

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Examining Self-directed Multicellular Organization and Morphogenesis of Human Induced Pluripotent Stem Cells

Permalink

<https://escholarship.org/uc/item/3f16d7r1>

Author

Libby, Ashley Rachel Greeley

Publication Date

2019

Peer reviewed|Thesis/dissertation

Examining Self-directed Multicellular Organization and Morphogenesis of Human Induced Pluripotent Stem Cells

by
Ashley RG Libby

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Developmental and Stem Cell Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

Susan Fisher

Susan Fisher

3D7FD5EAA89E411...

Chair

DocuSigned by:

Todd C McDevitt

Todd C McDevitt

DocuSigned by:

Bruce Conklin

Bruce Conklin

DocuSigned by:

Zev Gartner

Zev Gartner

6E99B613A5754DD...

Committee Members

Copyright 2019
By
Ashley R.G. Libby

Dedicated to the Carl Family.

Ohana means family. Family means no one gets left behind.

ACKNOWLEDGEMENTS

I have had the unique opportunity to pursue my PhD in a truly supportive environment without which I would not have made it through the rollercoaster of science, emotion, and at times mind numbing work that it took to complete my degree. As such, I have a long list of mentors, fellows, trainees, friends and family to thank for my success.

First, I would like to thank the UCSF Developmental Stem Cell Biology Program as a whole. I moved to California with no contacts in the area and the program and the people within it immediately became my first source of refuge, friendships, and learning. Specifically, my class (Bonnie, Akela, Hanna, Daniel, and Buddy) provided unlimited support early on, staying late at night in the RMB boardroom completing our genetics homework, eating at Kiki's Sushi every week, and collectively trying to figure out what T_{regs} were at our first BMS retreat. Beyond my class the DSCB community is one of both support and adventure, from tubing trips, to house parties, to camping in Yosemite, I have loved it all, and I will miss the time I got to spend with my fellow students.

I would like to thank my co-advisors Dr. Bruce Conklin and Dr. Todd McDevitt, for their unmatched support of my work over the past 5 years. At the start of my PhD, Bruce accepted me into his lab with the recommendation that I check out Todd's lab because he thought I would love the collaboration. I cannot thank Bruce enough for not only recognizing my interests, but also providing the push to seek out scientific advisors in areas outside of my expertise. In tangent, I must also thank Todd for never questioning my sudden appearance in his lab and welcoming me with open arms despite at first feeling out of place as a biologist in a lab full of bioengineers. Because of Todd's unwavering trust, I had the opportunity to start the early development working group in the lab, recruit graduate students to join the "dark side" of developmental biology, and pursue a project that veered away from the lab's expertise. Bruce and Todd made a perfect co-mentor match, where Bruce loved to think of big grandiose picture agendas, Todd was willing to painstakingly go over the nitty gritty details of my experiments, fellowship, and presentations.

Further, Todd set up collaborations (Demarcus Briers and Calin Belta) that I never would have dreamed of and forced me to step outside of my comfort zone. Over the past five years, under the mentorship of Bruce and Todd, I have learned many key lessons on how to present my science, write clearly, manage my time, and never forget the big picture goals of the science we do in lab. I am truly thankful to have been trained by Bruce and Todd.

I would like to thank my committee members Dr. Susan Fisher and Dr. Zev Gartner for their intellectual support of my thesis project and future career. Susan and Zev have provided key advice on how to focus my experiments and develop my project. Furthermore, they have been excellent advisors in my future career path, providing seemingly endless letters of recommendation for my seemingly endless fellowships.

I have found throughout my PhD that environment is one of the most important things to the success of a graduate career, and I have had a fantastic one in the Conklin and McDevitt labs and at the Gladstone as a whole. From the beginning of my PhD, I had an insane number of senior graduate students and outside mentors to shape my writing and line of scientific questioning. Specifically, I must thank McDevitt lab senior members (Jenna Wilson, Josh Zimmermann, Tracy Hookway, and Jessie Butts) for covering all my proposals in red ink, painfully teaching me how to write with “strong” language. I must thank fellow graduate students Bonnie Cole and Ariel Kauss for being in my qual's boat and making sure it didn't sink. I want to thank Dr. Polin So in the Conklin lab for answering every question, reading every document, and providing endless amounts of support and advice. Polin has played an unprecedented roll in my development as a developmental biologist and has always been a firm point of reference for the type of scientist I aim to become.

I would like to thank my collaborators Drs. Calin Belta, Iman Haghghi, and Demarcus Briers. Specifically, Demarcus for spending hours every week on skype calls with me as we perfected the ins and outs of our computational model. It is because of this collaboration that I understand the basis of machine learning and might try my own hand at computational modeling

in the future. Despite being across the country from one another, only ever meeting in person three times, and Demarcus and Iman graduating, we made it work. Thank you for teaching me how to be a good collaborator. It is a vital skill that I will take with me and it was an honor to learn from Demarcus.

As my PhD continued the McDevitt lab (McMurder) was a continued resource of support and fun far beyond just academic science. I know intricate details of baking pie crust (David Joy), crafting the perfectly toasted crème brûlée (Mike Kang), and endless recipes for a variety of Indian dishes (Vaishaali Natarajan). While my stomach was full, Oriane Matthys ensured that I always worked it off between spin class or swimming, and Ana Silva made sure that I always had a song to hum even if I couldn't dance like her. Further, everyone's enthusiasm for Halloween (my favorite holiday) has painted my PhD with so many happy moments, even if I will never forgive Ivana Vasic for skipping our Halloween party.

I would like to thank all of the interns (Jocelyn Cervantes), research associates (Nik Mendoza-Camacho and Martina (Marty) Krakora) and graduate students (Ivana Vasic, Jasmine King, Emily Bulger, Nick Elder) that for whatever reason Todd thought it was a good idea to let me train. Jocelyn Cervantes, my first intern, whose time in lab was marked by sticky notes strewn across my desk, taught me how to teach science and manage two people's time. Nik with his magic cloning hands and Marty with her crazy coding skills taught me that a little bit of sass can go a long way. All the graduate students that I have helped train in lab, have taught me the necessity of peripheral vision as they hovered by my desk with their latest problem. They constantly kept me on my toes with infinite questions ranging from "where is X reagent?" to "Do we really believe the French Flag Model?" to "Ashley, why is my differentiation breaking?". I am truly sorry that I will not get to see the continuation of their development as scientists, but it has been an honor training them to this point. Overall, mentorship and helping others become excited about science has been one of the highlights of my PhD and solidified my drive to become a professor in the future.

I would like to specifically thank the early development team in the McDevitt lab (David Joy and Ivana Vasic) who took the leap of faith and joined me in a new section of the lab. Technically, David was the first student I ever trained, which for anyone that knows David now, it seems laughable that he was ever at intro level in anything. It has been an incredible experience working with David, who in return for cell culture knowledge taught me how to code, grown my appreciation for dinosaur t-shirts, and tackled every side project Todd threw at us, all while complaining about how we need more robots in lab. He is the most talented person I know, and I can't express the gratitude I have for his constant support in lab. I would not have a PhD without him. I also want to thank Ivana, who completed our early development family. Anyone that meets Ivana will immediately be taken in by her bubbly presence. She forced David and I to leave behind our introverted nature as every day we were dragged into the conference room to once again argue over the details of embryonic development. Her frustration with the limits of our knowledge and drive to overcome them has been one of the highlights of working with her. All in all, there are no words that truly capture how important working in this team has been to me over the course of my PhD, and I am truly grateful to both David and Ivana.

Finally, I have to thank those who have kept me going outside of lab. I have relied on weekly DND sessions (Tawny, Hanna, Nate, Bonnie, Steph, Rebecca, Dan, and Natanya) to both keep my imagination alive and give me something to look forward to at the end of the day. I want to thank Black Hound Equestrian for giving me a place to sneak off to with fuzzy animals to pat. Particularly, my trainer Kate Jutagir and fellow riders Katie and Mareva, with whom I spent every Saturday morning riding for the past 5 years. I want to thank my family, who have supported me from across the country, providing endless emotional support even if they had no idea what I was doing. I want to thank the Kirby family for taking me in every Thanksgiving and providing an adoptive family away from home. I need to thank my high school friends Lynn Pincus, Elizabeth Aspinwall, and Marissa Edwards for always cheering me on and cheering me up. And finally, and most importantly, I want to thank the Carl Family: Bonnie Cole and Elijah Martin. Bonnie and Elijah

have not only been my roommates but have also truly become my family. They have welcomed me home every day and lived through the highs and lows of this degree with me. We have gone on endless adventures in the Bay Area (Mt. Tamalpais, Jane bread, Cool Patch Pumpkins, and Kirby Cove), California (Yosemite, Big Sur, and San Diego), and beyond (Hawaii, Japan (coming soon)). They have given me refuge, reminded me to laugh, and encouraged me to reach for the stars. I will always consider them my Ohana.

Contributions

The texts of Chapter 2, 3, and 4 are reprints of the material as it appears in “Spatiotemporal mosaic self-patterning of pluripotent stem cells using CRISPR interference”, “Automated Design of Pluripotent Stem Cell Self-Organization”, and “Axial Elongation of Caudalized Human PSC Organoids Mimics Neural Tube Development”. The co-author listed in these publications directed and supervised the research that forms the basis of this dissertation/thesis.

Examining Self-directed Multicellular Organization and Morphogenesis of Human Induced Pluripotent Stem Cells

Ashley Rachel Greeley Libby

Abstract

Embryonic morphogenesis is a critical determinant of tissue generation and function, yet many of the mechanisms that regulate morphogenic processes remain elusive due to the complex and dynamic nature of multicellular interactions and the limited tools to manipulate these systems at single cell resolution. Similarly, morphogenesis of human induced pluripotent stem-cell (hiPSC) derived organoids proceeds largely through self-organized pattern formation that crudely mimics organogenesis. The ability to control specific morphogenic processes would greatly enhance our ability to create bona vide human tissue and organ structures; however, robustly directing organoid morphogenesis requires development of novel control approaches. The four studies described in this dissertation developed systems to study pattern emergence and lineage fate specifications in hiPSC 2D and 3D cultures. First, the effects of mechanical regulation on population emergence and lineage fate specification was interrogated by mosaic knockdown of two mechanical regulators associated with cortical tension and cell-cell adhesion: Rho-associated kinase-1 (ROCK1) and E-cadherin (CDH1). Mosaic knockdown induced symmetry breaking (a pre-requisite of morphogenesis), which resulted in differential patterns of cell sorting and multicellular organization without disrupting pluripotency state. These results describe the spatiotemporal dynamics of multicellular self-organization that occur with the emergence of heterotypic cell populations within pluripotency allowing for population segregation before lineage fate commitment. Second, using machine learning, experimental parameters were optimized to yield pre-specified spatial patterns before successful validation *in vitro* with hiPSCs. Furthermore, asymmetric pre-patterning of iPSCs differentially impacted initial germ layer specification within differentiating colonies. This study demonstrated demonstrate that parameter optimization

achieved by automated machine learning could be used to efficiently predict and control hiPSC self-organization. Third, the regulation of CDH1 on organization and lineage fate was examined in 3D gastruloid cultures, where both microenvironment and CDH1 knockdown triggered the emergence of extra embryonic lineages. This study indicated that microenvironment is essential for specific lineage emergence in the early embryo. Fourth in the final study, population emergence was examined in neuronal organoid cultures that displayed axial extensions similar to neural tube formation and extension. Extending aggregates were dependent on the emergence of a neuromesodermal progenitor population alongside neuronal populations. This model of neural tube extension provides a platform to probe how disruption of population emergence affects morphological phenotypes. Overall, these studies provided avenues to elucidate the underlying multicellular interactions that enable the development of complex patterning events within the developing human embryo on multiple scales and across multiple developmental periods.

TABLE OF CONTENTS

Chapter 1 : Introduction	1
1.1 Overview.....	1
1.2 Symmetry Breaking in Early Development.....	4
1.2.1 Compaction.....	4
1.2.2 Gastrulation and Axis Formation.....	5
1.2.3 Spinal Cord Development.....	6
1.3 Stem Cell Modeling of Embryonic Development.....	8
1.3.1 Tri-lineage Monolayer Differentiations.....	9
1.3.2 Organoids.....	10
1.3.3 CRISPRi as a Tool to Induce Asymmetry.....	11
1.4 Mechanical Regulation of Patterning.....	13
1.4.1 Cell-cell Adhesion Regulation of Morphogenesis.....	13
1.4.2 Cortical Tension Regulation of Morphogenesis.....	17
1.5 Computational Modeling of Cellular Behaviors.....	19
1.5.1 Agent Based Models.....	20
1.5.2 Cellular Potts Model.....	21
1.5.3 Pattern Classification, Machine Learning, and Optimized Pattern Prediction.....	22
1.6 Concluding Remarks.....	23
1.7 Bibliography.....	25
Chapter 2 : Spatiotemporal mosaic self-patterning of pluripotent stem cells using CRISPR interference	45
2.1 Introduction.....	45
2.2 Materials and Methods.....	46
2.2.1 Human iPSC line generation and culture.....	46
2.2.2 Generation of mixed colonies.....	49

2.2.3 Western blot.....	49
2.2.4 RNA isolation and qPCR.....	49
2.2.5 Atomic Force Microscopy.....	50
2.2.6 Time-lapse imaging.....	50
2.2.7 Immunofluorescence staining.....	51
2.2.8 EDU Incorporation Assay.....	51
2.2.9 FLOW Analysis.....	52
2.2.10 FACS.....	52
2.2.11 Fluidigm 96.96 Array.....	52
2.2.12 Differentiations.....	55
2.2.13 Computational Image Analysis.....	56
2.2.14 Statistics.....	56
2.3 Results.....	57
2.3.1 CRISPRi KD in human iPSCs modulates epithelial morphology.....	57
2.3.2 Mosaic CRISPRi silencing results in multicellular organization.....	61
2.3.3 Mosaic human iPSC colonies retain a pluripotent phenotype.....	66
2.3.4 Mixed populations of KD cells display transient gene expression changes in coordination with emergence of patterns.....	68
2.3.5 Mixed populations allow for restricted germ lineage emergence.....	73
2.4 Discussion.....	74
2.5 Conclusion.....	81
2.6 Bibliography.....	82
Chapter 3 : Automated Design of Pluripotent Stem Cell Self-Organization.....	87
3.1 Introduction.....	87
3.2 Materials and Methods.....	90
3.2.1 Cell Lines.....	90

3.2.2 Generation of CRISPRi knockdown iPSC lines.....	90
3.2.3 Mixed Colony Generation	92
3.2.4 Immunofluorescence Staining and Imaging.....	92
3.2.5 Protein Quantification.....	93
3.2.6 mRNA quantification.....	93
3.2.7 Time Lapse Imaging.....	95
3.2.8 Comparison of in vitro and in silico Spatial Patterns.....	95
3.2.9 BMP4 differentiations.....	98
3.2.10 Cellular Potts Model Environment.....	98
3.2.11 Cellular Potts Model Dynamics.....	100
3.2.12 Physical Units and Other Cellular Phenomena of Cellular Potts Model.....	103
3.2.13 Model Fitting to Empirical Data.....	103
3.2.14 TSSL Scoring and Pattern Optimization.....	110
3.2.15 Quantification and Statistical Analysis.....	118
3.2.16 Software.....	118
3.2.17 Data Availability.....	119
3.3 Results.....	119
3.3.1 Pattern Synthesis: in silico prediction and automated discovery of spatial behaviors.....	119
3.3.2 Data-Driven Computational Model of Human iPSC Self-Organization.....	120
3.3.3 Formulating Parameters for Design Automation.....	124
3.3.4 Quantitative Pattern Classification.....	126
3.3.5 Automated Discovery of Pattern Producing Conditions.....	127
3.3.6 <i>In Silico</i> Model Accurately Predicts <i>In Vitro</i> Experimental Validation.....	131
3.3.7 Colony Organization Impacts Patterns of iPSC Differentiation.....	137
3.4 Discussion.....	138

3.5 Conclusion.....	141
3.6 Bibliography.....	142
Chapter 4 : Loss of E-cadherin promotes extraembryonic fates accompanying multilineage differentiation.....	148
4.1 Introduction.....	148
4.2 Materials and Methods.....	149
4.2.1 Human Induced Pluripotent Stem Cell Culture.....	149
4.2.2 Encapsulated Mixed Aggregate Generation.....	150
4.2.3 Un-encapsulated Mixed Aggregate Generation.....	150
4.2.4 Immunofluorescence Staining and Imaging.....	151
4.2.5 Whole Mount Lightsheet Imaging.....	151
4.2.6 Single Cell RNA Sequencing Sample and Library Preparation.....	152
4.2.7 Genome Annotation, RNA-seq Read Mapping, and Estimation of Gene and Isoform Expression.....	153
4.2.8 Data Acquisition, Processing, and Merging.....	153
4.2.9 Cluster Analysis.....	154
4.3 Results.....	154
4.3.1 Loss of CDH1 Promotes Protrusion Morphology when Encapsulated.....	154
4.3.2 Unencapsulated and Encapsulated CDH1(-) Aggregates Display Differing Morphologies.....	155
4.3.3 Emergence of Extraembryonic-like Cell Fates in Encapsulated Aggregates.....	158
4.3.4 Human iPSC Aggregates Have Similar Transcriptomes to Pre-implantation Human Embryos.....	159
4.4 Discussion.....	161
4.5 Conclusion.....	164

4.6 Bibliography.....	164
Chapter 5: Axial Elongation of Caudalized Human PSC Organoids Mimics	
Neural Tube Development.....	170
5.1 Introduction.....	170
5.2 Materials and Methods.....	171
5.2.1 Human Induced Pluripotent Stem Cell Line Generation and Culture.....	171
5.2.2 Organoid Differentiation.....	172
5.2.3 Organoid Elongation Imaging and Quantification.....	173
5.2.4 Real Time Quantitative Polymerase Chain Reaction.....	175
5.2.5 Histology, Immunocytochemistry and Imaging.....	176
5.2.6 Whole Mount Lightsheet Imaging.....	177
5.2.7 Single Cell RNA Sequencing Sample and Library Preparation.....	177
5.2.8 Genome Annotation, RNA-seq read mapping, and Estimation of Gene and Isoform Expression.....	178
5.2.9 Cluster Analysis.....	179
5.2.10 RNAScope.....	179
5.2.12 Statistical Analysis.....	180
5.3 Results.....	180
5.3.1 Emergence of Axial Extensions from Neuronal Organoids.....	180
5.3.2 Wnt Mediated Increase in Extensions and SOX2(+)T(+) Progenitor Populations.....	185
5.3.3 Regionalized HOX Gene Expression in Extending Aggregates.....	185
5.3.4 Manipulation of Endogenous BMP Signaling by CRISPR Interference.....	189
5.4 Discussion.....	190
5.5 Conclusion.....	191

5.6 Bibliography.....	192
Chapter 6: Future Considerations.....	197
6.1 CRISPRi Colony Pattern Formation Screens	198
6.2 Expanding CRISPRi Perturbations in the Cellular Potts Model.....	199
6.3 Interrogating Hippo Signaling in Pluripotent Stem Cell Transitions.....	201
6.4 Examining Neuromesodermal Progenitor Fate Regulation.....	202
6.5 Conclusions.....	204
6.6 Bibliography.....	205

LIST OF FIGURES

Figure 2.1: Maintenance of normal human iPSC karyotypes.....	48
Figure 2.2: CRISPRi of ROCK1 and CDH1 modulate physical properties of the cell.....	59
Figure 2.3: Protein knockdown time course for ROCK1 and CDH1.....	60
Figure 2.4: Morphology of ROCK1 knockdown human iPSCs.....	60
Figure 2.5: Maintenance of Tight junctions in knockdown human iPSCs.....	61
Figure 2.6: Cell-autonomous pattern emergence in mixed population colonies.....	63
Figure 2.7: Pattern emergence of variable ratios of mixed populations.....	64
Figure 2.8: Cell population change over time.....	65
Figure 2.9: Maintenance of nuclear pluripotency markers and epithelial phenotype.....	67
Figure 2.10: Maintenance of cell junction localized β -catenin in knockdown cells.....	68
Figure 2.11: Pluripotent and early germ layer gene expression in knockdown cells.....	69
Figure 2.12: Transient gene expression changes in mixed populations.....	71
Figure 2.13: Gene expression changes in WT-GFP cells mixed with CRISPRi cells.....	72
Figure 2.14: Germ lineage differentiations of mixed colonies.....	75
Figure 2.15: Inducible pattern emergence through the KD of molecules that affect human iPSC physical properties.....	76
Figure 3.1: Overview of Synthesis of Spatial Patterns.....	89
Figure 3.2: Karyotype of CRISPRi Gen2 CDH1 hiPS Cell line.....	91
Figure 3.3: Knockdown of CDH1 quantified by immunofluorescence imaging.....	94
Figure 3.4: Characterizing <i>in vitro</i> morphological metrics.....	96
Figure 3.5: Segmentation workflow.....	97
Figure 3.6: CRISPRi cell line mRNA knockdowns.....	109
Figure 3.7: Pairwise experiments to characterize dynamic changes in spatiotemporal behaviors.....	125

Figure 3.8: Quantitative Pattern Classifier with TSSL.....	128
Figure 3.9: TSSL increase with pattern formation.....	131
Figure 3.10: Computational synthesis of <i>de novo</i> spatial patterns and <i>in vitro</i> validation...	134
Figure 3.11: In silico vs. in vitro comparative metrics of generated target patterns.....	135
Figure 3.12: Titration of gene knockdown mixing ratios <i>in vitro</i> and <i>in silico</i>	136
Figure 3.13: Differentiation of Bullseye colonies.....	139
Figure 4.1: Encapsulated 50 cell human iPSC aggregate morphologic changes.....	156
Figure 4.2: Unencapsulated human iPSC 50 cell aggregate morphologies.....	157
Figure 4.3: Single cell sequencing analysis of encapsulated and unencapsulated aggregates.....	160
Figure 4.4: Lineage trajectories of 50 cell aggregates.....	161
Figure 4.5: Comparison of <i>in vitro</i> aggregate transcriptome with preimplantation human embryos.....	162
Figure 5.1: Emergence of axial extensions from neuronal organoids.....	182
Figure 5.2: Aggregate extension phenotypes.....	183
Figure 5.3: Wnt mediated increase in extensions and SOX2(+)T(+) progenitor population.....	184
Figure 5.4: Progenitor and HOX gene expression in extending and non-extending aggregates.....	187
Figure 5.5: Gene expression analysis of extending and non-extending aggregates by single cell RNA sequencing.....	188
Figure 5.6: Manipulation of endogenous BMP signaling by CRISPR interference.....	190

LIST OF TABLES

Table 2.1: CRISPRi gRNA sequences.....	47
Table 2.2: qPCR primer sequences.....	53
Table 2.3: Fluidigm primer sequences.....	53
Table 3.1: List of design parameters that map to experimental perturbations.....	104
Table 3.2: Model fitting parameters for the Space vs. Time, Protein Expression vs. Time, and the Protein Expression vs. Space characterizations.....	104
Table 5.1: CRISPRi guides.....	172
Table 5.2: qPCR primers.....	175
Table 5.3: Antibodies.....	177
Table 5.4: RNAscope Probes.....	179

Chapter 1: Introduction

1.1 Overview

The goal of this dissertation was to interrogate mechanisms behind the symmetry breaking events that regulate population emergence and organization within the developing embryo. Symmetry breaking is an overarching term used in developmental biology to describe instances in the development of complex organisms when organizational events occur that divide what was once a homogeneous population into separate heterogeneous populations. The emergence of heterogeneity can take the form of changes in gene expression, changes in cell shape, physical movement into patterns, etc. However, often developmental symmetry breaking events involve a wide variety of complex genetic, behavioral, and morphologic changes at once, making it difficult to interrogate the cause of the initial emergent heterogeneity. Therefore, this dissertation outlines multiple methods in the form of four studies to interrogate aspects of symmetry breaking events using human induced pluripotent stem cells (iPSCs) as a model for early embryonic development.

The goal of the first study presented in this dissertation was to develop a system in human iPSCs that would enable robust control over symmetry breaking events. Thus, an inducible CRISPR interference (CRISPRi) system was used in mixed populations of human iPSCs where only a portion of the colony contained the CRISPRi machinery. Then, with CRISPRi induction, a relatively homogeneous human iPSC population would undergo controlled symmetry breaking as only a subpopulation experienced genetic knockdown. This platform was used to interrogate mechanical regulators of population organization: the cell-cell adhesion molecule E-cadherin (CDH1) and the tension regulator Rho-associated coiled-coil kinase (ROCK1). Interestingly, mosaic knockdown triggered population organizations within human iPSC colonies without disrupting pluripotency. When patterned colonies were then differentiated, cells lacking CDH1

preferentially differentiated to the mesendoderm when in mixed colonies, while cells lacking ROCK1 were able to differentiate to all germ lineages equally. This study provided a unique method to control symmetry breaking events and demonstrated the importance of adhesion driven symmetry breaking in the regulation of lineage fate decisions.

With the development of a system to genetically control sub-population emergence, the goal of the second study was to apply the behaviors observed in the symmetry breaking *in vitro* studies to a computational model capable of predicting novel patterning events and organizations of human iPSCs within colonies. Combining a Cellular Potts-based computational model and a pattern recognition framework, a computational system to predict specific experimental parameters was created to derive patterning events upon induction of symmetry breaking. Executing the *in silico* derived experimental set up *in vitro* resulted in multicellular population organization that remarkably reflected the *in silico* predictions in both frequency and extent of pattern formation. Furthermore, differentiation of patterned human iPSC colonies resulted in divergent patterned cell fate commitment upon morphogen exposure, indicating that directed multicellular organization impacts lineage co-emergence. This study demonstrated that morphogenic dynamics could be predicted *in silico* to accurately manipulate human iPSCs *in vitro* to produce desired morphogenetic events, a critical first step towards robustly engineering more complex tissues.

