UNIVERSITY OF CALIFORNIA

Santa Barbara

A biophysical and bioinformatic approach to the analysis of canonical and non-canonical signaling dynamics involved in gastrulation

A Thesis submitted in partial satisfaction of the

requirements for the degree

Master of Science

in

Molecular, Cellular, and Developmental Biology

by

Chongxu Qiu

Committee in charge:

Professor Max Wilson, Chair

Professor Anthony De Tomaso

Professor Siddharth Dey

June 2023

The thesis of Chongxu Qiu is approved.

_____

Anthony De Tomaso

_____

Siddharth Dey

_____

Max Wilson, Committee Chair

June 2023

A biophysical and bioinformatic approach to the analysis of canonical and non-canonical signaling dynamics involved in gastrulation

by

Chongxu Qiu

ACKNOWLEDGEMENTS

Thank you to my advisor, Dr. Max Wilson, for enabling me to embark on this multidisciplinary journey through computational and developmental biology. Thank you for taking a leap of faith and allowing a graduating senior to join your lab three years ago when I didn't have the grades, experience, or qualifications. You gave me the freedom to explore and try any project I wanted to, and this work would have not been possible without your guidance. You gave me the opportunity of a lifetime, and I owe much of my current success to you. I hope I've made you proud and I hope that you will continue to watch what I do moving forward in my scientific career.

Thank you to Dr. Dasol Han, for giving me the opportunity to work under you despite my complete lack of laboratory experience. Thank you for the friendship, patience, and guidance you have shown me. Thank you for challenging me and pushing me out of my comfort zone. You taught me nearly everything I know, and your mentorship had the utmost influence on how I think and approach science. Thank you for molding me into the scientist I am today.

Thank you to my committee members, Dr. Tony De Tomaso and Dr. Siddharth Dey. Your advice has been invaluable and every conversation was a fantastic learning experience.

Thank you to Dr. Sanza Kazadi, for being the mentor that jump started my love and fascination for science. You were the first to show me the rigor, standard, and intensity necessary to be successful in science. Many of my learning and

teaching principles were derived from you and thank you for constantly being available to chat about anything and everything.

Thank you to Dr. Joesph Rufo, Robert Piscopio, Jose Armando Nunez, Mickolaj Godzik, Naomi Baxter, Surenna Pecchia, Sam Rosen, Erik Hopkins, Ian Matthews, Alex Maynard, for the advice, help, and support you have provided me during my time here as a graduate student. Your kindness made the late hours in lab much more manageable.

Lastly, I would like to thank my mother who has been an unwavering pillar of support, without whom none of this would have been possible. Thank you for all the sacrifices and for carrying the burden of two parents. Thank you for constantly pushing me to become a better person and for the constant advice.

ABSTRACT

A biophysical and bioinformatic approach to the analysis of canonical and non-canonical signaling dynamics involved in gastrulation.

by

Chongxu Qiu

There are numerous signaling dynamics involved in gastrulation. Using biophysical modelling, we find that liquid-liquid phase separation significantly improves the processing of transcriptional factor β-catenin in the canonical WNT signaling cascade. Using a finite element method, we find that a nucleating region can facilitate a nearly two-fold increase in β-catenin processing. Moving away from microscale biophysical dynamics, we employed a large-scale loss of function drug screen on 2D cultured embryo bodies and found that the 2D system can be used as a powerful, cheap, and reproducible teratogen screening platform and discovered the surprising role of inflammatory and immunological signaling on human gastrulation.

TABLE OF CONTENTS

Section 1

A biophysical modeling approach to the destruction complex of WNT signaling

Chapter 1 Introduction

1.1 Canonical WNT signaling

The canonical Wnt signaling pathway is a conserved, morphogenetic pathway that is essential for embryonic development, maintenance of adult tissue homeostasis, and induces malignant cancer when dysregulated [1-4]. Wnt signaling is facilitated by a protein assembly called the destruction complex (DC), which adjusts the cellular availability of β-catenin (β-cat), the pathway's central transcriptional effector, by tuning its interactions with Wnt kinases, CK1α and GSK3β, and the ubiquitinase, β-TRCP, which directs β-cat to the ubiquitin-mediated proteolysis machinery.

In a WNT active state, the kinase machinery that maintains low β-cat levels becomes dissociated, which allows β-cat to localize to the nucleus and induce gene transcription. The Wnt ligand binds to the extracellular domain of Frizzled, a plasma membrane spanning receptor, along with lipoprotein receptor-related protein 5/6 (LRP5/6). Upon signal activation, translocation of Axin and the DC to the nucleus occurs and phosphorylation induces binding of Axin to the cytoplasmic tail of LRP5/6. Disheveled (Dsh), is activated via phosphorylation and deactivates GSK3β, allowing for β-cat levels to increase. [5,21] β-cat level increases lead to nuclear localization where it recruits transcriptional coactivators. In the WNT inactive state, β-cat is continually degraded by the kinase constituents of the DC, where

ubiquitination of β-cat sends it to the proteasome to be digested, maintaining tissue homeostasis. [6]

There are two prevailing interpretations of the structural dynamics of canonical WNT OFF signaling. The first follows a signaling dynamic pipeline where β-cat and its kinase regulators colocalize, with β-cat being processed in an assembly-line-like model. The second interpretation involves the formation of a liquid-liquid phase separating (LLPS) droplet consisting of β-cat and its kinase interactors on Axin and Adenomatous Polyposis Coli (APC) with processing being improved by this biophysical dynamic.[8]

1.2 Structural Dynamics of the WNT

The extracellular signaling dynamics of WNT involves the selective recruitment of DC components to a biomolecular condensate on the plasma membrane consisting of Frizzled/LRP5/6 clusters [5]. Recent work has demonstrated that optogenetic clustering of LRP5/6 is sufficient to decrease β-cat levels, suggesting the formation of protein clusters at the Wnt receptor level is necessary in activating the pathway [7]. Less is known about the DC's native structure and how it maintains β-cat levels in the Wnt OFF state. Recently, DC scaffolds Axin and APC, and Disheveled have been shown to undergo LLPS both in vitro and in vivo (when overexpressed) [8-10]. It has been further demonstrated that cancer-causing mutations that eliminate Axin or APC LLPS correlate with uninhibited accumulation of β-cat [11-12]. This suggests that DC LLPS promotes β-cat degradation in Wnt OFF conditions [7] via concentration of DC kinases (CK1α, GSK3β) in DC scaffold (Axin and APC)

condensates to increase the rate of β-cat processing. LLPS-mediated DC concentration is distinct from models suggesting that the DC acts as an ordered, assembly-line-like manner. Conditions that alter the phase behavior of scaffolds and partitioning of clients are predicted to regulate the stability of β-cat and Wnt signal transmission. Building on results demonstrating that β-cat, Axin1 , and APC localize to the centrosome[16-19], we find that nucleation drives efficient degradation of β-cat and we utilize a Cahn-Hilliard-based simulation of DC droplet formation and enzyme kinetics to predict how nucleation and affinity of DC components promotes efficient β-cat processing.

Research Question

We are interested in exploring the role of liquid-liquid phase separation on enzymatic processing. A robust, modified Cahn-Hilliard based simulation was developed to explore the enzymatic dynamics in a theoretical cell environment.

Chapter 2: Methods

2.1 Finite Element Method and Program Pipeline

We used the python-based FEniCS computing environment to solve the modified reactive, multicomponent, Cahn-Hilliard partial differential equations (PDEs) using the Finite Element Method (FEM). The FEM is designed to solve differential equations that arise in engineering problems and can be expanded to structural analysis, heat transfer, and in our case, fluid dynamics. FEM is a general numerical method used for solving partial differential equations in two- or three-dimensional spaces. The operating principle of the FEM is that it separates a large system into smaller constituents called finite elements. This is achieved by discretization of the space by constructing a mesh of the object. By implementing a system of algebraic equations and approximating these finite elements into the larger whole, we can derive the behavior of the larger system.

Several reasons motivated our choice to use the FEM. Firstly, considering that cells can have variations in the local environment; namely the difference in the local dynamics of the cytoplasm and the nucleating region, the ability to define a mesh facilitates the inclusion of dissimilar material properties. Moreover, small-scale local effects like diffusion and accumulation can be captured. Additionally, the FEM has already been extensively used to develop various Cahn Hilliard models.

The functional pipeline of our code can be described as follows. First, all relevant parameters are defined (χ, λ, dt, and M). A square grid mesh with closed boundary conditions is established to mimic the closed system within a cell. To account for the nucleating region, a circular mesh is introduced. For each simulated

component, a layer is generated and +-5% noise of the initial value is added to induce inhomogeneities.

We found that for a system to phase separate using the Cahn Hilliard method, inhomogeneities are necessary. The Laplacian found in the Cahn Hilliard model, which intuitively acts as a mean finder in a field, requires noise for movement of components to occur. A uniform field results in a uniform mean, resulting in no change. Similarly, a uniform plane will result in a uniform Laplacian, and no phase separation will be observed.

Two functions are then defined, the first defines each DC component as a variable and builds the chemical potential equation. The second equation defines the Cahn-Hilliard, which takes input from the chemical potential equation to evolve the mesh over time. These two equations are the backbone of the modeling system and are highly modular, allowing for the user to define all tunable parameters. The final calculation is the built in Newtonian solver that iterates through a predefined number of steps to reduce the chemical potential, which is calculated at each time step before being fed into the Cahn Hilliard function. The final output of the equation is a series of meshes that represent the changing cellular system. These mesh files are then rendered and visualized using Paraview software.

2.2 Simulation Methods

In our simulation, we represent the volume fraction of each DC protein, ϕ_i, as an incompressible volume such that $\sum_{i=1}^{n} \phi = 1$ and approximate the reaction rates with spatially dependent analogues to well-mixed reactions using the simplified, non-state

dependent description of the second order rate $R_i = \pm k_{i,j}\phi_i\phi_j$,with production and consumption denoted by the sign of $k_{i,j}$The Cahn-Hilliard equation, in its general form, is a parabolic equation with first-order time derivatives, and second- and fourth-order spatial derivatives. To solve this equation using a standard Lagrange finite element basis the equation is recast as two coupled second-order equations:

$$\frac{\partial \phi_i}{\partial t} = \nabla \cdot M\big(\nabla(\mu_i)\big) \pm R\big(k_{i,j}\phi_i\phi_j\big)$$

$$\mu_i = \frac{df}{d\phi_i} - \lambda\nabla^2\phi_i$$

Where M is the mobility constant, with all DC components having the same diffusion rate, λ is the surface energy parameter that dictates the length of transition regions between domains, and is the polynomial double-well description of the free energy:

$$F = \sum_{i=1}^{N-1}\sum_{j=2}^{N}\chi_{i,j}\phi_i^2\phi_j^2$$

Where, $\chi_{i,j}$ describes interaction strength between DC proteins, the cytoplasm, and the centrosome. We modeled centrosomal nucleation as a region in the simulation with increased interaction strength as has been done previously to describe nucleation sites. To determine the size of this nucleation region we measured the relative volume of centrosomal-localized DC kinases and β-cat. And R is the added reaction term with such that

$R_i(k_{i,j}, \phi_i, \phi_j) = k_{i,j}\phi_i\phi_j$ for the creation of ϕ_i

and

$R_i(k_{i,j}, \phi_i, \phi_j) = -k_{i,j}\phi_i\phi_j$ for the consumption of ϕ_i.

The system is time discretized according to established methods. Assuming that the total free energy of the system decreases to a minimum with time, we use the built-in Newtonian solver in the FEniCS environment to approximate the forward evolution of the system in time. To represent the enzyme activities in the DC, we model only clients, with scaffolds existing implicitly as the interaction parameters between system components.

