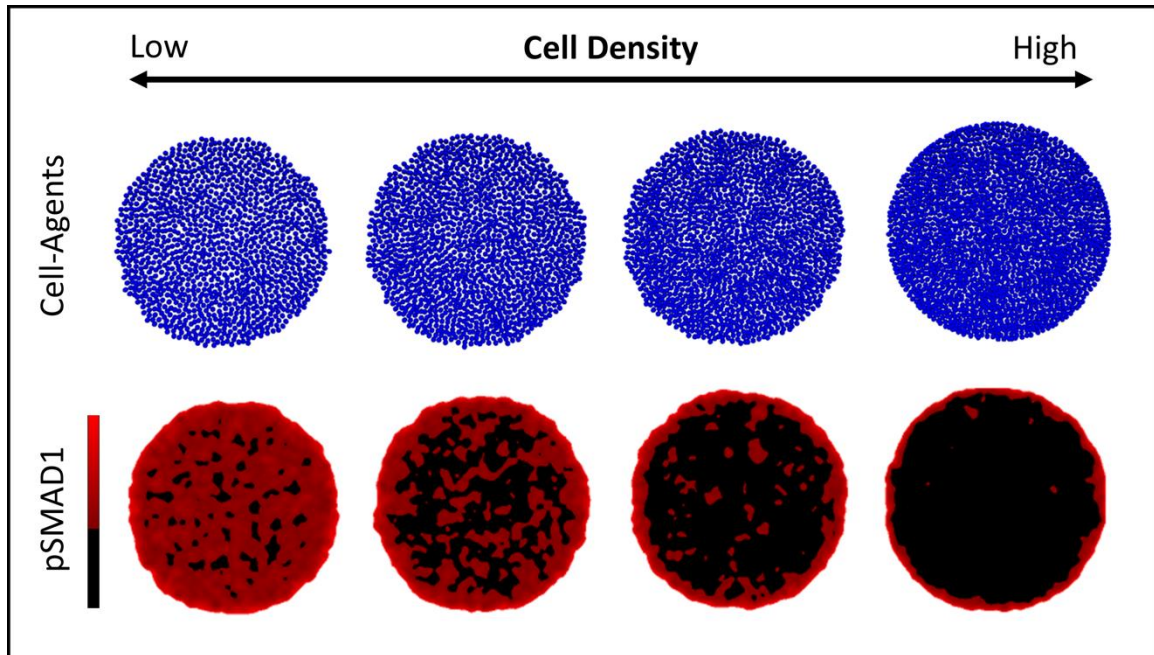


**Figure 42: The fifth frame "Compile & Run" for MsCAMgen. A real-time update on the colony morphology is shown in the window and a progress bar fills as the simulation proceeds.**

## 5.2. Application Example of MsCAMgen

As a proof of concept, MsCAMgen was used to recapitulate a recently published model on the radial self-organization of the three germ layers [226]. During gastrulation, epiblast cells differentiate into the three germ layers in a spatially defined manner. A similar patterning can be achieved *in vitro* by growing hESCs in the presence of bone morphogenetic protein 4 (BMP4) within a circular micropattern [228]. The confined geometry of the micropattern in conjunction with the BMP4, mimicking BMP4 secreted by extraembryonic tissue, creates a radial pattern of the three germ layers with

trophectoderm forming the outermost ring. The formation of this radial pattern is largely dependent on the spatial distribution of pSMAD1 activation in response to BMP4, which occurs along the outer edge of the colony. In the paper, Etoc et al. demonstrated the importance of high cell density for attaining this pSMAD1 localization [226]. As the cell density was decreased, larger regions in the center of the colony became pSMAD1 positive. The model they developed used the cell density and distance from the center of the colony to estimate pSMAD1 localization. Here, MsCAMgen was used to simulate PSC colonies of varying density constrained by a circular growth boundary. A single extracellular molecule, representing BMP4, was included with a constant degradation rate for each cell in the simulation. In this way, a pSMAD1 patterning could be attained that was similar to the model and experimental data (Figure 43).



**Figure 43: Simulated BMP4 activity, indicated by pSMAD1, as a function of cell density using a MsCAMgen generated model.**

### 5.3. Discussion

Biological systems and their respective processes are inherently complex and consist of multiple coordinated systems. In many cases, the functionality of these systems

would be impossible to intuit from knowledge of only the individual parts. To study the emergence of this behavior and identify the interactions involved requires a tool such as AB modeling that allows for the integration of various processes in a single simulation. However, biologists are not typically exposed to any substantial amount of programming or computational modeling. As such, there is a need for software that can allow biologists to study these emergent phenomena without requiring a significant amount of time deciphering computational syntax. MsCAMgen is a simple solution that provides biologists with a method of studying emergence in a large variety of contexts and with a significantly lower time investment.

The novelty of MsCAMgen over existing tools is that it was designed as an interface for generating models rather than an environment for model development. The distinction is important because an environment for model development merely provides a set of tools and instructions on their independent use. It is the responsibility of the user to implement these tools correctly and to organize their interactions in the form of a working model. Therefore, to develop a model, the user must become comfortable with the syntax of the respective programming language, define and code the order of operations, and validate that the implemented processes behave appropriately. In contrast, MsCAMgen generates a working model that incorporates pre-coded and pre-validated features. The user can add or remove these features within the graphical interface without requiring prior knowledge of the Python language and the model will be compiled correctly at the click of a button. This format provides multiple advantages over a typical framework for model development. Every model is created independently of previous models which prevents the possibility of a single error propagating throughout model versions. Furthermore, the addition or removal of features is automatically accounted for within every model script, effectively removing any requirement for debugging during model design. The only drawback, in terms of user difficulty, associated with the addition of features is an increase in the number of parameter values. However, there are constraints in the utility of

MsCAMgen. Models are currently constrained to 2D and simulations with over 1000 cells can slow down dramatically when extra- or intercellular transport is being considered. As such, MsCAMgen is mostly applicable for analyzing smaller niche environments.

The plug-and-play nature of MsCAMgen is a unique feature that not only significantly increases accessibility but also makes the addition or removal of model complexity a trivial task. The provided framework contains a plethora of options for studying largely variable systems. For example, tumors have vast heterogeneity in their composition and behavior. Using MsCAMgen, any 2D structure, such as that obtained from a histological section, can be drawn and converted into a cellular network within seconds. The discussed example illustrated the ease of modulating cell density but just as easily one could modulate the colony size, colony asymmetry, or cell heterogeneity. Overall, MsCAMgen is an adaptable tool that can provide insight into these highly variable systems.

Finally, the successful design of ELS relies on the coordination of multicellular populations to achieve functionality. However, as ELS aspire to higher functionality they also require tighter control over the cell-environmental and cell-cell interactions within their designs. To optimize these systems necessitates an iterative process. MsCAMgen facilitates an accelerated design method by enabling the analysis of multiscale communication within cell populations of any morphology or organization. The acceleration is unique to the presented interface because the implementation of these various changes to the format and design of a model introduce no potential for errors or time consequence. Furthermore, the tool described in this chapter could be coupled with other ABM optimization schemes, such as described in [154, 230], to determine the biologically relevant parameters that are best targeted to achieve desired macro-level features.

## CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS

### 6.1. Conclusions

In this work, I have explored the role of intercellular communication in neural differentiation. Building upon prior work in the Kemp and McDevitt labs that established transitions between discrete pattern classes during stem cell differentiation [137], the mechanism of non-secreted molecular transfer across membranes of adjacent cells was investigated in the context of morphogenic pattern trajectories through latent space. An agent-based model developed from experimentally-derived parameters was able to recapitulate and predict spatiotemporal trajectories of differentiation by using intercellular communication as the differentiation mechanism. Aided by computational simulations, I demonstrated that intercellular communication was modulated by the cell cycle to yield complex, dynamic networks of molecular information transmission. Heterogeneity emerged within the network in response to the natural asynchrony of cell cycle states within a population. Furthermore, this work identified a transient gain in intercellular connectivity as a typical feature during the loss of pluripotency. Perturbation to the intercellular network through chemical and genetic methods elucidated a strong temporal delay in differentiation in response to the decreased connectivity. In addition, a graphical user interface was developed for the production of multiscale communication agent-based models that provides a new resource for biological researchers. This platform allows for hypothesis testing of how emergent features arise in multicellular systems and will provide bioengineers the opportunity to iterate designs for engineered living systems.

#### Cell Cycle Modulation of Intercellular Transport

The initial goal of this work was to characterize intercellular communication in the context of multicellular systems. Previous work had established that cell-cell interactions could accurately depict the differentiation and patterning of mESC aggregates using a top-



down modeling approach. The question was posed whether a mechanistic representation of intercellular transport would be capable of manifesting spatial patterns in a similar manner. In answer, the model framework from White et al.[137] was adapted to include intercellular communication. Our quantification of intercellular transport rates identified the cell cycle as a potent source of heterogeneity in the intercellular network. Diffusion typically acts as an equalizer and, for spatial heterogeneity to be maintained, there needs to be a source that can disrupt the equalizing influence of the diffusion. In a reaction-diffusion system, the disruption is caused by the non-linear reactions. Here, the oscillations in GJ permeability caused by the cell cycle were sufficient to cause spatial heterogeneity in the diffusion patterns. In fact, this heterogeneity was a significant factor in producing the spatial patterning that was obtained in Chapter 3. Therefore, our results indicate that accounting for the influence of cell cycle asynchrony is important to accurately depict intercellular communication at a multicellular level.

### Spatial Analysis of Neural Differentiation

We combined network analysis and dimension reduction techniques to assess the progression of differentiation patterning. In this way, we were able to characterize temporal and spatial aspects of the differentiation using three latent variables. We found that neural differentiation begins with individual cells across the colony in a relatively random pattern. However, once a critical number of cells differentiate, there is a rapid gain of organization as clusters of differentiated cells form. These clusters predominantly occurred along the outside of the colony before propagating inwards. The intercellular model, gated by cell cycle, was able to mimic this spatiotemporal trajectory and visualize the similarities in latent space.

### Delaying Differentiation via Intercellular Communication

In this work, we analyzed the effect of multiple perturbations to intercellular communication on differentiation. In a remarkably consistent manner, each perturbation to the intercellular network caused a delay in the rate of differentiation. The consistency in this response, across multiple cell lines and species, suggests that intercellular communication has a vital role in the differentiation process. The capacity for intercellular communication to slow the rate of differentiation has interesting implications for directing differentiation. In one potential case, the blockade of intercellular communication decreases the rate of differentiation generically for all cell types. In that scenario, we can use an analogy to a feedback controller where intercellular communication is the “gain” of the differentiation process. By inhibiting intercellular communication, it takes more time to reach the target phenotype (i.e. the set point) but the control is more precise. In another case, decreased intercellular communication only decreases the rate of differentiation towards a subset of lineages. Therefore, it would specifically act as an inhibitory agent for those lineages and indirectly promote differentiation towards the unaffected lineages.

Additionally, we found that two neural differentiation protocols performed in two different pluripotent cell lines from different species both caused an increase in the level of Cx43 during the early stages of differentiation. The ability of different protocols, which enhance neural differentiation through separate mechanisms, to produce a similar modulation of Cx43 suggests that the effect is an underlying feature of the state transition rather than a product of the treatments. The concurrence with the loss of Oct4 in both species, despite differentiation occurring at different rates, further corroborates the hypothesis of a genetic mechanism. When considered in combination with the delayed differentiation from the perturbation studies, these results suggest that the increased connectivity is involved in the propagation and coordination of differentiation within the population. Thus, intercellular communication has a vital role during differentiation and the propagation of differentiation signals.

### MsCAMgen as a Design Tool

We developed a graphical interface for the generation of agent-based models containing various modes of communication. The interface, MsCAMgen, has unique applications for enhancing the design of engineered living systems. The spatial distribution of a secreted molecule could be simulated for a set design, the optimal seeding density of two cell types with different growth rates could be determined, or the effect of mixing a wildtype cell with a mutant. For example, MsCAMgen could create an intercellular model for the hiPSC system and simulate a mixed population of WT and knocked down cells. Experimentally this system has significant degrees of freedom: both cell density and time of inhibition can be adjusted. It would be painstaking and time consuming to attempt every combination but can be tested computationally with relative ease. To further illustrate the benefit of MsCAMgen, I will provide a simple comparison. The original intercellular model for mESC differentiation optimistically took one year to design. A synonymous model for the hiPSC system could be generated in less than one hour and parameterized in one week if data were available.

## **6.2. Future Directions**

### **6.2.1 Agent-Based Bioelectricity**

Various molecules can be transported through gap junction channels. In this work, we presented an intercellular diffusion framework both in the differentiation model in Chapter 3 and later as a module of MsCAMgen in Chapter 5. While the framework is capable of modeling any designated molecule, large diffusion constants have a significant computational cost, even when optimized as a C-compiled function. In most circumstances this is not an issue because the majority of molecules of interest diffuse at about the same rate as cAMP, but ionic diffusion is significantly faster. Additionally, it is usually desired to model multiple ions simultaneously to gauge membrane potential. Therefore, to model

ionic diffusion and bioelectricity, a tactic other than discrete diffusion would have to be employed. One amazing template for this is the BioElectric Tissue Simulation Engine which depicts cells using a Voronoi representation [218]. Transport is designated at each interface and is solved as a set of linear equations. At first glance this may seem incompatible with the dynamics of an agent-based simulation. However, given the exceptionally fast diffusion of ions compared to cellular movement it is possible to assume it is a static system that reaches steady-state at each time point. The Voronoi connections would need to be updated and other perturbations accounted for at each time point as well. The inclusion of bioelectricity seems a natural progression because membrane potential is a common GJ gating mechanism [32]. It would be interesting to see if bioelectricity and cell cycle effects would produce similar or divergent intracellular patterns.

## **6.2.2 Computational Improvements**

### **6.2.2.1 MsCAMgen Spatial Analysis**

In this work, we described a pipeline for analyzing spatial patterns within our intercellular agent-based model. A similar pipeline could be directly coupled to the MsCAMgen interface that allows the user to select from a list of network metrics and perform PCA on generated models to visualize the spatiotemporal trajectory. The extension would include three options: (i) a training set is automatically generated using the network data from the designed model and variations of the pattern classes defined in Chapter 3, (ii) the user uploads their own training set of defined pattern classes, or (iii) the trajectory is analyzed without the use of a training set. In the first case, the user can rapidly visualize the trajectory without requiring additional information. By using the network structures from the designed model, the pattern analysis is adaptable to various morphologies. However, a larger number of pattern classes would need to be defined to account for systems containing three cell types. Also, this pattern classification is predominantly focused on the organization of cell types within a population rather than the morphology

of the population. The second option provides flexibility but requires knowledge of the system/desired patterns in advance. The difficulty of this option is to ensure either that every user creates a training set of a specified format or that multiple formats can be uploaded and analyzed successfully. The third option removes many of the limitations of the first two, but the lack of training set makes the interpretation of the results in latent space more complicated. The inclusion of a training set of pattern classes allows the user to efficiently associate the latent variable axes with macroscopic properties of the system. It is still possible to identify the properties that each latent variable represents by examining the loads of each contributing metric but this can become cumbersome, especially if using network-based metrics that are not easily related to a biological phenomenon. Nonetheless, this option is useful for examining the temporal evolution of metric variability within a given model. For instance, if all of the simulation time points are clustered together for each latent variable, then none of the selected metrics exhibit variability over that time period. However, if a latent variable exhibits a bimodal distribution then the metrics that are diverging can be identified as well as the time at which the divergence occurs. The divergence time would be particularly useful for identifying when major changes are occurring in the system and thus when to apply perturbations to prevent or maximize those changes.