After establishing control over symmetry breaking events and subsequent human iPSC organization in monolayer culture, the goal of the third study was to interrogate the regulation of symmetry breaking through changes in adhesions in a 3D space to better mimic multi-dimensional organization of embryonic development. Specifically, compaction of the embryo from the morula to the early blastocyst and subsequent gastrulation involves a wide variety of multi-dimensional population organizations and migrations that are difficult to recapitulate in monolayer cultures. Therefore, the previously described mosaic knockdown of CDH1 was initiated in 50 cell human iPSC 3D aggregates encapsulated in an alginate hydrogel. With mosaic CDH1 knockdown

population segregation observed and further complete knockdown of CDH1 triggered emergent morphologies reminiscent of embryonic stages. Transcriptome analysis of encapsulated and unencapsulated aggregates revealed emergence of all three embryonic germ lineages without exposure to external morphogens. Moreover, the presence of an extraembryonic-like population in only the encapsulated aggregates indicated that micro-environment may play a role in generating cues for lineage emergence. Overall, this study demonstrates an organoid culture system mimicking aspects of both morphogenic arrangements and cellular lineage fates that arise in the early embryo.

Moving beyond the initial symmetry breaking events in the embryo such as compaction and gastrulation, the final study used the three dimensional culture techniques from the third study to interrogate more complex symmetry breaking in the form of body axis extension. An *in vitro* organoid model of axial extension was developed from a previously reported human iPSC interneuron differentiation protocol. The resultant organoids recapitulated a large number of morphologic and gene expression features found in *in vivo* axial elongation, forming internal tube-like epithelial compartments. Additionally, these organoids contained cells simultaneously expressing brachyury(T) and SOX2 suggesting the emergence of neuromesodermal progenitors (NMPs), which *in vivo* contribute to both the closing neural tube and paraxial mesoderm. Elongations increased in a dose dependent manner with the addition of the small molecule Wnt agonist CHIR, recapitulating the *in vivo* role of Wnt signaling in caudalization and NMP emergence. Single cell RNA sequencing revealed increased MIXL1(+) paraxial mesoderm populations in extending aggregates. HOX gene expression was regionalized in extending aggregates, with hindbrain (HOXB1) expression at the base, while brachial (HOXC6) and thoracic (HOXB9) were expressed in the extensions. To explore dorsoventral polarity in the organoids, knockdowns of BMP inhibitors Noggin and Chordin were generated using CRISPR interference, yielding increased expression of dorsal neural fates in inhibited lines due to loss of endogenous BMP inhibition. This study is the first to report robust neural tube axial extension in an organoid

system. Further, it demonstrates the power of human iPSC organoid models to elucidate many of the underlying multicellular interactions that drive symmetry breaking enabling the development of complex human tissues.

This dissertation provides the first description of controlled symmetry breaking events in human iPSC culture, a quantitative computational model capable of predicting human iPSC patterning events, and a description of multiple unique organoid systems that highlight morphogenic symmetry breaking events. Furthermore, it interrogates mechanisms by which adhesions regulate lineage emergence in early embryogenesis, enabling the future study of the cumulative mechanisms by which robust morphogenesis is controlled. The following introduction provides the necessary background on topics critical to this work including, symmetry breaking in early embryonic development, stem cell modeling of embryonic development, mechanical regulation of patterning, and computational modeling of cellular dynamics.

1.2 Symmetry Breaking in Early Embryonic Development

Symmetry breaking in biological systems occurs with the emergence of a heterogeneous subpopulation, most commonly undergoing changes in cellular polarity, gene expression, replication, or cellular morphology. In embryonic development, symmetry breaking events characterize the repeated emergence of different cell fates and dictate the creation of the complex structures that comprise embryonic tissues. The following sections describe key events in early mammalian development that rely on symmetry breaking events and subsequent organization to regulate embryo morphogenesis as it transitions between stages.

1.2.1 Compaction

One of the earliest symmetry breaking events in the developing embryo occurs at compaction where a subpopulation of cells in the morula becomes transcriptionally distinct,

leading to morphogenic reorganization to form the blastocyst. The blastocyst is divided into two structurally distinct tissues: an outer layer of polarized trophectoderm cells which surrounds a cavity called the blastocoel, and inner layer of non-polar cells called the inner cell mass (Gueth-Hallonet and Maro 1992). The emergence of the trophectoderm and distinction from the inner cell mass depends on the co-expression of the transcription factors CDX2 and GATA3 (Ralston et al. 2010, 2; Niwa et al. 2005; Strumpf et al. 2005, 2) regulated by the farther upstream transcription factor TEAD4 where embryos lacking TEAD4 fail to produce trophectoderm or develop the blastocoel cavities (Nishioka et al. 2008; Yagi et al. 2007).

In preimplantation mouse embryos, changes in TEAD4 signaling are largely regulated by the Hippo pathway where nuclear accumulation of Yes-associated protein 1 (YAP1) directs TEAD4's localization within the genome, regulating its transcriptional activity (Hirate et al. 2012; Nishioka et al. 2009). Interestingly, changes in YAP1 nuclear accumulation are linked to changes in cell-cell adhesions and subsequent cellular polarity as the embryo grows in size and cell number (Hirate et al. 2013; Mihajlović and Bruce 2016). Furthermore, YAP1 nuclear accumulation has been shown to be mechanically responsive; changes in ECM stiffness as well as tension across cellular junctions can drive nuclear localization of YAP1 (Benham-Pyle, Pruitt, and Nelson 2015; Dupont et al. 2011), suggesting that the physical rearrangement of cells during compaction may contribute to trophoblast fate specification. However, the exact relationship between changing adhesions and trophoblast specification is unclear.

1.2.2 Gastrulation and Axis Formation

The next symmetry breaking event that occurs in development is formation of the epiblast and the primitive endoderm, followed by gastrulation of the late blastocyst, which also relies on changes in polarity and adhesion (Montero and Heisenberg 2004). The late blastocyst is composed of two compartments: the epiblast, which will form the embryo proper, and the primitive endoderm, which will eventually become the yolk sac. During gastrulation, the epiblast transitions

from a relatively homogeneous population of epithelial pluripotent cells to a spatially organized, multicellular tissue of distinct progenitor cells. This occurs when cells within the epithelial layer begin to express mesenchymal transcription factors like SNAIL and Brachyury, delaminate from the epithelial layer, and invade into the surrounding tissue forming the primitive streak (Tada, Concha, and Heisenberg 2002; Montero and Heisenberg 2004; Tam and Behringer 1997). The delamination of cells not only delineates a change in fate to different mesodermal populations, but it is also accompanied by changes in adhesion as cells change from epithelial to mesenchymal phenotypes and migrate across the embryo. With movements associated with gastrulation comes the formation of the body axis where the primitive streak delineates the dorsal side of the body forming the dorsal-ventral axis as well as establishing the anterior-posterior axis with its direction of extension and inherently the left-right body axis (Beddington and Robertson 1999). However, the specific dynamics of these processes are not well established due to gastrulation occurring post implantation in many model mammalian systems making its direct examination difficult.

1.2.3 Spinal Cord Development

As the embryo continues to develop past gastrulation, multiple tissues arise and begin to morphologically shape the embryo, leading to anisotropic morphologies. For example, the spine, a crucial structure that both enables physical support as well as protection of essential neural projections that connect the body to the brain, begins to develop at gastrulation. Dysregulation of the key processes leading to its formation results in several congenital abnormalities (Kaplan, Spivak, and Bendo 2005). The generation of the spine patterns and elongates the anterior-posterior body axis and specifies the tissues that contribute to the spinal cord (Steventon et al. 2016; Wilson, Olivera-Martinez, and Storey 2009; Schiffmann 2006; Yamaguchi 2001). Although neural tube formation and axial extension have long been studied in model organisms such as chick and amphibians, it is complicated to study the multi-cellular interactions that drive this axial extension at high spatiotemporal resolution in mammalian embryos because it occurs post-

implantation (Viebahn 1999; Beddington and Robertson 1999). The following subsection outlines our current understanding on how neural tube patterning is established.

The process of neural tube formation, or neurulation, is characterized by thickening and flattening of the neural plate in response to signals from the notochord, followed by folding inward ventrally and finally pinching off at the top of the fold to close the tube (J. L. Smith and Schoenwolf 1989; Tam and Behringer 1997). Coincident with neural tube closure is extension of the anterior-posterior axis as the entire embryo breaks symmetry and elongates. Concomitant with elongation is the establishment of different progenitor zones within the neural tube creating both ventral-dorsal and rostral-caudal patterning of the neural tube. These patterning events rely on a combination of morphogen gradients across the embryo that establish the proper signaling milieu to determine progenitor fate based on dorsal-ventral or rostral-caudal location within the developing neural tube. Dorsal-ventral gradients differences in sonic hedgehog (Shh) released ventrally from the notochord and bone morphogenic proteins (BMP) released dorsally from the roof plate (Ericson et al. 1997; Liem et al. 1995; McMahon et al. 1998). Patterning the rostral-caudal axis of the embryo is the Wingless-type MMTV integration site family (Wnts), retinoic acid (RA) released from the neighboring somites, and fibroblast growth factors (FGFs), which dictate convergent extension, progenitor proliferation, and have been shown to regulate HOX gene profiles (Janssens et al. 2010; Carpenter 2002; Corral and Storey 2004; del Corral et al. 2003; Bel-Vialar, Itasaki, and Krumlauf 2002).

As the neural tube continues to extend, an additional axial pool of progenitors, called neuromesodermal progenitors (NMPs), contributes to both the neural tube and the surrounding paraxial mesoderm that eventually becomes that somites (Henrique et al. 2015). This axial stem cell pool of NMPs and are marked by high expression of the transcription factors SOX2 and Brachyury (Henrique et al. 2015; Gouti et al. 2014) and reside in the caudal lateral epiblast just below the closing neural tube. Interestingly, this population of cells is hypothesized to regulate the balanced production of neural tissue and surrounding somites that eventually form the spinal

cord and vertebrae. However, due to NMP location in the post implantation embryo making visualization difficult, the dynamics of NMP symmetry breaking that regulate the production of two different tissues are largely unknown, leaving multiple unanswered questions about their maintenance, differentiation, patterning, and subsequent regulation over axis extension.

Overall, spinal cord development provides multiple examples of symmetry breaking. First, morphogenic symmetry breaking occurs with physical elongation of the embryonic axis through neural tube closure and extension. Second, symmetry breaking is observed in coordinated population emergence and patterning of neuronal subtypes specific to location within the developing neural tube. Finally, the maintained progenitor pools like NMPs continuously demonstrate regulation over population emergence as the embryo continues to develop providing a constant source of new tissue through symmetry breaking lineage fate decisions.

1.3 Stem Cell Modeling of Embryonic Development

The repeated relevance of symmetry breaking events in the development of the early embryo and the health implications of their dysregulation in human congenital diseases (congenital heart disease, spinal malformations, etc.) highlights the importance of mechanistically understanding these basic processes across embryonic development (Kaplan, Spivak, and Bendo 2005; Srivastava 2000). However, the dynamics of development have typically been difficult to interrogate *in vivo* due to the physical restrictions, optical opacity, and complex signaling milieu inherent to the developing embryo. Furthermore, human specific embryonic processes are largely unknown as using human embryos to study morphogenic events presents ethical dilemmas. Therefore, to study symmetry breaking events and morphogenesis, it is essential to establish an *in vitro* human system that promotes the coincident development of analogous heterogeneous populations.

Human pluripotent stem cells (hPSCs) provide an unlimited source of cells that can mimic developmental differentiation processes and maintain the ability to self-organize into tissue-like structures, such as optic cups, gut organoids, or stratified cortical tissues (Eiraku et al. 2008; 2011; Spence et al. 2010). Human pluripotent stem cells are split into two main categories: embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). ESCs are isolated from the inner cell mass of embryos and maintain a cellular state of pluripotency, meaning that they have the ability to both renew themselves and differentiate to all three germ lineages (Trounson 2001). Similarly, iPSCs, which are cells reprogramed from fibroblasts by activating the pluripotent gene regulatory network, demonstrate the ability to both renew and differentiate to all three germ lineages (Yamanaka 2012; Takahashi et al. 2007). However, despite the inherent ability of both ESCs and iPSCs to form multiple cell types in 2D culture or in 3D organoids, stem cell differentiations are intrinsically variable (Bredenoord, Clevers, and Knoblich 2017). Historically, differentiations that mimic symmetry breaking at gastrulation, allowing for the emergence of all three lineages, have been in the form of embryoid bodies. These models are inherently random, difficult to robustly reproduce, and do not recapitulate body plan formation during embryonic development (Kurosawa 2007; Carpenedo, Sargent, and McDevitt 2007). As a result of this variability and the lack of alternative human models that faithfully promote asymmetric emergence, many of the mechanisms that control and coordinate human morphogenesis remain undefined. Therefore, new approaches to reliably control the emergence and organization of multiple cell types would greatly advance tissue modeling and organ developmental studies.

1.3.1 Tri-lineage monolayer differentiations

Monolayer stem cell differentiations have been developed for a wide variety of cell types and range in their purity and reproducibility (Butts et al. 2017; Lian et al. 2012; Spence et al. 2010; Gouti et al. 2014). Often, these differentiations rely on the use of either morphogens or small molecules added to the cell culture medium to activate certain cell signaling pathways and direct

cell lineage fate decisions to the desired cell type. However, even with external signals directing cell fate, differentiations in monolayer culture often produce a heterogeneous mixture of cells that remain largely unorganized, despite the emergence of multiple cell types. These differentiations do not reflect the organized tissue patterning observed within the developing embryo. To address this, recent techniques have been developed that allow for both multi-population emergence as well as organization of multiple cell types in the form of tri-lineage differentiations which mimic aspects of both population emergence and patterning as in the early embryo.

Patterned tri-lineage differentiations can be achieved by stimulating monolayer cultures of PSCs on micropatterned surfaces that restrict colony boundaries (Warmflash et al. 2014). When morphogens present at gastrulation (BMPs, Wnts, Activin) are added to cultures with such boundary conditions, radial patterns of specification emerge: first an outer ring of CDX2(+) cells forms, followed by an inner ring of brachyury(+) cells, and a center of SOX2(+); marking an extraembryonic-like population, a mesendoderm population and a ectoderm population, respectively (Warmflash et al. 2014; Tewary et al. 2017; James et al. 2005; Britton et al. 2019). These studies demonstrate that PSCs have an intrinsic ability to undergo symmetry breaking and self-organize when presented with boundary conditions. However, despite the ability to generate patterns of cells, these systems only produce radial organization events, leaving many unanswered questions about how a radially symmetrical embryo at the morula or blastocyst stage can robustly generate anisotropic patterning events such as the primitive streak or axial elongation.

1.3.2 Organoids

An additional example of symmetry breaking and patterning *in vitro*, is the creation of organoid models where PSC differentiations are conducted in three dimensional environments, yielding multiple cell types that spatially organize into structures reminiscent of embryonic tissues (Bredenoord, Clevers, and Knoblich 2017). Such organoid models have been established for

multiple organ systems such as gut, brain, and liver (Lancaster and Knoblich 2014; Clevers 2016). In these systems, PSCs or progenitors are exposed to signaling cues which drive lineage fates toward specific tissues; a signaling milieu which is reinforced by endogenous signaling between cells. A combination of these exogenous and endogenous signals allows for the emergence of the many cell types involved in that tissue's structure and function. Organoids offer potential in both therapeutic research and basic scientific discovery as they can be used to recapitulate elements of development and disease, offering opportunities that range from mechanistic interrogation of biological processes to drug testing. However, a major limitation of organoid models is the wide variety of phenotypes that are generated from a single differentiation. Structural organization within organoid systems is difficult to both control and repeat (Bredenoord, Clevers, and Knoblich 2017). This lack of precision is highly contrast to the robustness of embryonic development and organogenesis where generation of body plan and tissue structure is tightly regulated. Therefore, engineered platforms that control symmetry breaking and organization would facilitate the mechanistic study of the biological processes that regulate morphogenesis.

1.3.3 CRISPRi as a Tool to Induce Asymmetry

A hallmark of symmetry breaking is the asymmetric co-emergence of distinct cell populations that then can self-organize to form developmental patterns, multicellular structures, and ultimately, functional tissues and organs (Bronner 2016; Lancaster and Knoblich 2014; Sasai 2013). Controlling cellular heterogeneity *in vitro* is often achieved by independent differentiation of hPSCs followed by re-combination of distinct cell types, which fails to mimic parallel cell-type emergence (Matthys, Hookway, and McDevitt 2016). Attempts to engineer *in vitro* systems that yield controlled emergence of spatial organization often rely on extrinsic physical restriction of cells to direct subsequent multicellular pattern formation (Warmflash et al. 2014; Hsiao et al. 2009). Physical constraints allow the observational study of cell-cell interactions within defined

regions, but artificially restrict cell behaviors by limiting the degrees of freedom in which morphogenic phenomena can occur. Additionally, current tools to interrogate gene function, such as genetic knockouts or siRNA (Boettcher and McManus 2015), cannot selectively perturb gene expression of subpopulations of cells *in situ*, which is required to generate controlled asymmetry analogous to embryonic morphogenesis.

Several of these limitations can be addressed with inducible CRISPR interference (CRISPRi) systems in mammalian cells (Larson et al. 2013; Mandegar et al. 2016; Gilbert et al. 2014). CRISPRi uses the gene editing tools of CRISPR (Qi et al. 2013; Jinek et al. 2013; Cong et al. 2013) but prevents double stranded breaks by using a catalytically dead Cas9 (dCas9). With additional fusion of the dCas9 to a Krüppel associated box (KRAB) domain allows for gene repression by guide RNA mediated binding of the dCas9-KRAB to the transcription start site of the gene or interest, preventing polymerase binding and subsequent transcription (Larson et al. 2013). Furthermore, inducible CRISPRi systems enable temporal regulation over specific genetic targets with limited off-target effects (Boettcher and McManus 2015). Thus, with temporal knockdown constraints, precisely-controlled biological systems can be engineered to induce well-defined genetic perturbation at explicit times and within defined populations of cells, mimicking developmental symmetry-breaking events. Moreover, applying such a technology to stem cell populations will facilitate mechanistic study of symmetry breaking in a wide variety of tissues as the same PSC line containing a CRISPRi system can be used in trilineage differentiations or in multiple types of organoids systems, allowing for robust genetic control over population emergence. Overall, the advent of CRISPR technologies presents multiple avenues to regulate population emergence *in vitro* which not only engineers control over heterogenous systems, but also provide methods by which to robustly interrogate morphogenic processes dependent of population emergence and symmetry breaking.

1.4 Mechanical Regulation of Patterning

Symmetry breaking events can occur not only at the gene expression level, but also at the physical morphogenic level. Morphogenic asymmetries arise from reorganization of cells due to local changes in mechanical tissue stiffness and cell adhesions that facilitate physical organization of developing embryos (Krieg et al. 2008; Maître et al. 2012; Heisenberg and Bellaïche 2013). Mechanical rearrangement is necessary for many aspects of morphogenesis, including cell polarity, collective movement, multicellular organization, and organ size regulation (Arboleda-Estudillo et al. 2010; Maître 2017). Differential adhesion (Foty and Steinberg 2004; 2005) and cortical tension (Essen 1997; Krieg et al. 2008) are critical determinants of mechanically-driven cell sorting, in which both processes are known to contribute to tissue organization (Lecuit and Lenne 2007). In cortical tension-dominated sorting, variable actin cytoskeleton-generated cortex tension stimulates sorting of individual cells, whereas differential adhesion sorting promotes segregation of cell populations due to intercellular homophilic adhesions. In this thesis specific molecules regulating either cell-cell adhesion or cortical tension are manipulated to introduce physical asymmetries within populations of iPSCs. However, although the regulation of cellular organization by physical forces has been highly studied, how changing adhesions and cortex tension regulate the coordination of lineage fate decisions with emergence of patterns is not completely understood.

1.4.1 Cell-cell Adhesion Regulation of Morphogenesis

The regulation of changing adhesions is required for many developmental morphogenic processes including mesoderm formation, neural crest formation, heart field migration, and heart septa formation (Heasman et al. 1994; Montero and Heisenberg 2004; Radice et al. 1997; Xu, Baribault, and Adamson 1998). These distinct morphogenic events that rely on epithelial to mesenchymal transitions (EMT) are often characterized by subpopulation organization,

delamination from an epithelial layer, and subsequent migration across the embryo (C.A. Burdsal, C.H. Damsky, and R.A. Pedersen 1993); all of which rely on changing intercellular adhesion. Consequently, the regulation of changing intercellular contacts in coordination with lineage transitions is essential for robust morphogenic processes. Despite much research into how changing intercellular adhesion molecules regulate cell population movement, there has been little work mechanistically examining connections between changing intercellular adhesions and lineage fate decisions.

An important family of intercellular adhesion molecules involved in morphogenesis are cadherins. Cadherins are single pass Ca^{2+} -dependent transmembrane glycoproteins that mediate cell-cell adhesions. E-cadherin (CDH1), a type I classical cadherin, is widely associated with early developmental morphogenesis (Ringwald et al. 1987). Unsurprisingly, the reduction of CDH1 levels is essential to EMT and its regulation of subsequent morphogenic processes (Heasman et al. 1994; Przybyla, Lakins, and Weaver 2016). Consequently, CDH1 has been implicated in the control of downstream lineage fate decisions that help to regulate morphogenesis in a range of animal models (Przybyla, Lakins, and Weaver 2016; Li et al. 2010; C.A. Burdsal, C.H. Damsky, and R.A. Pedersen 1993). For example, prevention of its down regulation in chick embryos results in wide spread developmental defects including lack mesenchymal mesoderm and neural crest formation (Nieto et al. 1994). Additionally, *in vivo* over-expression of CDH1 in *Xenopus laevis* prevents the induction of the mesoderm lineage during gastrulation (Heasman et al. 1994). Furthermore, changes in CDH1 contacts coordinate with *in vitro* stem cell lineage decisions. These studies suggest that CDH1 may play a larger role in lineage specification during morphogenesis, in addition to regulating the self-organization and migration of cell populations.

Overlap between CDH1 control over adhesion and potential control over lineage specification, lies with the intracellular domain of the CDH1 complex. The intracellular CDH1 complex involves multiple scaffolding proteins: β -catenin, α -catenin, γ -catenin (plakoglobin) and p120-catenin. These proteins serve as mediators that attach the cytoplasmic domain of CDH1 to

the actin cytoskeleton, allowing for physical anchorage and stabilization of intercellular adhesions (Leckband and Rooij 2014; Steinberg and McNutt 1999; Huber et al. 2001). Interestingly the CDH1 complex proteins have been characterized to have multiple interactions within the cell outside of their role in adhesion within the CDH1 complex (A. L. Smith et al. 2012; Munemitsu et al. 1995; Przybyla, Lakins, and Weaver 2016; Heuberger and Birchmeier 2010). In particular, the protein β -catenin has the ability to translocate to the nucleus and drive transcriptional regulation.

Within the nucleus, canonical β -catenin transcriptional regulation is key to the Wnt signaling (Dale 1998). Wnt ligand binding disrupts the β -catenin destruction complex from degrading cytoplasmic β -catenin, allowing for β -catenin accumulation, phosphorylation, and eventual translocation to the nucleus. In the nucleus, β -catenin interacts with LEF/TCF transcription factors, promoting the transcription of downstream targets of the Wnt pathway essential to lineage specification (Dale 1998). Consequently, the inhibition of β -catenin prevents mesoderm induction in *Xenopus* embryos (Heasman et al. 1994). However, the TCF transcription factors involved in the Wnt pathway are not the only interactors reported for β -catenin within the nucleus. In fact, a large body of scientific literature indicates the overlap of the Hippo, TGF β , and Wnt signaling pathways by varying transcription factor interactions with β -catenin (B. Zhou et al. 2012; Benham-Pyle, Pruitt, and Nelson 2015; Dale 1998; Attisano and Wrana 2013). β -catenin can interact with the YAP/TAZ transcription factor complex as a part of the Hippo pathway as well as with SMAD2/3 nuclear complex involved in the TGF β signaling pathway (B. Zhou et al. 2012; Benham-Pyle, Pruitt, and Nelson 2015). This poses additional intrigue into the regulation of lineage fate decisions by β -catenin because the YAP/TAZ pathway has been shown to help establish and maintain pluripotent expansion of stem cells (Ohgushi, Minaguchi, and Sasai 2015). And the Activin and Nodal pathways, a sub-set of the TGF β family, are associated with continued maintenance of the transcriptional pluripotency network in human stem cells (James et al. 2005).

Interestingly, feedback between CDH1 adhesion and β -catenin signaling has been demonstrated in all of these transcriptional control pathways. Mechanical perturbations of CDH1

have been shown to result in nuclear localization of β -catenin, and conversely canonical-Wnt signaling activates Snail and Slug, transcription factors that down-regulate CDH1 expression (Heuberger and Birchmeier 2010; Cano et al. 2000; B. P. Zhou et al. 2004; Bolós et al. 2003; Benham-Pyle, Pruitt, and Nelson 2015). Mechanical strain via CDH1 contacts results in nuclear localized Yap1 and β -catenin, which have been shown to complex and regulate epithelial cell cycle entry and consequently tissue size (Benham-Pyle, Pruitt, and Nelson 2015, 1; Heallen et al. 2011). Moreover, activation of the TGF β /Activin/Nodal pathway in human embryonic stem cells can be induced by mechanical strain and contributes to prevention of differentiation (Saha et al. 2008; James et al. 2005; Saha et al. 2006). The combination of these findings provides a strong scientific premise that CDH1's mechanical activity at the plasma membrane may regulate how β -catenin interacts with these downstream pathways during differentiation. The changing adhesions in EMT mediated by reduction CDH1 contacts may result in the release of membrane-bound β -catenin that could activate and amplify the canonical-Wnt pathway as well as interact with the Activin/Nodal and YAP/TAZ pathways to help facilitate a transition out of pluripotency (Przybyla, Lakins, and Weaver 2016). Consequently, *in vivo* mesodermal defects with CDH1 overexpression, similar to β -catenin depletion, may be due to the recruitment of β -catenin to over expressed CDH1 junctions, reducing the cytoplasmic and nuclear pool of signaling-competent β -catenin (Heuberger and Birchmeier 2010). In fact, in pathological situations where CDH1 is down-regulated, β -catenin once held at CDH1 junctions can translocate to the nucleus, resulting in EMT (Kam and Quaranta 2009; Gottardi, Wong, and Gumbiner 2001). Despite the evidence pointing toward a transcriptional regulation of lineage fate via CDH1 regulated β -catenin availability and downstream signaling, there has been little mechanistic interrogation of how changing adhesions associated with developmental morphogenesis and EMT play a larger role in lineage specification. More over the vast majority of our understanding of CDH1 comes from the study of animal models, which have different transcriptional networks in the control of pluripotency maintenance and lineage transitions (Kattman et al. 2011; Dalton 2013; Takashima et al. 2014),

making conclusions about human CDH1 developmental mechanisms difficult to solidify. As a result, fundamental questions about how CDH1 coordinates human morphogenic events and lineage fate have yet to be answered and have led to contradictory literature on CDH1's role in human pluripotency maintenance.