Component Variable	Component Name	CH Equation
ϕ_1	GSK3β	$\frac{\partial \phi_1}{\partial t} = M_1 \nabla^2 \mu_1$
ϕ_2	CK1α	$\frac{\partial \phi_2}{\partial t} = M_2 \nabla^2 \mu_2$
ϕ_3	β-Catenin	$\frac{\partial \phi_3}{\partial t} = M_3 \nabla^2 \mu_3 - \left(k_{2,3} \cdot \phi_2 \cdot \phi_3\right)$
ϕ_4	Phospho-S45 β-Catenin	$\frac{\partial \phi_4}{\partial t} = M_4 \nabla^2 \mu_4 + (k_{2,3} \cdot \phi_2 \cdot \phi_3) - (k_{1,4} \cdot \phi_1 \cdot \phi_4)$
ϕ_5	Phospho-S33/S37/S45/T41 β-Catenin	$\frac{\partial \phi_5}{\partial t} = M_5 \nabla^2 \mu_5 + (k_{1,4} \cdot \phi_1 \cdot \phi_4)$
ϕ_6	Cytoplasm	$\frac{\partial \phi_6}{\partial t} = M_6 \nabla^2 \mu_6$
ϕ_7	Nucleator	$\frac{\partial \phi_7}{\partial t} = M_7 \nabla^2 \mu_7$

Table 1: Describes each of the DC kinase components modelled in the system along with their respective Cahn Hilliard equations and reactive systems

2.3 Interaction Parameter

One of the key factors that tunes system behavior is the interaction parameter χ. Assuming a system with constant temperature and pressure, the interaction parameter determines the free energy of the system. The Cahn Hilliard model is partially based on the Flory-Huggins solution theory, where $\chi = z\left(2\omega_{ij} - \omega_{ii} - \omega_{jj}\right)/(2k_bT)$, and ω are the pairwise repulsive interaction energies between components i and j. When χ is positive between two components, the system will tend to de-mix. If χ is negative between two components, they will tend to mix. Lastly, if χ is neutral, the two components are interaction-less. For simplicity, we limited interactions to one of three types: binding (χ ~ -0.1), neutral (χ ~ 0), and separating (χ ~ 2). We had to choose a small binding constant because higher values resulted in a loss temporal simulation resolution as the Newtonian solver would iterate through the chemical potential too quickly. As noted above we represent the binding action of DC scaffolds implicitly. Given that the APC/Axin interacts with the DC proteins, the following interaction constants were selected for the system with implicit Axin. We set mixing = 2.0, neutral = 0.0, and de-mixing = -0.1.

Interaction	Behavior	Source
GSK/Axin-CK1/Axin	Neutral	Ref. 27, 31
GSK/Axin-βCat/Axin	Neutral	Ref. 27, 31
GSK/Axin-P1βCat/Axin	Neutral	Ref. 27, 31
GSK/Axin-P4βCat/Axin	Neutral	Ref. 27, 31
GSK/Axin-Cytoplasm	Separating	Ref. 30
GSK/Axin-Centrosome	Binding	Ref. 30
CK1/Axin-βCat/Axin	Neutral	Ref. 27, 31
CK1/Axin-P1βCat/Axin	Neutral	Ref. 27, 31
CK1/Axin-P4βCat/Axin	Neutral	Ref. 27, 31
CK1/Axin-Cytoplasm	Separating	Ref. 27, 31
CK1/Axin-Centrosome	Binding	Ref. 27, 31
βCat/Axin-P1βCat/Axin	Neutral	Ref. 27, 31
βCat/Axin-P4βCat/Axin	Neutral	Ref. 27, 31
βCat/Axin-Cytoplasm	Separating	Ref. 27, 31
βCat/Axin-Centrosome	Binding	Ref. 27, 31
P1βCat/Axin-P4βCat/Axin	Neutral	Ref. 27, 31
P1βCat/Axin-Cytoplasm	Separating	Ref. 27, 31
P1βCat/Axin-Centrosome	Binding	Ref. 27, 31
P4βCat/Axin-Cytoplasm	Neutral	Ref. 27, 31
P4βCat/Axin-Centrosome	Neutral	Ref. 27, 31
Cytoplasm-Centrosome	Separating	Ref. 27, 31

Table 1.2 Binding actions of DC components and derived citations

Chapter 3. Results

3.1 Simulation Results

Building on previous work that demonstrated LLPS of DC components can be induced by the centrosome acting as a nucleating agent, we simulated the phosphorylation of β-cat by DC kinases, using a multi-step reactive, multi-component, Cahn-Hilliard system [24-26]. We chose to represent the function of DC scaffolds implicitly through the interaction parameters between kinases and β-cat.

To test the effects of nucleation on β-cat processing, we compared simulations in the presence and absence of a nucleation region as well as systems where DC components were overexpressed. In the absence of a nucleating agent, which mimicked endogenously expressed conditions, the systems did not undergo spontaneous phase separation. However, when a nucleating agent was present, DC components localized into a single droplet surrounding the nucleation region. When we artificially increased the amount of DC components, we saw aberrant droplet formation, which mimicked overexpressed conditions. (Fig 1.1 C)

We also found that the nucleated system processed β-cat and its intermediates at an accelerated rate over a wide range of nucleator sizes. Notably, the nucleated system accelerated β-cat processing, increasing pathway efficiency. This efficiency gain was maintained over a large range of reaction rates. We believe that cells utilize phase separation to induce localized improvements in processing and in a perfect system where proteins are processed upon synthesis. As predicted, in systems with high reaction rates, the processing improvement of nucleated phase separation is no longer observed. (Fig 1.1 D-F)

Given our findings that nucleation drives efficient processing of β-cat, we hypothesized that the interaction parameter that drives phase separation is a control parameter for β-cat processing. To determine the relationship between processing efficiency and the interaction strength parameter, we systematically decreased the χ between DC components and the cytoplasm. We found that reducing condensation amount and rate on the nucleator by altering χ decreased the speed and efficiency of β-cat processing. Together, these results demonstrate that nucleation of DC components has the potential to increase β-cat processing and that a tunable control parameter of this process is the free energy of mixing.

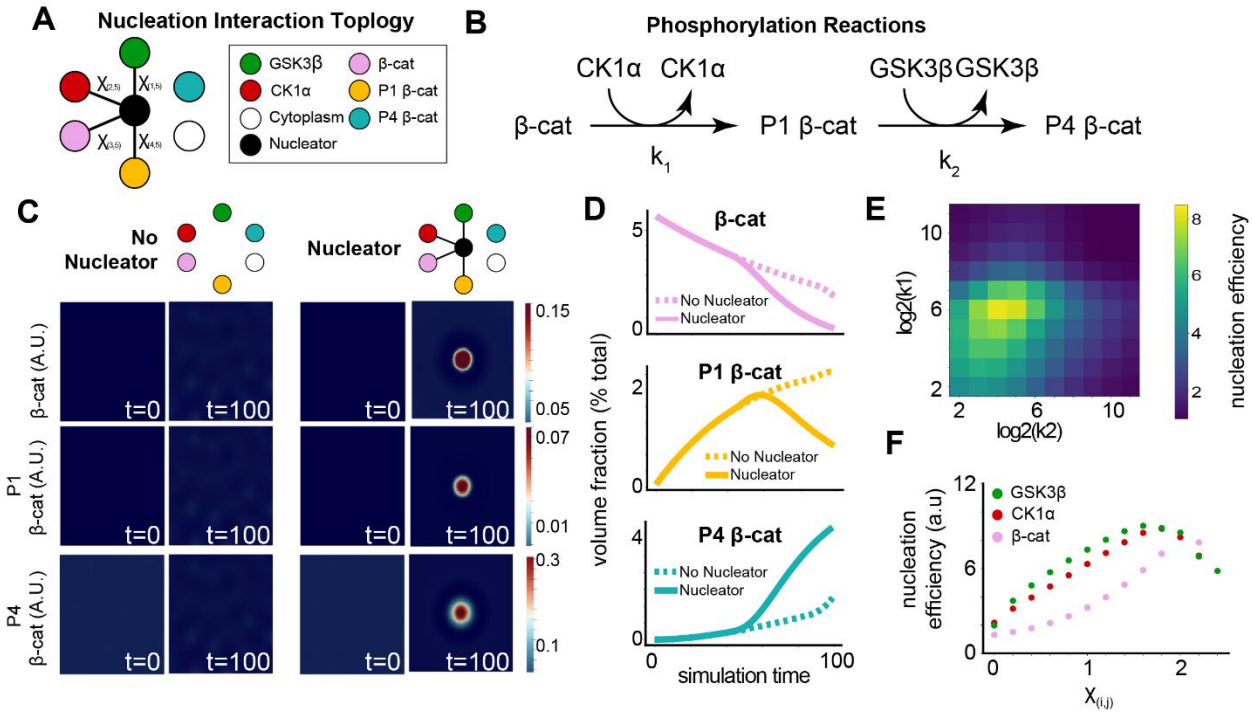

Figure 1.1: In silico modeling of β-cat processing efficiency from a nucleated liquid droplet. (A) Nucleation interaction topology that describes the pairwise interactions between each component of the simulation. Connected components minimize free energy by mixing, and unconnected components either de-mix or remain in a noninteracting neutral state. (B) Schema describing the phosphorylation reactions and rates modeled in the simulation. (C) Simulation at steps 0 and 100 comparing a system with and without a centrosome. (D) Quantification of each form of β-cat with and without a centrosome. (E) Nucleation efficiency as a function of both rate parameters k_1 and k_2. (F) Nucleation efficiency in simulations as a function of the interaction parameters between a single client and the cytoplasm.

3.2 Robustness Test: Diffusion Rate, Reaction Rate, Interaction Constants

After generating the initial set of simulations, we assessed the robustness of the modeling system. Specifically, we performed scans on the diffusion rate M, reaction rate $k_{i,j}$, and interaction constants $\chi_{i,j}$. We theorized that once the reaction rates of the system reach sufficiently high levels, the effect of LLPS will have no noticeable effect on the enzymatic activity of the system.

To investigate this, we ran a linear scan on $k_{i,j}$ on a 2^n scale from 2 to 16,384. The results demonstrated that phase separation drastically improves phosphorylation activity at intermediate rates of reaction. However, phase separation was shown to no longer affect reaction speed in systems with high reaction rates. In an environment with high reaction rates, the system is composed of diffusion-limited enzymes, with enzyme activity occurring upon contact.

Considering that both diffusion rates and interaction constants are intrinsically linked to phase separation behavior, we theorized that increasing these values would result in increased b-cat processing speed. The results demonstrated that processing speed is impacted by these constants. By increasing diffusion rates and interaction constants, the DC constituents can form interaction centers at an accelerated rate, so although the overall efficiency doesn't change, the rate of reaction does.

3.3 Robustness Test: Reaction Rate vs Interaction Constant

We conducted a comprehensive 10x10 scan where we altered the reaction rates and interaction constants. The first and second β-catenin processing rates were identical, ranging from a scale from 4 to 2048. For the interaction constant, we increased by increments of 0.2 starting from 0.6 to 2.4. Following the same procedure as the reaction rate scan, we generated efficiency matrices to assess the impact of these variations. The difference in efficiency shows that increased phase separation also boosts enzymatic activity at low reaction rates. (Fig 1.2 A) Examining the efficiency comparison matrix, we notice a hotspot in the region of high phase separation and intermediate reaction rate which further demonstrates that phase separation is used by the cell to increase localized enzymatic processing. (Fig 1.2 B)

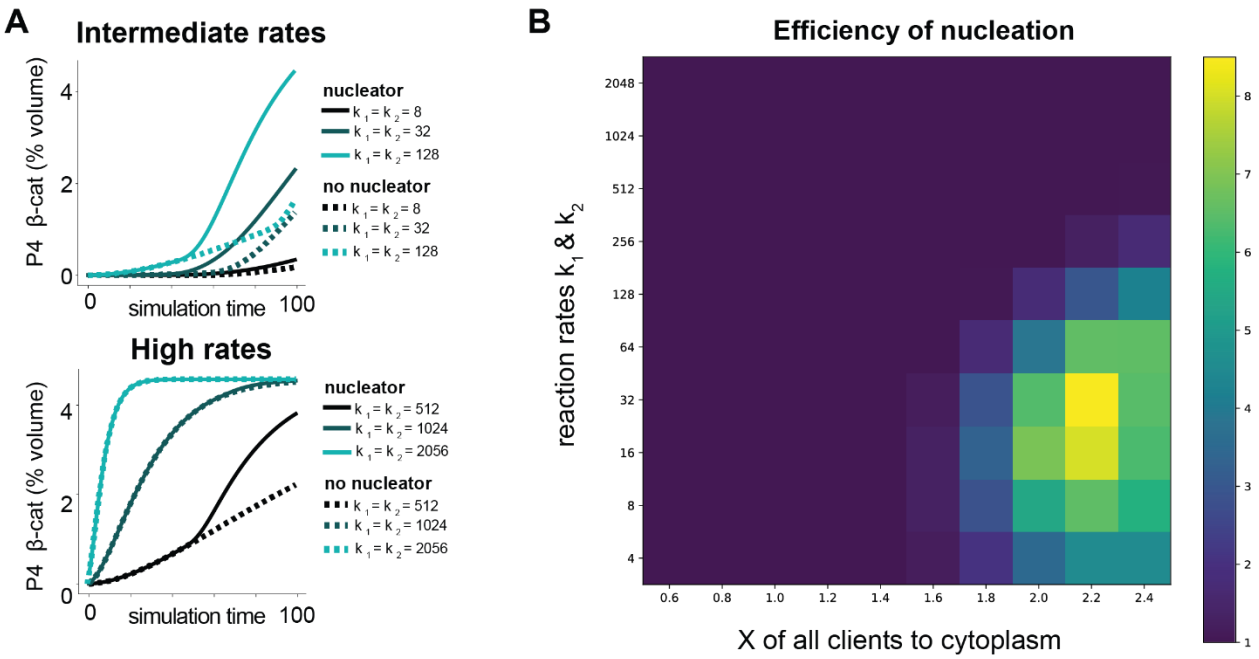

Figure 1.2: (A) P4 β b-catenin accumulation in the log2 scan of kinase reaction

rates. (B) Nucleation efficiency as function of reaction rate and interaction constants $\chi_{i,j}$