The spatial analysis extension described here would substantially enhance the quantitative capability of MsCAMgen. One of the major limitations of ABMs is the difficulty associated with quantitatively validating the model results. As described in Figure 5, the combination of multivariate metric extraction and dimension reduction techniques allows for the quantitative analysis of spatial features and the comparison of those features between experimental and simulation results. Thus, by including an option for the simultaneous analysis of experimental images using the PCA transform defined from training set (or simulation data in Option 3), the user can both design and validate their model within MsCAMgen. The extension also benefits users that are designing

multicellular constructs in a number of ways. The goal of many engineered living systems is to attain and maintain a pattern or cellular organization that is associated with a desired functionality. Visualization of the spatiotemporal trajectory through PCA provides invaluable information such as which patterns/organizations develop, the time ranges where organizational transitions occur, and the duration of time that the system maintains a specific organization. All of this information can be employed to guide future modifications in an efficient manner to achieve a desired pattern and the conditions necessary to retain it.

#### 6.2.2.2 GPU-Acceleration

As mentioned in the Background Chapter, the use of the GPU to accelerate agent-based modeling has many potential benefits but also some limitations. Many of the limitations revolve around a loss of flexibility for defining interactions while the benefits lie with the capacity for exceptionally fast parallelization. Therefore, a middle ground is suggested. The most computationally expensive actions within an ABM are diffusion and collision detection/optimization. In both of these cases, the values can be populated within a separate matrix, independent from the agent objects, and passed to the GPU to solve and return the updated values to populate the objects. In this way, it is possible to achieve acceleration while maintaining the flexibility of object-based agents. Of note, this would only be a benefit if the number of iterations was sufficiently high ( $\sim 10^6$  iterations) because of the computational overhead costs involved in transferring information between the CPU and GPU. This caveat is overcome when considering multiple extracellular and intercellular molecules in MsCAMgen. Assuming the GPU hardware contains sufficient memory, the data for each molecule could be transferred simultaneously to the GPU and then solved in parallel. The overhead cost would remain approximately the same while the number of iterations would dramatically increase. Thus, we would be able to obtain the

maximum possible acceleration and further enhance the ease of testing various permutations of model design.

### **6.2.3 hiPSC Cx43-KD Mixed Populations**

One of the major findings of this work was that inhibition of intercellular communication was able to delay differentiation. The next step for this work is investigating how heterogeneous modulation of intercellular communication will affect the dynamics of differentiation within a population. One approach to achieve this is through mixed populations of the LBC2-GJA1 cell line with the parent cell line, WTC11. A major advantage of the LBC2-GJA1 cell line is that it has a constitutively expressed GFP-Lamin-B reporter as an identifier. As such, it is possible to identify the location of cells lacking intercellular communication within the mixed population and assess how the spatial arrangement affects the differentiation patterning both locally and globally. It would be possible to attain better spatial control utilizing optogenetic rather than TET induction but to attain this spatial control would require a high-resolution laser, which limits data collection. Furthermore, constant laser exposure would be required to maintain the knockdown for any significant period of time due to the rapid turnover rate of connexins. For these reasons, the CRISPRi knockdown provides the most robust method for interrogating spatially heterogeneous perturbations to the intercellular network during differentiation protocols. The TET-induced CRISPR system is also capable of scaling to 3D with minimal technical difficulty and technologies currently exist that could be employed to gain control over the initial spatial arrangement of the two cell types within the population.

The information attained from these studies has enormous utility for the derivation of spatially and temporally controlled differentiation protocols. By understanding the behavior of cells with hindered cell-cell communication, both individually and in clusters, it is possible to identify cellular arrangements that lead to specific differentiation patterns.

Further control could be gained by additional studies that investigate the cross-talk between extracellular and intercellular means of communication. In particular, it is possible that intercellular perturbation can modulate the transduction of extracellular signals through a cellular population. One advantage of modulating intercellular communication is the degree of spatial resolution that can be obtained. The intercellular communication of a single cell can be completely abrogated whereas it is infeasible to generate extracellular gradients at a similar resolution or with a similar extent of spatial heterogeneity. Thus, there is tremendous promise for exploiting the intercellular network to direct spatiotemporal differentiation, both directly and through the propagation of signals that originate from other means of communication.



## REFERENCES

1. Lv, S., et al., *Glucagon-induced extracellular cAMP regulates hepatic lipid metabolism*. J. Endocrinol., 2017. **234**(2): p. 73-87.
2. Paulmann, N., et al., *Intracellular serotonin modulates insulin secretion from pancreatic beta-cells by protein serotonylation*. PLoS Biol, 2009. **7**(10): p. e1000229.
3. Gairhe, S., et al., *Serotonin passes through myoendothelial gap junctions to promote pulmonary arterial smooth muscle cell differentiation*. Am J Physiol Lung Cell Mol Physiol, 2012. **303**(9): p. L767-77.
4. Almaca, J., et al., *Human Beta Cells Produce and Release Serotonin to Inhibit Glucagon Secretion from Alpha Cells*. Cell Rep, 2016. **17**(12): p. 3281-3291.
5. Burnside, A.S. and P. Collas, *Induction of Oct-3/4 expression in somatic cells by gap junction-mediated cAMP signaling from blastomeres*. Eur J Cell Biol, 2002. **81**(11): p. 585-91.
6. Chandrasekhar, A., et al., *Intercellular redistribution of cAMP underlies selective suppression of cancer cell growth by connexin26*. PLoS One, 2013. **8**(12): p. e82335.
7. Kameritsch, P., et al., *Gap junctional communication promotes apoptosis in a connexin-type-dependent manner*. Cell Death & Disease, 2013. **4**(4): p. e584-e584.
8. Camp, J.G., et al., *Multilineage communication regulates human liver bud development from pluripotency*. Nature, 2017. **546**(7659): p. 533-538.
9. Thomson, J.A., et al., *Embryonic Stem Cell Lines Derived from Human Blastocysts*. Science, 1998. **282**: p. 1145-1147.
10. Robertson, E.J., *Derivation and Maintenance of Embryonic Stem Cell Cultures*. Methods Mol Biol., 1990. **5**: p. 223-236.
11. Inlay, M.A., et al., *Identification of multipotent progenitors that emerge prior to hematopoietic stem cells in embryonic development*. Stem Cell Reports, 2014. **2**(4): p. 457-72.
12. Parfitt, D.E. and M.M. Shen, *From blastocyst to gastrula: gene regulatory networks of embryonic stem cells and early mouse embryogenesis*. Philos Trans R Soc Lond B Biol Sci, 2014. **369**(1657).
13. Thompson, D.W., *On Growth and Form*. 1942, Cambridge, UK: Cambridge University Press.
14. Turing, A.M., *The Chemical Basis of Morphogenesis*. Phil. Trans R. Soc. Lond., 1952. **237**(641): p. 37-72.
15. Meinhardt, H. and A. Gierer, *Applications of a Theory of Biological Pattern Formation Based on Lateral Inhibition*. J Cell Science, 1974. **15**: p. 321-346.
16. Wolpert, L., *Chapter 6 Positional Information and Pattern Formation*, in *Current Topics in Developmental Biology Volume 6*. 1971. p. 183-224.
17. Kears, M. "The Potency of the First Two Cleavage Cells in Echinoderm Development. Experimental Production of Partial and Double Formations" (1891-1892), by Hans Driesch. 2012.
18. Wolpert, L., *Positional information and the spatial pattern of cellular differentiation*. J. Theor. Biol., 1969. **25**: p. 1-47.
19. Wolpert, L., *Positional Information*, in *eLS*. 2016. p. 1-5.