1.4.2 Cortical Tension Regulation of Morphogenesis

As the embryo begins to form complex tissue shapes, in parallel with the changing adhesions that regulate processes such as EMT, changes in tensile forces also help to shape epithelial tissues within the early embryo. In particular, the tension across the actin cytoskeleton and adhesions effectively couple cells within an epithelial sheet (Brodland 2002; Hočevár Brezavšček et al. 2012). This tension can generate a wide range of morphogenic changes where the mechanically steady state of the tissue and thus its subsequent shape are dictated by generation of force across the tissues surface or the cortex or each individual cell (Lecuit and Lenne 2007; Lecuit and Yap 2015; Brodland 2002).

Individual cells generate tension through the activation of the actin cytoskeleton contractile machinery. Specifically, non-muscle myosin II motors are actin binding proteins that localize to organized bundles of actin filaments at the cortex of cells. Here they act as linkers between actin filaments and can drive constriction of the cellular cortex by movement along parallel actin filaments (Vicente-Manzanares et al. 2009; Henn and Cruz 2005). Interestingly, the asymmetric distribution of actin filaments and myosin motors drives several processes in early developmental morphogenesis, such as inner cell partitioning of the early embryo (Samarage et al. 2015; Ducibella and Anderson 1975), gastrulation (Krieg et al. 2008), axis extension (Tahinci and Symes 2003; Butler et al. 2009; Dawes-Hoang et al. 2005), and neural plate folding (Rolo, Skoglund, and Keller 2009).

Non-muscle motor myosin II is a hexameric protein formed of two heavy chains, four light chains where its conformation is regulated by phosphorylation of the light chains by myosin

phosphatase (Vicente-Manzanares et al. 2009; Somlyo and Somlyo 2000). Myosin phosphatase itself is regulated by both fluctuations in intracellular calcium as well as Rho family GTPases and downstream Rho kinases where RhoA activates Rho-associated coiled-coil containing protein kinase (ROCK1) which then in turn activates myosin phosphatase through phosphorylation. This complicated signaling cascade allows for multiple cellular signaling mechanisms to regulate myosin activation and subsequent generation of tension. For example, the wnt planar cell polarity pathway activates actin cytoskeleton reorganization via activation of RhoA, initiating the signaling cascade that results in contractility (Tada, Concha, and Heisenberg 2002; Matthews et al. 2008; Kim and Han 2005).

Beyond cortical tension's roles in cellular and tissue organization, the changing of tension has also been implicated in lineage fate decisions within the embryo as well as PSC differentiations. At gastrulation, the embryo generate three separate germ lineages which have been shown to display differences in cortical tension both *in vivo* as well as in *in vitro* culture (Krieg et al. 2008; Sliogeryte et al. 2014). Organ stratification of the epidermis has been shown to be reinforced by changes in cortical tension that lead to differences in proliferation and differentiation as cells move upward through the stratified tissue (Miroshnikova et al. 2018). Furthermore, uneven distribution of myosin leads to asymmetric cell division in multiple progenitors as they undergo differentiation (Ou et al. 2010; Lechler and Fuchs 2005), suggesting that symmetry breaking in tension generation at the individual cell level can also regulate lineage fate decisions. In addition, the reduction of asymmetries in cortex tension is associated with stabilization of cellular fates. For example, ROCK inhibition and thus reduction in tension generation is often used in human pluripotent cell culture and has been implicated in pluripotency maintenance (Ohgushi, Minaguchi, and Sasai 2015; McBeath et al. 2004). However, how changes in cortex tension might mechanistically regulate cell fate as well as changes in morphology is not well understood.

1.5 Computational Modeling of Cellular Behaviors

Computational modeling offers a method to systematically study the regulation of symmetry breaking events and morphogenesis without the limitations presented by *in vitro* or *in vivo* experimentation. Specifically, patterns generated by biological systems during morphogenesis often resemble solutions to mathematical models such as non-linear reaction diffusion systems from chemistry or magnetic spin model systems from solid state physics (Heisenberg 2017; Turing 1952; Marcon and Sharpe 2012; Graner and Glazier 1992). As a result, the interplay between *in silico* mathematical modeling of biological systems and *in vivo* and *in vitro* experiments interrogating such models has shaped our understanding of developmental morphogenesis. However, such experiments are limited by the ability to measure characteristic properties of the biological system and perturb those properties in a spatio-temporally controlled way. Where computational models can control all aspects of an experiment, often both *in vivo* and *in vitro* systems have limited manipulations that can be performed to test a hypothesis, limiting the ability to match a computational model with a biological one with high fidelity.

For example, although computational approaches can test general principles of biology *in silico*, it is often difficult to directly map these models to specific *in vitro* mechanisms and perturbations, making it challenging to systematically synthesize experimentally tractable perturbations *in silico* that can be accurately reproduced *in vitro*. Previously, several groups (M. Molitoris et al. 2016; Tewary et al. 2017; Warmflash et al. 2014) have induced radial organization of differentiated germ layers by restricting hPSC colonies to micropatterned islands, or have used molecular engineering of cell surface and/or substrate properties to extrinsically control cell location and subsequent multicellular patterning *in vitro* (M. Molitoris et al. 2016; Hsiao et al. 2009; Chandra Ravi A. et al. 2005; L. MacKay, Sood, and Kumar 2014; Toda et al. 2018). However, the resulting patterns that arise spontaneously afford limited control of precise multicellular organization or override the intrinsic mechanisms that regulate cell-mediated

morphogenic assembly. Theoretical *in silico* frameworks have been developed to computationally model multicellular organization (Bartocci et al. 2016; Briers et al. 2016; Sharpe 2017) and automate the design of non-spatial cellular logic (Nielsen et al. 2016). However, despite the independent development of these *in vitro* and *in silico* frameworks for multicellular patterning, the ability of the *in silico* framework to predict a set of manipulations to generate *de novo* multicellular organization *in vitro* has yet to be fully demonstrated. In Chapter 3 of this dissertation, we discuss the development of an extended Cellular Potts computational model that successfully relates *in silico* predictions of pattern formation with attainable *in vitro* outputs through the use of machine learning and mathematical optimization.

1.5.1 Agent Based Models

Agent based models (ABMs) offer a unique opportunity to computationally model collective cellular behaviors while maintaining individual cell identity. ABMs consist of agents that maintain autonomous decision making processes in reaction to a given set of inputs (Bonabeau 2002). By representing cells as individual agents, ABMs allow for discrete models of cell behavior to environmental stimuli where individual cells make cell state decisions (Van Liedekerke et al. 2015; Thorne et al. 2007). ABMs are incredibly dynamic and can incorporate environment, time, and interactions between agents, making them particularly efficient at representing spatial patterning as a result of multiple interactions (Glen, Kemp, and Voit 2019). Additionally, ABMs are often designed to relate measurable metrics of cellular behavior directly to the established rules for decision making in the model, allowing for simulations that go beyond theory and are based on biological measurements (Thorne et al. 2007; Glen, Kemp, and Voit 2019). Furthermore, because ABMs allow for individual cellular behaviors to be established, they allow for the examination of emergent properties at the tissue level that result from distinct single cell decisions.

1.5.2 Cellular Potts Model

There are several established ABM frameworks that can be used to model symmetry breaking and cellular population emergence such as vertex models (Fletcher et al. 2014), spring force models (Güdükbay, Özgüç, and Tokad 1997), or the Cellular Potts model (Voss-Böhme 2012; Marée, Grieneisen, and Hogeweg 2007). The large differences between these ABM frameworks are based on their ability to incorporate the movement of cells, where the spring force models do not allow for cell movement (Gusev 2004; Güdükbay, Özgüç, and Tokad 1997), vertex based models allow for epithelial sheet movement (Hočevar Brezavšček et al. 2012), and Cellular Potts models allow for dynamic cellular movements (Scianna, Preziosi, and Wolf 2012; Szabó and Merks 2013). Because, early mammalian development relies on a wide variety of cellular movements in coincidence with fate emergence (Montero and Heisenberg 2004), the Cellular Potts model provides a rich, computationally-efficient framework to interrogate multicellular morphogenic patterning as a result of symmetry breaking events at the single cellular level.

The Cellular Potts Model (CPM) is an extension of Ising models (Yang 1952) and represents the spatial environment of cells grown in a monolayer using a 2D square lattice grid (Marée, Grieneisen, and Hogeweg 2007). Each square region in the grid (i.e. a lattice site) represents a partial region of a cell's membrane or the medium surrounding a cell. Therefore, cells are represented in the model as combinations of connecting lattice sites. A cell ID is assigned to each lattice site to identify the region of a cell that occupies a lattice site. For example, 100 connected lattice sites each having a cell ID equal to 1 represent a single cell. Complex behaviors such as preferential cell-cell adhesions, cortical tension, and cell migration, are achieved by probabilistically copying lattice sites to adjacent regions, which in the CPM is known as a copy attempt. Each copy attempt is accepted with a probability related to a Hamiltonian function which represents the sum of competing forces influencing intracellular behaviors and cell interactions with the environment. For example, conservation of cell area, locally polarized cell migration, cell-cell adhesion, and cortical tension, and chemotaxis. Every competing force is represented by a

score and a weight, where the score represents a reward or penalty depending on the divergence of a cell from its target behavior, while the weight represents the relative importance of the respective cell behavior. Every time a simulation is run the model aims to reach an equilibrium state dictated by the probabilistic rewards and penalties for individual cellular behaviors. Thus, the model predicts cellular organization and migration behaviors based off of the given set of rules for every simulated cell.

1.5.3 Pattern Classification, Machine learning, and Optimized Pattern Prediction

Both *in vitro* biological systems and *in silico* computational models of pattern generation generate large data sets where individual patterns each need to be classified. However, manual classification for such data sets is not tractable due to length of time to process every sample and the possibility that a human classifier may miss subtle patterns in the data that are biologically relevant. Thus, there is a need for a systematic classification system of said patterns, where similar emergent patterns can be grouped together in an unbiased manner. Machine learning offers a high throughput unbiased approach to classify emergent patterns and their similarity with one another. In particular, rule based classification algorithms that use decision trees allow for the generation of hierarchical partitioning of a data set and can generate a similarity relation or metric between patterns (Cohen 1995; Aggarwal 2014; Bartocci et al. 2016). By treating each simulated configuration or empirical colony image as an individual sample in a data set, a rules-based classifier creates a simple set of rules to categorize the different configurations of cells available to the *in silico* simulation and *in vitro* pattern generation system.

One way to generate these rules-based classifiers, first a training data set is generated with a diverse sample of the multicellular patterns generated by the model. From this training data, a set of rules are derived that represent the complexity of the image as a set of higher-order properties useful for classification. Finally, these generated rules are applied to novel images and patterns to enable comparison between patterns and allow ranking patterns when optimizing

towards a target configuration (Coenen and Leng 2007). In the case of spatial patterning, Tree Spatial Superposition Logic can be used to classify pattern types (Haghighi et al. 2015; Bartocci et al. 2016; Briers et al. 2016). This algorithm uses quadtree logic to represent the spatial relationships within a complex image (Bartocci et al. 2016; Finkel and Bentley 1974; Jackins and Tanimoto 1983) where every image generates a unique quad tree representation that maps the complexity of the image to a reduced vector of values reflecting higher level image structure. Thus, the TSSL creates a reduced version of the image that allows for rapid comparison of image patterns and yields a score of pattern similarity. Despite the rapid advances in classification algorithms, these machine learning techniques are often used as endpoint analyses of biological systems, leaving many potential opportunities to employ feedback from the rules gleaned from classification algorithms to predict and modulate the biological phenomena that produce such patterns.

1.6 Concluding Remarks

In conclusion, embryonic development is marked by repeated symmetry breaking events and subsequent morphogenic reorganizations. The emergence of heterogeneity can range from changes in transcription to differences in physical cell properties across the embryo. However, methods to dynamically interrogate these processes *in vivo* are limited due to the inaccessibility of the embryo at such timepoints, lack of robust tools, and complexity of the embryonic system. Therefore, pluripotent stem cells models paired with computational modeling offer a unique opportunity to develop robust and controllable models for symmetry breaking events in early developmental processes.

This thesis presents a unique combination of bio-engineering, developmental biology, and computational modeling to overcome these limitations, interrogating adhesion regulation of

heterogeneity and lineage fate decisions. The work described here sets the stage for future studies examining the role of population heterogeneity on tissue formation eventually highlighting the series of events necessary to build an organ and enabling the creation of functional therapeutic tissues outside of the body.

1.7 Bibliography

- Aggarwal, Charu C. 2014. *Data Classification: Algorithms and Applications*. CRC Press.
- Arboleda-Estudillo, Yohanna, Michael Krieg, Jan Stühmer, Nicholas A. Licata, Daniel J. Muller, and Carl-Philipp Heisenberg. 2010. "Movement Directionality in Collective Migration of Germ Layer Progenitors." *Current Biology* 20 (2): 161–69. <https://doi.org/10.1016/j.cub.2009.11.036>.
- Attisano, Liliana, and Jeffrey L. Wrana. 2013. "Signal Integration in TGF- β , WNT, and Hippo Pathways." *F1000Prime Reports* 5 (June). <https://doi.org/10.12703/P5-17>.
- Bartocci, E., E. Aydin Gol, I. Haghghi, and C. Belta. 2016. "A Formal Methods Approach to Pattern Recognition and Synthesis in Reaction Diffusion Networks." *IEEE Transactions on Control of Network Systems* PP (99): 1–1. <https://doi.org/10.1109/TCNS.2016.2609138>.
- Beddington, Rosa S. P., and Elizabeth J. Robertson. 1999. "Axis Development and Early Asymmetry in Mammals." *Cell* 96 (2): 195–209. [https://doi.org/10.1016/S0092-8674\(00\)80560-7](https://doi.org/10.1016/S0092-8674(00)80560-7).
- Bel-Vialar, Sophie, Nobue Itasaki, and Robb Krumlauf. 2002. "Initiating Hox Gene Expression: In the Early Chick Neural Tube Differential Sensitivity to FGF and RA Signaling Subdivides the HoxB Genes in Two Distinct Groups." *Development* 129 (22): 5103–15.
- Benham-Pyle, Blair W., Beth L. Pruitt, and W. James Nelson. 2015. "Mechanical Strain Induces E-Cadherin–Dependent Yap1 and β -Catenin Activation to Drive Cell Cycle Entry." *Science* 348 (6238): 1024–27. <https://doi.org/10.1126/science.aaa4559>.
- Boettcher, Michael, and Michael T. McManus. 2015. "Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR." *Molecular Cell* 58 (4): 575–85. <https://doi.org/10.1016/j.molcel.2015.04.028>.
- Bolós, Victoria, Hector Peinado, Mirna A. Pérez-Moreno, Mario F. Fraga, Manel Esteller, and Amparo Cano. 2003. "The Transcription Factor Slug Represses E-Cadherin Expression

- and Induces Epithelial to Mesenchymal Transitions: A Comparison with Snail and E47 Repressors.” *Journal of Cell Science* 116 (3): 499–511. <https://doi.org/10.1242/jcs.00224>.
- Bonabeau, Eric. 2002. “Agent-Based Modeling: Methods and Techniques for Simulating Human Systems.” *Proceedings of the National Academy of Sciences* 99 (suppl 3): 7280–87. <https://doi.org/10.1073/pnas.082080899>.
- Bredenoord, Annelien L., Hans Clevers, and Juergen A. Knoblich. 2017. “Human Tissues in a Dish: The Research and Ethical Implications of Organoid Technology.” *Science* 355 (6322): eaaf9414. <https://doi.org/10.1126/science.aaf9414>.
- Briers, D., I. Haghighi, D. White, M. L. Kemp, and C. Belta. 2016. “Pattern Synthesis in a 3D Agent-Based Model of Stem Cell Differentiation.” In *2016 IEEE 55th Conference on Decision and Control (CDC)*, 4202–7. <https://doi.org/10.1109/CDC.2016.7798907>.
- Britton, George, Idse Heemskerk, Rachel Hodge, Amina A. Qutub, and Aryeh Warmflash. 2019. “A Novel Self-Organizing Embryonic Stem Cell System Reveals Signaling Logic Underlying the Patterning of Human Ectoderm.” *Development* 146 (20): dev179093. <https://doi.org/10.1242/dev.179093>.
- Brodland, G. Wayne. 2002. “The Differential Interfacial Tension Hypothesis (DITH): A Comprehensive Theory for the Self-Rearrangement of Embryonic Cells and Tissues.” *Journal of Biomechanical Engineering* 124 (2): 188–97. <https://doi.org/10.1115/1.1449491>.
- Bronner, Marianne E. 2016. “How Inhibitory Cues Can Both Constrain and Promote Cell Migration.” *The Journal of Cell Biology* 213 (5): 505–7. <https://doi.org/10.1083/jcb.201605074>.
- Butler, Lucy C., Guy B. Blanchard, Alexandre J. Kabla, Nicola J. Lawrence, David P. Welchman, L. Mahadevan, Richard J. Adams, and Benedicte Sanson. 2009. “Cell Shape Changes Indicate a Role for Extrinsic Tensile Forces in Drosophila Germ-Band Extension.” *Nature Cell Biology* 11 (7): 859–64. <https://doi.org/10.1038/ncb1894>.

- Butts, Jessica C., Dylan A. McCreedy, Jorge Alexis Martinez-Vargas, Frederico N. Mendoza-Camacho, Tracy A. Hookway, Casey A. Gifford, Praveen Taneja, Linda Noble-Haeusslein, and Todd C. McDevitt. 2017. "Differentiation of V2a Interneurons from Human Pluripotent Stem Cells." *Proceedings of the National Academy of Sciences* 114 (19): 4969–74. <https://doi.org/10.1073/pnas.1608254114>.
- C.A. Burdsal, C.H. Damsky, and R.A. Pedersen. 1993. "The Role of E-Cadherin and Integrins in Mesoderm Differentiation and Migration at the Mammalian Primitive Streak | Development." *Development*. July 1, 1993. <http://dev.biologists.org/content/118/3/829.abstract>.
- Cano, Amparo, Mirna A. Pérez-Moreno, Isabel Rodrigo, Annamaria Locascio, María J. Blanco, Marta G. del Barrio, Francisco Portillo, and M. Angela Nieto. 2000. "The Transcription Factor Snail Controls Epithelial–Mesenchymal Transitions by Repressing E-Cadherin Expression." *Nature Cell Biology* 2 (2): 76–83. <https://doi.org/10.1038/35000025>.
- Carpenido, Richard L., Carolyn Y. Sargent, and Todd C. McDevitt. 2007. "Rotary Suspension Culture Enhances the Efficiency, Yield, and Homogeneity of Embryoid Body Differentiation." *STEM CELLS* 25 (9): 2224–34. <https://doi.org/10.1634/stemcells.2006-0523>.
- Carpenter, Ellen M. 2002. "Hox Genes and Spinal Cord Development." *Developmental Neuroscience* 24 (1): 24–34. <https://doi.org/10.1159/000064943>.
- Chandra Ravi A., Douglas Erik S., Mathies Richard A., Bertozzi Carolyn R., and Francis Matthew B. 2005. "Programmable Cell Adhesion Encoded by DNA Hybridization." *Angewandte Chemie* 118 (6): 910–15. <https://doi.org/10.1002/ange.200502421>.
- Clevers, Hans. 2016. "Modeling Development and Disease with Organoids." *Cell* 165 (7): 1586–97. <https://doi.org/10.1016/j.cell.2016.05.082>.

- Coenen, Frans, and Paul Leng. 2007. "The Effect of Threshold Values on Association Rule Based Classification Accuracy." *Data & Knowledge Engineering, Web data and schema management*, 60 (2): 345–60. <https://doi.org/10.1016/j.datak.2006.02.005>.
- Cohen, William W. 1995. "Fast Effective Rule Induction." In *Machine Learning Proceedings 1995*, 115–123. Elsevier.
- Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *Science* 339 (6121): 819–23. <https://doi.org/10.1126/science.1231143>.
- Corral, Ruth Diez del, Isabel Olivera-Martinez, Anne Goriely, Emily Gale, Malcolm Maden, and Kate Storey. 2003. "Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension." *Neuron* 40 (1): 65–79. [https://doi.org/10.1016/S0896-6273\(03\)00565-8](https://doi.org/10.1016/S0896-6273(03)00565-8).
- Corral, Ruth Diez del, and Kate G. Storey. 2004. "Opposing FGF and Retinoid Pathways: A Signalling Switch That Controls Differentiation and Patterning Onset in the Extending Vertebrate Body Axis." *BioEssays* 26 (8): 857–69. <https://doi.org/10.1002/bies.20080>.
- Dale, C. Trevor. 1998. "Signal Transduction by the Wnt Family of Ligands." *Biochemical Journal* 329 (2): 209–23. <https://doi.org/10.1042/bj3290209>.
- Dalton, Stephen. 2013. "Signaling Networks in Human Pluripotent Stem Cells." *Current Opinion in Cell Biology* 25 (2): 241–46. <https://doi.org/10.1016/j.ceb.2012.09.005>.
- Dawes-Hoang, Rachel E., Kush M. Parmar, Audrey E. Christiansen, Chris B. Phelps, Andrea H. Brand, and Eric F. Wieschaus. 2005. "Folded Gastrulation, Cell Shape Change and the Control of Myosin Localization." *Development* 132 (18): 4165–78. <https://doi.org/10.1242/dev.01938>.
- Ducibella, Thomas, and Everett Anderson. 1975. "Cell Shape and Membrane Changes in the Eight-Cell Mouse Embryo: Prerequisites for Morphogenesis of the Blastocyst." *Developmental Biology* 47 (1): 45–58. [https://doi.org/10.1016/0012-1606\(75\)90262-6](https://doi.org/10.1016/0012-1606(75)90262-6).

- Dupont, Sirio, Leonardo Morsut, Mariaceleste Aragona, Elena Enzo, Stefano Giulitti, Michelangelo Cordenonsi, Francesca Zanconato, et al. 2011. "Role of YAP/TAZ in Mechanotransduction." *Nature* 474 (7350): 179–83. <https://doi.org/10.1038/nature10137>.
- Eiraku, Mototsugu, Nozomu Takata, Hiroki Ishibashi, Masako Kawada, Eriko Sakakura, Satoru Okuda, Kiyotoshi Sekiguchi, Taiji Adachi, and Yoshiki Sasai. 2011. "Self-Organizing Optic-Cup Morphogenesis in Three-Dimensional Culture." *Nature* 472 (7341): 51. <https://doi.org/10.1038/nature09941>.
- Eiraku, Mototsugu, Kiichi Watanabe, Mami Matsuo-Takasaki, Masako Kawada, Shigenobu Yonemura, Michiru Matsumura, Takafumi Wataya, Ayaka Nishiyama, Keiko Mugeruma, and Yoshiki Sasai. 2008. "Self-Organized Formation of Polarized Cortical Tissues from ESCs and Its Active Manipulation by Extrinsic Signals." *Cell Stem Cell* 3 (5): 519–32. <https://doi.org/10.1016/j.stem.2008.09.002>.
- Ericson, J., J. Briscoe, P. Rashbass, V. van Heyningen, and T. M. Jessell. 1997. "Graded Sonic Hedgehog Signaling and the Specification of Cell Fate in the Ventral Neural Tube." *Cold Spring Harbor Symposia on Quantitative Biology* 62 (January): 451–66. <https://doi.org/10.1101/SQB.1997.062.01.053>.
- Essen, David C. Van. 1997. "A Tension-Based Theory of Morphogenesis and Compact Wiring in the Central Nervous System." *Nature* 385 (6614): 313–18. <https://doi.org/10.1038/385313a0>.
- Finkel, R. A., and J. L. Bentley. 1974. "Quad Trees a Data Structure for Retrieval on Composite Keys." *Acta Informatica* 4 (1): 1–9. <https://doi.org/10.1007/BF00288933>.
- Fletcher, Alexander G., Miriam Osterfield, Ruth E. Baker, and Stanislav Y. Shvartsman. 2014. "Vertex Models of Epithelial Morphogenesis." *Biophysical Journal* 106 (11): 2291–2304. <https://doi.org/10.1016/j.bpj.2013.11.4498>.

- Foty, Ramsey A., and Malcolm S. Steinberg. 2004. "Cadherin-Mediated Cell-Cell Adhesion and Tissue Segregation in Relation to Malignancy." *International Journal of Developmental Biology* 48 (5–6): 397–409. <https://doi.org/10.1387/ijdb.041810rf>.
- . 2005. "The Differential Adhesion Hypothesis: A Direct Evaluation." *Developmental Biology* 278 (1): 255–63. <https://doi.org/10.1016/j.ydbio.2004.11.012>.
- Gilbert, Luke A., Max A. Horlbeck, Britt Adamson, Jacqueline E. Villalta, Yuwen Chen, Evan H. Whitehead, Carla Guimaraes, et al. 2014. "Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation." *Cell* 159 (3): 647–61. <https://doi.org/10.1016/j.cell.2014.09.029>.
- Glen, Chad M., Melissa L. Kemp, and Eberhard O. Voit. 2019. "Agent-Based Modeling of Morphogenetic Systems: Advantages and Challenges." *PLOS Computational Biology* 15 (3): e1006577. <https://doi.org/10.1371/journal.pcbi.1006577>.
- Gottardi, C. J., E. Wong, and B. M. Gumbiner. 2001. "E-Cadherin Suppresses Cellular Transformation by Inhibiting Beta-Catenin Signaling in an Adhesion-Independent Manner." *The Journal of Cell Biology* 153 (5): 1049–60.
- Gouti, Mina, Anestis Tsakiridis, Filip J. Wymeersch, Yali Huang, Jens Kleinjung, Valerie Wilson, and James Briscoe. 2014. "In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity." *PLOS Biology* 12 (8): e1001937. <https://doi.org/10.1371/journal.pbio.1001937>.
- Graner, François, and James A. Glazier. 1992. "Simulation of Biological Cell Sorting Using a Two-Dimensional Extended Potts Model." *Physical Review Letters* 69 (13): 2013–16. <https://doi.org/10.1103/PhysRevLett.69.2013>.
- Güdükbay, Uğur, Bülent Özgüç, and Yılmaz Tokad. 1997. "A Spring Force Formulation for Elastically Deformable Models." *Computers & Graphics, Computer Graphics in China*, 21 (3): 335–46. [https://doi.org/10.1016/S0097-8493\(97\)00011-3](https://doi.org/10.1016/S0097-8493(97)00011-3).