3.4 Nucleator Size Scan

To further investigate the effects of the nucleating region on enzymatic activity, we utilized the model to adjust the nucleating region's size. From the image analysis of live cells, the nucleating region roughly occupies ~0.005% of the cell. Considering that our mesh is built on a 1 by 1 grid, we can approximate this area as representative of a cell. We proceeded to measure the enzymatic activity of a nucleating region at various sizes: 0.001, 0.005, 0.01, 0.05, and 0.1. Surprisingly, we found that having a nucleating region that was 0.5% of the system generated the highest level of enzymatic activity.

Interestingly, we observed a surprising phenomenon where larger nucleating regions would initially form smaller punctum before merging into a single larger punctum. (Fig 1.3 A) On the other hand, the smaller nucleating regions would immediately form a single punctum. This phenomenon led to notable differences in processing efficiency. The smaller nucleator demonstrated nearly a threefold increase in efficiency compared to a system with a larger nucleating region.

The reason behind this enhancement in enzymatic activity with a smaller nucleation region can be attributed to the rapid formation of a more concentration interaction center. When the nucleating region is smaller, the limited space prompts the components of the DC to converge sooner, enabling the initiation of

phosphorylation, and reduces the amount of random non-reactive interactions. Compared to a larger nucleating region, the separate puncta can have highly variable DC components in each droplet, resulting in reduced β -catenin processing. (Fig 1.3 C-D)

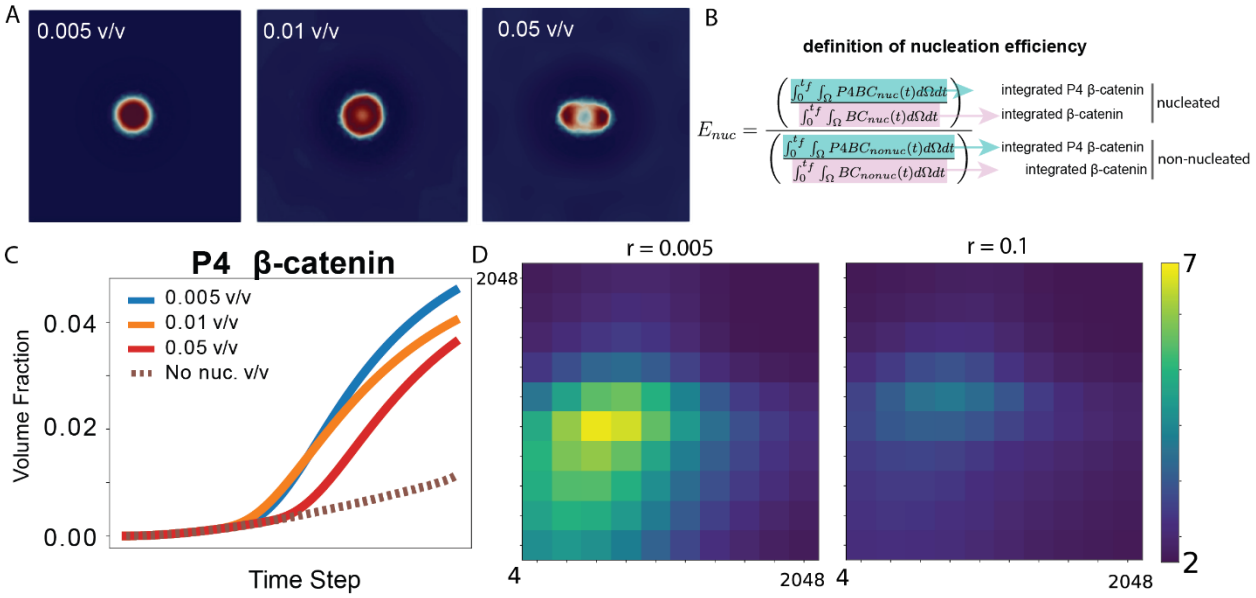

Figure 1.3: (A) Demonstration of the effects of nucleator size on system nucleation behavior. Nucleator size affects how DC components phase separate onto the nucleating agent and show sub-structure formation. (B) Definition of nucleation efficiency as the ratio of the quotient of P4 β -catenin and β catenin in a nucleated vs unnucleated system. (C) Measuring the gain in processing with a smaller nucleator. (D) Calculation of the efficiency gain across a log2 scale of reaction rates comparing a smaller nucleator to a larger one.

Chapter 4. Discussion

Based on the recent work that suggested LLPS plays a role in the destruction complex structure, we successfully demonstrated that LLPS facilitates enzymatic dynamics by utilizing a robust biophysical model. Using this model, recent work has been done using optogenetic clustering on the centrosome via an optogenetic GSK3 β tool. The optogenetic GSK3 β induces clustering of GSK β and other WNT DC components and demonstrates an increase in β-catenin processing when cells expressing the optogenetic protein are exposed to light. OptoGSK-β has been shown to also localize to the centrosome, a potential nucleating region, which further proves the validity of the biophysical simulation. [74]

Although the simulation is robust, we do not account for a variety of other factors that can also affect processing. As mentioned before, it has been shown that the DC components tend to localize onto the centrosome. The centrosome also plays a key role in cell division and its structural dynamics can change depending on cell cycle. We do not account for how cell cycle and cell division events can affect the overall diffusion rates within a cell as well how the changes in centrosomal structure can impact how the DC components LLPS onto the centrosome. Furthermore, the simulation is largely constrained due to the fact we only model one specific dynamic within the cell, while there are a huge variety of different factors that can impact processing. In addition, the model is a Cahn-Hilliard model, one of the many fluid dynamic simulation methods. Further work using different models is necessary to further validate the increase in enzymatic processing. Lastly, there are

many known interactors that affect LLPS within cells. This has a potential lead in therapeutic implications since further work can be conducted in seeing how LLPS affects disease progression and how changing this dynamic can improve the LLPS related malignancies.

Section 2

The exploration of non-canonical signaling pathways during early human gastrulation

Introduction

Chapter 1: Human Embryonic Development

1.1 Early Prenatal Development (Embryogenesis)

Fertilization marks the beginning of embryonic development. The fusion of the ovum and sperm gives rise to the zygote, which consists of totipotent blastomeres. To prevent additional implantation, a glycoprotein shell called the zona pellucida forms around the zygote. During this time, blastomeres undergo cleavage, a specific method of embryonic mitosis, where they divide and generate new cells without increasing the overall volume of the embryo. At the 16-cell stage, known as the morula, the size remains comparable to that of the fertilized ovum. The compaction of blastomeres is vital for the formation of tight junctions and cell polarization, which are crucial for the subsequent development of the blastocyst, facilitated by the zona pellucida. During this stage, blastomeres differentiate for the first time, giving rise to two distinct populations of cells: the inner cell mass (ICM) and the outer layer of cells known as trophoblasts. [34]

At this point, the cells form a uniform ball-like structure. The ICM will eventually develop into the embryo, while the trophoblasts will play a major role in forming the placenta, which supplies nourishment to the developing embryo. The establishment of the trophoblastic layer initiates cavitation, a process in which sodium pump activity creates an osmotic gradient, resulting in the formation of a

fluid-filled cavity called the blastocoel.[35] This stage is known as the blastocyst. Tight junction formation becomes crucial during this phase, as leakage from blastocoel formation can lead to implantation failure. Implantation serves as an early checkpoint for embryonic development. Typically, the blastocyst will embed itself in the endometrial stroma of the uterine wall, and the levels of human chorionic gonadotropin in maternal serum and urine increase. These hormone levels serve as early indicators of pregnancy, although approximately 30 percent of pregnancies may fail at this stage. However, if the blastocyst successfully implants, further rounds of differentiation occur. [36,37]

The trophoblast differentiates into two distinct layers: the outer trophoblast (syncytiotrophoblast) and the inner trophoblast (cytotrophoblast). This differentiation event allows the blastocyst to implant into the uterine wall by enzymatically digesting the uterine tissue, facilitating implantation. Simultaneously, the ICM expands and forms a flattened tissue layer, differentiating into the epiblast and hypoblast. This process establishes the dorsal and ventral axes and represents a precursor event to gastrulation. [38]

1.2 Gastrulation

Gastrulation occurs after the establishment of the dorsal and ventral axes in the embryo. It is marked by the emergence of a groove at the caudal end of the epiblast known as the primitive streak. The primitive streak establishes the cranial and caudal axis. At the cranial end of the primitive streak, epiblast cells undergo invagination and form the primitive pit.[39] The primitive streak and pit will elongate

across the gastrula and eventually create a primitive node that organizes the primordia. Following the formation of the primitive node, waves of cell migration and differentiation occur. Epithelial cells will undergo differentiation into mesenchymal cells and migrate down the primitive streak via epithelial-mesenchymal transition (EMT). These mesenchymal cells will integrate into the hypoblast layer and further differentiate into endodermal cells. A secondary wave of differentiation and migration results in the formation of the mesoderm. The remaining epiblast will then differentiate into ectoderm. [40]

Each of these germ layers corresponds to primordia that is crucial to organogenesis, the endoderm forms the gastrointestinal and respiratory tract, the mesoderm forms the mesenchyme, mesothelium, and blood cells, and the ectoderm forms the skin and neural tissue. Furthermore, gastrulation also provides the mechanisms necessary to establish a multileveled body plan, left/right symmetry, and loss of bilateral symmetry. Failure in coordination of proper germ layer formation can result in congenital anomalies post birth, some detectable at birth and some apparent during the postnatal year. In severe cases, malformed embryos will be spontaneously aborted during the first weeks of development. Given the intricacy of this process, it is important to understand the signaling dynamics that are crucial in ensuring that this pattern formation occurs properly. By unraveling the molecular and signaling dynamics of this process, we can gain insights into the regulatory processes that govern embryonic development. [41-43]

1.3 Signaling Dynamics of Gastrulation

Gastrulation is a highly complex and dynamic process that involves a series of intricate cellular and molecular events, encompassing cell morphogenesis, movement, and signaling. Multiple factors contribute to gastrulation, including transcriptional factors, morphogenetic gradients, and differential gene expression, with spatial and temporal aspects playing crucial roles. Traditionally, the study of gastrulation has focused on three key signaling pathways: Bone Morphogenic Protein (BMP), wingless/int-1 (WNT), and Activin/Nodal. These pathways form a positive feedback loop that establishes embryo polarity and guides the development of the three germ layers.[46-50]

A well-established cascade of signaling events has been elucidated, beginning with BMP signaling in the extraembryonic ectoderm, which initiates WNT signaling in the epiblast and endoderm. WNT signaling, in turn, activates NODAL signaling, which provides a positive feedback loop and sustains BMP signaling. This circularized signaling dynamic is instrumental in initiating the formation of the primitive streak and germ layers. Through the induction of epithelial-mesenchymal transition, cells undergo a transformation that leads to migration from the epiblast to the hypoblast, facilitating the establishment of germ layers.[51]

Due to the ethical constraints on studying human embryos, particularly the widely accepted 14-day rule that prohibits experimentation beyond this stage, researchers have primarily relied on mouse and zebrafish embryo models to infer the mechanisms of human gastrulation. However, recent advances in micropattern

technologies and the isolation of human embryonic tissues have paved the way for the development of highly repeatable and accurate human models.[52]

1.4 Human Embryonic Stem Cells

Human embryonic stem cells (hESCs) have revolutionized scientific research since their successful isolation from the ICM of the blastocyst. In 1998, James A. Thomson successfully isolated the first hESCs using immunosurgery techniques. The process involved the use of Pronase to eliminate the zona pellucida, followed by an anti-human whole serum antibody to remove any extraembryonic tissue. This breakthrough opened a new era of research possibilities, particularly in the realm of human development and stem cell therapeutics. [53]

The defining characteristic of hESCs is their pluripotency, which refers to their ability to differentiate into the three principal germ layers: ectoderm, mesoderm, and endoderm. These germ layers play a crucial role in the development of various organs in the body. A key distinction between totipotent and pluripotent cells lies in the former's unique capacity to form extraembryonic tissue.