20. Rubin, L. and J.W. Saunders, *Ectodermal-Mesodermal Interactions in the Growth of Limb Buds in the Chick Embryo: Constancy and Temporal Limits of the Ectodermal Induction*. Development Biology, 1972. **28**: p. 94-112.
21. Wolpert, L., C. Tickle, and A. Martinez Arias, *Principles of Development*. 2015, Oxford: Oxford University Press.
22. Pietak, A. and M. Levin, *Bioelectric gene and reaction networks: computational modelling of genetic, biochemical and bioelectrical dynamics in pattern regulation*. J R Soc Interface, 2017. **14**(134).
23. Ashe, H.L. and J. Briscoe, *The interpretation of morphogen gradients*. Development, 2006. **133**(3): p. 385-94.
24. Green, J.B. and J. Sharpe, *Positional information and reaction-diffusion: two big ideas in developmental biology combine*. Development, 2015. **142**(7): p. 1203-11.
25. Salazar-Ciudad, I., J. Garcia-Fernandez, and R.V. Sole, *Gene networks capable of pattern formation: from induction to reaction-diffusion*. J Theor Biol, 2000. **205**(4): p. 587-603.
26. Izaguirre, J.A., et al., *CompuCell, a multi-model framework for simulation of morphogenesis*. Bioinformatics, 2004. **20**(7): p. 1129-37.
27. Miura, T. and P.K. Maini, *Periodic pattern formation in reaction-diffusion systems: An introduction for numerical simulation*. Anatomical Science International, 2004. **79**: p. 112-123.
28. Muller, P., et al., *Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning system*. Science, 2012. **336**(6082): p. 721-4.
29. Mou, C., et al., *Cryptic patterning of avian skin confers a developmental facility for loss of neck feathering*. PLoS Biol, 2011. **9**(3): p. e1001028.
30. Sheth, R., et al., *Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism*. Science, 2012. **338**(6113): p. 1476-80.
31. Guo, Y., et al., *Branching patterns emerge in a mathematical model of the dynamics of lung development*. J Physiol, 2014. **592**(2): p. 313-24.
32. Levin, M., *Molecular bioelectricity: how endogenous voltage potentials control cell behavior and instruct pattern regulation in vivo*. Mol Biol Cell, 2014. **25**(24): p. 3835-50.
33. Emmons-Bell, M., et al., *Gap Junctional Blockade Stochastically Induces Different Species-Specific Head Anatomies in Genetically Wild-Type Girardia dorotocephala Flatworms*. Int J Mol Sci, 2015. **16**(11): p. 27865-96.
34. Urdy, S., *On the evolution of morphogenetic models: mechano-chemical interactions and an integrated view of cell differentiation, growth, pattern formation and morphogenesis*. Biol Rev Camb Philos Soc, 2012. **87**(4): p. 786-803.
35. Mercker, M., D. Hartmann, and A. Marciniak-Czochra, *A mechanochemical model for embryonic pattern formation: coupling tissue mechanics and morphogen expression*. PLoS One, 2013. **8**(12): p. e82617.
36. Gierer, A. and H. Meinhardt, *A theory of biological pattern formation*. Kybernetik, 1972. **12**: p. 30-39.
37. Schnackenberg, J., *Simple chemical reaction systems with limit cycle behavior*. J. Theor. Biol, 1979. **81**: p. 389-400.
38. Murray, J.D., *Biomathematics*. Vol. 19. 1990: Springer-Verlag Berlin Heidelberg.

39. Chaturvedi, R., et al., *On multiscale approaches to three-dimensional modelling of morphogenesis*. J R Soc Interface, 2005. **2**(3): p. 237-53.
40. Okuda, S., et al., *Combining Turing and 3D vertex models reproduces autonomous multicellular morphogenesis with undulation, tubulation, and branching*. Sci Rep, 2018. **8**(1): p. 2386.
41. Miura, T., et al., *Mixed-mode pattern in Doublefoot mutant mouse limb--Turing reaction-diffusion model on a growing domain during limb development*. J Theor Biol, 2006. **240**(4): p. 562-73.
42. Seirin Lee, S., E.A. Gaffney, and R.E. Baker, *The dynamics of Turing patterns for morphogen-regulated growing domains with cellular response delays*. Bull Math Biol, 2011. **73**(11): p. 2527-51.
43. Marcon, L., et al., *High-throughput mathematical analysis identifies Turing networks for patterning with equally diffusing signals*. Elife, 2016. **5**.
44. Pearson, J.E., *Pattern formation in a (2+1)-species activator-inhibitor-immobilizer system*. Physica A, 1992. **188**: p. 178-189.
45. Korvasova, K., et al., *Investigating the Turing conditions for diffusion-driven instability in the presence of a binding immobile substrate*. J Theor Biol, 2015. **367**: p. 286-295.
46. Hillenbrand, P., U. Gerland, and G. Tkacik, *Beyond the French Flag Model: Exploiting Spatial and Gene Regulatory Interactions for Positional Information*. PLoS One, 2016. **11**(9): p. e0163628.
47. Tkacik, G., et al., *Positional information, positional error, and readout precision in morphogenesis: a mathematical framework*. Genetics, 2015. **199**(1): p. 39-59.
48. Summerbell, D., J.H. Lewis, and L. Wolpert, *Positional Information in Chick Limb Morphogenesis*. Nature, 1973. **244**: p. 492-496.
49. Gregor, T., et al., *Probing the limits to positional information*. Cell, 2007. **130**(1): p. 153-64.
50. Lawrence, P.A., J. Casal, and G. Struhl, *Cell interactions and planar polarity in the abdominal epidermis of Drosophila*. Development, 2004. **131**(19): p. 4651-64.
51. Ma, D., et al., *Fidelity in planar cell polarity signaling*. Nature, 2003. **421**(6922): p. 543-47.
52. Michaelis, U.R., *Mechanisms of endothelial cell migration*. Cell Mol Life Sci, 2014. **71**(21): p. 4131-48.
53. Bedau, M.A., *Emergent Models of Supple Dynamics in Life and Mind*. Brain and Cognition, 1997. **34**: p. 5-27.
54. Coronado, D., et al., *A short G1 phase is an intrinsic determinant of naive embryonic stem cell pluripotency*. Stem Cell Res, 2013. **10**(1): p. 118-31.
55. Cheeseman, B.L., D.F. Newgreen, and K.A. Landman, *Spatial and temporal dynamics of cell generations within an invasion wave: a link to cell lineage tracing*. J Theor Biol, 2014. **363**: p. 344-56.
56. Lake, J.I. and R.O. Heuckeroth, *Enteric nervous system development: migration, differentiation, and disease*. Am J Physiol Gastrointest Liver Physiol, 2013. **305**(1): p. G1-24.
57. Nagy, N. and A.M. Goldstein, *Enteric nervous system development: A crest cell's journey from neural tube to colon*. Semin Cell Dev Biol, 2017. **66**: p. 94-106.