- Gueth-Hallonet, Catherine, and Bernard Maro. 1992. "Cell Polarity and Cell Diversification during Early Mouse Embryogenesis." *Trends in Genetics* 8 (8): 274–79. [https://doi.org/10.1016/0168-9525\(92\)90253-Z](https://doi.org/10.1016/0168-9525(92)90253-Z).
- Gusev, Andrei A. 2004. "Finite Element Mapping for Spring Network Representations of the Mechanics of Solids." *Physical Review Letters* 93 (3): 034302. <https://doi.org/10.1103/PhysRevLett.93.034302>.
- Haghighi, Iman, Austin Jones, Zhaodan Kong, Ezio Bartocci, Radu Gros, and Calin Belta. 2015. "SpaTeL: A Novel Spatial-Temporal Logic and Its Applications to Networked Systems." In *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, 189–198. HSCC '15. New York, NY, USA: ACM. <https://doi.org/10.1145/2728606.2728633>.
- Heallen, Todd, Min Zhang, Jun Wang, Margarita Bonilla-Claudio, Ela Klysik, Randy L. Johnson, and James F. Martin. 2011. "Hippo Pathway Inhibits Wnt Signaling to Restrain Cardiomyocyte Proliferation and Heart Size." *Science* 332 (6028): 458–61. <https://doi.org/10.1126/science.1199010>.
- Heasman, Janet, Aaron Crawford, Kim Goldstone, Peggy Garner-Hamrick, Barry Gumbiner, Pierre McCrea, Chris Kintner, Chikako Yoshida Noro, and Chris Wylie. 1994. "Overexpression of Cadherins and Underexpression of β -Catenin Inhibit Dorsal Mesoderm Induction in Early *Xenopus* Embryos." *Cell* 79 (5): 791–803. [https://doi.org/10.1016/0092-8674\(94\)90069-8](https://doi.org/10.1016/0092-8674(94)90069-8).
- Heisenberg, Carl-Philipp. 2017. "D'Arcy Thompson's 'on Growth and Form': From Soap Bubbles to Tissue Self-Organization." *Mechanisms of Development*, On Growth and Form at 100, 145 (June): 32–37. <https://doi.org/10.1016/j.mod.2017.03.006>.
- Heisenberg, Carl-Philipp, and Yohanns Bellaïche. 2013. "Forces in Tissue Morphogenesis and Patterning." *Cell* 153 (5): 948–62. <https://doi.org/10.1016/j.cell.2013.05.008>.

- Henn, Arnon, and Enrique M. De La Cruz. 2005. "Vertebrate Myosin VIIIb Is a High Duty Ratio Motor Adapted for Generating and Maintaining Tension." *Journal of Biological Chemistry* 280 (47): 39665–76. <https://doi.org/10.1074/jbc.M507667200>.
- Henrique, Domingos, Elsa Abranches, Laure Verrier, and Kate G. Storey. 2015. "Neuromesodermal Progenitors and the Making of the Spinal Cord." *Development* 142 (17): 2864–75. <https://doi.org/10.1242/dev.119768>.
- Heuberger, Julian, and Walter Birchmeier. 2010. "Interplay of Cadherin-Mediated Cell Adhesion and Canonical Wnt Signaling." *Cold Spring Harbor Perspectives in Biology* 2 (2): a002915. <https://doi.org/10.1101/cshperspect.a002915>.
- Hirate, Yoshikazu, Katie Cockburn, Janet Rossant, and Hiroshi Sasaki. 2012. "Tead4 Is Constitutively Nuclear, While Nuclear vs. Cytoplasmic Yap Distribution Is Regulated in Preimplantation Mouse Embryos." *Proceedings of the National Academy of Sciences* 109 (50): E3389–90. <https://doi.org/10.1073/pnas.1211810109>.
- Hirate, Yoshikazu, Shino Hirahara, Ken-ichi Inoue, Atsushi Suzuki, Vernadeth B. Alarcon, Kazunori Akimoto, Takaaki Hirai, et al. 2013. "Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos." *Current Biology* 23 (13): 1181–94. <https://doi.org/10.1016/j.cub.2013.05.014>.
- Hočevar Brezavšček, Ana, Matteo Rauzi, Maria Leptin, and Primož Ziherl. 2012. "A Model of Epithelial Invagination Driven by Collective Mechanics of Identical Cells." *Biophysical Journal* 103 (5): 1069–77. <https://doi.org/10.1016/j.bpj.2012.07.018>.
- Hsiao, Sonny C., Betty J. Shum, Hiroaki Onoe, Erik S. Douglas, Zev J. Gartner, Richard A. Mathies, Carolyn R. Bertozzi, and Matthew B. Francis. 2009. "Direct Cell Surface Modification with DNA for the Capture of Primary Cells and the Investigation of Myotube Formation on Defined Patterns." *Langmuir* 25 (12): 6985–91. <https://doi.org/10.1021/la900150n>.

- Huber, Andrew H., Daniel B. Stewart, Douglas V. Laurents, W. James Nelson, and William I. Weis. 2001. "The Cadherin Cytoplasmic Domain Is Unstructured in the Absence of β -Catenin A POSSIBLE MECHANISM FOR REGULATING CADHERIN TURNOVER." *Journal of Biological Chemistry* 276 (15): 12301–9. <https://doi.org/10.1074/jbc.M010377200>.
- Jackins, C. L., and S. L. Tanimoto. 1983. "Quad-Trees, Oct-Trees, and K-Trees: A Generalized Approach to Recursive Decomposition of Euclidean Space." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-5 (5): 533–39. <https://doi.org/10.1109/TPAMI.1983.4767433>.
- James, Daylon, Ariel J. Levine, Daniel Besser, and Ali Hemmati-Brivanlou. 2005. "TGF β /Activin/Nodal Signaling Is Necessary for the Maintenance of Pluripotency in Human Embryonic Stem Cells." *Development* 132 (6): 1273–82. <https://doi.org/10.1242/dev.01706>.
- Janssens, Sylvie, Tinneke Denayer, Tom Deroo, Frans Van Roy, and Kris Vleminckx. 2010. "Direct Control of Hoxd1 and Irx3 Expression by Wnt/Beta-Catenin Signaling during Anteroposterior Patterning of the Neural Axis in *Xenopus*." *International Journal of Developmental Biology* 54 (10): 1435–42. <https://doi.org/10.1387/ijdb.092985sj>.
- Jinek, Martin, Alexandra East, Aaron Cheng, Steven Lin, Enbo Ma, and Jennifer Doudna. 2013. "RNA-Programmed Genome Editing in Human Cells." Edited by Detlef Weigel. *ELife* 2 (January): e00471. <https://doi.org/10.7554/eLife.00471>.
- Kam, Yoonseok, and Vito Quaranta. 2009. "Cadherin-Bound β -Catenin Feeds into the Wnt Pathway upon Adherens Junctions Dissociation: Evidence for an Intersection between β -Catenin Pools." *PLOS ONE* 4 (2): e4580. <https://doi.org/10.1371/journal.pone.0004580>.
- Kaplan, Kevin M., Jeffrey M. Spivak, and John A. Bendo. 2005. "Embryology of the Spine and Associated Congenital Abnormalities." *The Spine Journal* 5 (5): 564–76. <https://doi.org/10.1016/j.spinee.2004.10.044>.

- Kattman, Steven J., Alec D. Witty, Mark Gagliardi, Nicole C. Dubois, Maryam Niapour, Akitsu Hotta, James Ellis, and Gordon Keller. 2011. "Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines." *Cell Stem Cell* 8 (2): 228–40. <https://doi.org/10.1016/j.stem.2010.12.008>.
- Kim, Gun-Hwa, and Jin-Kwan Han. 2005. "JNK and ROK α Function in the Noncanonical Wnt/RhoA Signaling Pathway to Regulate *Xenopus* Convergent Extension Movements." *Developmental Dynamics* 232 (4): 958–68. <https://doi.org/10.1002/dvdy.20262>.
- Krieg, M., Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. J. Müller, and C.-P. Heisenberg. 2008. "Tensile Forces Govern Germ-Layer Organization in Zebrafish." *Nature Cell Biology* 10 (4): 429–36. <https://doi.org/10.1038/ncb1705>.
- Kurosawa, Hiroshi. 2007. "Methods for Inducing Embryoid Body Formation: In Vitro Differentiation System of Embryonic Stem Cells." *Journal of Bioscience and Bioengineering* 103 (5): 389–98. <https://doi.org/10.1263/jbb.103.389>.
- Lancaster, Madeline A., and Juergen A. Knoblich. 2014. "Organogenesis in a Dish: Modeling Development and Disease Using Organoid Technologies." *Science* 345 (6194): 1247125. <https://doi.org/10.1126/science.1247125>.
- Larson, Matthew H., Luke A. Gilbert, Xiaowo Wang, Wendell A. Lim, Jonathan S. Weissman, and Lei S. Qi. 2013. "CRISPR Interference (CRISPRi) for Sequence-Specific Control of Gene Expression." *Nature Protocols* 8 (11): 2180–96. <https://doi.org/10.1038/nprot.2013.132>.
- Lechler, Terry, and Elaine Fuchs. 2005. "Asymmetric Cell Divisions Promote Stratification and Differentiation of Mammalian Skin." *Nature* 437 (7056): 275–80. <https://doi.org/10.1038/nature03922>.
- Leckband, D. E., and J. de Rooij. 2014. "Cadherin Adhesion and Mechanotransduction." *Annual Review of Cell and Developmental Biology* 30 (1): 291–315. <https://doi.org/10.1146/annurev-cellbio-100913-013212>.

- Lecuit, Thomas, and Pierre-François Lenne. 2007. "Cell Surface Mechanics and the Control of Cell Shape, Tissue Patterns and Morphogenesis." *Nature Reviews Molecular Cell Biology* 8 (8): nrm2222. <https://doi.org/10.1038/nrm2222>.
- Lecuit, Thomas, and Alpha S. Yap. 2015. "E-Cadherin Junctions as Active Mechanical Integrators in Tissue Dynamics." *Nature Cell Biology* 17 (5): 533–39. <https://doi.org/10.1038/ncb3136>.
- Li, Dong, Jiayi Zhou, Lu Wang, Myung Eun Shin, Pei Su, Xiaohua Lei, Haibin Kuang, et al. 2010. "Integrated Biochemical and Mechanical Signals Regulate Multifaceted Human Embryonic Stem Cell Functions." *The Journal of Cell Biology* 191 (3): 631–44. <https://doi.org/10.1083/jcb.201006094>.
- Lian, Xiaojun, Cheston Hsiao, Gisela Wilson, Kexian Zhu, Laurie B. Hazeltine, Samira M. Azarin, Kunil K. Raval, Jianhua Zhang, Timothy J. Kamp, and Sean P. Palecek. 2012. "Robust Cardiomyocyte Differentiation from Human Pluripotent Stem Cells via Temporal Modulation of Canonical Wnt Signaling." *Proceedings of the National Academy of Sciences* 109 (27): E1848–57. <https://doi.org/10.1073/pnas.1200250109>.
- Liem, Karel F, Gabi Tremml, Henk Roelink, and Thomas M Jessell. 1995. "Dorsal Differentiation of Neural Plate Cells Induced by BMP-Mediated Signals from Epidermal Ectoderm." *Cell* 82 (6): 969–79. [https://doi.org/10.1016/0092-8674\(95\)90276-7](https://doi.org/10.1016/0092-8674(95)90276-7).
- L. MacKay, Joanna, Anshum Sood, and Sanjay Kumar. 2014. "Three-Dimensional Patterning of Multiple Cell Populations through Orthogonal Genetic Control of Cell Motility." *Soft Matter* 10 (14): 2372–80. <https://doi.org/10.1039/C3SM52265B>.
- Maître, Jean-Léon. 2017. "Mechanics of Blastocyst Morphogenesis." *Biology of the Cell* 109 (9): 323–38. <https://doi.org/10.1111/boc.201700029>.
- Maître, Jean-Léon, Hélène Berthoumieux, Simon Frederik Gabriel Krens, Guillaume Salbreux, Frank Jülicher, Ewa Paluch, and Carl-Philipp Heisenberg. 2012. "Adhesion Functions in Cell Sorting by Mechanically Coupling the Cortices of Adhering Cells." *Science* 338 (6104): 253–56. <https://doi.org/10.1126/science.1225399>.

- Mandegar, Mohammad A., Nathaniel Huebsch, Ekaterina B. Frolov, Edward Shin, Annie Truong, Michael P. Olvera, Amanda H. Chan, et al. 2016. "CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs." *Cell Stem Cell* 18 (4): 541–53. <https://doi.org/10.1016/j.stem.2016.01.022>.
- Marcon, Luciano, and James Sharpe. 2012. "Turing Patterns in Development: What about the Horse Part?" *Current Opinion in Genetics & Development*, Genetics of system biology, 22 (6): 578–84. <https://doi.org/10.1016/j.gde.2012.11.013>.
- Marée, Athanasius F. M., Verônica A. Grieneisen, and Paulien Hogeweg. 2007. "The Cellular Potts Model and Biophysical Properties of Cells, Tissues and Morphogenesis." In *Single-Cell-Based Models in Biology and Medicine*, edited by Alexander R. A. Anderson, Mark A. J. Chaplain, and Katarzyna A. Rejniak, 107–136. Basel: Birkhäuser Basel. https://doi.org/10.1007/978-3-7643-8123-3_5.
- Matthews, Helen K., Lorena Marchant, Carlos Carmona-Fontaine, Sei Kuriyama, Juan Larraín, Mark R. Holt, Maddy Parsons, and Roberto Mayor. 2008. "Directional Migration of Neural Crest Cells in Vivo Is Regulated by Syndecan-4/Rac1 and Non-Canonical Wnt Signaling/RhoA." *Development* 135 (10): 1771–80. <https://doi.org/10.1242/dev.017350>.
- Matthys, Oriane B., Tracy A. Hookway, and Todd C. McDevitt. 2016. "Design Principles for Engineering of Tissues from Human Pluripotent Stem Cells." *Current Stem Cell Reports* 2 (1): 43–51. <https://doi.org/10.1007/s40778-016-0030-z>.
- McBeath, Rowena, Dana M Pirone, Celeste M Nelson, Kiran Bhadriraju, and Christopher S Chen. 2004. "Cell Shape, Cytoskeletal Tension, and RhoA Regulate Stem Cell Lineage Commitment." *Developmental Cell* 6 (4): 483–95. [https://doi.org/10.1016/S1534-5807\(04\)00075-9](https://doi.org/10.1016/S1534-5807(04)00075-9).
- McMahon, Jill A., Shinji Takada, Lyle B. Zimmerman, Chen-Ming Fan, Richard M. Harland, and Andrew P. McMahon. 1998. "Noggin-Mediated Antagonism of BMP Signaling Is Required

- for Growth and Patterning of the Neural Tube and Somite.” *Genes & Development* 12 (10): 1438–52. <https://doi.org/10.1101/gad.12.10.1438>.
- Mihajlović, Aleksandar I., and Alexander W. Bruce. 2016. “Rho-Associated Protein Kinase Regulates Subcellular Localisation of Angiotensin and Hippo-Signalling during Preimplantation Mouse Embryo Development.” *Reproductive BioMedicine Online* 33 (3): 381–90. <https://doi.org/10.1016/j.rbmo.2016.06.028>.
- Miroshnikova, Yekaterina A., Huy Q. Le, David Schneider, Torsten Thalheim, Matthias Rübsam, Nadine Bremicker, Julien Polleux, et al. 2018. “Adhesion Forces and Cortical Tension Couple Cell Proliferation and Differentiation to Drive Epidermal Stratification.” *Nature Cell Biology* 20 (1): 69–80. <https://doi.org/10.1038/s41556-017-0005-z>.
- M. Molitoris, Jared, Saurabh Paliwal, Rajesh B. Sekar, Robert Blake, JinSeok Park, Natalia A. Trayanova, Leslie Tung, and Andre Levchenko. 2016. “Precisely Parameterized Experimental and Computational Models of Tissue Organization.” *Integrative Biology* 8 (2): 230–42. <https://doi.org/10.1039/C5IB00270B>.
- Montero, Juan-Antonio, and Carl-Philipp Heisenberg. 2004. “Gastrulation Dynamics: Cells Move into Focus.” *Trends in Cell Biology* 14 (11): 620–27. <https://doi.org/10.1016/j.tcb.2004.09.008>.
- Munemitsu, S., I. Albert, B. Souza, B. Rubinfeld, and P. Polakis. 1995. “Regulation of Intracellular Beta-Catenin Levels by the Adenomatous Polyposis Coli (APC) Tumor-Suppressor Protein.” *Proceedings of the National Academy of Sciences* 92 (7): 3046–50. <https://doi.org/10.1073/pnas.92.7.3046>.
- Nielsen, Alec A. K., Bryan S. Der, Jonghyeon Shin, Prashant Vaidyanathan, Vanya Paralanov, Elizabeth A. Strychalski, David Ross, Douglas Densmore, and Christopher A. Voigt. 2016. “Genetic Circuit Design Automation.” *Science* 352 (6281): aac7341. <https://doi.org/10.1126/science.aac7341>.

- Nieto, M. A., M. G. Sargent, D. G. Wilkinson, and J. Cooke. 1994. "Control of Cell Behavior during Vertebrate Development by Slug, a Zinc Finger Gene." *Science* 264 (5160): 835–39. <https://doi.org/10.1126/science.7513443>.
- Nishioka, Noriyuki, Ken-ichi Inoue, Kenjiro Adachi, Hiroshi Kiyonari, Mitsunori Ota, Amy Ralston, Norikazu Yabuta, et al. 2009. "The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass." *Developmental Cell* 16 (3): 398–410. <https://doi.org/10.1016/j.devcel.2009.02.003>.
- Nishioka, Noriyuki, Shinji Yamamoto, Hiroshi Kiyonari, Hiroko Sato, Atsushi Sawada, Mitsunori Ota, Kazuki Nakao, and Hiroshi Sasaki. 2008. "Tead4 Is Required for Specification of Trophectoderm in Pre-Implantation Mouse Embryos." *Mechanisms of Development* 125 (3–4): 270–83. <https://doi.org/10.1016/j.mod.2007.11.002>.
- Niwa, Hitoshi, Yayoi Toyooka, Daisuke Shimosato, Dan Strumpf, Kadue Takahashi, Rika Yagi, and Janet Rossant. 2005. "Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation." *Cell* 123 (5): 917–29. <https://doi.org/10.1016/j.cell.2005.08.040>.
- Ohgushi, Masatoshi, Maki Minaguchi, and Yoshiki Sasai. 2015. "Rho-Signaling-Directed YAP/TAZ Activity Underlies the Long-Term Survival and Expansion of Human Embryonic Stem Cells." *Cell Stem Cell* 17 (4): 448–61. <https://doi.org/10.1016/j.stem.2015.07.009>.
- Ou, Guangshuo, Nico Stuurman, Michael D'Ambrosio, and Ronald D. Vale. 2010. "Polarized Myosin Produces Unequal-Size Daughters During Asymmetric Cell Division." *Science* 330 (6004): 677–80. <https://doi.org/10.1126/science.1196112>.
- Przybyla, Laralynne, Johnathon N. Lakins, and Valerie M. Weaver. 2016. "Tissue Mechanics Orchestrate Wnt-Dependent Human Embryonic Stem Cell Differentiation." *Cell Stem Cell* 19 (4): 462–75. <https://doi.org/10.1016/j.stem.2016.06.018>.
- Qi, Lei S., Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin, and Wendell A. Lim. 2013. "Repurposing CRISPR as an RNA-Guided

- Platform for Sequence-Specific Control of Gene Expression." *Cell* 152 (5): 1173–83.
<https://doi.org/10.1016/j.cell.2013.02.022>.
- Radice, Glenn L., Helen Rayburn, Hiroaki Matsunami, Karen A. Knudsen, Masatoshi Takeichi, and Richard O. Hynes. 1997. "Developmental Defects in Mouse Embryos Lacking N-Cadherin." *Developmental Biology* 181 (1): 64–78.
<https://doi.org/10.1006/dbio.1996.8443>.
- Ralston, Amy, Brian J. Cox, Noriyuki Nishioka, Hiroshi Sasaki, Evelyn Chea, Peter Rugg-Gunn, Guoji Guo, Paul Robson, Jonathan S. Draper, and Janet Rossant. 2010. "Gata3 Regulates Trophoblast Development Downstream of Tead4 and in Parallel to Cdx2." *Development* 137 (3): 395–403. <https://doi.org/10.1242/dev.038828>.
- Ringwald, M, R Schuh, D Vestweber, H Eistetter, F Lottspeich, J Engel, R Dölz, F Jähnig, J Epplen, and S Mayer. 1987. "The Structure of Cell Adhesion Molecule Uvomorulin. Insights into the Molecular Mechanism of Ca²⁺-Dependent Cell Adhesion." *The EMBO Journal* 6 (12): 3647–53.
- Rolo, Ana, Paul Skoglund, and Ray Keller. 2009. "Morphogenetic Movements Driving Neural Tube Closure in *Xenopus* Require Myosin IIB." *Developmental Biology* 327 (2): 327–38.
<https://doi.org/10.1016/j.ydbio.2008.12.009>.
- Saha, Somen, Lin Ji, Juan J. de Pablo, and Sean P. Palecek. 2006. "Inhibition of Human Embryonic Stem Cell Differentiation by Mechanical Strain." *Journal of Cellular Physiology* 206 (1): 126–37. <https://doi.org/10.1002/jcp.20441>.
- . 2008. "TGFβ/Activin/Nodal Pathway in Inhibition of Human Embryonic Stem Cell Differentiation by Mechanical Strain." *Biophysical Journal* 94 (10): 4123–33.
<https://doi.org/10.1529/biophysj.107.119891>.
- Samarage, Chaminda R., Melanie D. White, Yanina D. Álvarez, Juan Carlos Fierro-González, Yann Henon, Edwin C. Jesudason, Stephanie Bissiere, Andreas Fouras, and Nicolas

- Plachta. 2015. "Cortical Tension Allocates the First Inner Cells of the Mammalian Embryo." *Developmental Cell* 34 (4): 435–47. <https://doi.org/10.1016/j.devcel.2015.07.004>.
- Sasai, Yoshiki. 2013. "Cytosystems Dynamics in Self-Organization of Tissue Architecture." *Nature* 493 (7432): 318–26. <https://doi.org/10.1038/nature11859>.
- Schiffmann, Yoram. 2006. "Symmetry Breaking and Convergent Extension in Early Chordate Development." *Progress in Biophysics and Molecular Biology* 92 (2): 209–31. <https://doi.org/10.1016/j.pbiomolbio.2005.10.002>.
- Scianna, Marco, Luigi Preziosi, and Katarina Wolf. 2012. "A Cellular Potts Model Simulating Cell Migration on and in Matrix Environments." *Mathematical Biosciences & Engineering* 10 (1): 235. <https://doi.org/10.3934/mbe.2013.10.235>.
- Sharpe, James. 2017. "Computer Modeling in Developmental Biology: Growing Today, Essential Tomorrow." *Development* 144 (23): 4214–25. <https://doi.org/10.1242/dev.151274>.
- Sliogeryte, Kristina, Stephen D. Thorpe, David A. Lee, Lorenzo Botto, and Martin M. Knight. 2014. "Stem Cell Differentiation Increases Membrane-Actin Adhesion Regulating Cell Blebability, Migration and Mechanics." *Scientific Reports* 4 (December): 7307. <https://doi.org/10.1038/srep07307>.
- Smith, Andrew L., Michael R. Dohn, Meredith V. Brown, and Albert B. Reynolds. 2012. "Association of Rho-Associated Protein Kinase 1 with E-Cadherin Complexes Is Mediated by P120-Catenin." *Molecular Biology of the Cell* 23 (1): 99–110. <https://doi.org/10.1091/mbc.E11-06-0497>.
- Smith, Jodi L., and Gary C. Schoenwolf. 1989. "Notochordal Induction of Cell Wedging in the Chick Neural Plate and Its Role in Neural Tube Formation." *Journal of Experimental Zoology* 250 (1): 49–62. <https://doi.org/10.1002/jez.1402500107>.
- Somlyo, Andrew P., and Avril V. Somlyo. 2000. "Signal Transduction by G-Proteins, Rho-Kinase and Protein Phosphatase to Smooth Muscle and Non-Muscle Myosin II." *The Journal of Physiology* 522 (2): 177–85. <https://doi.org/10.1111/j.1469-7793.2000.t01-2-00177.x>.

- Spence, Jason R., Christopher N. Mayhew, Scott A. Rankin, Matthew F. Kuhar, Jefferson E. Vallance, Kathryn Tolle, Elizabeth E. Hoskins, et al. 2010. "Directed Differentiation of Human Pluripotent Stem Cells into Intestinal Tissue *in Vitro*." *Nature* 470 (7332): 105. <https://doi.org/10.1038/nature09691>.
- Srivastava, Deepak. 2000. "Congenital Heart Defects." *Circulation Research* 86 (9): 917–18. <https://doi.org/10.1161/01.RES.86.9.917>.
- Steinberg, Malcolm S, and Patrick M McNutt. 1999. "Cadherins and Their Connections: Adhesion Junctions Have Broader Functions." *Current Opinion in Cell Biology* 11 (5): 554–60. [https://doi.org/10.1016/S0955-0674\(99\)00027-7](https://doi.org/10.1016/S0955-0674(99)00027-7).
- Steventon, Ben, Fernando Duarte, Ronan Lagadec, Sylvie Mazan, Jean-François Nicolas, and Estelle Hirsinger. 2016. "Species-Specific Contribution of Volumetric Growth and Tissue Convergence to Posterior Body Elongation in Vertebrates." *Development* 143 (10): 1732–41. <https://doi.org/10.1242/dev.126375>.
- Strumpf, Dan, Chai-An Mao, Yojiro Yamanaka, Amy Ralston, Kallayanee Chawengsaksophak, Felix Beck, and Janet Rossant. 2005. "Cdx2 Is Required for Correct Cell Fate Specification and Differentiation of Trophectoderm in the Mouse Blastocyst." *Development* 132 (9): 2093–2102. <https://doi.org/10.1242/dev.01801>.
- Szabó, András, and Rmh Merks. 2013. "Cellular Potts Modeling of Tumor Growth, Tumor Invasion, and Tumor Evolution." *Frontiers in Oncology* 3 (April): 87. <https://doi.org/10.3389/fonc.2013.00087>.
- Tada, Masazumi, Miguel L Concha, and Carl-Philipp Heisenberg. 2002. "Non-Canonical Wnt Signalling and Regulation of Gastrulation Movements." *Seminars in Cell & Developmental Biology* 13 (3): 251–60. [https://doi.org/10.1016/S1084-9521\(02\)00052-6](https://doi.org/10.1016/S1084-9521(02)00052-6).
- Tahinci, Emilios, and Karen Symes. 2003. "Distinct Functions of Rho and Rac Are Required for Convergent Extension during *Xenopus* Gastrulation." *Developmental Biology* 259 (2): 318–35. [https://doi.org/10.1016/S0012-1606\(03\)00206-9](https://doi.org/10.1016/S0012-1606(03)00206-9).