Since their initial isolation, hESCs have rapidly emerged as a powerful tool in the field of therapeutics and regenerative medicine. Significant progress has been made in inducing tissue differentiation to obtain pure populations of specific cell types. Furthermore, hESCs have recently been employed to develop model systems that enable the study of human gastrulation, shedding light on the early stages of embryonic development. These advancements hold immense potential for improving our understanding of human biology and advancing medical treatments.[54]

1.5 Micropatterning and germ layer markers

Under standard culturing conditions, human embryonic stem cell colonies grow in a wide range of sizes and shapes. When these unorganized colonies are treated with BMP4, typically used to induce gastrulation, spatial patterning in these colonies vary drastically. However, if the colonies' growth is constrained on lithographically micropatterned extracellular matrices, germ layer pattern formation occurs consistently. SOX2 expression is specifically maintained during ectodermal differentiation and reduced in mesodermal differentiation, Brachyury (BRA) is specifically maintained in mesoderm formation, and GATA3/CDX2 is maintained in endodermal differentiation. During pattern formation on micropatterned colonies, SOX2 is expressed at the center of the colony, a ring of BRA is expressed followed by an outermost ring of GATA3. There is slight co-expression of BRA and GATA3 in regions where mesoderm and endoderm overlap, but not expression of SOX2 in either of the outer two layers. Given the consistency of this pattern formation, we chose to use this model to explore a simple question, what other signaling pathways are involved in early gastrulation? Although it is demonstrated that 2D embryoid bodies are a good representation of the 3D embryo, there are still many key dynamics that are overlooked when doing so, thus the development of 3D gastruloids is also crucial for exploring development. [55-56]

Research Question

We are interested in exploring overlooked signaling pathways in the context of human gastrulation. BMP, WNT, NODAL, and TGF are frequently studied in conjunction, but it is likely that there are also other surface receptors on human embryonic stem cells that sense the microenvironment during gastrulation that help the system decide pattern formation, cell commitment to a specific fate, and termination. Thus, we utilized a high-throughput drug screening method as a loss of function experiment to explore this question.

Chapter 2: Materials and Methods

2.1 Cell Culture:

All experiments were performed with the H9 (WA09) human embryonic stem cells from WiCell (Cat No. WB66446). For routine cell culture maintenance, the H9 cells are grown on Corning Matrigel hESC-Qualified Matrix, LDEV-free (Cat No. 354277) coated dishes and cultured in mTeSR Plus feeder free maintenance medium (StemCell Technologies Cat No. 100-0276). Matrigel dishes were coated overnight at 4 C then warmed at room temperature for 1 hour prior to cell seeding. Cells passages were performed either manually via colony picking or using ReLeSR (Stemcell technologies). We used an optimized protocol where the cells were treated with ReLeSR for 30 seconds at room temperature, the ReLeSR was aspirated then, and the treated cells were placed in 37 C for 3 minutes before being lifted with mTeSR. Pipette pumps limited the number to keep colonies integrity and cell media was changed the next day.

For micropattern cell culture, custom ordered 500 uM micro-patterned 96 well glass bottom dishes (CYTOO) were coated with CellAdhere Laminin 521 (STEMCELL Technologies Cat No. 200-0117) at a concentration of 10 ug/mL for 2 hours at room temperature or optionally overnight at 4 C. The wells are serially diluted with ice-cold Ca2+ and Mg2+ free PBS. Cells already resuspended in growth medium with 5 uM ROCK inhibitor (STEMCELL Technologies Cat No. Y-27632) are then immediately plated upon removal of PBS.

For cell seeding onto the micropattern, cells growing on MatrigeI are lifted using Accutase (STEMCELL Technologies Cat No. 07920). Cells are treated for 5 minutes in Accutase before lifted and diluted into prewarmed mTeSR. Cells were centrifuged for 5 minutes at 600 g and 5x10^5 cells per well were resuspended in Rock-I containing media and added to the micropatterned plate. After 4 to 8 hours the medium was replaced with mTeSR plus without ROCK inhibitor and incubated overnight.

For the drug conditions, we used the stem cell signaling compound library available from Selleckchem (Cat No. L2100). Manufacturer suggestions were followed where a 100 uM intermediate concentration was prepared before further dilution to 5 uM. During the drug screen, all conditions were treated with 50 ng/mL BMP4 and drug for 48 hours before being stained and fixed.

2.2 Immunofluorescence

Wells are fixed with 4% paraformaldehyde and rinsed twice with PBS then permeabilized with -80 C ethanol for two minutes. Wells are then incubated with primary antibody at 4 C overnight on a rocker, washed three times in PBS for 5 minutes each, and incubated with secondary antibody and DAPI nuclear counterstain for 1 hour at room temperature on a rocker before being washed 3 times with PBS. The follow table describes the antibodies used

Target	Host	Manufacturer
BRA	Goat	R&D Systems AF2085
GATA3	Mouse	ThermoFisher MA1-028
SOX2	Rabbit	Cell Signaling 3579

Table 2.1: Primary Antibodies

Intensity	Target	Host	Manufacturer
Alexa Fluor 647	Goat	Donkey	ThermoFisher A21447
Alexa Fluor 568	Mouse	Donkey	ThermoFisher A10037
Alexa Fluor 488	Rabbit	Donkey	ThermoFisher A21206

Table 2.2: Secondary Antibodies

2.3 Imaging

Fixed cell imaging experiments were conducted using a Nikon W2 SoRa spinning-disk confocal microscope. Four channels were utilized corresponding to DAPI, Alexa488, Alexa555 and Alexa647 conjugated antibodies. To capture comprehensive staining information, each colony was imaged five times, with the Z stacks incremented at 10 um per stack.

Exposure settings were optimized such that each channel was set to 400 ms, with the summation and averaging factors for each set to 4x to enhance nuclear imaging. To ensure robustness and solid representation of each drug condition, we took an average of eight images per drug condition. Furthermore, positional data for each colony was recorded to ease in image to drug condition assignment.

2.4 Image Processing (Fiji)

We developed two FIJI macros for large-scale image processing. The first macro performs Z Projection on each image colony by channel. The second goes a step further and compiles all the channels into a single image. Each imaged colony was max projected in FIJI using the following settings: Image > Stacks > Z Project… > Max Intensity.

2.5 Nuclei Detection and Quantification

We developed MATLAB software to accurately extract nuclear intensity and positional intensity of imaged colonies. We first extract the DAPI image from the stack and convert it into a binary image. To ensure reliable nuclei detection, we

apply a background noise clean up algorithm using a morphological structuring element method. The morphological structuring element is a fundamental tool in nuclei detection algorithms because it can enhance detection and allow user specification of nuclei features. We perform a series of dilation and filling operations to obtain a grain free mask of the DAPI channel. We lastly use MATLAB's connected component detector (*bwconncomp*) to extract all positional and density information. Intensity data is extracted by masking the generated segmentation across the imaged channels and finding the intensity values for each respective germ layer stain. To determine colony edge information, we calculate the centroid information of all the nuclei before using a boundary detection algorithm to detect the edges of the colony along with each cell's nearest border cell. Utilizing the location data, we establish 50 bins that span from the center of the colony to the edge and fill each with the average intensity of cells for each channel. Lastly, we export both the raw intensity information and the binned intensity information for further analysis.

2.6 Uniform Manifold Approximation and Projection, K-means clustering algorithm

UMAP is a topological based dimensionality reduction technique commonly used in machine learning and data analysis. UMAP operates via gradient descent, starting by constructing a graph representation of the data where data points are represented as nodes and edges are established to connect nearby points. UMAP then optimizes the graph representation to find a lower dimensional embedding that

preserves the local and global graph structure of the data via cross-entropy minimization from a higher dimensional space to a lower one. The optimization process creates a fuzzy simplicial set that includes weights that indicate the strength of relationship between neighboring points. This representation is continually refined based on two main factors, maintenance of the local structure and dimensional reduction. We chose to use this dimensionality reduction method due to the versatility and speed of the UMAP algorithm. UMAP can effectively capture the complex spatial dynamics of gastrulation data and represent it into a lower-dimensional space, facilitating improvements in visualization and analysis. [59]

To exact trends and patterns generated from the UMAP algorithm, we employed K-Means clustering, a commonly used machine algorithm designed to partition distinct clusters in data. K-means was chosen due to its scalability and simplicity given our need to extract general trends in gastrulation patterning. In addition, the size and complexity of the screen results were already overwhelming, and our aim was to uncover broad perturbations in pattern formation phenotypes.

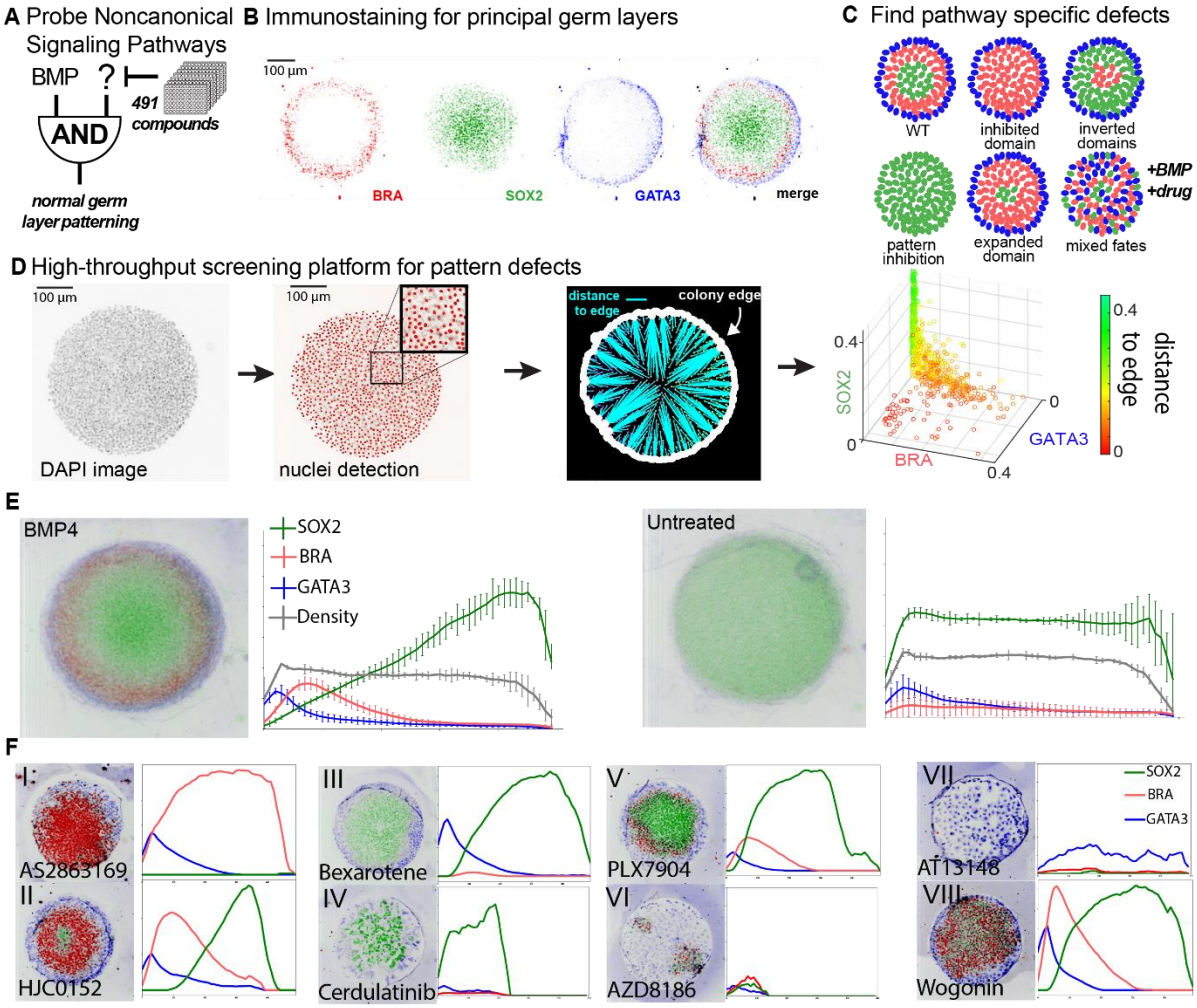

Figure 2.1: (A) Utilizing a stem cell differentiation oriented drug library, we co-treat micropatterned hESCs to explore the non-canonical signaling dynamics that may interact with early gastrulation initiation. (B) Immunostaining representation of each germ layer where BRA represents mesoderm, SOX2 represents ectoderm, and GATA3 represents endoderm. (C) Potential pathway specific defects. (D) Visual representation of the nuclei detection pipeline. (E) A 2D bin-wise representation of the controls shows expected pattern formation in BMP4 treated colonies. (F) The screen revealed a plethora of phenotype perturbations.