58. Zhang, D., et al., *Neural crest regionalisation for enteric nervous system formation: implications for Hirschsprung's disease and stem cell therapy*. Dev Biol, 2010. **339**(2): p. 280-94.
59. Zhang, D., et al., *Stochastic clonal expansion of "superstars" enhances the reserve capacity of enteric nervous system precursor cells*. Dev Biol, 2018.
60. Cheeseman, B.L., et al., *Cell lineage tracing in the developing enteric nervous system: superstars revealed by experiment and simulation*. J R Soc Interface, 2014. **11**(93): p. 20130815.
61. Peurichard, D., et al., *Simple mechanical cues could explain adipose tissue morphology*. J Theor Biol, 2017. **429**: p. 61-81.
62. Tang, J., et al., *Irradiation of juvenile, but not adult, mammary gland increases stem cell self-renewal and estrogen receptor negative tumors*. Stem Cells, 2014. **32**(3): p. 649-61.
63. Binder, B.J. and K.A. Landman, *Exclusion processes on a growing domain*. J Theor Biol, 2009. **259**(3): p. 541-51.
64. Newgreen, D.F., et al., *Simple rules for a "simple" nervous system? Molecular and biomathematical approaches to enteric nervous system formation and malformation*. Dev Biol, 2013. **382**(1): p. 305-19.
65. Simpson, M.J., K.A. Landman, and B.D. Hughes, *Multi-species simple exclusion processes*. Physica A: Statistical Mechanics and its Applications, 2009. **388**(4): p. 399-406.
66. Simpson, M.J., et al., *Cell proliferation drives neural crest cell invasion of the intestine*. Dev Biol, 2007. **302**(2): p. 553-68.
67. Montero, J.A. and J.M. Hurle, *Sculpturing digit shape by cell death*. Apoptosis, 2010. **15**(3): p. 365-75.
68. Handyside, A.H. and S. Hunter, *Cell division and death in the mouse blastocyst before implantation*. Roux's Arch Dev Biol, 1986. **195**: p. 519-526.
69. Grant, M.R., et al., *Simulating properties of in vitro epithelial cell morphogenesis*. PLoS Comput Biol, 2006. **2**(10): p. e129.
70. Rejniak, K.A. and A.R. Anderson, *A computational study of the development of epithelial acini: I. Sufficient conditions for the formation of a hollow structure*. Bull Math Biol, 2008. **70**(3): p. 677-712.
71. Kim, S.H., et al., *A computational approach to resolve cell level contributions to early glandular epithelial cancer progression*. BMC Syst Biol, 2009. **3**: p. 122.
72. Tang, J., et al., *Phenotypic transition maps of 3D breast acini obtained by imaging-guided agent-based modeling*. Integr Biol (Camb), 2011. **3**(4): p. 408-21.
73. Barker, N., et al., *Identification of stem cells in small intestine and colon by marker gene Lgr5*. Nature, 2007. **449**(7165): p. 1003-7.
74. Baker, A.M., et al., *Quantification of crypt and stem cell evolution in the normal and neoplastic human colon*. Cell Rep, 2014. **8**(4): p. 940-7.
75. Bravo, R. and D.E. Axelrod, *A calibrated agent-based computer model of stochastic cell dynamics in normal human colon crypts useful for in silico experiments*. Theor Biol Med Model, 2013. **10**(66).
76. Tabata, T. and Y. Takei, *Morphogens, their identification and regulation*. Development, 2004. **131**(4): p. 703-12.

77. Rogulja, D. and K.D. Irvine, *Regulation of cell proliferation by a morphogen gradient*. Cell, 2005. **123**(3): p. 449-61.
78. Leung, M.C., et al., *Computational modeling and simulation of genital tubercle development*. Reprod Toxicol, 2016. **64**: p. 151-61.
79. Visvader, J.E. and G.J. Lindeman, *Mammary stem cells and mammaryogenesis*. Cancer Res, 2006. **66**(20): p. 9798-801.
80. D.L., P., et al., *Radiation effects on breast cancer risk: a pooled analysis of eight cohorts*. Radiat Res., 2002. **158**(2): p. 220-35.
81. Beyer, T. and M. Meyer-Hermann, *Multiscale modeling of cell mechanics and tissue organization*. IEEE Eng Med Biol Mag, 2009. **28**(2): p. 38-45.
82. Ingham-Dempster, T., D.C. Walker, and B.M. Corfe, *An agent-based model of anoikis in the colon crypt displays novel emergent behaviour consistent with biological observations*. R Soc Open Sci, 2017. **4**(4): p. 160858.
83. Galea, G.L., et al., *Biomechanical coupling facilitates spinal neural tube closure in mouse embryos*. Proc Natl Acad Sci U S A, 2017. **114**(26): p. E5177-E5186.
84. Szabo, A. and R. Mayor, *Cell traction in collective cell migration and morphogenesis: the chase and run mechanism*. Cell Adh Migr, 2015. **9**(5): p. 380-3.
85. Hynes, R.O., *The extracellular matrix: not just pretty fibrils*. Science, 2009. **326**(5957): p. 1216-9.
86. Gattazzo, F., A. Urciuolo, and P. Bonaldo, *Extracellular matrix: a dynamic microenvironment for stem cell niche*. Biochim Biophys Acta, 2014. **1840**(8): p. 2506-19.
87. Nakazawa, N., et al., *Laminin-1 promotes enteric nervous system development in mouse embryo*. Pediatr Surg Int, 2013. **29**(11): p. 1205-8.
88. Akbareian, S.E., et al., *Enteric neural crest-derived cells promote their migration by modifying their microenvironment through tenascin-C production*. Dev Biol, 2013. **382**(2): p. 446-56.
89. Boucaut, J.C., *Biologically active synthetic peptides as probes of embryonic development: a competitive peptide inhibitor of fibronectin function inhibits gastrulation in amphibian embryos and neural crest cell migration in avian embryos*. The Journal of Cell Biology, 1984. **99**(5): p. 1822-1830.
90. Darribere, T. and J.E. Schwarzbauer, *Fibronectin matrix composition and organization can regulate cell migration during amphibian development*. Mechanisms of Development, 2000. **92**: p. 239-250.
91. Straus, A.H., et al., *Mechanism of fibronectin-mediated cell migration: dependence or independence of cell migration susceptibility on RGDS-directed receptor (integrin)*. Exp Cell Res., 1989. **183**(1): p. 126-39.
92. Bonnans, C., J. Chou, and Z. Werb, *Remodelling the extracellular matrix in development and disease*. Nat Rev Mol Cell Biol, 2014. **15**(12): p. 786-801.
93. Robertson, S.H., et al., *Multiscale computational analysis of Xenopus laevis morphogenesis reveals key insights of systems-level behavior*. BMC Syst Biol, 2007. **1**: p. 46.
94. Cooper, J.A., *A mechanism for inside-out lamination in the neocortex*. Trends Neurosci, 2008. **31**(3): p. 113-9.