- Takahashi, Kazutoshi, Koji Tanabe, Mari Ohnuki, Megumi Narita, Tomoko Ichisaka, Kiichiro Tomoda, and Shinya Yamanaka. 2007. "Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors." *Cell* 131 (5): 861–72. <https://doi.org/10.1016/j.cell.2007.11.019>.
- Takashima, Yasuhiro, Ge Guo, Remco Loos, Jennifer Nichols, Gabriella Ficz, Felix Krueger, David Oxley, et al. 2014. "Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human." *Cell* 158 (6): 1254–69. <https://doi.org/10.1016/j.cell.2014.08.029>.
- Tam, Patrick P. L, and Richard R Behringer. 1997. "Mouse Gastrulation: The Formation of a Mammalian Body Plan." *Mechanisms of Development* 68 (1): 3–25. [https://doi.org/10.1016/S0925-4773\(97\)00123-8](https://doi.org/10.1016/S0925-4773(97)00123-8).
- Tewary, Mukul, Joel Ostblom, Laura Prochazka, Teresa Zulueta-Coarasa, Nika Shakiba, Rodrigo Fernandez-Gonzalez, and Peter W. Zandstra. 2017. "A Stepwise Model of Reaction-Diffusion and Positional-Information Governs Self-Organized Human Peri-Gastrulation-like Patterning." *Development*, January, dev.149658. <https://doi.org/10.1242/dev.149658>.
- Thorne, Bryan C., Alexander M. Bailey, Douglas W. DeSimone, and Shayn M. Peirce. 2007. "Agent-Based Modeling of Multicell Morphogenic Processes during Development." *Birth Defects Research Part C: Embryo Today: Reviews* 81 (4): 344–53. <https://doi.org/10.1002/bdrc.20106>.
- Toda, Satoshi, Lucas R. Blauch, Sindy K. Y. Tang, Leonardo Morsut, and Wendell A. Lim. 2018. "Programming Self-Organizing Multicellular Structures with Synthetic Cell-Cell Signaling." *Science*, May, eaat0271. <https://doi.org/10.1126/science.aat0271>.
- Trounson, Alan O. 2001. "The Derivation and Potential Use of Human Embryonic Stem Cells." *Reproduction, Fertility and Development* 13 (8): 523–32. <https://doi.org/10.1071/rd01101>.

- Turing, A. M. 1952. "The Chemical Basis of Morphogenesis." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237 (1): 37–72. <https://doi.org/10.1007/BF02459572>.
- Van Liedekerke, P., M. M. Palm, N. Jagiella, and D. Drasdo. 2015. "Simulating Tissue Mechanics with Agent-Based Models: Concepts, Perspectives and Some Novel Results." *Computational Particle Mechanics* 2 (4): 401–44. <https://doi.org/10.1007/s40571-015-0082-3>.
- Vicente-Manzanares, Miguel, Xuefei Ma, Robert S. Adelstein, and Alan Rick Horwitz. 2009. "Non-Muscle Myosin II Takes Centre Stage in Cell Adhesion and Migration." *Nature Reviews Molecular Cell Biology* 10 (11): 778–90. <https://doi.org/10.1038/nrm2786>.
- Viebahn, Christoph. 1999. "3 The Anterior Margin of the Mammalian Gastrula: Comparative and Phylogenetic Aspects of Its Role in Axis Formation and Head Induction." In *Current Topics in Developmental Biology*, edited by Roger A. Pedersen and Gerald P. Schatten, 46:63–103. Academic Press. [https://doi.org/10.1016/S0070-2153\(08\)60326-9](https://doi.org/10.1016/S0070-2153(08)60326-9).
- Voss-Böhme, Anja. 2012. "Multi-Scale Modeling in Morphogenesis: A Critical Analysis of the Cellular Potts Model." *PLOS ONE* 7 (9): e42852. <https://doi.org/10.1371/journal.pone.0042852>.
- Warmflash, Aryeh, Benoit Sorre, Fred Etoc, Eric D. Siggia, and Ali H. Brivanlou. 2014. "A Method to Recapitulate Early Embryonic Spatial Patterning in Human Embryonic Stem Cells." *Nature Methods* 11 (8): 847–54. <https://doi.org/10.1038/nmeth.3016>.
- Wilson, Valerie, Isabel Olivera-Martinez, and Kate G. Storey. 2009. "Stem Cells, Signals and Vertebrate Body Axis Extension." *Development* 136 (10): 1591–1604. <https://doi.org/10.1242/dev.021246>.
- Xu, W., H. Baribault, and E. D. Adamson. 1998. "Vinculin Knockout Results in Heart and Brain Defects during Embryonic Development." *Development (Cambridge, England)* 125 (2): 327–37.

- Yagi, Rieko, Matthew J. Kohn, Irina Karavanova, Kotaro J. Kaneko, Detlef Vullhorst, Melvin L. DePamphilis, and Andres Buonanno. 2007. "Transcription Factor TEAD4 Specifies the Trophectoderm Lineage at the Beginning of Mammalian Development." *Development* 134 (21): 3827–36. <https://doi.org/10.1242/dev.010223>.
- Yamaguchi, Terry P. 2001. "Heads or Tails: Wnts and Anterior–Posterior Patterning." *Current Biology* 11 (17): R713–24. [https://doi.org/10.1016/S0960-9822\(01\)00417-1](https://doi.org/10.1016/S0960-9822(01)00417-1).
- Yamanaka, Shinya. 2012. "Induced Pluripotent Stem Cells: Past, Present, and Future." *Cell Stem Cell* 10 (6): 678–84. <https://doi.org/10.1016/j.stem.2012.05.005>.
- Yang, C. N. 1952. "The Spontaneous Magnetization of a Two-Dimensional Ising Model." *Physical Review* 85 (5): 808–16. <https://doi.org/10.1103/PhysRev.85.808>.
- Zhou, Beiyun, Yixin Liu, Michael Kahn, David K. Ann, Arum Han, Hongjun Wang, Cu Nguyen, et al. 2012. "Interactions Between β -Catenin and Transforming Growth Factor- β Signaling Pathways Mediate Epithelial-Mesenchymal Transition and Are Dependent on the Transcriptional Co-Activator CAMP-Response Element-Binding Protein (CREB)-Binding Protein (CBP)." *Journal of Biological Chemistry* 287 (10): 7026–38. <https://doi.org/10.1074/jbc.M111.276311>.
- Zhou, Binhua P., Jiong Deng, Weiya Xia, Jihong Xu, Yan M. Li, Mehmet Gunduz, and Mien-Chie Hung. 2004. "Dual Regulation of Snail by GSK-3 β -Mediated Phosphorylation in Control of Epithelial–Mesenchymal Transition." *Nature Cell Biology* 6 (10): 931–40. <https://doi.org/10.1038/ncb1173>.

Chapter 2: Spatiotemporal mosaic self-patterning of pluripotent stem cells using CRISPR interference

2.1 Introduction

Embryos begin as a collection of relatively heterogeneous cells that transition through stages that involve the emergence of multiple cell populations that then self-organize to form embryonic tissues (Montero & Heisenberg, 2004). This embryonic morphogenesis involves interactions of these emergent asymmetric cell populations to form complex multicellular patterns and structures comprised of distinct cell types. Furthermore, cellular rearrangement coordinates with mechanical changes in cellular cortical tension as well as dynamic changes in cell-cell adhesions (Arboleda-Estudillo, n.d.; Krieg et al., 2008; Lecuit & Lenne, 2007; Schäfer, Narasimha, Vogelsang, & Leptin, 2014). However, current methods to model morphogenic events lack control over cell-type co-emergence and offer little capability to selectively perturb specific cell subpopulations.

In this study, we explored whether mechanical manipulation of human induced pluripotent stem cells (iPSCs) sub-populations results in controllable cell driven self-organization into repeatable patterns. We employed an inducible CRISPR interference (CRISPRi) system in iPSCs (Mandegar et al., 2016) to silence key proteins that regulate cell adhesion (CDH1) and cortical tension (ROCK1). We genetically induced controlled symmetry-breaking events within human iPSC populations by creating mixed populations of human iPSCs with and without the CRISPRi system and then induced mosaic knockdown (KD). Mosaic KD was employed to interrogate how the creation of physical asymmetries in an otherwise homogeneous population leads to multicellular organization and pattern formation. We show that induction of mosaic KD of ROCK1

or CDH1 results in a “bottom-up” cell-driven pattern formation of human iPSC colonies while preserving pluripotency.

2.2 Materials and Methods

2.2.1 Human iPSC line generation and culture

All work with human iPSC lines was approved by the University of California, San Francisco Human Gamete, Embryo and Stem Cell Research (GESCR) Committee. Human iPSC lines were cultured in feeder-free media conditions on growth factor-reduced Matrigel (BD Biosciences) and fed daily with mTeSR™-1 medium (STEMCELL Technologies)(Ludwig et al., 2006). Accutase (STEMCELL Technologies) was used to dissociate iPSCs to single cells during passaging. Cells were passaged at a seeding density of 12,000 cells per cm² and the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10μM; Selleckchem) was added to the media upon passaging to promote survival (Park, Kim, Jung, & Roh, 2015; Watanabe et al., 2007).

The generation of the ROCK1 CRISPRi line was previously created and described by Mandegar et al. (Mandegar et al., 2016). For the generation of the CDH1 CRISPRi lines, five CRISPRi gRNAs were designed to bind within 150bp of the TSS of CDH1 and cloned into the gRNA-CKB vector using BsmBI ligation following the previously described protocol (Mandegar et al., 2016) (Table 2.1). gRNA expression vectors were nucleofected into the CRISPRi-Gen1C human iPSC line from the Conklin Lab using the Human Stem Cell Nucleofector Kit 1 solution with the Amaxa nucleofector 2b device (Lonza). Nucleofected cells were seeded into 3 wells of a 6-well plate (~7400 cell/cm²) in mTeSR™-1 media with Y-27632 (10μM) for 2 days and treated with blasticidin (10μg/ml) for a selection period of 7 days. Surviving colonies were pooled and passaged in mTeSR™-1 with blasticidin and Y-27632 for a single day then

Table 2.1: Table 2.1 displays guide RNA sequences used for CRISPRi knockdown.

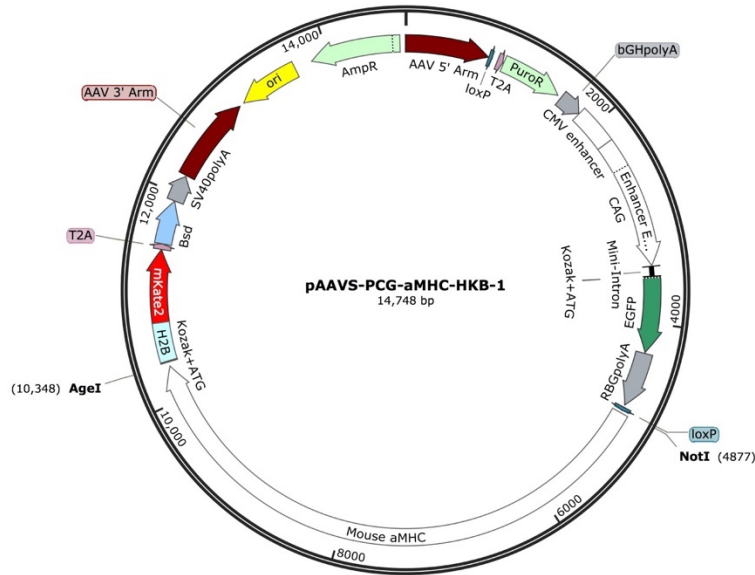
Guide RNA (gRNA) Target	Location to TSS	Sequence
CDH1	-141	TCACCGCGTCTATGCGAGGC
CDH1	-161	CACCCGGCCTCGCATAGACG
CDH1	-68	CCCGTACCGCTGATTGGCTG
CDH1	-46	TCAGCCAATCAGCGGTACGG
CDH1	-6	GCAGTTCCGACGCCACTGAG
ROCK1 (Mandegar et al., 2016)	+11	CGGGGCGCGGACGCTCGGAA
KCNH2 (Off Target Guide)	-175	TTCTGGGCGCGCGAGTCCCA

transitioned to mTeSR™-1 media only. Once stable polyclonal populations of CDH1 CRISPRi iPSCs for each of the five guides were established, the cells were incubated with doxycycline (2µM) for 96 h. KD efficiency was evaluated by mRNA collection and subsequent qPCR, comparing levels of transcript with a time-matched control of the same line without CRISPRi induction. The CRISPRi CDH1 cell line with the guide producing the best KD was selected (gRNA -6).

To generate the WT-GFP line, 2 million WTC clone11 human iPSCs were nucleofected as previously described with the knock-in plasmid containing a CAG promoter-driven EGFP and AAVS1 TALEN pair vectors (Figure 2.1A). After cell recovery, puromycin (0.5 µg/ml) was added to the media for a selection period of 7 days. Individual stable EGFP-expressing colonies were picked using an EVOS FL microscope (Life Technologies) and transferred to individual wells of a 24-well plate in mTeSR media with Y-27632 (10µM) and subsequently expanded into larger vessels.

All cell lines were karyotyped by Cell Line Genetics and were deemed karyotypically normal before proceeding with experiments (Figure 2.1B)

A



B

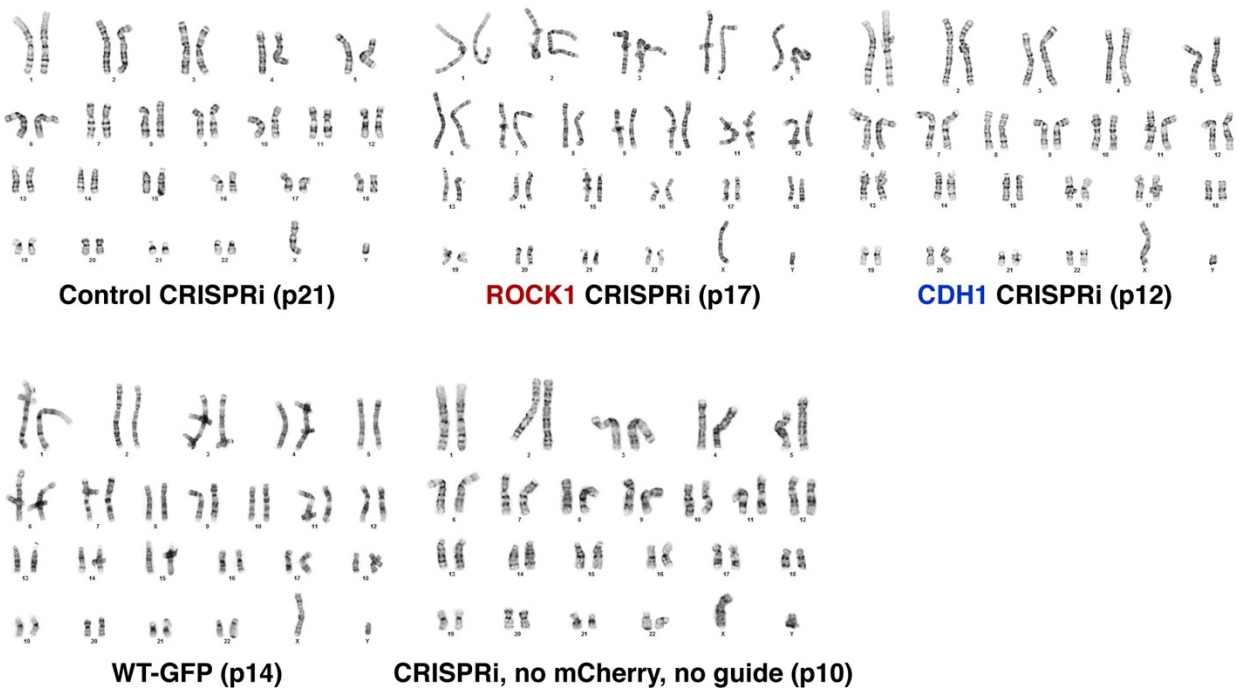


Figure 2.1: Maintenance of normal human iPSC karyotypes. (A) Vector map of constitutive GFP cloned into WT human iPSCs to create the WT-GFP line. **(B)** Karyotypes of all cell lines used in experiments displayed no chromosomal defects. Passage numbers are indicated by p#.

2.2.2 Generation of mixed colonies

Cell aggregates of ~100 cells were created using 400 X 400 µm PDMS microwell inserts in 24-well plates (~975 microwell per well) similar to previously published protocols (Hookway, Butts, Lee, Tang, & McDevitt, 2015; Ungrin, Joshi, Nica, Bauwens, & Zandstra, 2008). Dissociated iPSC cultures were resuspended in mTeSR™-1 supplemented with Y-27632 (10 µM), mixed at proper ratios and concentration (100 cells/well), added to microwells, and centrifuged (200 rcf). After 18 h of formation, 100 cell aggregates were transferred in mTeSR™-1 to Matrigel-coated 96-well plates (~15 aggregates/cm²) and allowed to spread into 2D colonies.

2.2.3 Western blot

Human iPSCs were washed with cold PBS, incubated for 10 min on ice in RIPA Buffer (Sigma-Aldrich), and supernatant collected. Three replicates were used for each condition. The supernatant protein content was determined using a Pierce BCA Protein Assay kit (ThermoFisher Scientific) colorimetric reaction and quantified on a SpectraMax i3 Multi-Mode Platform (Molecular Devices). Subsequently, 20 µg of protein from each sample was resolved by SDS-PAGE and transferred to a nitrocellulose membrane (Invitrogen). The membranes were incubated overnight at 4 °C with primary antibodies: anti-ROCK1 (AbCAM 1:200), anti-CDH1 (AbCAM 1:200), anti-GAPDH, (Invitrogen 1:10,000), followed by incubation (30 min at room temperature) with infrared secondary antibodies: IRDye 800CW and IRDye 680CW (LI-COR 1:13,000), and imaged on the Odyssey Fc Imaging System (LI-COR Biosciences). Protein levels were quantified using Image Studio Lite (LI-COR Biosciences).

2.2.4 RNA isolation and qPCR

Total RNA isolation was performed using a RNeasy Mini Kit (QIAGEN) according to manufacturer's instructions and quantified with a Nanodrop 2000c Spectrometer (ThermoFisher Scientific). cDNA was synthesized by using an iScript cDNA Synthesis kit (BIORAD) and the

reaction was run on a SimpliAmp thermal cycler (Life Technologies). To quantify individual genes, qPCR reactions were run on a StepOnePlus Real-Time PCR system (Applied Biosciences) and detected using Fast SYBR Green Master Mix (ThermoFisher Scientific). Relative gene expression was determined by normalizing to the housekeeping gene 18S rRNA, using the comparative threshold (C_T) method. Gene expression was displayed as fold change of each sample (ROCK1 CRISPRi or CDH1 CRISPRi) versus the off-target guide control (KCNH2 CRISPRi). The primers were designed using the NCBI Primer-BLAST website and are listed in Table 2.2. Statistical analysis was conducted using a two-tailed unpaired *t*-test between any two groups ($p < 0.05$, $n = 3$).

2.2.5 Atomic Force Microscopy

All AFM indentations were performed using an MFP3D-BIO inverted optical atomic force microscope (Asylum Research) mounted on a Nikon TE2000-U inverted fluorescent microscope. Silicon nitride cantilevers were used with spring constants ranging from 0.04 to 0.06 N/m and borosilicate glass spherical tips 5 μm in diameter (Novascan Tech). Each cantilever was calibrated using the thermal oscillation method prior to each experiment. Samples were indented at 1 $\mu\text{m/s}$ loading rate, with a maximum force of 4 nN. Force maps were typically obtained as a 6x6 raster series of indentations utilizing the FMAP function of the IGOR PRO build supplied by Asylum Research, for a total of 36 data points per area of interest measured every 5 microns. Two 5 micron by 5 micron areas of interest were sampled for each sample. The Hertz model was used to determine the elastic modulus of the sample at each point probed. Samples were assumed to be incompressible and a Poisson's ratio of 0.5 was used in the calculation of the Young's elastic modulus.

2.2.6 Time-lapse imaging

Human iPSC colonies were imaged in 96-well plates (ibidi) on an inverted AxioObserver Z1 (Zeiss) with an ORCA-Flash4.0 digital CMOS camera (Hamamatsu). Using ZenPro software,

colony locations were mapped and a single colony was imaged daily for 6 days. To obtain time-lapse movies, a single colony was imaged over the course of 12 h at a rate of one picture taken every 30 min.

2.2.7 Immunofluorescence staining

Human iPSC colonies were fixed for 30 min in 4% paraformaldehyde (VWR) and washed 3X with PBS. Fixed colonies were permeabilized with 0.3% Triton X-100 (Sigma Aldrich) throughout blocking and antibody incubation steps. Samples were incubated in primary antibodies over night at 4 °C, subsequently washed with PBS and incubated in secondary antibodies for an hour at room temperature. Primary antibodies used were: anti-OCT4 (SantaCruz 1:400), anti-SOX2 (AbCAM 1:400), anti-Zo1 (LifeTechnologies 1:400), NANOG (AbCAM 1:300), anti- β -catenin (BD Biosciences 1:200), anti-EpCAM (Millipore 1:200). All secondary antibodies were used at 1:1000 and purchased from Life Technologies.

2.2.8 EDU Incorporation Assay

Pure populations of WT, CRISPRi control, CRISPRi ROCK1, and CRISPRi CDH1 were treated with DOX (2 μ M) for 5 days in series. Cultures were then pulsed with EDU by supplementing Click-It EDU (10 μ M) to the media for 6 hours. Cultures were then washed 3X with PBS and fixed with 4% paraformaldehyde (VWR) for 15 min and subsequently washed with PBS. Samples were then permeabilized in PBS with 0.5% Triton-X 100 (Sigma Aldrich) for 20 minutes. Samples were then incubated with Click-It EDU detection kit as per manufacturer's instructions (ThermoFisher Scientific). Samples were then analyzed via FLOW on a BD LSR-FLOW Cytometer. Analysis was done on a minimum of 10,000 events.

2.2.9 FLOW Analysis

Mixed human iPSC populations and pure population controls were dissociated from tissue culture plates and washed with PBS. Cells were fixed for 15 minutes with 4% paraformaldehyde (VWR) and washed 3X with PBS. Samples were incubated in Hoechst stain for 30 minutes and run on a LSR-II analyzer (BD Biosciences) to detect ratio of WT-GFP(+) to CRISPRi mCherry(+) populations as well as EDU incorporation. Analysis was done on a minimum of 10,000 events.

2.2.10 FACS

Mixed human iPSC populations and pure population controls were dissociated from tissue culture plates and washed 3X with PBS. A LIVE/DEAD stain (ThermoFisher Scientific) was used per manufacture instructions. Prior to sorting, cells were suspended in PBS supplemented with Y-27632(10 μ M) and kept on ice. A BD FACSAria II cell sorter (BD Biosciences) was used to isolate pure populations of WT-GFP and CRISPRi iPSCs by first identifying the live cells via the LIVE/DEAD(350) stain and subsequently sorting the mCherry(+) GFP(-) populations from the mCherry(-)GFP(+) populations directly into TRIzol LS Reagent (ThermoFisher Scientific). Samples were then stored at -80 °C until subsequent mRNA extraction.

2.2.11 Fluidigm 96.96 Array

Sorted human iPSCs stored in TRIzol LS were thawed on ice and mRNA was extracted using a Direct-zol RNA MiniPrep Plus kit (ZYMO Research) following the manufacturer's instructions. RNA was converted to cDNA using the iScript cDNA synthesis kit (Bio-Rad). Forward and reverse primers for genes were designed using NCBI's Primer-BLAST (Sup. Table 3). Primers were pooled to 500 nM to enable specific-target amplification and cDNA was amplified with PreAmp Master Mix (Fluidigm) and pooled primers for 15 cycles. Pre-amplified samples mixed with 2X SsoFast EvaGreen Supermix with low ROX(Bio-Rad) and 20X DNA Binding Dye Sample Loading Reagent (Fluidigm), and 10 μ M primer sets were mixed with 2X Assay Loading

Reagent (Fluidigm). 5 µl of diluted cDNA and primers and were loaded onto the IFC chip per manufacturer's instructions and loaded into the chip using the IFC Controller HX (Fluidigm). qPCR was run for 40 cycles in the IFC chip using the BioMark HD in the BioMark HD Data Collection Software. Resulting data was analyzed in the Real-Time PCR Analysis Software. All instruments and software involved with the IFC chip were manufactured by Fluidigm. Gene expression levels were calculated with respect to time-matched pure populations of WT iPSCs, and hierarchically clustered and plotted using Genesis software (Institute for Genomics and Bioinformatics, Graz University of Technology).

Table 2.2: qPCR primers. Table 2.2 outlines primer sequences used for gene expression analysis by quantitative PCR.