3. Results

3.1 Ternary plot representations elucidate drug conditions that general pure cell commitment populations

Prior to imaging the drug conditions, we generated plate maps for all drug panels and screened for missed hits, defined as drug conditions that result in cell death. We developed a simple sorting algorithm to extract drug conditions with live cells that sorted the list to match the meandering pattern the confocal microscope images the colonies. A grid representing the centers of each well was generated and the acquired location data of each colony was used to extrapolate each colony's position to a specific drug condition. All intensity and positional information were then output and stored into a pandas dataframe for further processing.

We utilized ternary plots to represent the 3D nature of our signaling information (BRA, SOX2, GATA, signaling intensities) in a 2D manner while retaining signaling intensity information. Previous studies have shown that micropatterned human gastruloids exhibit cell fates outside of the presumed three germ layer fates. [29] Furthermore, available literature has shown that cell fate commitment confusion, better referred to as cell plasticity, is observed during gastrulation, typically owing to factors such as incomplete or asynchronous signaling, tight-junction disruption, and other environmental influences. [63-65] It has also further been demonstrated that this plasticity is necessary for endoderm formation during gastrulation. In the context of germ layer patterning in our 2D organoids, we observed mixed staining in these cell types. Therefore, we defined two possible conditions a cell can have during

gastrulation: a germ layer committed cell fate and a germ layer plastic cell fate. Using the devised ternary plotting scheme, corners of the plot represent cell fate and contrast of hue representing signaling intensity. For instance, a cell that is committed to a specific cell fate will have both high hue contrast and will plot close to a corner, conversely a cell with high cell fate plasticity will plot in the center of ternary plot and have low hue. (Fig 2.2 A-B)

We developed an evaluation system to quantify the extent of cell commitment and differentiation within colonies. Each cell was represented as a vector such denoted as

$$\overrightarrow{c_j} = \left[I_{bra}, I_{sox}, I_{gata}\right]$$

where I_j represents the normalized intensity of each germ layer marker. The standard deviation of the vector values was used to measure a cell's commitment to a specific fate, where a higher standard deviation indicating a committed cell and a low deviation indicating a plastic cell. To evaluate the overall cell commitment in a given drug condition, we calculated the average standard deviation across all cells, σ_i,, and normalized it to rank drug conditions that result in high or low cell commitment. The following equations describe this ranking system.

$$\sigma_i = \frac{\sum_{j=1}^{N} \sqrt{\frac{\sum_{i=1}^{3}\left(I_i - \overline{c_j}\right)}{3}}}{N}$$

$$ranking = \frac{\sigma_i - min(\sigma)}{max(\sigma) - min(\sigma)}$$

In a theoretical colony where all cells are differentiated to a specific fate, the ranking value is 1.0 whereas a theoretical colony with cells that are confused has a ranking value of 0.0. (Fig 2.2 C) Once again, consistent with the results of Minn et. al.[61], BMP4 differentiated colonies have a low commitment ranking. Exploration of outliers with high cell commitment reveals ternary plots that are skewed towards one cell fate or have very definitive patterning and outliers with low cell commitment show scattered intensity plots. Exploring the targeted drug pathways after this analysis yields several interesting indicators. In drug conditions that result in highly committed colony cell fates, we see that a majority of the drugs are WNT signal activators, like CHIR, which result in the formation of pure mesodermal populations. (Fig 2.2 D)

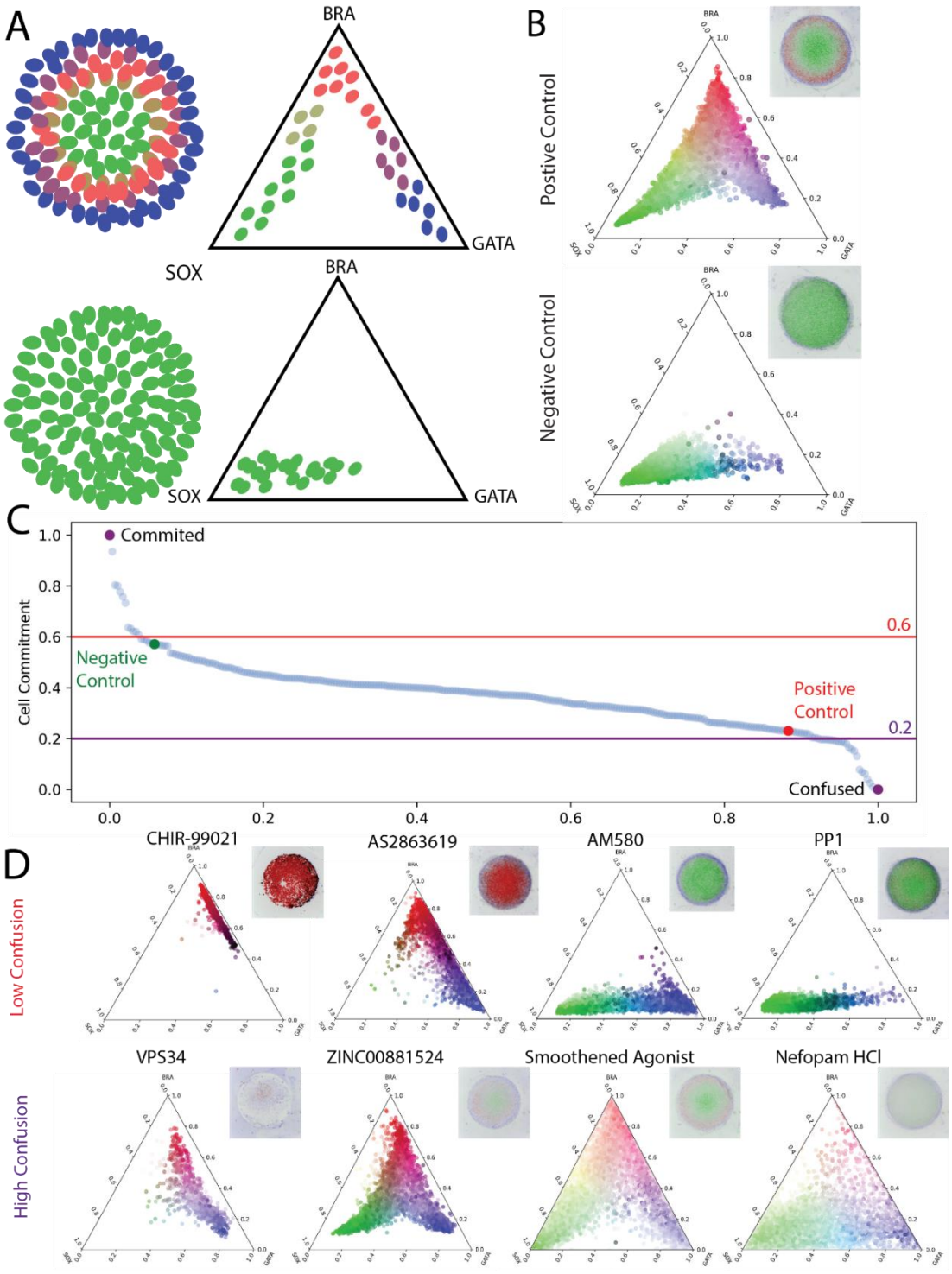

Figure 2.2 (A) Graphical representation of expected control conditions if plotted on ternary plots. (B) Ternary representation showed mixed staining in BMP4

differentiated colonies. (C) Ranking colony commitments based on our correlation-based metric. (D) Low confusion hits show drugs that generate pure populations. High confusion hits demonstrate highly plastic cell fates.

3.2 UMAP and K-means clustering reveals 5 distinct pattern forming phenotypes

The results of the screen demonstrate pattern forming phenotypes and non-pattern phenotypes. To categorize these phenotypes we used the stored dataframe from the aforementioned nuclei detection algorithm to extrapolate pairwise germ layer signaling correlation in imaged microcolonies. Uniform Manifold Approximation and Projection (UMAP) was used in conjunction with K-means clustering, to generate trends in colony pattern morphology. [59] The UMAP generated five distinct pattern phenotypes; standard pattern formation, expanded mesoderm, colony wide differentiation inhibition, mesoderm formation inhibition with endoderm formation, and complete pattern loss. (Fig 2.3) While drugs can have off target effects making the determination of a functional outcome of the exact target pathway difficult to conclude we theorized that if multiple chemically distinct drugs known to hit different nodes in the same pathway resulted in the same morphological outcome it would point to that specific pathway as the underlying cause of a particular morphological defect.

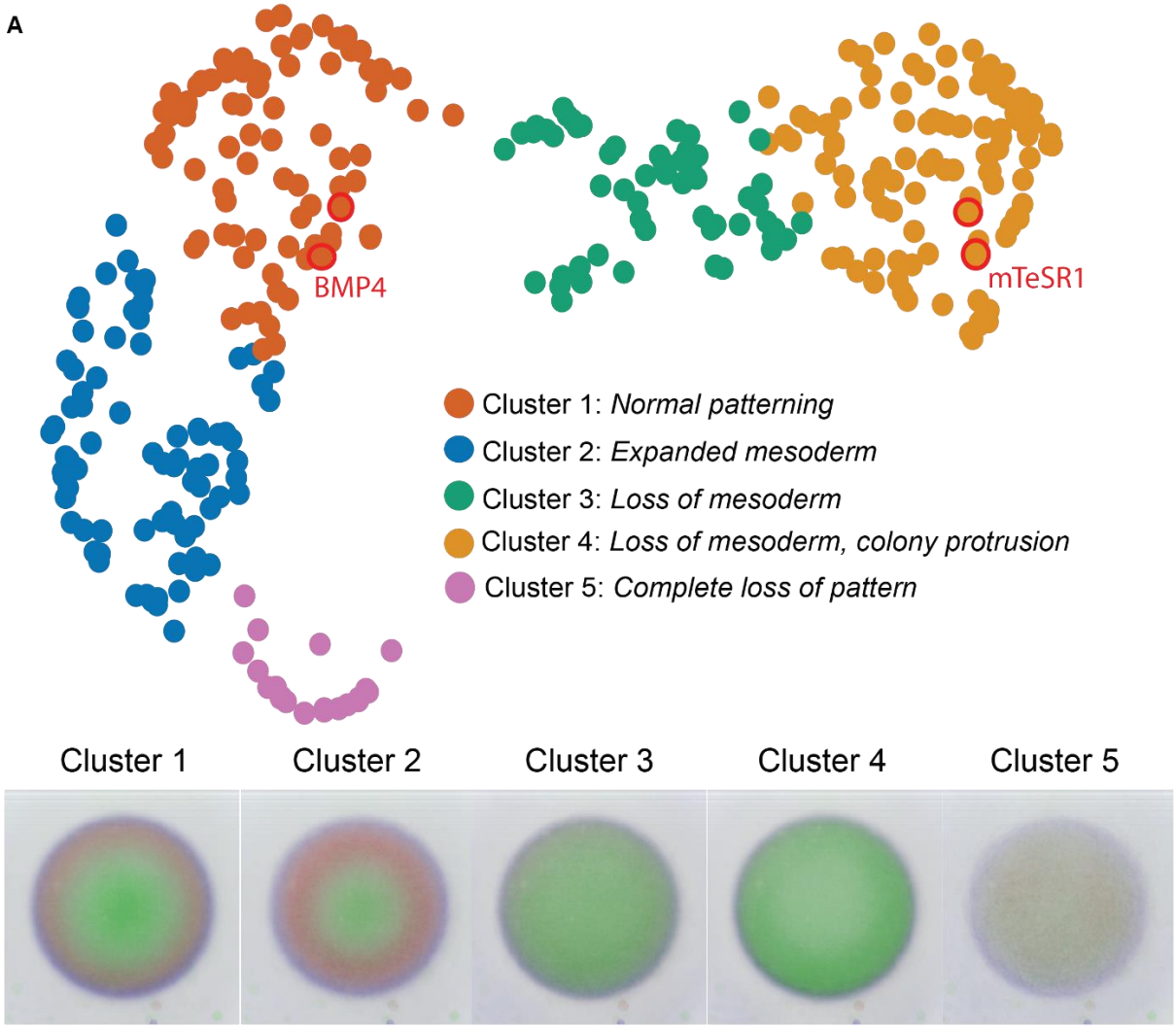

Figure 2.3 (A)UMAP compression of the imaged colonies. Cluster analysis reveals 5 distinct colony phenotypes. We see normal patterning, expanded mesoderm, loss of mesoderm, colony protrusion loss of mesoderm, and complete loss of pattern