95. Kriegstein, A.R. and S.C. Noctor, *Patterns of neuronal migration in the embryonic cortex*. Trends Neurosci, 2004. **27**(7): p. 392-9.
96. Lambert de Rouvroit, C. and A.M. Goffinet, *Neuronal Migration*. Mechanisms of Development, 2001. **105**: p. 47-56.
97. D'Arcangelo, G., et al., *Reelin Is a Secreted Glycoprotein Recognized by the CR-50 Monoclonal Antibody*. The Journal of Neuroscience, 1997. **17**(1): p. 23-31.
98. Quattrocchi, C.C., et al., *Reelin is a serine protease of the extracellular matrix*. J Biol Chem, 2002. **277**(1): p. 303-9.
99. Rice, D.S. and T. Curran, *Role of the Reelin Signaling Pathway in Central Nervous System Development*. Annu. Rev. Neurosci., 2001. **24**: p. 1005-039.
100. Caviness, V.S., *Patterns of Cell and Fiber Distribution in the Neocortex of the Reeler Mutant Mouse*. J. Comp. Neur., 1976. **170**: p. 435-448.
101. Caffrey, J.R., et al., *An in silico agent-based model demonstrates Reelin function in directing lamination of neurons during cortical development*. PLoS One, 2014. **9**(10): p. e110415.
102. Olson, E.C., S. Kim, and C.A. Walsh, *Impaired neuronal positioning and dendritogenesis in the neocortex after cell-autonomous Dab1 suppression*. J Neurosci, 2006. **26**(6): p. 1767-75.
103. Setty, Y., et al., *How neurons migrate: a dynamic in-silico model of neuronal migration in the developing cortex*. BMC Syst Biol, 2011. **5**: p. 154.
104. Wynn, M.L., P.M. Kulesa, and S. Schnell, *Computational modelling of cell chain migration reveals mechanisms that sustain follow-the-leader behaviour*. J R Soc Interface, 2012. **9**(72): p. 1576-88.
105. Wang, Z., et al., *An Observation-Driven Agent-Based Modeling and Analysis Framework for C. elegans Embryogenesis*. PLoS One, 2016. **11**(11): p. e0166551.
106. Lambert, B., et al., *Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis*. J Math Biol, 2018.
107. Varner, V.D. and C.M. Nelson, *Computational models of airway branching morphogenesis*. Semin Cell Dev Biol, 2017. **67**: p. 170-176.
108. Qutub, A.A. and A.S. Popel, *Elongation, proliferation & migration differentiate endothelial cell phenotypes and determine capillary sprouting*. BMC Syst Biol, 2009. **3**: p. 13.
109. Walpole, J., et al., *Agent-based model of angiogenesis simulates capillary sprout initiation in multicellular networks*. Integr Biol (Camb), 2015. **7**(9): p. 987-97.
110. Menshykau, D. and D. Iber, *Kidney branching morphogenesis under the control of a ligand-receptor-based Turing mechanism*. Phys Biol, 2013. **10**(4): p. 046003.
111. Hirashima, T., Y. Iwasa, and Y. Morishita, *Dynamic modeling of branching morphogenesis of ureteric bud in early kidney development*. J Theor Biol, 2009. **259**(1): p. 58-66.
112. Perfahl, H., et al., *3D hybrid modelling of vascular network formation*. J Theor Biol, 2017. **414**: p. 254-268.
113. Scianna, M., C.G. Bell, and L. Preziosi, *A review of mathematical models for the formation of vascular networks*. J Theor Biol, 2013. **333**: p. 174-209.
114. Liu, G., et al., *Module-based multiscale simulation of angiogenesis in skeletal muscle*. Theor Biol Med Model, 2011. **8**: p. 6.

115. Kleinstreuer, N., et al., *A computational model predicting disruption of blood vessel development*. PLoS Comput Biol, 2013. **9**(4): p. e1002996.
116. Garbey, M., M. Rahman, and S.A. Berceli, *A Multiscale Computational Framework to Understand Vascular Adaptation*. J Comput Sci, 2015. **8**: p. 32-47.
117. Walpole, J., et al., *Agent-based computational model of retinal angiogenesis simulates microvascular network morphology as a function of pericyte coverage*. Microcirculation, 2017. **24**(8).
118. Merks, R.M., et al., *Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling*. Dev Biol, 2006. **289**(1): p. 44-54.
119. Merks, R.M. and J.A. Glazier, *Dynamic mechanisms of blood vessel growth*. Nonlinearity, 2006. **19**(1): p. C1-C10.
120. Merks, R.M., et al., *Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth*. PLoS Comput Biol, 2008. **4**(9): p. e1000163.
121. Kohn-Luque, A., et al., *Early embryonic vascular patterning by matrix-mediated paracrine signalling: a mathematical model study*. PLoS One, 2011. **6**(9): p. e24175.
122. Watson, M.G., et al., *Dynamics of angiogenesis during murine retinal development: a coupled in vivo and in silico study*. J R Soc Interface, 2012. **9**(74): p. 2351-64.
123. Bentley, K., H. Gerhardt, and P.A. Bates, *Agent-based simulation of notch-mediated tip cell selection in angiogenic sprout initialisation*. J Theor Biol, 2008. **250**(1): p. 25-36.
124. Turner, D.A., et al., *An interplay between extracellular signalling and the dynamics of the exit from pluripotency drives cell fate decisions in mouse ES cells*. Biol Open, 2014. **3**(7): p. 614-26.
125. Nichols, J., et al., *Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4*. Cell, 1998. **95**: p. 379-391.
126. Mitsui, K., et al., *The Homeoprotein Nanog is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells*. Cell, 2003. **113**: p. 631-642.
127. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function*. Genes Dev, 2003. **17**: p. 126-140.
128. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells*. Cell, 2005. **122**(6): p. 947-56.
129. Thomson, M., et al., *Pluripotency factors in embryonic stem cells regulate differentiation into germ layers*. Cell, 2011. **145**(6): p. 875-89.
130. Warmflash, A., B.L. Arduini, and A.H. Brivanlou, *The molecular circuitry underlying pluripotency in embryonic stem cells*. Wiley Interdiscip Rev Syst Biol Med, 2012. **4**(5): p. 443-56.
131. Nazareth, E.J., et al., *High-throughput fingerprinting of human pluripotent stem cell fate responses and lineage bias*. Nat Methods, 2013. **10**(12): p. 1225-31.
132. Wang, Z., et al., *Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells*. Cell Stem Cell, 2012. **10**(4): p. 440-54.
133. Smith, A., et al., *Differentiation inhibiting activity (DIA/LIF) and mouse development*. Dev Biol, 1992. **151**: p. 339-351.
134. Niwa, H., et al., *A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells*. Nature, 2009. **460**(7251): p. 118-22.

135. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos*. Nature, 1981. **292**: p. 154-156.
136. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells*. Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7634-7638.
137. White, D.E., et al., *Spatial pattern dynamics of 3D stem cell loss of pluripotency via rules-based computational modeling*. PLoS Comput Biol, 2013. **9**(3): p. e1002952.
138. White, D.E., et al., *Quantitative multivariate analysis of dynamic multicellular morphogenic trajectories*. Integr Biol (Camb), 2015. **7**(7): p. 825-33.
139. Nissen, S.B., et al., *Four simple rules that are sufficient to generate the mammalian blastocyst*. PLoS Biol, 2017. **15**(7): p. e2000737.
140. Buske, P., et al., *A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt*. PLoS Comput Biol, 2011. **7**(1): p. e1001045.
141. Hester, S.D., et al., *A multi-cell, multi-scale model of vertebrate segmentation and somite formation*. PLoS Comput Biol, 2011. **7**(10): p. e1002155.
142. Setty, Y., et al., *A model of stem cell population dynamics: in silico analysis and in vivo validation*. Development, 2012. **139**(1): p. 47-56.
143. Atwell, K., et al., *Mechano-logical model of C. elegans germ line suggests feedback on the cell cycle*. Development, 2015. **142**(22): p. 3902-11.
144. Fogel, J.L., et al., *A minimally sufficient model for rib proximal-distal patterning based on genetic analysis and agent-based simulations*. Elife, 2017. **6**.
145. Van Scoy, G.K., et al., *A cellular automata model of bone formation*. Math Biosci, 2017. **286**: p. 58-64.
146. Kaul, H., et al., *Synergistic activity of polarised osteoblasts inside condensations cause their differentiation*. Sci Rep, 2015. **5**: p. 11838.
147. Vasiev, B., et al., *Modeling gastrulation in the chick embryo: formation of the primitive streak*. PLoS One, 2010. **5**(5): p. e10571.
148. Herberg, M., et al., *Image-Based Quantification and Mathematical Modeling of Spatial Heterogeneity in ESC Colonies*. Cytometry A, 2015. **87A**: p. 481-490.
149. Herberg, M., et al., *Dissecting mechanisms of mouse embryonic stem cells heterogeneity through a model-based analysis of transcription factor dynamics*. J R Soc Interface, 2016. **13**(117).
150. Setty, Y., et al., *Four-dimensional realistic modeling of pancreatic organogenesis*. PNAS, 2008. **105**(51): p. 20374-20379.
151. Voss-Böhme, A., *Multi-Scale Modeling in Morphogenesis: A Critical Analysis of the Cellular Potts Model*. PLoS One, 2012. **7**(9).
152. Gallaher, J., et al., *Hybrid approach for parameter estimation in agent-based models*. bioRxiv Preprint, 2017.
153. Bartocci, E., et al., *A Formal Methods Approach to Pattern Recognition and Synthesis in Reaction Diffusion Networks*. IEEE Trans Control Netw. Syst., 2016. **5**(1): p. 308-320.
154. Briers, D., et al., *Pattern Synthesis in a 3D Agent-Based Model of Stem Cell Differentiation*. 55th IEEE Conference on Decision and Control, 2016.