Gene Description	Gene Symbol	Forward Primer	Reverse Primer
18S ribosomal RNA	18S	CTTCCACAGGAGGCCTACAC	CTTCGGCCACACCCTTAAT
Rho associated coiled-coil containing protein kinase 1	ROCK1	GTTCCCCTCCGAGCGTC	TGTCGGCCTTCTGTTCAAA
E-cadherin	CDH1	CGGGAATGCAGTTGAGGATC	AGGATGGTGAAGCGATGGC
POU domain, class 5, transcription factor 1	POU5F1	ATGCATTCAAACCTGAGGTGCCT	AACTTCACCTTCCCTCCAACCA
SRY box 2	SOX2	TCAGGAGTTGTCAAGGCAGAG	GCCGCCGCCGATGATTTGTTAT
Homeobox protein nanog	NANOG	CAATGGTGTGACGCAGGGAT	TGCACCAGGTCTGAGTGTTTC
Brachyury	BRA (T)	TTTCCAGATGGTGAGAGCCG	CCGATGCCTCAAACCTCCAG
SRY box 9	SOX9	GAGCGAGCGGTGCATTTG	TGGTGTCTGAGAGGCACAG

Table 2.3: Fluidigm primer sequences. Table 2.3 shows gene targets and primer sequences used for the Fluidigm 96.96 array gene expression analysis.

Gene Name	Gene Symbol	Forward Primer	Reverse Primer
18S ribosomal RNA	18S	CTTCCACAGGAGGCCTACAC	CTTCGGCCACACCCTTAAT
Polyubiquitin C precursor	UBC	AGTAGTCCCTTCTCGGCGAT	GACGATCACAGCGATCCACA
POU domain, class 5, transcription factor 1	POU5F1	ATGCATTCAAACCTGAGGTGCCT	AACTTCACCTTCCCTCCAACCA
SRY box 2	SOX2	TTTGTCCGAGACGGAGAAGC	TAAGTGTCCATGCGCTGGTT
Homeobox protein nanog	NANOG	CAATGGTGTGACGCAGGGAT	TGCACCAGGTCTGAGTGTTTC
C-myc protein	MYC	CAAGAGGCGAACACACAACG	GTCGTTTCCGCAACAAGTCC
Fibroblast growth factor 8	FGF8	GCGCATCCCTAGTGAAGGAG	CCGTCTCCAGGATGAGCTTT
Fibroblast growth factor 4	FGF4	AGTACCCCGGCATGTTTCATC	TCATCCGAAGAAAGTGCACCA
Fibroblast growth factor receptor 1	FGFR1	GTCTGCTGACTCCAGTGCAT	ACGGTTGGTTTGTCTTGT
Fibroblast growth factor receptor 2	FGFR2	ACAGTTTCCGGTCTGAGTCCAG	CATGACCACTTGCCCAAAGC
Fibroblast growth factor receptor 3	FGFR3	AGGAGCTCTTCAAGCTGCTG	ACAGGTCCAGGTAAGTCTCGTCG
Fibroblast growth factor receptor 4	FGFR4	GAGGAGGACCCACATGGA	TACTACCTGGCCAAAGCAGC
Epidermal growth factor receptor	EGFR	CTAAGATCCCCTCCATCGCC	GGAGCCCAGCACTTTGATCT
Epidermal growth factor	EGF	GTCTTGACTCTACTCCACCCC	CTCGGTAAGTACTGACTCGCTCC
Nodal	NODAL	CTGGAGGTGCTGCTTTCAGG	CCCATCCACTGCCACATCTT

Gene Name	Gene Symbol	Forward Primer	Reverse Primer
Bone morphogenic protein 4	BMP4	CGGAAGCTAGGTGAGTGTGG	CATAGGTCCCTGCAGTAGCG
Bone morphogenic protein 7	BMP7	CACTCGAGCTTCATCCACCG	GGTACTGAAGACGGCCTTG
Cerberus 1	CER1	TTTGCTTTGGGAAATGCGGG	CTCCGTCTTACCTTGCACCT
Wnt family, member 2	WNT2	GCTACGACACCTCCCATGTC	GGTCATGTAGCGGTTGTCCA
Wnt family, member 3	WNT3	CACAACACGAGGACGGAGAA	GCTTCCCATGAGACTTCGCT
Wnt family, member 5A	WNT5A	TCTGGCTCCACTTGTGCTC	CGACCACCAAGAATTGGCTTC
Wnt family, member 11	WNT11	TCTTTGGGGTGGCACTTCTC	TGCCGAGTTCACCTTGACGAG
Inhibin A	INHBA	GCTCAGACAGCTCTACCACA	CCTCTCAGCCAAAGCAAGGG
Transforming growth factor beta 1	TGFB1	CTGTCCAACATGATCGTGCG	TGACACAGAGATCCGCAGTC
Transforming growth factor beta 2	TGFB2	CGACGAAGAGTACTACGCCA	TCAAGGTACCCACAGAGCAC
SMAD family, member 1	SMAD1	TGTATTCGTGAGTTTCGCGGT	CCAAATGCAAAAGGACAGCAGA
SMAD family, member 2	SMAD2	GCTCCCTCCGTCTTCCATAC	CTTGTATCGAACCTCCCGGC
Dickkopf-related protein 1	DKK1	GGGTCTTTGTGCGGATGGTA	CTGGTACTTATCCCGCCCG
Dickkopf-related protein 2	DKK2	GGTACTCGGCACAGAGATCG	CCCTGATGGAGCACTGGTTT
Left right determination factor 1	LEFTY1	GCCATCGAGGGACTTGACTT	AAACTGAGCAAGGGCTCTCC
CAMP responsive element binding protein 1	CREB1	CTCAGCCGGGTACTACCATTC	CATGTTACCATCTTCAAAGTACG
Fos proto-oncogene	FOS	CCGAGCTGGTGCATTACAGA	ACACACTCCATGCGTTTTGC
Jun proto-oncogene	JUN	GTGCCGAAAAGGAAGCTGG	CTGCGTTAGCATGAGTTGGC
Notch 1	NOTCH1	GCAAGAACGCCGGGACA	GGCTGGCAGGATTTCCCTGA
Notch 2	NOTCH2	GATACAGATGCCAGTGTGTCCC	AGACAAATGCCCTGGATGGAAAA
Notch 3	NOTCH3	GGACGTCACTGTGAATCCT	GAACCTCCCTGCCAGGTTGG
Notch 4	NOTCH4	GAGACGTGCCAGTTTCTGA	GAGGCAAGTGCAACAAGAAGC
GATA binding protein 4	GATA4	ACCTGGGACTTGGAGGATAGCAA	CCATCAGCGTGTAAAGGCATCTGA
Delta like canonical notch ligand 1	DLL1	GGAGGCACTGTGACGACAA	GCACACTCGCACACATAGC
Chordin	CHRD	TATGCCTTGGACGAGACGTG	GGTTGGGCACTCTGGTTTGA
Caudal type homeobox 2	CDX2	GCAGCCAAGTGAAAACCAGG	TTCTCTCCTTTGCTCTGCG
Actin, alpha 2	ACTA2	AAAGCAAGTCCTCCAGCGTT	TAGTCCCGGGATAGGCAAAA
GATA binding protein 6	GATA6	TCTCCATGTGCATTGGGGAC	AAGGAAATCGCCCTGTTTCTG
SRY-box 17	SOX17	GGACCGCACGGAATTTGAAC	TAATATACCGCGCAAGTGGC
Eomesdoermin	EOMES	AGTCACCTTCTCCAGCGTG	CCTCTTCCGAGGGGAAGGTA
Alpha fetoprotein	AFP	CTGCTGCAGCCAAAGTGAAG	ATAGCGAGCAGCCCAAAGAA
Hes family BHLH transcription factor 1	HES1	AAAAATTCTCGTCCCGGTT	ATGCCGCGAGCTATCTTTCT
Brachyury transcription factor	T	TTTCCAGATGGTGAGAGCCG	CCGATGCCTCAACTCTCCAG
Snail family transcriptional repressor 1	SNAI1	CGAGTGTTCTTCTGCGCTA	GGGCTGCTGGAAGGTAAACT
Snail family transcriptional repressor 2	SNAI2	GCTACCCAATGGCCTCTCTC	CTTCAATGGCATGGGGTCT
Mesoderm posterior BHLH transcription factor 1	MESP1	GACCCATCGTTCTCTGTAC	CTGAAGAGCGGAGATGAG
Paired box 6	PAX6	GAGCGAGCGGTGCATTTG	TCAGATTCTATGCTGATTGGTAT
Nestin	NES	CCACCCTGCAAAGGGAATCT	GGTGAAGCTTGGGCACAAAAG
E-cadherin	CDH1	GCTGGACCGAGAGAGTTTCC	CAAAATCCAAGCCCCTGGTG
N-cadherin	CDH2	CATTGCCATCCTGCTCTGCATC	GCGTCTTTATCCCGGCGTTTC
P-cadherin	CDH3	GACACCATGTACCGTCCTC	TTCTGCGCAACAGAGAACA
Lymphoid enhancer binding factor 1	LEF1	CCCGTGAAGAGCAGGCTAAA	AGGCAGCTGTCACTTCTGGA
Beta catenin	CTNNB1	GCGCCATTTTAAAGCCTCTCG	GAGTAGCCATTGTCCACGCT
Junction plakoglobin	CTNNG	CCCATACTCAGTAGCCACG	CATCCTCCTCCATGATGCC
Alpha catenin	CTNNA1	TCGGGCCTCTGGAATTTAGC	CAGCCAAAACATGGGCCTTC
Protocadherin 8	PCDH8	GCTGATCGTCATCATCGTGC	AAGGTGAGCACGTCGAACAT
Protocadherin 1	PCDH1	GGGACTGACTGCTTGTGG	CTCCCAATGAGGGTGTGGG
Gap junction protein gamma 1	GJC1	CCCGTGCTACAATGGACCAA	TCTAGCAGGCGAGTCAGGAA
Gap junction protein alpha 1	GJA1	AGCCACTAGCCATTGTGGAC	CCACCTCCACCGGATCAAAA
Integrin subunit alpha v	ITGA5	ACAAATGCTCCTAGGCACCC	GCGGGTAGAAGACCAGTCAC
Integrin subunit alpha 4	ITGA4	CAGGTTTAAAGCATGGCCACA	TGGCATTGGCATTGTGTACC
Integrin subunit beta 3	ITGB3	ACCAGTAACTGCGGATTGG	TCCGTGACACACTCTGCTTC
Integrin subunit beta 1	ITGB1	GCCGCGCGAAAAGATG	ACATCGTGCAGAAGTAGGCA
Integrin subunit alpha 1	ITGA1	GGCAGCACAATTCATGCACA	AAATGTACACAGCTCCCCCG

Gene Name	Gene Symbol	Forward Primer	Reverse Primer
Integrin subunit alpha 3	ITGA3	TAGGCAGCTTCATCCTGCAC	GTCCTGCCACCCATCATTGT
Integrin subunit alpha 5	ITGA5	TTACGGGACTCAACTGCACC	AGCCTGAAACACTCAGCCTC
Integrin subunit alpha 6	ITGA6	ACACAGCATTGTATATGTGAAGCA	CCGAATCCCATTGCTTTGGC
Integrin subunit beta 2	ITGB2	CAACCCACCACCTCCTCCAAG	ACTTCGTGCACTCCTGAGAGA
Integrin subunit alpha 2	ITGA2	CGGTTATTCAGGCTCACCGA	CCTCGGGGCCTTCAAGAAAT
Epithelial cell adhesion molecule	EPCAM	GAACACTGCTGGGGTCAGAA	TCCTTCTGAAGTGCAGTCCG
Rho associated coiled-coil containing protein kinase 1	ROCK1	GTTTGAACAGGAAGGCGGAC	ATGCCCCATGGAGACTTAGC
Rho associated coiled-coil containing protein kinase 2	ROCK2	TCCCGATAACCACCCCTCTT	TGCCTTCATCTGTAGACCTCTG
Myosin heavy chain 9	MYH9	TGGTTTACCTGCACCGTTGA	TTCGGCAACCAGTGTAGACC
Myosin heavy chain 10	MYH10	GGTGGGTTTGGGACTGAGG	ACAGCCCTGTCCACAAAGAG
Myosin light chain 9	MYL9	TTGACAAGGAGGACCTGCAC	GC GTTGCGAATCACATCCTC
Myosin light chain 12B	MYL12B	CCCTGTGCCCAACACTATCC	ACACATTGGATGTTGCACGC
Myosin light chain 12A	MYL12A	GCACTTGGTCAATACCACGC	CGAGAATCCGAGCACCTCTC
Ras homolog family member A	RHOA	ACACACCAGGCGCTAATTCA	CCCCAGAGCTATGCCAACAA
Protein tyrosine kinase 2	PTK2	TGGGCGGAAAGAAATCCTGC	GCCCGTCACATTCTCGTACA
Mitogen-activated protein kinase 8	MAPK8	CTGAAGCAGAAGCTCCACCA	CACCTAAAGGAGAGGGCTGC
EPH receptor A2	EPHA2	TTCCATTAAGGACTCGGGGC	ATCAGGTCCCCTTCTTGC
EPH receptor A3	EPHA3	GCTGAAGACGGCACTAGGAC	AGGGCAGTGAGAGGAGCATA
EPH receptor A4	EPHA4	ACGCGTGCTCATCTTGTGTA	GTGAGGAGAGGACAGGTCT
EPH receptor B2	EPHB2	CAGCATTACCCTGTCGTGGT	GGGGTCACTTCTGTGATGGT
EPH receptor B4	EPHB4	GTCCCGCGCGGAGTATC	CCTGAGGGAATGTCACCCAC
Ephrin A5	EFNA5	CTTCTCGCTCCTACCCCT	ACACACATCCAGAGACCAG
Ephrin B2	EFNB2	GGAAGTACTGCTGGGGTGT	GTGCATCTGTCTGCTTGGTCT
Beta-actin	ACTB	ACAGAGCCTCGCCTTTGCC	GAGGATGCCTCTCTTGCTCTG
Death associated protein kinase 3	DAPK3	GAGAATCTGAGGAGCTGGGT	GAACTTGGCTGCGTACTCCT

2.2.12 Differentiations

For the Dual SMAD and CHIR germ lineage differentiations. 100 cell mixed colonies were generated as previously described, cultured in mTeSR™-1 medium (STEMCELL Technologies), and allowed to form patterns for 5 days. On the 5th day either SB 435142 (10µM; Sigma-Aldrich) and LDN 193189 (0.2µM; Sigma-Aldrich) or CHIR 99021 (12µM; Selleckchem) were supplemented into the media. CHIR was pulsed for 24 hour periods on day 5 and day 8 of the mesendoderm directed differentiation. Dual SMAD inhibition was kept constant for 6 days by maintaining SB and LDN in the media to direct germ lineage to an ectodermal fate. 6 days after the induction of differentiation, colonies were washed 3X with PBS and fixed for staining with 4% paraformaldehyde (VWR) as previously described.

2.2.13 Computational Image Analysis

For the radial position computational analysis fluorescent images were split into single RGB channel images using the python module scikit-image (Walt et al., 2014) where the red channel denoted CRISPRi cells, the green channel denoted WT cells, and the blue channel denoted DAPI staining of the entire colony. A mask of each channel was created by thresholding, removing small objects, and removing small holes. The radius(r) of each colony was calculated using the DAPI mask and the ratio of inner red cell area vs. outer red cell area was calculated by taking the logical AND of the red channel mask and the DAPI mask above $\frac{3}{4}r$ vs the logical AND of the red channel mask and the DAPI mask below $\frac{1}{4}r$ and normalizing to the total red mask area. To ensure accuracy the inner vs. outer red ratio was averaged with the inverse of the inner vs. outer green ratio (calculated in the same manner using the green channel mask).

For the differentiation analysis fluorescent images were split into single RGB channel images where the red channel denoted CRISPRi cells, the green channel denoted WT cells, and the blue channel denoted either PAX6 or EOMES positive cells. The pixel area of cell types was determined by thresholding the WT, CRISPRi, or EOMES/PAX6+ images, removing small objects and removing small holes to create a mask of the area covered by each individual cell type. The EOMES/PAX6+ population was calculated by taking the logical AND of either the WT and EOMES/PAX6+ mask or the CRISPRi and the EOMES/PAX6+ masks and then normalizing to total area of the WT or CRISPRi mask respectively. The ratio of EOMES/PAX6+ cells in CRISPRi cell compared to WT was calculated by dividing the normalized EOMES/PAX6+ area of the CRISPRi mask by the normalized EOMES/PAX6+ area of the WT mask.

2.2.14 Statistics

To ensure random sampling of colonies in all cell mixing experiments, 10 colonies were chosen at random on day 0 before pattern formation and traced throughout the experiment using time lapse imaging. Each experiment was performed with at least three biological replications

unless otherwise mentioned. For immunostaining experiments, 5 randomly selected fields of view were imaged per colony, with the exception of the differentiation experiments (Figure 2.14) where 25 fields of view were imaged per colony to account for any increased variation in lineage markers after differentiation. Unpaired T-tests were used to compare two groups with a false discovery rate of 1% (Benjamini–Hochberg–Yekutieli method). One-way analysis of variance (ANOVA) was used to compare three or more groups, followed by post-hoc pairwise comparisons by Tukey’s tests. In gene expression analysis, three replicates were used for each condition, and all gene expression was normalized to control mixed populations (off-target guide without knockdown) to control for any gene expression changes due to mixing or the process of FACS sorting. In all comparisons, significance was specified as $p \leq 0.05$.

2.3 Results

2.3.1 CRISPRi KD in human iPSCs modulates epithelial morphology.

To establish an inducible CRISPRi KD of ROCK1 or CDH1, we used a doxycycline (DOX)-inducible CRISPRi human iPSC line (CRISPRi-Gen1C) (Mandegar et al., 2016) (Figure 2.2). Guide RNA (gRNA) sequences designed to target the transcription start site of ROCK1 or CDH1 (Table 2.1) were introduced into CRISPR-Gen1C human iPSCs and KD was induced by the addition of DOX (2 μ M) into cell culture media. After three days of KD induction, ROCK1 mRNA levels were <30% of iPSCs without DOX treatment, and CDH1 mRNA levels in CDH1 KD human iPSCs were <10% compared to untreated controls (Figure 2.2C). Protein KD followed a similar trend where KD populations compared to untreated controls resulted in <20% ROCK1 protein and <10% of CDH1 protein by day three of DOX treatment, and reduced protein levels were maintained through day six of CRISPRi induction (Figure 2.2C, Figure 2.3).

Both the ROCK1 KD cells and the CDH1 KD cells retained epithelial human iPSC morphologies without single cell migration away from the colonies (Figure 2.2D). However, CDH1

KD iPSCs displayed irregular colony shapes that maintained smooth peripheral edges, but contained regions lacking cells within colonies (Figure 2.2D). Conversely, ROCK1 KD iPSCs displayed round colony shapes (similar to wildtype human iPSCs) but individual cells along the border of ROCK1 KD colonies extended protrusions out away from the colony (Figure 2.2D). As expected, human iPSCs treated with the small-molecule ROCK inhibitor Y-27632 yielded a similar morphology to the ROCK1 KD human iPSCs with extended cell protrusions at the colony borders (Figure 2.4).

To further confirm the physical effects of knocking down CDH1 or ROCK1 selectively in human iPSCs, we performed immunofluorescent (IF) staining of CDH1 expression. After 5 days of DOX treatment, CDH1 KD iPSCs exhibited a complete loss of CDH1 expression, as expected, whereas the ROCK1 KD iPSCs and the control iPSCs (with off-target CRISPRi guide) maintained robust expression of CDH1 along the plasma membrane (Figure 2.2E). To interrogate cell cortical tension, the contact angles between cells were measured based on IF of zona occluden-1 (ZO1), a protein associated with tight junctions (Figure 2.5A). Contact angles were not statistically different in either the ROCK1 KD or CDH1 KD cells compared to the control, but all populations displayed a subtle reduction in mean contact angle with DOX addition that was not significantly different between any of the groups (Figure 2.5B). However, when direct measurements of iPSC elasticity were taken using atomic force microscopy after 6 days of KD, ROCK1 KD cells displayed a two-fold higher cortical stiffness than the control and CDH1 KD populations, whereas the latter groups did not differ from one another (Figure 2.2E). Therefore, CRISPRi silencing of targeted genes associated with cellular mechanical properties resulted in distinct physical differences between the otherwise similar cell populations.

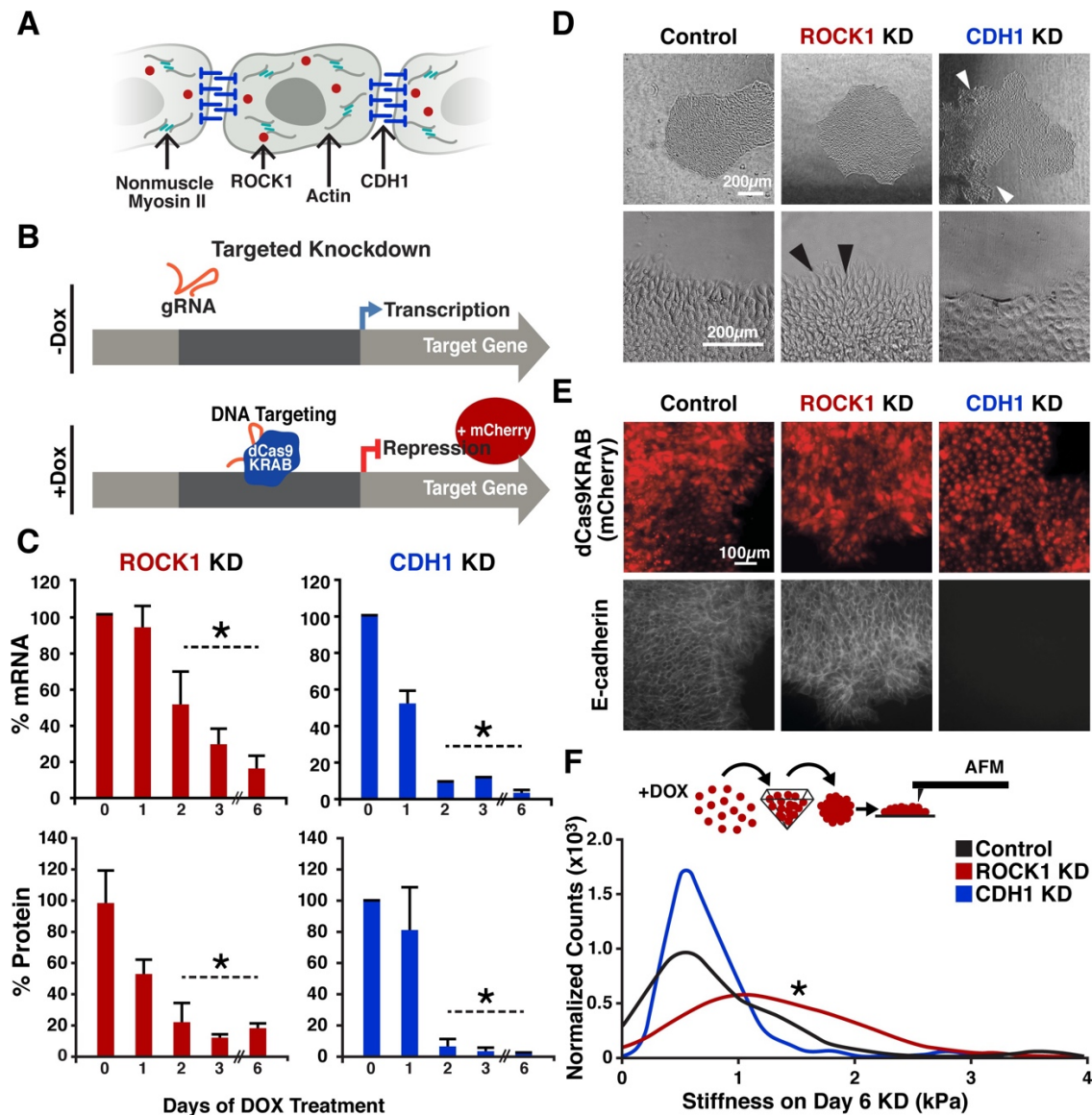


Figure 2.2: CRISPRi of ROCK1 and CDH1 modulate physical properties of the cell. (A) Schematic of ROCK1 and CDH1 within a cell. CDH1 is a trans-membrane adhesion molecule that locates to the borders of cells and ROCK1 is a cytoplasmic kinase that acts upon non-muscle myosin II. (B) Schematic of the CRISPRi system. Doxycycline addition to the human iPSC culture media leads to the expression of mCherry and dCas9-KRAB to induce knockdown of target gene. (C) qPCR and western blot quantification of knockdown timing; knockdown of both mRNA and protein were achieved by day three of DOX treatment when compared to untreated human iPSCs ($p < 0.05$, $n = 3$, data represent mean \pm SD). (D) Brightfield imaging of knockdown human iPSCs indicated morphological differences in colony shape (white arrows) and cell extensions (black arrows) at colony borders. (E) Live reporter fluorescence for dCas9-KRAB expression (red) and immunostaining for CDH1 (gray) demonstrated loss of CDH1 in induced CDH1 CRISPRi human iPSCs, but maintenance of CDH1 contacts in the off-target control and ROCK1 KD human iPSCs. (F) Atomic force microscopy (AFM) of knockdown populations exhibited a twofold increase in Young's elastic modulus of ROCK1 knockdown cells compared to control and CDH1 knockdown cells ($p < 0.05$, $n = 36, 65, 72$ force points for Control, ROCK1 KD, and CDH1 KD, respectively, area under curve = 1).