3.3 Cell density impacts mesoderm and endoderm formation

As aforementioned, the key dynamics during gastrulation are cell density, cell migration, and differentiation. Despite the plethora of phenotypes that arise from the drug screen, there was another consistent phenotype where if a colony had high

cellular density, mesoderm pattern formation was constrained. We thus hypothesized that cellular density, which implies an increase in mechanical strain, acts as a secondary regulator for mesoderm differentiation. Thus, we devised a program to analyze this phenomena. We extracted all mesoderm forming conditions from the UMAP and performed a thresholded based analysis on a cell's commitment to BRA. We recalculated the intensity information for each channel for each cell as a percentage of expression for a specific germ layer. We then set a threshold of .4 to indicate whether a cell differentiated to a specific fate. We then calculated the total percentage of cells per colony that were committed to a specific germ layer. We then plotted these ratios with respect to cellular density on a per cluster basis. The results demonstrated that colonies with higher cellular density in mesoderm forming drug conditions had a significantly lower amount of mesoderm differentiated cells, whereas drug conditions that had a lower cellular density resulted in a significantly higher percentage of cells that were differentiated to mesoderm. (Fig 2.4)

Another cell density based dynamic that we observed was that in colonies that stained primarily for the endodermal marker, cell density was extremely sparse. Further analysis demonstrated that the common signaling pathway in cells with low density and uniform endodermal staining resulted in a series of Rho-kinase (ROCK) inhibitors. We theorize that this effect arises due the failure of the stem cells to compact. On 2d micropatterned cultures, BMP induced cultures are affected by the “edge effect”, where increased cell density in the center of the colony results in the relocalization of BMP receptors to lateral surfaces below tight junctions, limiting exposure to BMP4. [66] Furthermore, Rho-kinases are known to inhibit the

depolymerization of actin filaments.[67] It is well documented that 2d stem cell organoids will contract and expand after exposure to ROCK inhibitors are limited. [66] Thus we theorize that the ROCK inhibitor prevents contract and cytoskeletal recovery, allowing for more BMP receptors to be exposed. The continuous signaling from BMP4 eventually results in cells that exhibit uniform expression of endodermal markers.

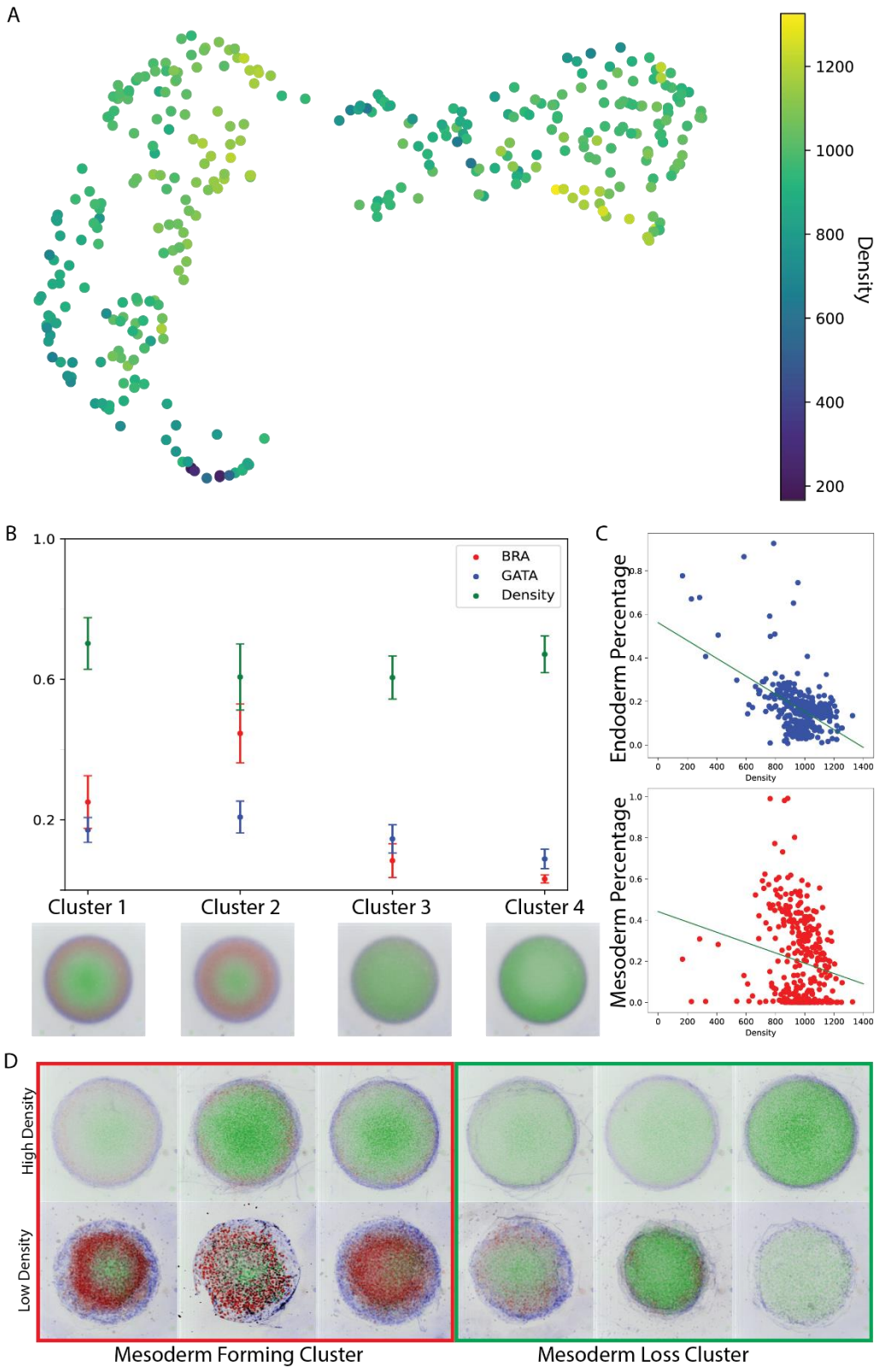
A
1200
1000
800
600
400
200
Density
B
1.0
0.6
0.2
BRA
GATA
Density
Cluster 1
Cluster 2
Cluster 3
Cluster 4
C
Endoderm Percentage
Density
Mesoderm Percentage
Density
D
High Density
Low Density
Mesoderm Forming Cluster
Mesoderm Loss Cluster

Figure 2.4 (A) Density based coloring of the UMAP shows obvious differences, namely that colonies with high density in mesodermal conditions also lead to lower mesoderm percentages. (B) Cluster based analysis shows that mesoderm is the germ layer that varies the highest throughout each cluster. (C) Endoderm and mesoderm percentages plotted with respect to colony density shows that density is correlated with the ratio of mesoderm and endoderm formation. (D) Representative images of colonies in mesoderm and non-mesoderm forming colonies based on density.

3.4 Pattern analysis reveals known and novel teratogens

Current methods for human treatment screens are mainly conducted in animal models, namely mice, or on stem cells in culture. Recently, there has been growing interest in the usage of embryoid models like micropatterns and 3d models in suspension. However, none of them have been thoroughly used to screen for teratogens. Given the highly reproducible and relatively cheap nature of the micropattern organoids, this provides an ideal platform for teratogen screening. [68]

In the screen, several drug conditions are known teratogens and they result in aberrant pattern changes which include inversion of germ layer fate patterning, breakage of radial symmetry, a phenomenon yet to be demonstrated in recent literature, and formation of 3d structures. Given this, we theorize that pregnancy failures may not only be due to genetic, morphological, or pattern forming defects during early embryogenesis, but may be linked to diet, even in seemingly healthy pregnancies. (Fig 2.5)

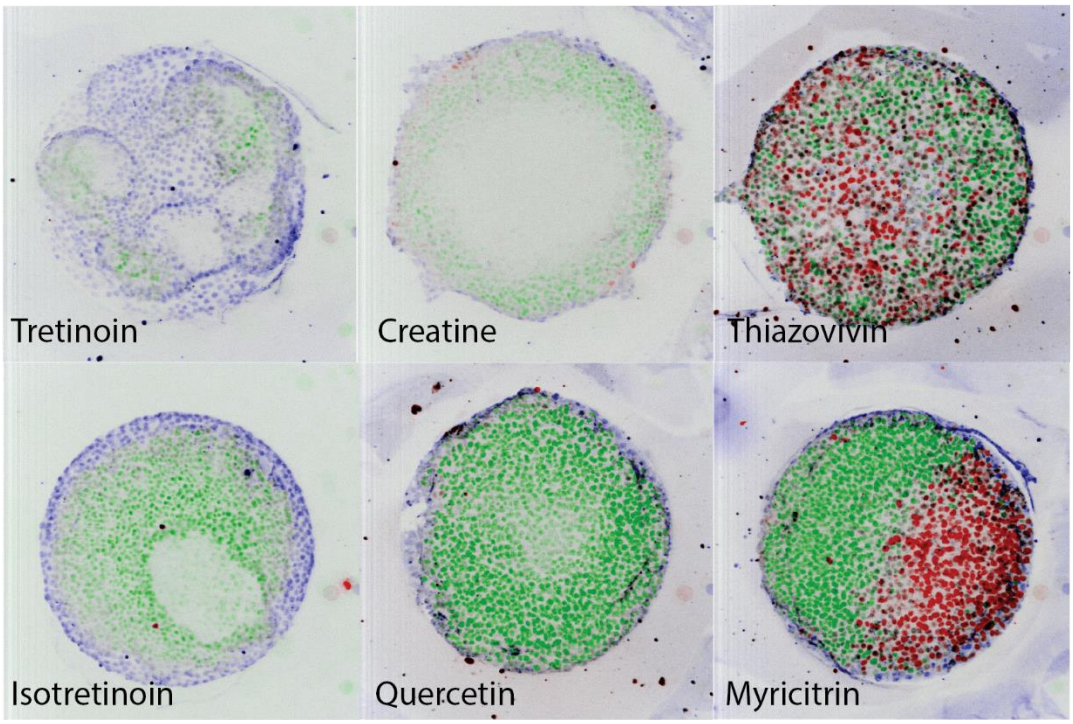

Figure 2.5 Tretinoin, commonly known as Accutane is well documented teratogen. We found compounds, Creatine, Thiazovivin, Isotretinoin, Quercetin, and Myricitrin induced drastic changes in gastrulation pattern formation. Creatine is a commonly taken supplement, and Quercetin and myricitrin are commonly available flavonoids.