155. Cockrell, R.C., et al., *Towards anatomic scale agent-based modeling with a massively parallel spatially explicit general-purpose model of enteric tissue (SEGMENT\_HPC)*. PLoS One, 2015. **10**(3): p. e0122192.
156. Richmond, P., et al., *High performance cellular level agent-based simulation with FLAME for the GPU*. Brief Bioinform, 2010. **11**(3): p. 334-47.
157. Jeannin-Girardon, A., P. Ballet, and V. Rodin, *Large Scale Tissue Morphogenesis Simulation on Heterogenous Systems Based on a Flexible Biomechanical Cell Model*. IEEE/ACM Trans Comput Biol Bioinform, 2015. **12**(5): p. 1021-33.
158. Nicholson, B.J., et al., *The molecular basis of selective permeability of connexins is complex and includes both size and charge*. Braz J Med Biol Res, 2000. **33**(4).
159. Weber, P.A., et al., *The permeability of gap junction channels to probes of different size is dependent on connexin composition and permeant-pore affinities*. Biophys J, 2004. **87**(2): p. 958-73.
160. Spagnol, G., et al., *Secondary structural analysis of the carboxyl-terminal domain from different connexin isoforms*. Biopolymers, 2016. **105**(3): p. 143-62.
161. Brink, P., *Gap Junction Voltage Dependence: A Clear Picture Emerges*. J. Gen. Physiol., 2000. **116**: p. 11-12.
162. Morley, G.E., et al., *Structure of connexin43 and its regulation by pHi*. J. Cardiovasc. Electrophys., 1997. **8**: p. 939-951.
163. Contreras, J.E., et al., *Metabolic inhibition induces opening of unapposed connexin 43 gap junction hemichannels and reduces gap junctional communication in cortical astrocytes in culture*. Proc Natl Acad Sci U S A, 2002. **99**(1): p. 495-500.
164. King, T.J. and L. P.D., *Temporal Regulation of Connexin Phosphorylation in Embryonic and Adult Tissues*. Biochem Biophys Acta, 2005. **1719**(1-2): p. 24-35.
165. Lemcke, H. and S.A. Kuznetsov, *Involvement of connexin43 in the EGF/EGFR signalling during self-renewal and differentiation of neural progenitor cells*. Cell Signal, 2013. **25**(12): p. 2676-84.
166. Zhang, Y.W., et al., *A Novel Route for Connexin 43 to Inhibit Cell Proliferation: Negative Regulation of S-Phase Kinase-associated Protein (Skp 2)*. Cancer Research, 2003. **63**: p. 1623-1630.
167. Johnson, R.G., et al., *Gap junction assembly: roles for the formation plaque and regulation by the C-terminus of connexin43*. Mol Biol Cell, 2012. **23**(1): p. 71-86.
168. Scemes, E., D.C. Spray, and P. Meda, *Connexins, pannexins, innexins: novel roles of "hemi-channels"*. Pflugers Arch, 2009. **457**(6): p. 1207-26.
169. Stout, C., D.A. Goodenough, and D.L. Paul, *Connexins: functions without junctions*. Curr Opin Cell Biol, 2004. **16**(5): p. 507-12.
170. Malarkey, E.B. and V. Parpura, *Mechanisms of glutamate release from astrocytes*. Neurochem Int., 2008. **52**(1-2): p. 142-154.
171. Berthoud, V.M., et al., *Roles and regulation of lens epithelial cell connexins*. FEBS Lett, 2014. **588**(8): p. 1297-303.
172. Mandal, A., M. Shahidullah, and N.A. Delamere, *Calcium entry via connexin hemichannels in lens epithelium*. Exp Eye Res, 2015. **132**: p. 52-8.
173. Novak, A., et al., *Proteomics profiling of human embryonic stem cells in the early differentiation stage*. Stem Cell Rev, 2012. **8**(1): p. 137-49.
174. Boucher, S. and S.A.L. Bennett, *Differential Connexin Expression, Gap Junction Intercellular Coupling, and Hemichannel Formation in NT2/D1 Human Neural*

- Progenitors and Terminally Differentiated hNT Neurons*. J Neuroscience, 2003. **72**: p. 393-404.
175. Abrams, C.K. and J.E. Rash, *Connexins in the Nervous System*. 2009: p. 323-357.
  176. Dermietrel, R., et al., *Gap Junctions between Cultured Astrocytes: Immunocytochemical, Molecular, and Electrophysiological Analysis*. J Neuroscience, 1991. **11**(5): p. 1421-1432.
  177. Froes, M.M., et al., *Gap-junctional coupling between neurons and astrocytes in primary central nervous system cultures*. Proc Natl Acad Sci U S A, 1999. **96**: p. 7541-7546.
  178. Nagy, J.I. and J.E. Rash, *Connexins and gap junctions of astrocytes and oligodendrocytes in the CNS*. Brain Research Reviews, 2000. **32**: p. 29-44.
  179. Fang, J.S., et al., *Connexin45 regulates endothelial-induced mesenchymal cell differentiation toward a mural cell phenotype*. Arterioscler Thromb Vasc Biol, 2013. **33**(2): p. 362-8.
  180. Hartfield, E.M., et al., *Connexin 36 expression regulates neuronal differentiation from neural progenitor cells*. PLoS One, 2011. **6**(3): p. e14746.
  181. Rinaldi, F., et al., *Cross-regulation of Connexin43 and beta-catenin influences differentiation of human neural progenitor cells*. Cell Death Dis, 2014. **5**: p. e1017.
  182. Santiago, M.F., et al., *The carboxyl-terminal domain of connexin43 is a negative modulator of neuronal differentiation*. J Biol Chem, 2010. **285**(16): p. 11836-45.
  183. Shi, Q., et al., *Connexin Controls Cell-Cycle Exit and Cell Differentiation by Directly Promoting Cytosolic Localization and Degradation of E3 Ligase Skp2*. Dev Cell, 2015. **35**(4): p. 483-96.
  184. Weider, K., et al., *Altered differentiation and clustering of Sertoli cells in transgenic mice showing a Sertoli cell specific knockout of the connexin 43 gene*. Differentiation, 2011. **82**(1): p. 38-49.
  185. Ressot, C., et al., *Connexin32 Mutations Associated with X-Linked Charcot–Marie–Tooth Disease Show Two Distinct Behaviors: Loss of Function and Altered Gating Properties*. J Neuroscience, 1998. **18**(11): p. 4063-4075.
  186. Gabriel, H., et al., *Transplacental Uptake of Glucose Is Decreased in Embryonic Lethal Connexin26-deficient Mice*. J Cell Bio., 1998. **140**: p. 1453-1461.
  187. Temme, A., et al., *High incidence of spontaneous and chemically induced liver tumors in mice deficient for connexin32*. Current Biology, 1997. **7**(9): p. 713-716.
  188. Simon, A.M., et al., *Female infertility in mice lacking connexin 37*. Nature, 1997. **385**: p. 525-529.
  189. Reaume, A.G., et al., *Cardiac malformation in neonatal mice lacking connexin43*. Science, 1995. **267**: p. 1831-1834.
  190. Todorova, M.G., B. Soria, and I. Quesada, *Gap junctional intercellular communication is required to maintain embryonic stem cells in a non-differentiated and proliferative state*. J Cell Physiol, 2008. **214**(2): p. 354-62.
  191. Boassa, D., et al., *Trafficking and recycling of the connexin43 gap junction protein during mitosis*. Traffic, 2010. **11**(11): p. 1471-86.
  192. Faherty, S., et al., *Self-renewal and differentiation of mouse embryonic stem cells as measured by Oct4 expression: the role of the cAMP/PKA pathway*. In Vitro Cell Dev Biol Anim, 2007. **43**(1): p. 37-47.