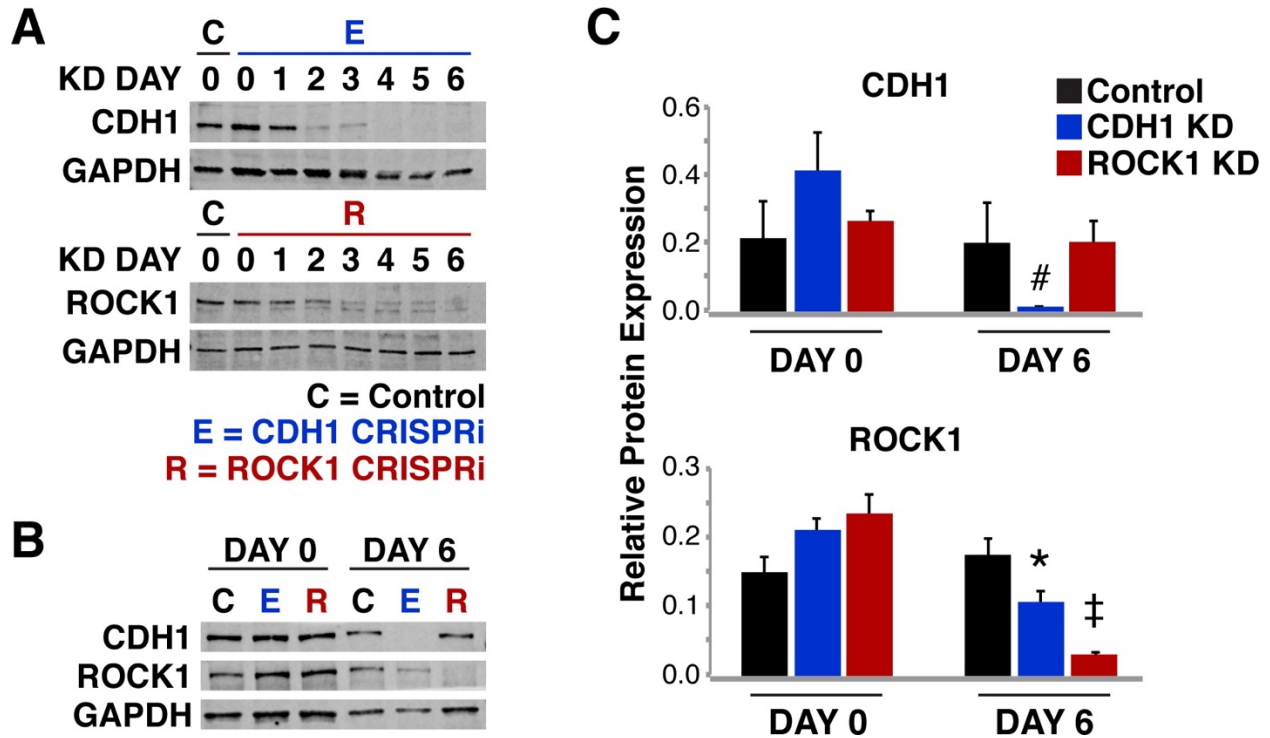


Figure 2.3: Protein knockdown time course for ROCK1 and CDH1. (A) Western blot reflecting KD time course of ROCK1 and CDH1 over 6 days. (B) Western blot of ROCK1 and CDH1 protein levels in KD cells. (C) Densitometry quantification of CDH1 and ROCK1 protein levels in both the ROCK KD cells and the CDH1 KD cells. CDH1 was only significantly different in cells where CDH1 was KD. ROCK1 protein levels exhibited a significant decrease in both CDH1 and ROCK1 ($p < 0.01$).

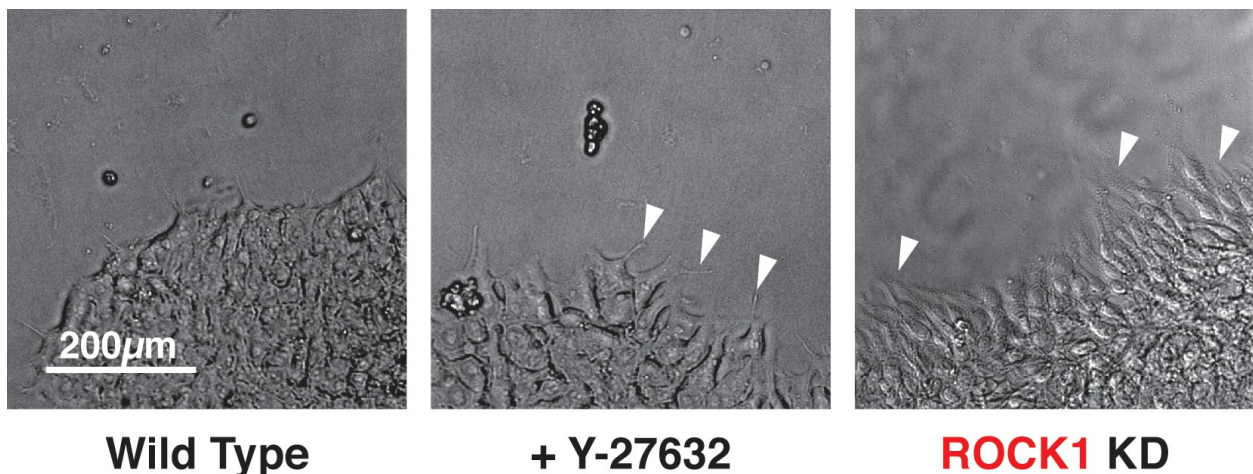


Figure 2.4: Morphology of ROCK1 knockdown human iPSCs. Morphology of ROCK1 KD cells compared to WT cells treated with small molecule inhibitor of ROCK1, Y-27632. The ROCK1 KD and the Y-27632 treated cells displayed cell protrusions (white arrowheads) at the borders of colonies when compared to uninhibited WT cells.

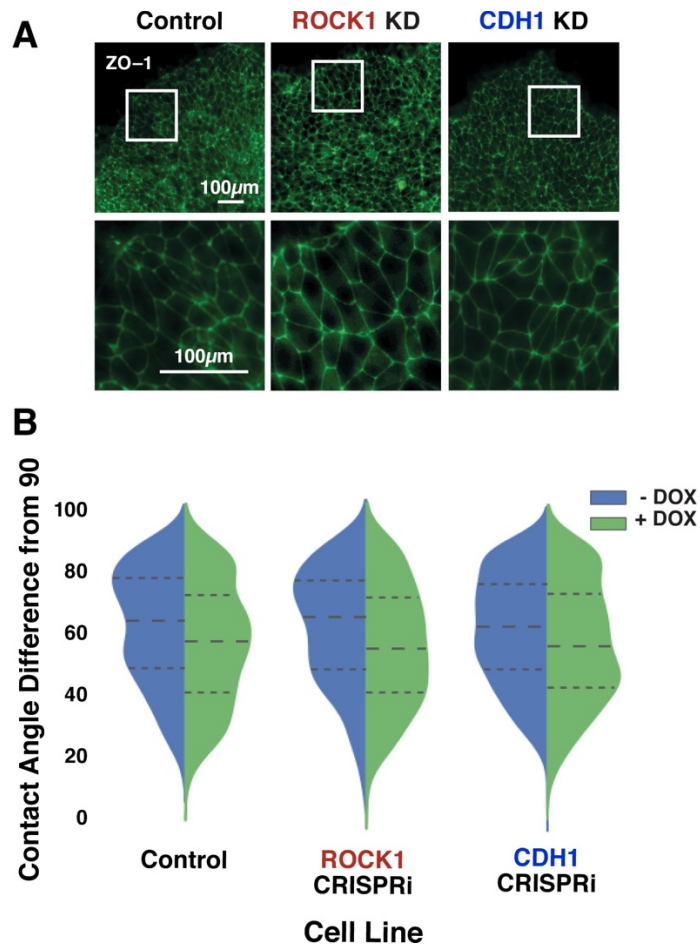


Figure 2.5: Maintenance of tight junctions in knockdown human iPSCs.

(A) Immunostaining of zona occludens 1 (ZO-1) in pure populations of human iPSCs after 6 days of induced KD of ROCK1 or CDH1. (B) Population distribution of contact angles before and after DOX treatment measured at ZO-1 junctions. The contact angles of KD cells did not differ significantly compared to the off-target CRISPRi guide control. Dotted lines delineate quartiles. ($p < 0.05$; $n = 15$ colonies, 50 angles per colony).

2.3.2 Mosaic CRISPRi silencing results in multicellular organization

To examine whether mosaic KD of a single molecule impacted human iPSC organization, ROCK1 or CDH1 CRISPRi populations were pretreated with DOX for 5 days and mixed with isogenic wildtype human iPSCs that constitutively expressed GFP (WT-GFP) at a 1:3 ratio. Forced aggregation of ROCK1 KD: WT-GFP iPSCs or CDH1 KD: WT-GFP iPSCs and subsequent re-plating were used to create individual colonies of randomly mixed ROCK1 KD iPSCs or CDH1 KD iPSCs with the WT-GFP cells (Figure 2.6A). After 5 days in mixed culture, ROCK1 KD cells sorted radially from the WT-GFP cells, clustering primarily at the edges of the colonies (Figure 2.6B-D). However, separation of the ROCK1 KD cells did not result in distinct smooth borders between the WT-GFP and ROCK1 KD human iPSC populations. In contrast,

CDH1 KD cells robustly separated from the GFP-WT population, forming sharp boundaries between populations irrespective of their spatial location within the colony (Figure 2.6B-D).

To determine whether pattern emergence was impacted by the relative proportion of mosaic KD within a colony, KD cells were mixed with control CRISPRi iPSCs lacking any gRNA or fluorescent protein at varying cell ratios of 1:1, 1:3, and 3:1. Clustering of ROCK1 KD cells was less apparent as the proportion of ROCK1 KD cells within a colony increased. In fact, increasing ROCK1 KD iPSCs to 75% of the colony resulted in the entire colony morphology displaying characteristics of a pure ROCK1 KD colony (Figure 2.7). On the other hand, the CDH1 CRISPRi cells separated from the colorless CRISPRi population, irrespective of cell ratio, indicating that the spatial organization occurred regardless of relative population size within an iPSC colony. The ability of both the ROCK1 and CDH1 CRISPRi KD populations to physically partition from otherwise identical CRISPRi-engineered iPSCs that lacked a gRNA confirms that the production of dCas9-KRAB is not responsible for the previously observed pattern formation when CRISPRi KD iPSCs were mixed with the WT-GFP cells, but rather that the segregation is a direct result of KD of the target gene (Figure 2.7).

Based on the sorting behaviors of ROCK1 KD: WT-GFP and CDH1 KD: WT-GFP colonies when the KD of ROCK1 or CDH1 was present at the time of mixing, we next examined whether induction of mosaic KD after mixing resulted in similar sorting patterns as previously observed. This scenario more accurately models the onset of initial symmetry-breaking events among homogeneous pluripotent cells during embryonic development. Non-induced CRISPRi populations were mixed with WT-GFP iPSCs (1:3 ratio), re-plated, and then treated with DOX to induce KD (Figure 2.6E). ROCK1 KD post-mixing within mosaic colonies did not result in noticeable radial segregation of ROCK1 KD human iPSCs from WT-GFP iPSCs (Figure 2.6F-H), as observed for pre-mixed colonies. Instead, the post-mixing ROCK1 mosaic KD colonies exhibited greater vertical stacking of ROCK1 KD cells and WT-GFP cells in the z-plane of the mixed colonies (Figure 2.6G), whereas the pre-induced mixed colonies remained segregated

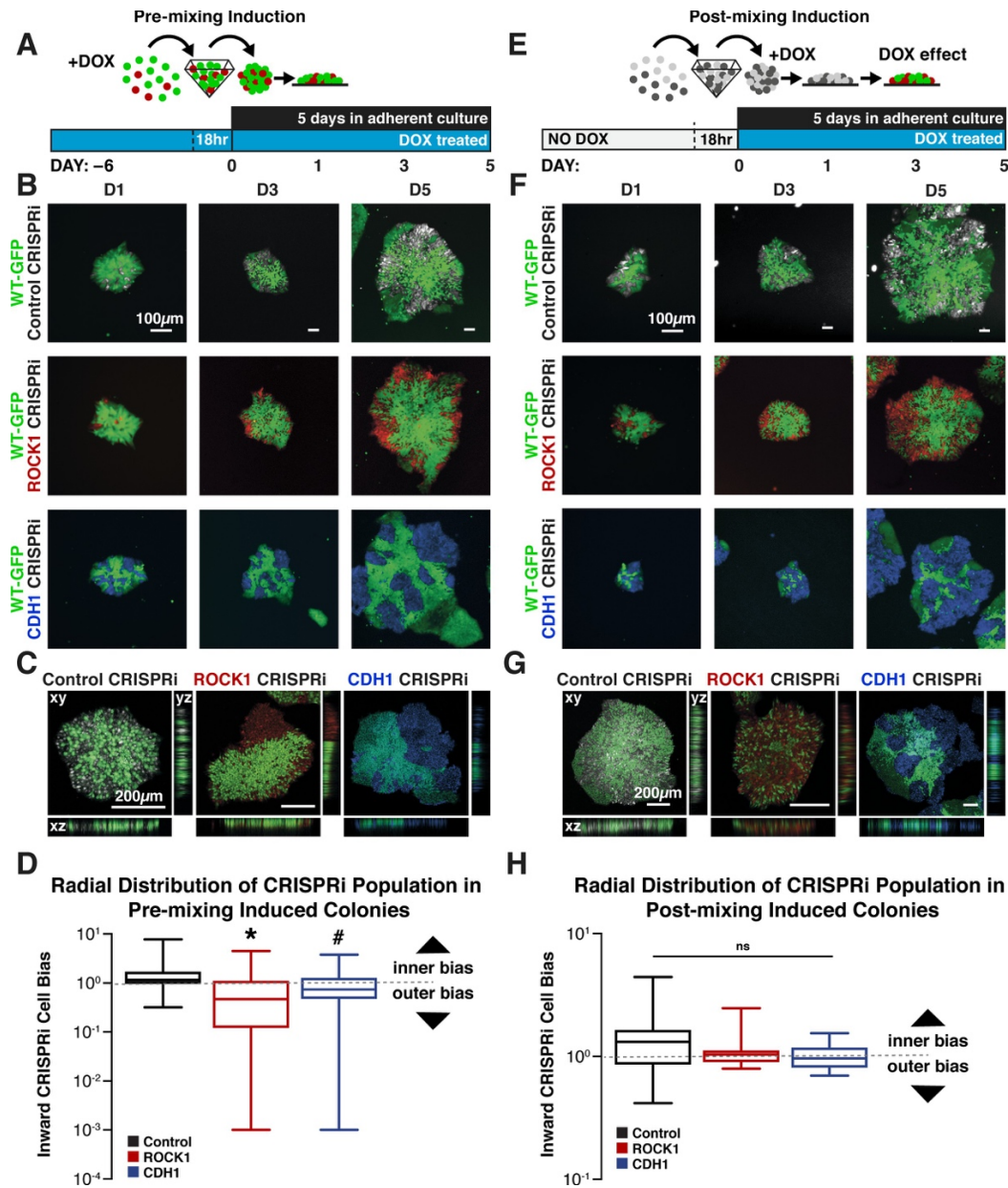


Figure 2.6: Cell-autonomous pattern emergence in mixed population colonies. (A) Schematic of experimental timeline. WT-GFP and ROCK1- or CDH1- CRISPRi human iPSCs were pretreated with doxycycline for 6 days before aggregation in pyramidal microwells and re-plating as mixed colonies. (B) Live cell imaging of pattern emergence over time from mixing colonies. Control populations remain mixed, ROCK1 KD human iPSCs cluster radially at borders of colonies, and CDH1 KD populations sort themselves from WT-GFP human iPSCs regardless of location within colony. (C) Confocal microscopy of patterned colonies of human iPSCs with KD induction prior to mixing. (D) Quantification of the radial distribution of KD cells in pre-induced mixed colonies. The ratio of inner cell area to outer cell area normalized to total cell area is displayed ($n = 25$, * and # indicate significance, $p < 0.05$). (E) Schematic of experimental timeline for WT-GFP and ROCK1- or CDH1-CRISPRi human iPSCs treated with doxycycline upon re-plating as mixed colonies. (F) Live cell imaging of pattern emergence in post-mixing induction colonies, where CRISPRi KD is induced after cell population mixing. (G) Confocal microscopy of patterned human iPSC colonies with KD induction upon mixing populations, where ROCK KD cells stack vertically with WT-GFP human iPSCs. (H) Quantification of the radial distribution of KD cells in post-induced mixed colonies. The ratio of inner cell area to outer cell area normalized to total cell area is displayed ($n = 20$).

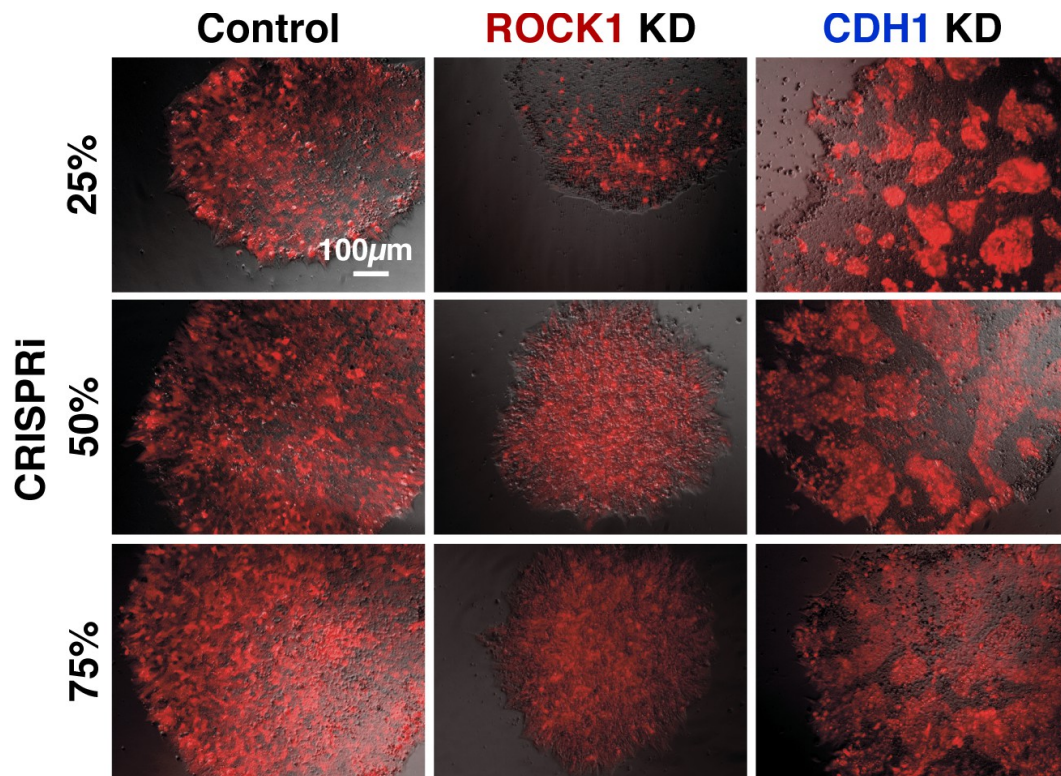


Figure 2.7: Pattern emergence of variable ratios of mixed populations. Live imaging on mixed colonies of CRISPRi KD human iPSCs with WT human iPSCs, where a picture was taken every 30 min for 12 hr, 4 days after the start of KD and mixing of two populations (no color = no guide, red = KD population). The displayed images were taken at the start of day four. The ratio of CRISPRi KD cells was varied at 25%, 50%, and 75% of the overall population.

primarily in a 2D planar manner (Figure 2.6C). In contrast, the mosaic silencing of CDH1 post-mixing maintained robust segregation of the CDH1 KD cells from the WT-GFP iPSCs, although the borders between cell populations lacking CDH1 contacts and neighboring WT-GFP cells were somewhat less distinct than the pre-induced CDH1 KD: WT-GFP mixed colonies. Overall, the inducible CRISPRi mixed colonies displayed the ability to mimic several different patterns of intrinsic symmetry-breaking events that resulted in distinct cell sorting and multicellular pattern formation.

In addition to the changes in organization within colonies, there were significant changes in the ratio of cells within the colonies from the original seeding density of 3:1 WT to CRISPRi. There was an increase in the CRISPRi proportion in the mixed colonies over time (Figure 2.8A).

To determine if this was in response to mixing with a WT population, EDU incorporation studies were performed on pure CRISPRi and WT populations. When EDU incorporation was examined over a period of DOX treatment for 5 days, the WT-GFP cells displayed a reduced replication rate compared to the CRISPRi lines (Figure 2.8B), however cell replication rate did not account for pattern formation as the CRISPRi control mixed colonies did not display a distinguishable pattern.

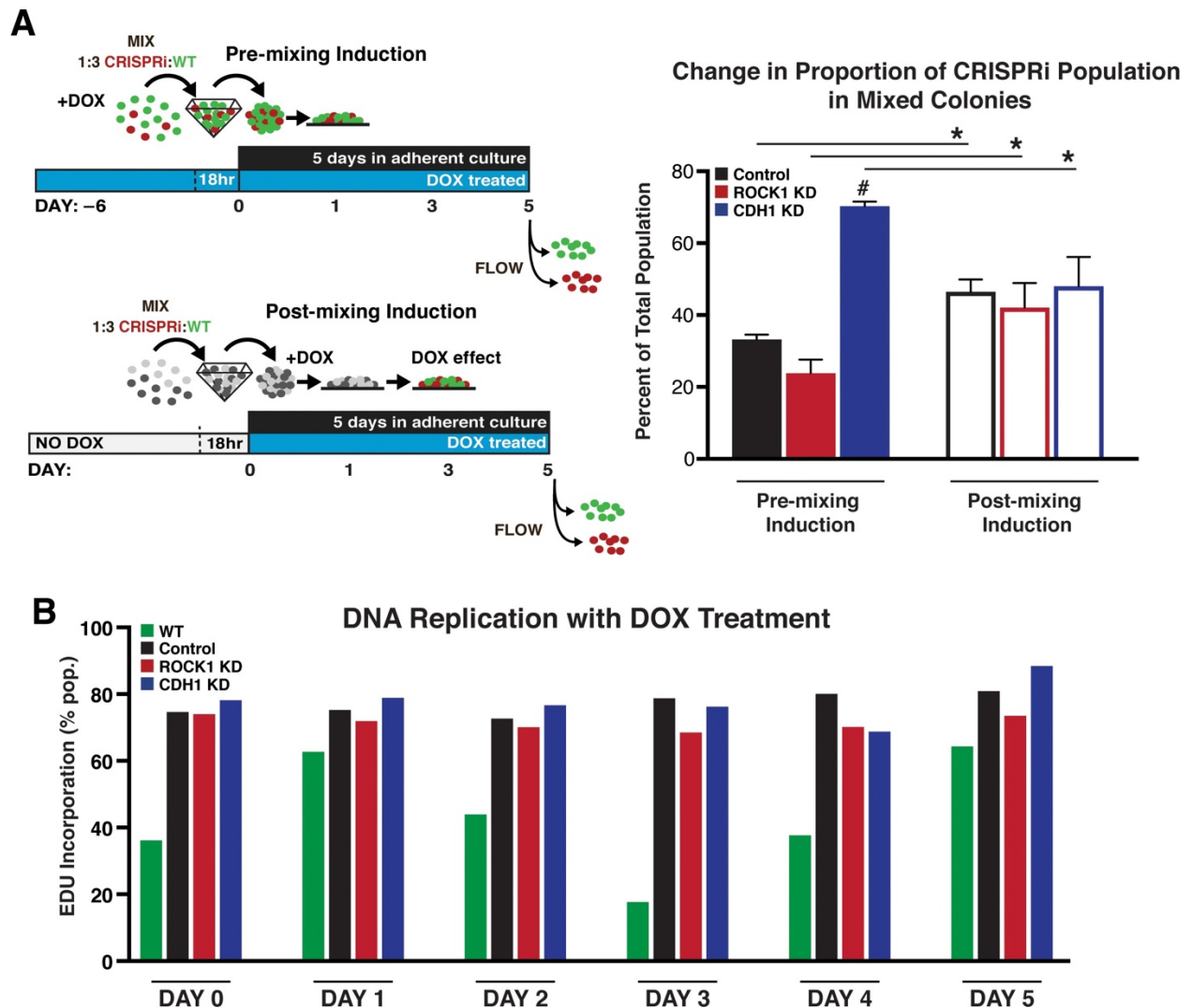


Figure 2.8: Cell population change over time. (A) Percent of CRISPRi KD cells after 5 days in mixed culture measured by flow cytometry. (n = 10,000 events, three biological reps; * indicated significance between pre and post mixed and # indicates significance from other KDs; p<0.05) (B) EdU incorporation overtime in pure populations as KD is induced by DOX treatment (n = 10,000 events).

2.3.3 Mosaic human iPSC colonies retain a pluripotent phenotype

Colony morphology and expression of epithelial markers, such as epithelial cell adhesion molecule (EpCAM), were examined to determine if the cells that lost CDH1 expression segregated from their WT-GFP neighbors due to delamination, or loss of the epithelial phenotype characteristic of iPSCs. ROCK1 KD: WT-GFP and CDH1 KD: WT-GFP colonies maintained an epithelial morphology throughout 6 days of CRISPRi silencing (Figure 2.9A) with no observed migration by CRISPRi-modulated cells away from the colonies. Both ROCK1 KD and CDH1 KD iPSCs within mixed colonies expressed EpCAM at cell-cell boundaries after 6 days of CRISPRi induction despite changes in cortical tension or intercellular adhesion due to loss of ROCK1 or CDH1, respectively (Figure 2.9A). Furthermore, ROCK1/CDH1 KD iPSCs displayed cell junction-localized β -catenin in pure colonies after 6 days of CRISPRi induction, suggesting maintenance of adherens junctions and epithelial colonies (Figure 2.10).

Since the decrease of CDH1 is commonly associated with loss of pluripotency in PSCs, the expression and localization of the common pluripotency transcription factors, OCT3/4 and SOX2, were examined. Both proteins maintained strong nuclear expression in pure ROCK1 KD or CDH1 KD colonies after 6 days of KD induction (Figure 2.11A). Moreover, despite the physical segregation of cells induced by KD in mixed populations, no pattern could be observed based on pluripotency marker expression (Figure 2.9B). Furthermore, the abundance of OCT3/4 and SOX2 transcripts was unchanged in pure colonies of CDH1 KD cells and though variable in pure colonies of ROCK1 KD iPSCs, was not significantly different (Figure 2.11B). However, because the transcription factors SOX2 and OCT4 were retained by cells for a period of time during the process of differentiation, genes associated with the primitive streak (Brachyury (BRA)) and the neural crest (SOX9) were interrogated in either ROCK1 or CDH1 KD cells over 6 days (Figure 2.11C). Both BRA and SOX9 were significantly increased on day 3 of KD in ROCK1 KD cells, however at day 6 the gene expression returned to levels comparable to day 0 before ROCK1 KD. Although the CDH1 KD cells did not display any significant trends, it is noteworthy that the samples had a

large variation with a standard deviation often 3 times higher than that of the ROCK1 KD cells. This high variation between biological replicates potentially indicates that the KD of CDH1 created a large variability in the gene regulation of these two lineage specific genes. The large variation in gene expression observed in both ROCK1 KD and CDH1 KD cells could indicate that the cells are in a transition state from pluripotency. However, these results indicate that the loss of ROCK1 or CDH1 is not sufficient to disrupt the pluripotent gene regulatory network and induce an exit from the pluripotent state.