3.5 Hypergeometric clustering reveals enrichment of immunological signaling process in mesoderm forming patterns

We employed hypergeometric distribution calculations to analyze the distribution of signaling pathways on the clustered UMAP. Hypergeometric distribution allows us to determine the discrete probability distribution of drugs occurring in specific clusters. Our analysis revealed two distinct distribution regimes: cluster-based enrichment of a specific signaling pathway target and nonspecific distributions of cluster phenotype. Notably, we observed an overrepresentation of drugs targeting immunological and inflammatory responses in clusters exhibiting regular and expanded mesoderm formation. (Fig 2.6A-B) Conversely, we found an enrichment of protease-targeting genes in phenotypes where no pattern formation occurred. To ascertain the relevance of these signaling pathways in gastrulation, we examined the human gastrula single-cell sequencing database [74] and discovered high expression of CXCR signaling in emergent mesoderm.

We imaged the colonies under different drug conditions for a period of 48 hours and made an intriguing observation. Colonies treated with CXCR inhibitors rapidly formed structures resembling a neural groove.(Fig 2.6C) Further examination revealed that cells in the endodermal region migrated along the boundary of the micropatterned colony in a circular manner while simultaneously proliferating. Cells at the center of the pattern exhibited rapid proliferation with less migration. These behaviors appeared to induce the formation of a neural groove-like structure. Existing literature supports the involvement of CXCR in cell migration during

gastrulation. Specifically, CXCR2 is expressed in the primitive streak, and studies on CXCR2-deficient mouse embryos have demonstrated defects in patterning within the primitive streak. Perturbations in CXCR4 have also been shown to impair cell migration and patterning [69-70]. However, limited research has explored these dynamics in the context of human gastrulation. Therefore, we designed a straightforward cell tracking experiment to quantify changes in cell migration during gastrulation. This involved imaging colonies with inhibited CXCR for 48 hours at 20-minute intervals and performing manual cell tracking, focusing on changes and consistencies in the edge morphologies of cell clusters. The results revealed that CXCR-inhibited colonies exhibited significant increases in cell migration, with cells not only proliferating at a faster rate but also displaying lateral migration. Over time, the cells began to aggregate, forming structures resembling neural tubes. CXCR-inhibited cells demonstrated nearly a two-fold increase in cell migration. (Fig 2.6D)

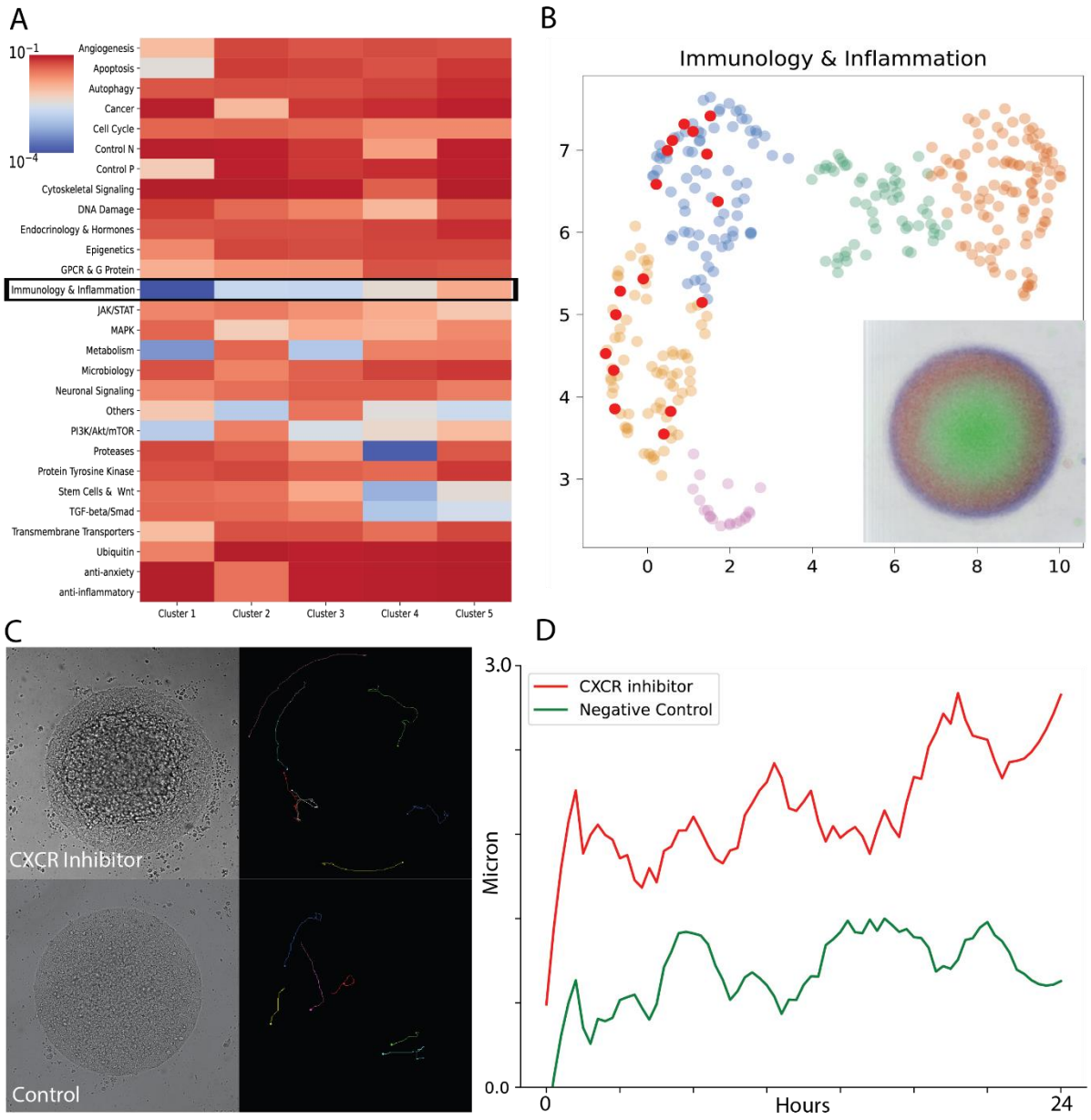

Figure 2.6: (A) Hypergeometric distribution calculations show that immunological signaling drugs are highly enriched in mesoderm forming colonies. (B) Highlighting immunological targeting drugs on the UMAP as well as the average colony morphology shows that mesoderm formation is increase. (C) Comparison between a CXCR inhibited colony and a non-CXCR inhibited colony shows drastic changes in

morphology. Cell tracks also show increased migration. (D) A quantification of cell tracking shows a near 2-fold increase in cell migration in CXCR inhibited conditions.

4. Discussion

In this study, we successfully utilized a high-throughput drug screen as an alternative method to induce pathway-specific perturbations without relying on CRISPR technology. Our findings revealed that gastrulation is not only influenced by non-canonical signaling pathways but is also highly susceptible to perturbations, particularly in the formation of the mesoderm and endoderm. We observed that a diverse range of signaling cascades and density-dependent dynamics must occur correctly for proper pattern formation during gastrulation. Furthermore, our study demonstrated the effectiveness of 2D micropatterns in detecting teratogens. We identified that flavonoids, when administered incorrectly, can significantly disrupt proper pattern formation. This highlights the potential impact of diet on pregnancy outcomes and suggests that many pregnancy failures could be attributed to the consumption of certain chemicals.

The investigation of immune signaling dynamics in pattern formation provided new insights into the facilitation of gastrulation. The interaction between the maternal and fetal immune systems is still not well understood, and our results suggest that the maternal immune system may play a role in embryo development. This finding has implications for understanding how the maternal immune system influences the embryo and how immune strain on the mother can affect development.
Additionally, our drug screen identified conditions that led to patterns with high levels of plasticity. However, since immunofluorescent signaling is not an ideal indicator for assessing plasticity, future studies should focus on differentiating 2D gastruloids

under these drug conditions and perform transcriptomic analysis using single-cell RNA sequencing.

Lastly, the density-dependent formation of the mesodermal and endodermal germ layers highlights the significance of mechanical forces during early embryonic development. Previous studies have shown that cellular tension and compression play crucial roles in cell fate differentiation or maintenance. To further understand the role of cell mechanics in human germ layer fate formation and patterning, we suggest employing traction force microscopy, which can provide insights into the forces exerted by cells during gastrulation.

From the results of this study, we propose using micropatterned human embryonic stem cells as a platform for exploring teratogens. The simplicity, repeatability, and cost-effectiveness of using micropatterns for teratogen screening can uncover previously overlooked chemicals that may contribute to the termination of otherwise healthy pregnancies. Additionally, conducting teratogenic screens using 3D embryonic models should be considered to validate the findings from 2D micropatterns.

In conclusion, our study sheds new light on the signaling dynamics and sensitivity of human gastrulation. It suggests that diet and exposure to certain chemicals can have profound effects on pattern formation and pregnancy outcomes. Furthermore, the role of immune signaling and cell mechanistics during gastrulation warrants further investigation. By leveraging the insights gained from this research, we can pave the way for advancements in understanding and potentially preventing developmental abnormalities and pregnancy failures.

References

1. A. Kusserow, *et al.*, Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* **433**, 156–160 (2005).
2. R. Nusse, H. Clevers, Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).
3. H. Clevers, Wnt/β-Catenin Signaling in Development and Disease. *Cell* **127**, 469–480 (2006).
4. V. S. W. Li, *et al.*, Wnt Signaling through Inhibition of β-Catenin Degradation in an Intact Axin1 Complex. *Cell* **149**, 1245–1256 (2012).
5. A. Cliffe, F. Hamada, M. Bienz, A Role of Dishevelled in Relocating Axin to the Plasma Membrane during Wingless Signaling. *Curr. Biol.* **13**, 960–966 (2003).
6. MacDonald BT, Tamai K, He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell. 2009 Jul;17(1):9-26. doi: 10.1016/j.devcel.2009.06.016. PMID: 19619488; PMCID: PMC2861485.
7. L. J. Bugaj, A. T. Choksi, C. K. Mesuda, R. S. Kane, D. V. Schaffer, Optogenetic protein clustering and signaling activation in mammalian cells. *Nat. Methods* **10**, 249–252 (2013).
8. D. M. Roberts, *et al.*, Defining components of the ßcatenin destruction complex and exploring its regulation and mechanisms of action during development. *PLoS One* **7** (2012).

9. K. N. Schaefer, *et al.*, Supramolecular assembly of the beta-catenin destruction complex and the effect of Wnt signaling on its localization, molecular size, and activity in vivo. *PLoS Genet.* **14**, 1–44 (2018).
10. W. Kan, *et al.*, Limited dishevelled/Axin oligomerization determines efficiency of Wnt/β-catenin signal transduction. *Elife* **9** (2020).
11. K. N. Schaefer, M. Peifer, Wnt/Beta-Catenin Signaling Regulation and a Role for Biomolecular Condensates. *Dev. Cell* **48**, 429–444 (2019).
12. D. M. Roberts, *et al.*, Deconstructing the βcatenin destruction complex: mechanistic roles for the tumor suppressor APC in regulating Wnt signaling. *Mol. Biol. Cell* **22**, 1845–1863 (2011).
13. K. Yamanishi, *et al.*, A direct heterotypic interaction between theDIX domains of Dishevelled and Axin mediates signaling to β-catenin. *Science Signaling* **12**, 1–8 (2019).
14. M. I. Pronobis, N. M. Rusan, M. Peifer, A novel GSK3-regulated APC:Axin interaction regulates Wnt signaling by driving a catalytic cycle of efficient βcatenin destruction. *Elife* **4**, 1–31 (2015).
15. S. F. Shimobayashi, P. Ronceray, D. W. Sanders, M. P. Haataja, C. P. Brangwynne, Nucleation landscape of biomolecular condensates. *Nature*, 1–4 (2021).
16. S. C. Weber, C. P. Brangwynne, Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.* **25**, 641–646 (2015).