193. Ponsioen, B., et al., *Direct measurement of cyclic AMP diffusion and signaling through connexin43 gap junctional channels*. Exp Cell Res, 2007. **313**(2): p. 415-23.
194. Turner, D.A., P. Baillie-Johnson, and A. Martinez Arias, *Organoids and the genetically encoded self-assembly of embryonic stem cells*. Bioessays, 2016. **38**(2): p. 181-91.
195. Borodinsky, L.N., et al., *Spatiotemporal integration of developmental cues in neural development*. Dev Neurobiol, 2015. **75**(4): p. 349-59.
196. Dessaud, E., A.P. McMahon, and J. Briscoe, *Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network*. Development, 2008. **135**(15): p. 2489-503.
197. Kutejova, E., J. Briscoe, and A. Kicheva, *Temporal dynamics of patterning by morphogen gradients*. Curr Opin Genet Dev, 2009. **19**(4): p. 315-22.
198. Worsdorfer, P., et al., *Abrogation of Gap Junctional Communication in ES Cells Results in a Disruption of Primitive Endoderm Formation in Embryoid Bodies*. Stem Cells, 2017. **35**(4): p. 859-871.
199. Jaffe, L.F., *The role of ionic currents in establishing developmental pattern*. Phil. Trans R. Soc. Lond., 1981. **295**: p. 553-566.
200. Mathews, J. and M. Levin, *Gap junctional signaling in pattern regulation: Physiological network connectivity instructs growth and form*. Dev Neurobiol, 2017. **77**(5): p. 643-673.
201. Bevans, C.G., et al., *Isoform Composition of Connexin Channels Determines Selectivity among Second Messengers and Uncharged Molecules*. J Biol Chem, 1998. **273**(5): p. 2808-2816.
202. Fukumoto, T., R. Blakely, and M. Levin, *Serotonin transporter function is an early step in left-right patterning in chick and frog embryos*. Dev Neurosci, 2005. **27**(6): p. 349-63.
203. Martinez-Ramirez, A.S., et al., *Nucleotides and nucleoside signaling in the regulation of the epithelium to mesenchymal transition (EMT)*. Purinergic Signal, 2017. **13**(1): p. 1-12.
204. Leavesley, S.J. and T.C. Rich, *Overcoming limitations of FRET measurements*. Cytometry A, 2016. **89**(4): p. 325-7.
205. Abbaci, M., et al., *Advantages and limitations of commonly used methods to assay the molecular permeability of gap junctional intercellular communication*. Biotechniques, 2008. **45**(1): p. 33-52, 56-62.
206. Yi, C., et al., *Monitoring gap junctional communication in astrocytes from acute adult mouse brain slices using the gap-FRAP technique*. J Neurosci Methods, 2018. **303**: p. 103-113.
207. Xu, J., et al., *Retinoic acid promotes neural conversion of mouse embryonic stem cells in adherent monoculture*. Mol Biol Rep, 2012. **39**(2): p. 789-95.
208. Smith, K.N., A.M. Singh, and S. Dalton, *Myc represses primitive endoderm differentiation in pluripotent stem cells*. Cell Stem Cell, 2010. **7**(3): p. 343-54.
209. Rela, L. and L. Szczupak, *Gap Junctions: Their Importance for the Dynamics of Neural Circuits*. Molecular Neurobiology, 2004. **30**(3): p. 341-357.
210. Oyamada, M., et al., *Connexin expression and gap-junctional intercellular communication in ES cells and iPS cells*. Front Pharmacol, 2013. **4**: p. 85.

211. Lemcke, H., et al., *Neuronal differentiation requires a biphasic modulation of gap junctional intercellular communication caused by dynamic changes of connexin43 expression*. Eur J Neurosci, 2013. **38**(2): p. 2218-28.
212. Banfalvi, G., *Cell Cycle Synchronization*. Methods Mol Biol, 2011.
213. Abbaci, M., et al., *Gap junctional intercellular communication capacity by gap-FRAP technique: a comparative study*. Biotechnol J, 2007. **2**(1): p. 50-61.
214. Solan, Joell L. and Paul D. Lampe, *Connexin43 phosphorylation: structural changes and biological effects*. Biochemical Journal, 2009. **419**(2): p. 261-272.
215. Zhao, Q., et al., *Rapid induction of cAMP/PKA pathway during retinoic acid-induced acute promyelocytic leukemia cell differentiation*. Leukemia, 2004. **18**(2): p. 285-92.
216. Mondal, A., et al., *Computational simulations of asymmetric fluxes of large molecules through gap junction channel pores*. J Theor Biol, 2017. **412**: p. 61-73.
217. Suhail, Y., et al., *Modeling intercellular transfer of biomolecules through tunneling nanotubes*. Bull Math Biol, 2013. **75**(8): p. 1400-16.
218. Pietak, A. and M. Levin, *Exploring Instructive Physiological Signaling with the Bioelectric Tissue Simulation Engine*. Front Bioeng Biotechnol, 2016. **4**: p. 55.
219. Stumpf, P.S., et al., *Stem Cell Differentiation as a Non-Markov Stochastic Process*. Cell Syst, 2017. **5**(3): p. 268-282 e7.
220. MacArthur, B.D. and I.R. Lemischka, *Statistical mechanics of pluripotency*. Cell, 2013. **154**(3): p. 484-9.
221. Mathieu, J. and H. Ruohola-Baker, *Metabolic remodeling during the loss and acquisition of pluripotency*. Development, 2017. **144**(4): p. 541-551.
222. Kallas, A., et al., *Nocodazole treatment decreases expression of pluripotency markers Nanog and Oct4 in human embryonic stem cells*. PLoS One, 2011. **6**(4): p. e19114.
223. Turcato, S. and L.H. Clapp, *Effects of the adenylyl cyclase inhibitor SQ22536 on iloprost-induced vasorelaxation and cyclic AMP elevation in isolated guinea-pig aorta*. Br. J. Pharm., 1999. **126**: p. 845-847.
224. Mandegar, M.A., et al., *CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs*. Cell Stem Cell, 2016. **18**(4): p. 541-53.
225. Chambers, S.M., et al., *Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling*. Nat Biotechnol, 2009. **27**(3): p. 275-80.
226. Etoc, F., et al., *A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization*. Dev Cell, 2016. **39**(3): p. 302-315.
227. Wilensky, U. *NetLogo*. 1999; Available from: <http://ccl.northwestern.edu/netlogo/>.
228. Warmflash, A., et al., *A method to recapitulate early embryonic spatial patterning in human embryonic stem cells*. Nat Methods, 2014. **11**(8): p. 847-54.
229. Tewary, M., et al., *A stepwise model of reaction-diffusion and positional information governs self-organized human peri-gastrulation-like patterning*. Development, 2017. **144**(23): p. 4298-4312.
230. Mehdipour, N., et al., *Spatial-Temporal Pattern Synthesis in a Network of Locally Interacting Cells*. IEEE Conference on Decision and Control, 2018.