17. S. M. Vora, J. S. Fassler, B. T. Phillips, Centrosomes are required for proper β-catenin processing and Wnt response. *Mol. Biol. Cell* **31**, 1951–1961 (2020).
18. K. Ruan, *et al.*, PLK1 interacts and phosphorylates Axin that is essential for proper centrosome formation. *PLoS One* **7**, e49184 (2012).
19. C. Lui, M. T. S. Mok, B. R. Henderson, Characterization of Adenomatous Polyposis Coli Protein Dynamics and Localization at the Centrosome. *Cancers* **8** (2016).
20. S. Munemitsu, *et al.*, The APC gene product associates with microtubules in vivo and promotes their assembly in vitro. *Cancer Res.* **54**, 3676–3681 (1994).
21. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
22. T. Schwarz-Romond, C. Metcalfe, M. Bienz, Dynamic recruitment of axin by Dishevelled protein assemblies. *J. Cell Sci.* **120**, 2402–2412 (2007).
23. Z. C. Elmore, R. X. Guillen, K. L. Gould, The kinase domain of CK1 enzymes contains the localization cue essential for compartmentalized signaling at the spindle pole. *Mol. Biol. Cell* **29**, 1664–1674 (2018).
24. B. C. Mbom, W. J. Nelson, A. Barth, B-catenin at the centrosome: Discrete pools of β-catenin communicate during mitosis and may co-ordinate centrosome functions and cell cycle progression. *Bioessays* **35**, 804–809 (2013).

25. D. T. McSwiggen, M. Mir, X. Darzacq, R. Tjian, Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
26. C. Liu, *et al.*, Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* **108**, 837–847 (2002).
27. C. J. Fiol, A. Wang, R. W. Roeske, P. J. Roach, Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates. *J. Biol. Chem.* **265**, 6061–6065 (1990).
28. M. I. Pronobis, N. Deuitch, V. Posham, Y. Mimori-Kiyosue, M. Peifer, Reconstituting regulation of the canonical Wnt pathway by engineering a minimal β-catenin destruction machine. *Mol. Biol. Cell* **28**, 41–53 (2016).
29. A. A. M. André, E. Spruijt, Liquid-Liquid Phase Separation in Crowded Environments. *Int. J. Mol. Sci.* **21** (2020).
30. T.-M. Li, *et al.*, Multivalent tumor suppressor adenomatous polyposis coli promotes Axin biomolecular condensate formation and efficient β-catenin degradation. *Sci. Rep.* **10**, 17425 (2020).
31. K. Fumoto, M. Kadono, N. Izumi, A. Kikuchi, Axin localizes to the centrosome and is involved in microtubule nucleation. *EMBO Rep.* **10**, 606–613 (2009).
32. E. Lee, A. Salic, R. Krü Ger, R. Heinrich, M. W. Kirschner, The roles of APC and axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* **1**, e10 (2003).
33. https://fenicsproject.org/olddocs/dolfin/1.3.0/python/demo/documented/cahn-hilliard/python/documentation.html#.

34. Zhang HT, Hiiragi T. Symmetry Breaking in the Mammalian Embryo. Annu Rev Cell Dev Biol. 2018 Oct 06;34:405-426
35. Câmara DR, Kastelic JP, Thundathil JC. Role of the Na+/K+-ATPase ion pump in male reproduction and embryo development. Reprod Fertil Dev. 2017 Aug;29(8):1457-1467.
36. Larsen EC, et al. New insights into mechanisms behind miscarriage. BMC Med. 2013;11:154.
37. Zinaman MJ, et al. Estimates of human fertility and pregnancy loss. Fertil Steril. 1996;65(3):503–9.
38. Niu Y, Sun N, Li C, Lei Y, Huang Z, Wu J, Si C, Dai X, Liu C, Wei J, Liu L, Feng S, Kang Y, Si W, Wang H, Zhang E, Zhao L, Li Z, Luo X, Cui G, Peng G, Izpisúa Belmonte JC, Ji W, Tan T. Dissecting primate early post-implantation development using long-term in vitro embryo culture. Science. 2019 Nov 15;366(6467)
39. Mikawa T, Poh AM, Kelly KA, Ishii Y, Reese DE. Induction and patterning of the primitive streak, an organizing center of gastrulation in the amniote. Dev Dyn. 2004 Mar;229(3):422-32.
40. Chhabra S, Liu L, Goh R, Kong X, Warmflash A. Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids. PLoS Biol. 2019 Oct 15;17(10):e3000498. doi: 10.1371/journal.pbio.3000498. PMID: 31613879; PMCID: PMC6814242.

41. Simunovic M, Metzger JJ, Etoc F, Yoney A, Ruzo A, Martyn I, Croft G, You DS, Brivanlou AH, Siggia ED. A 3D model of a human epiblast reveals BMP4-driven symmetry breaking. Nat Cell Biol. 2019 Jul;21(7):900-910.
42. Afouda, B. A., Ciau-Uitz, A. & Patient, R. 2005. GATA4, 5 and 6 mediate TGFβ maintenance of endodermal gene expression in Xenopus embryos. Development 132, 763– 774.
43. Fontaine, J. & Le Douarin, N. M. 1977. Analysis of endoderm formation in the avian blastoderm by the use of quail-chick chimaeras. The problem of the neurectodermal origin of the cells of the APUD series. J. Embryol. Exp. Morph. 41, 209– 222.
44. Bellairs, R. 1957. Studies on the development of the foregut in the chick embryo. 4. Mesodermal induction and mitosis. J. Embryol. Exp. Morph. 5, 340– 350.
45. Skromne I, Stern CD. Interactions between Wnt and Vg1 signalling pathways initiate primitive streak formation in the chick embryo. Development. 2001 Aug;128(15):2915-27.
46. Faure S, de Santa Barbara P, Roberts DJ, Whitman M. Endogenous patterns of BMP signaling during early chick development. Dev Biol. 2002 Apr 01;244(1):44-65.
47. Durston AJ. What are the roles of retinoids, other morphogens, and Hox genes in setting up the vertebrate body axis? Genesis. 2019 Jul;57(7-8):e23296.

48. Tam PP, Loebel DA. Gene function in mouse embryogenesis: get set for gastrulation. Nat Rev Genet. 2007 May;8(5):368-81.
49. Wei S, Wang Q. Molecular regulation of Nodal signaling during mesendoderm formation. Acta Biochim Biophys Sin (Shanghai). 2018 Jan 01;50(1):74-81.
50. Steventon B, Araya C, Linker C, Kuriyama S, Mayor R. Differential requirements of BMP and Wnt signalling during gastrulation and neurulation define two steps in neural crest induction. Development. 2009 Mar;136(5):771-9. doi: 10.1242/dev.029017. Epub 2009 Jan 28. PMID: 19176585; PMCID: PMC2685944.
51. Chen T, You Y, Jiang H, Wang ZZ. Epithelial-mesenchymal transition (EMT): A biological process in the development, stem cell differentiation, and tumorigenesis. J Cell Physiol. 2017 Dec;232(12):3261-3272. doi: 10.1002/jcp.25797. Epub 2017 Apr 10. PMID: 28079253; PMCID: PMC5507753.
52. Warnock M REPORT OF THE COMMITTEE OF INQUIRY INTO HUMAN FERTILISATION AND EMBRYOLOGY. London: Her Majesty's Stationery Office, 1984.
53. James A. Thomson et al.,Embryonic Stem Cell Lines Derived from Human Blastocysts.Science282,1145-1147(1998).DOI:10.1126/science.282.5391.1145
54. Vazin T, Freed WJ. Human embryonic stem cells: derivation, culture, and differentiation: a review. Restor Neurol Neurosci. 2010;28(4):589-603. doi: 10.3233/RNN-2010-0543. PMID: 20714081; PMCID: PMC2973558.

A. Warmflash, B. Sorre, F. Etoc, E. D. Siggia, and A. H. Brivanlou. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. 2014. Nature Methods. 11:847-854.

55. Heemskerk, I., & Warmflash, A. (2016). Pluripotent stem cells as a model for embryonic patterning: From signaling dynamics to spatial organization in a dish. Developmental Dynamics.
56. L. Najman and M. Schmitt. Watershed of a continuous function. In Signal Processing (Special issue on Mathematical Morphology.), Vol. 38 (1994), pages 99–112
57. Edward R. Dougherty, An Introduction to Morphological Image Processing, ISBN 0-8194-0845-X (1992)
58. Becht, E., McInnes, L., Healy, J. et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol 37, 38–44 (2019). https://doi.org/10.1038/nbt.4314
59. Bhattacharyya, A. (1943). "On a measure of divergence between two statistical populations defined by their probability distributions". Bulletin of the Calcutta Mathematical Society. 35: 99–109. MR 0010358
60. Xu RH, Chen X, Li DS, Li R, Addicks GC, Glennon C, Zwaka TP, Thomson JA. BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat Biotechnol. 2002 Dec;20(12):1261-4. doi: 10.1038/nbt761. Epub 2002 Nov 11. PMID: 12426580.
61. Kyaw Thu Minn, Yuheng C Fu, Shenghua He, Sabine Dietmann, Steven C George, Mark A Anastasio, Samantha A Morris, Lilianna Solnica-Krezel

(2020) High-resolution transcriptional and morphogenetic profiling of cells from micropatterned human ESC gastruloid cultures eLife 9:e59445

62. Arnold SJ, Robertson EJ. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol. 2009 Feb;10(2):91-103. doi: 10.1038/nrm2618. Epub 2009 Jan 8. PMID: 19129791.
63. Cockburn K, Rossant J. Making the blastocyst: lessons from the mouse. J Clin Invest. 2010 Apr;120(4):995-1003. doi: 10.1172/JCI41229. Epub 2010 Apr 1. PMID: 20364097; PMCID: PMC2846056.
64. ten Berge D, Koole W, Fuerer C, Fish M, Eroglu E, Nusse R. Wnt signaling mediates self-organization and axis formation in embryoid bodies. Cell Stem Cell. 2008 Nov 6;3(5):508-18. doi: 10.1016/j.stem.2008.09.013. PMID: 18983966; PMCID: PMC2683270.
65. Etoc F, Metzger J, Ruzo A, Kirst C, Yoney A, Ozair MZ, Brivanlou AH, Siggia ED. A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. Dev Cell. 2016 Nov 7;39(3):302-315. doi: 10.1016/j.devcel.2016.09.016. Epub 2016 Oct 13. PMID: 27746044; PMCID: PMC5113147.
66. Hartmann S, Ridley AJ and Lutz S (2015) The Function of Rho-Associated Kinases ROCK1 and ROCK2 in the Pathogenesis of Cardiovascular Disease. Front. Pharmacol. 6:276. doi: 10.3389/fphar.2015.00276
67. Hartmann S, Ridley AJ and Lutz S (2015) The Function of Rho-Associated Kinases ROCK1 and ROCK2 in the Pathogenesis of Cardiovascular Disease. Front. Pharmacol. 6:276. doi: 10.3389/fphar.2015.00276

68. Zeigler BM, Sugrue SP. Role of CXCR2 in the Formation of the Neural Tube in the Early Chick Embryo. Dev Dyn. 2009;238(5):1244-1253. doi:10.1002/dvdy.21951
69. Chan XY, Blackstone BN, Wu YI, et al. A Synthetic Chemokine Receptor Ligand Enhances the Efficiency of MSC-based Therapies. Sci Rep. 2018;8(1):9402. Published 2018 Jun 19. doi:10.1038/s41598-018-27689-w
70. Yusuf F, Rehimi R, Dai F, Brand-Saberi B. Expression of chemokine receptor CXCR4 during chick embryo development. Anat Embryol (Berl). 2005 Aug;210(1):35-41. doi: 10.1007/s00429-005-0013-9. Epub 2005 Jul 27. PMID: 16047188.
71. Nair S, Schilling TF. Chemokine signaling controls endodermal migration during zebrafish gastrulation. Science. 2008 Oct 3;322(5898):89-92. doi: 10.1126/science.1160038. Epub 2008 Aug 21. PMID: 18719251; PMCID: PMC2770598.
72. Chu, LF., Leng, N., Zhang, J. et al. Single-cell RNA-seq reveals novel regulators of human embryonic stem cell differentiation to definitive endoderm. Genome Biol 17, 173 (2016). https://doi.org/10.1186/s13059-016-1033-x
73. Tyser, R.C.V., Mahammadov, E., Nakanoh, S. et al. Single-cell transcriptomic characterization of a gastrulating human embryo. Nature 600, 285–289 (2021). https://doi.org/10.1038/s41586-021-04158-y

74. S. Lach et al., Nucleation of the destruction complex on the centrosome accelerates degradation of β-catenin and regulates Wnt signal transmission.Proc. Natl. Acad. Sci. U.S.A.119, e2204688119 (2022)