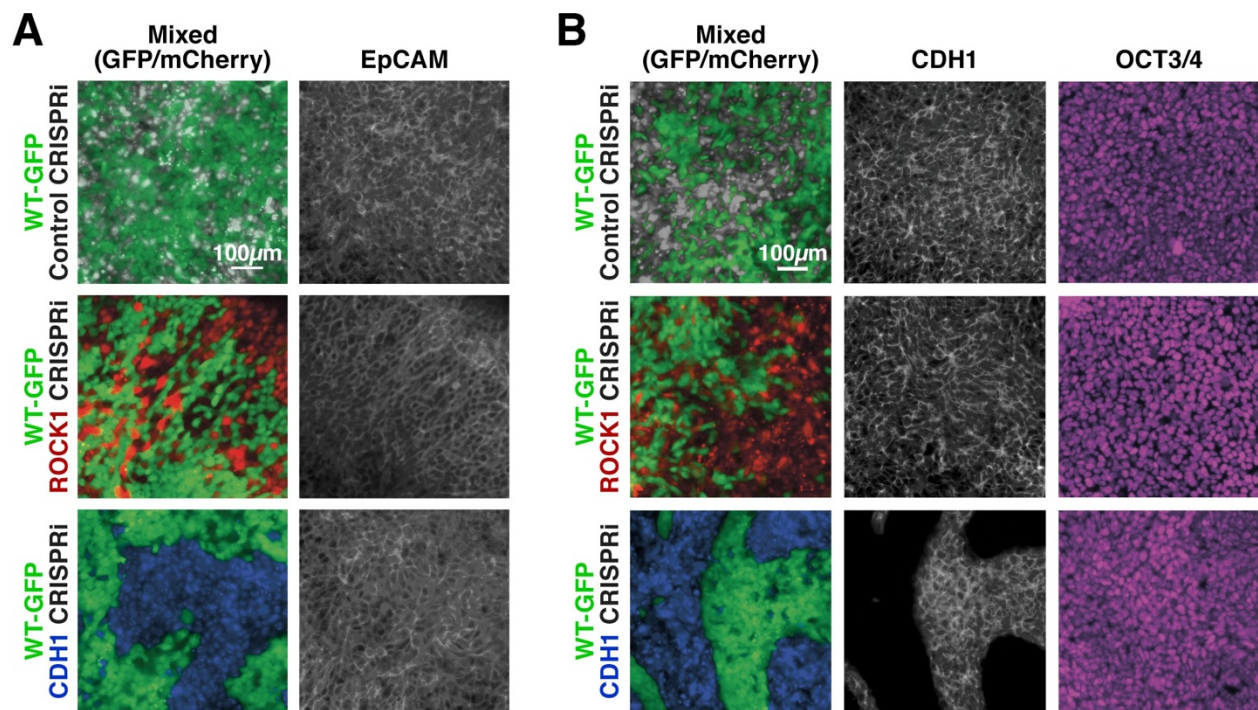


Figure 2.9: Maintenance of nuclear pluripotency markers and epithelial phenotype.(A) Immunostaining of EpCAM for mixed colonies displayed relatively uniform expression regardless of KD. (B) Immunostaining for E-cadherin (CDH1) and OCT3/4 in patterned human iPSC colonies demonstrating nuclear localized OCT3/4 throughout the mixed populations.

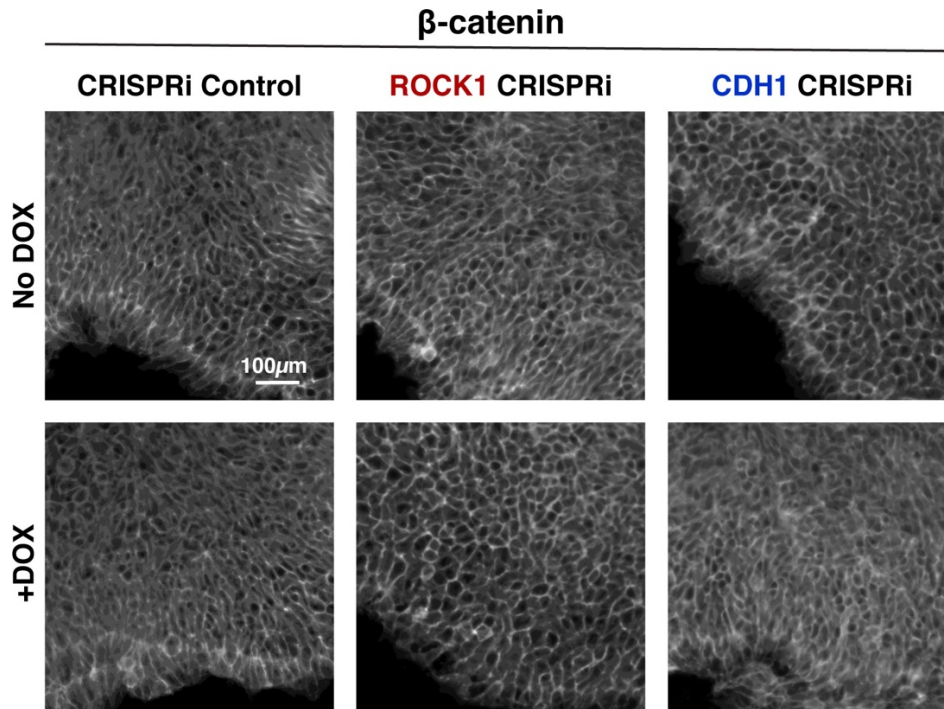


Figure 2.10: Maintenance of cell junction localized β-catenin in knockdown cells. Immunostaining of β-catenin pure human iPSC colonies on day six of KD. β-catenin remains localized to the cell junctions in human iPSCs before and after DOX treatment to induce KD of ROCK1 and CDH1, respectively.

2.3.4 Mixed populations of KD cells display transient gene expression changes in coordination with emergence of patterns

Since pluripotency markers appeared to be maintained irrespective of mosaic patterning, gene expression changes in pluripotency markers (SOX2, NANOG), mesendoderm markers (SOX17, BRA) and ectoderm markers (PAX6, SOX9) were examined with the induction of mosaic patterning at days 1, 3 and 6 after KD induction (Figure 2.12A). To take into account potential gene expression changes that result from mixing iPSC lines, un-induced mixed populations and un-induced pure populations were included as controls. BRA did not have any significant changes with induction of ROCK1 KD or CDH1 KD in a mixed population, however SOX9 increased in both ROCK1 KD and CDH1 KD cells (Figure 2.12C,D). Interestingly, similar to pure populations there was large amount of variance, often displaying over a fold change difference in gene expression between biological replicates in mixed colonies.

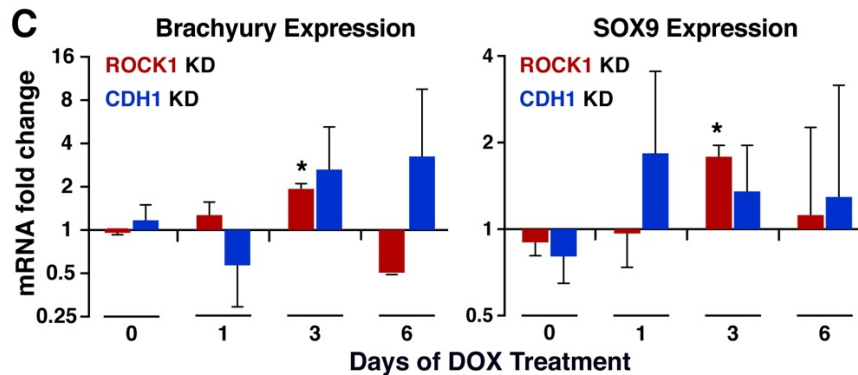
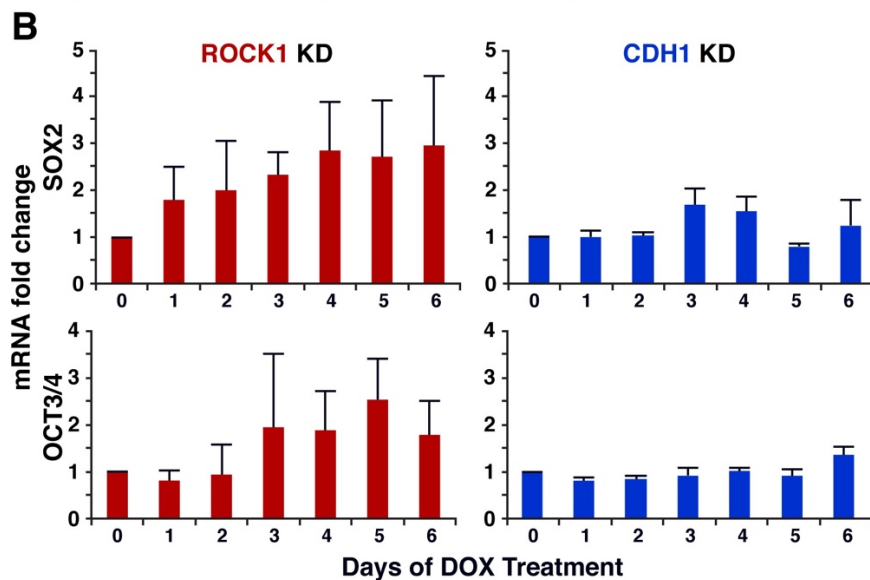
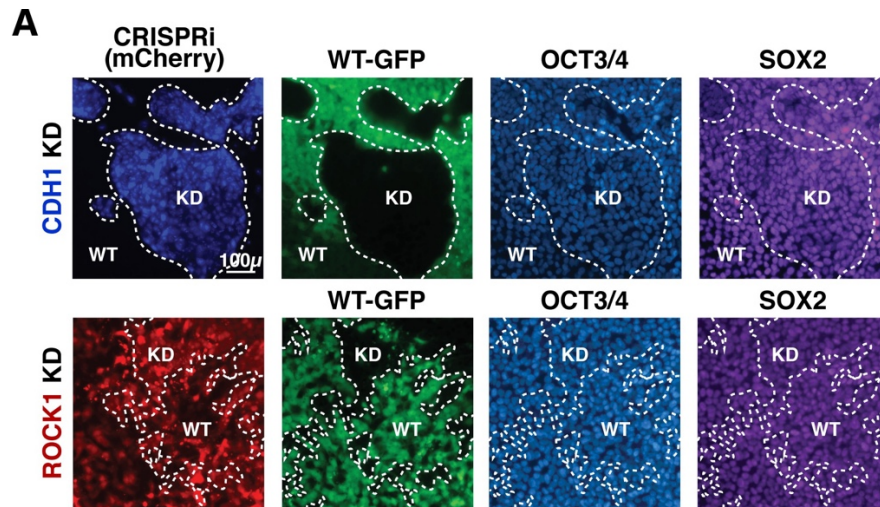


Figure 2.11: Pluripotent and early germ layer gene expression in knockdown cells.(A) Immunostaining of OCT3/4 and SOX2 in mixed colonies on day six of KD, where borders between cell populations are denoted by dotted lines. (B) Gene expression of OCT3/4 and SOX2 in ROCK1 KD and CDH1 KD human iPSCs quantified by mRNA fold change (n = 3). (C) Gene expression of markers of the primitive streak (Brachyury) and neural crest (SOX9) in pure populations of ROCK1 KD or CDH1 KD human iPSCs quantified by mRNA fold change (n = 3, * indicates significance; p<0.05).

Because there were gene expression changes unique to the mixed population mosaic KDs, whether these gene changes were due to just the mixing of two cell populations, or whether they were specific to the induction of a symmetry breaking event by mosaic KD of either ROCK1 or CDH1 was interrogated. A select set of genes involved in pluripotent stem cell signaling, early lineage fate transitions, and regulation of physical cell properties (Table 2.2) was examined in both pure KD populations and mixed KD populations. An ANOVA analysis was used to examine gene expression changes that can be attributed to mixing two different cell types (mixed populations without KD), to KD of ROCK1 or CDH1 in a pure population, and to mosaic KD or KD in the presence of a WT neighbor (Figure 2.12E,F). Overall, few changes in gene expression resulted from mixing un-induced CRISPRi populations with WT-GFP (Figure 2.13A), and therefore subsequent data was normalized to pure un-induced populations and then to mixed un-induced populations to minimize false positives that resulted from mixing of cell lines without induction of KD.

In ROCK1 KD cells mixed with WT, on day 1 the gene expression changes that could be attributed to solely to gene KD (Figure 2.12E, left column) were associated with primitive streak formation (Snail (SNAI1), SMAD1). On day 3, we observed an upregulation of adhesion molecules (EPHA4, ITGA4), as well as NANOG, Nestin (NES), and TGF β upregulation and FGFR2 downregulation. Interestingly, there were a large amount of changes in gene expression that were specific to the mosaic induction of ROCK1 KD in addition to those that were a direct response to ROCK1 KD (Figure 2.12E, right column). For example, the down regulation of both cell-cell adhesions as well as cell-ECM adhesions. Additionally, genes that were upregulated in the pure KD context were down regulated in the mosaic KD context, such as SNAI1, NES and NANOG. However, at day 6 of mosaic KD we did not observe any persistent significant changes in the examined genes (Figure 2.12E). CDH1 KD caused an upregulation in genes associated with cell-cell adhesion on day 1 that is exacerbated in a mosaic CDH1 KD (Figure 2.12F). Interestingly, both Wnt3 and downstream Wnt targets such as SNAI1 and SNAI2 are significantly upregulated

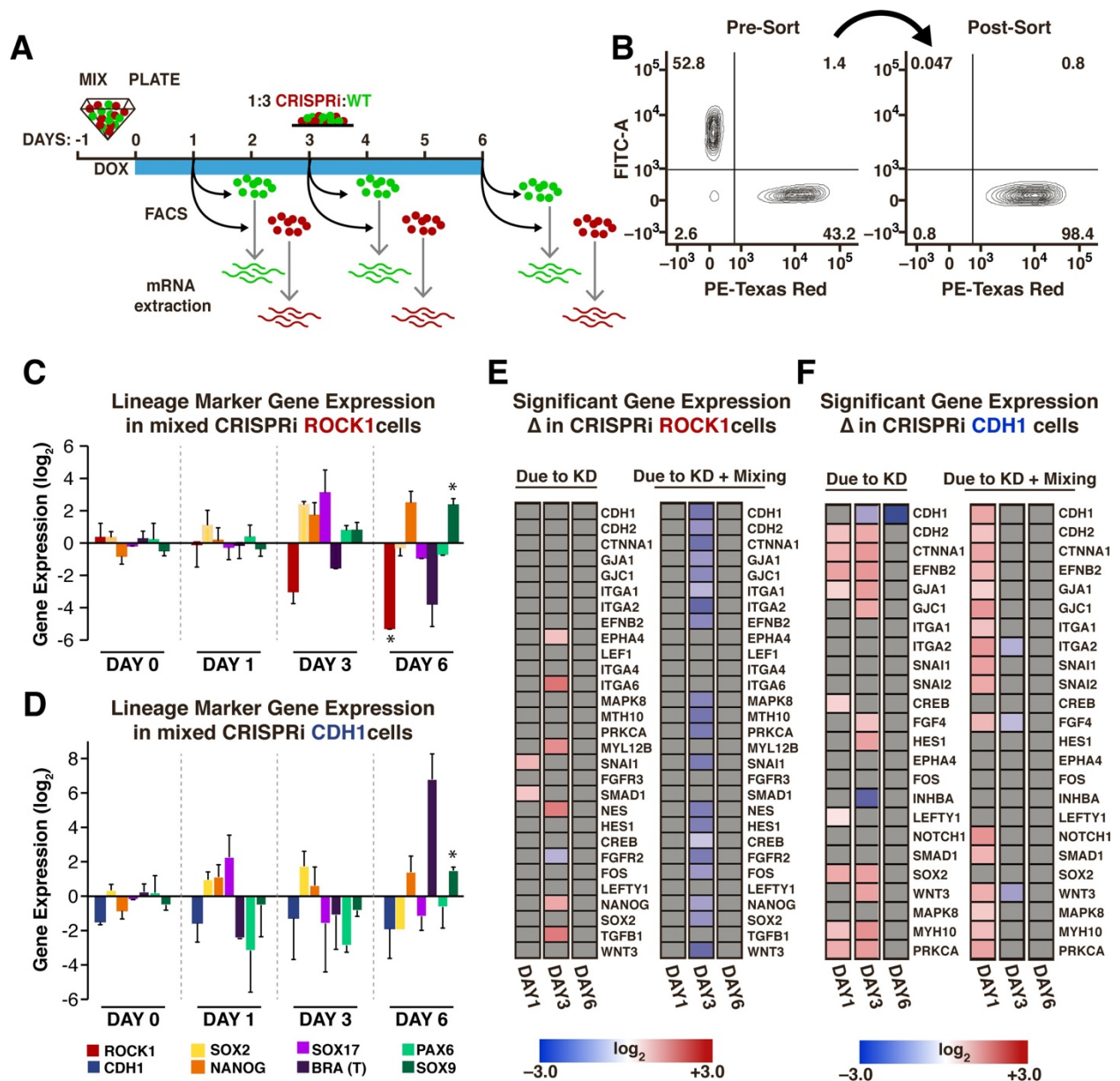


Figure 2.12: Transient gene expression changes in mixed populations.(A) Schematic of experimental timeline; WT-GFP and ROCK1- or CDH1-CRISPRi human iPSCs were mixed and re-plated prior to KD induction. Different cell populations were isolated by FACS for mRNA extraction on days 1, 3, and 6 after KD induction. ($n = 3$ per condition). (B) Representative scatter plot of a FACS-sorted population of mCherry + cells (indicating KD induction) with >98% purity. (C,D) Plots of specific mRNA expression changes at days 1, 3, and 6 in KD cell populations that have been mixed with WT. (* and # indicate significance, $p < 0.05$). (E,F) Heat maps display fold change expression of genes found to display significant changes in ROCK1 or CDH1 KD cells mixed with WT-GFP human iPSCs when compared to time-matched, off-target control human iPSCs. Grey color indicates non-significance. Significance ($p < 0.05$, $n = 3$) was determined using a one-way analysis of variance (ANOVA) followed by post-hoc pairwise comparisons by Tukey's tests to determine the effect of mixing populations, the effect of solely KD, and the effect of KD within a mixed population.

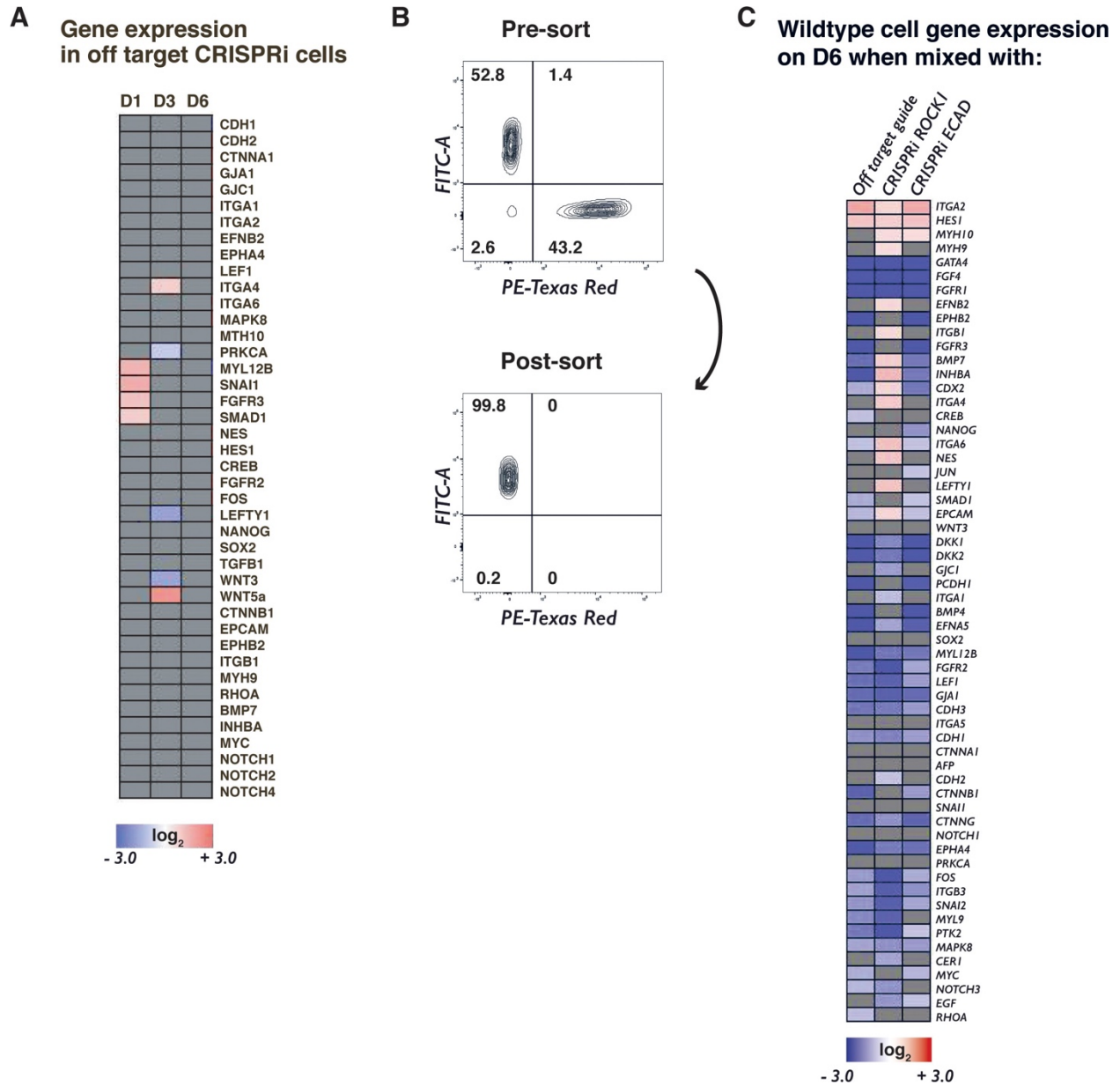


Figure 2.13: Gene expression changes in WT-GFP cells mixed with CRISPRi cells.(A) Heat map showing gene expression changes due to mixing two populations together without induction of CRISPRi in sub population where grey indicates non-significance. (B) Example panel of FACS-sorted population of GFP+ (i.e. WT) cells with >99% purity. (C) Heat map of significant gene expression fold changes in WT-GFP cells mixed for 6 days with OTG control, ROCK1 KD, or CDH1 KD human iPSCs when compared to time-matched, pure WT-GFP populations. Grey color indicates non-significance. Significance ($p < 0.05$, $n = 3$) was determined using an ANOVA taking into account the independent effects of mixing populations, KD alone, and KD within a mixed population.

specifically in a mosaic KD on day 1 of KD. Similar to the transient wave of gene expression changes observed in ROCK1 KD cells, mosaic CDH1 KD did not have any observed significant changes on day 6 of KD except for CDH1 (Figure 2.12F). This was consistent with our previous observation of the maintenance of pluripotency in the ROCK1 KD: WT-GFP and CDH1 KD: WT-GFP colonies (Figure 2.9). Furthermore, the recovery of homeostatic gene expression profiles closely followed the dynamics of distinct pattern establishment in the mixed populations.

In addition to examining the KD cells, we examined the gene expression profiles of the neighboring WT cells that constituted the majority of cells in each colony. On day six of KD induction, the WT-GFP cells that were mixed with CDH1 KD human iPSCs had gene expression patterns that resembled the WT-GFP cells mixed with the control CRISPRi populations, whereas the WT-GFP cells mixed with ROCK1 KD iPSCs exhibited a different expression profile. Interestingly, the WT-GFP cells mixed with ROCK1 CRISPRi iPSCs demonstrated changes in genes associated with cell sorting and movement, such as ephrins and integrins, and up-regulation in myosin proteins (MYH9, MYH10) (Figure 2.13B,C). Overall, the changes in the WT-GFP iPSC gene expression suggests that targeted manipulation of gene expression in an emerging sub-population can exert non-cell autonomous effects on the opposing population and may be influenced by the respective multicellular organization of the two populations.

2.3.5 Mixed populations allow for restricted germ lineage emergence

A great advantage in the controllable induction of two separate populations in pluripotent stem cells is potential co-emergence of multiple differentiated cell populations with predictable spatial organization. To examine how pre-patterning pluripotent stem cells could contribute to co-emergence of populations, two variations of germ lineage differentiations were utilized to direct the human iPSCs to either an ectodermal or mesendodermal fate (Figure 2.14A). Then the proportion of either PAX6+ cells indicating neuro-ectoderm lineage or Eomesodermin (EOMES)+ cells indicating a mesendoderm lineage in the WT and CRISPRi mixed populations was

examined. Interestingly, the ROCK1 KD population did not display a significant difference in PAX6⁺ or EOMES⁺ cells compared to the WT cells (Figure 2.14B,D). However, the CDH1 KD population displayed decreased PAX6⁺ cells in the ectoderm directed differentiation and increased EOMES⁺ cells in the mesendoderm directed differentiation (Figure 2.14C,D). Overall, these studies demonstrate the ability to control both patterning of stem cells as well as subsequent differentiation potential.

2.4 Discussion

In this study we examined the effect of inducing specific genetic KD in subpopulations of human iPSCs within an otherwise homogeneous population of pluripotent cells. Historically, small-molecule chemical inhibitors, antibodies, and homogeneous genetic knockouts are often used to interrogate the molecular mechanisms involved in morphogenesis (Lecuit & Lenne, 2007; McBeath, Pirone, Nelson, Bhadriraju, & Chen, 2004; Salbreux, Charras, & Paluch, 2012). However, these methods can't selectively discriminate between different cells, or they fail to address how the emergence of heterotypic interactions affects cell-cell organization. Here, we report that silencing of target genes by CRISPRi within only subpopulations provides multiple avenues to genetically control the emergence of asymmetric cell phenotypes and development of multicellular patterns. Specifically, we demonstrate that mosaic KD of target genes ROCK1 or CDH1 result in distinct patterning events wherein cell-driven segregation dictates colony organization without loss of pluripotency (Figure 2.15).

ROCK1 regulates actin-myosin contraction (McBeath et al., 2004) and facilitates expansion of PSCs (Ohgushi, Minaguchi, & Sasai, 2015; Park et al., 2015), and its acute inhibition by small molecules leads to a "relaxed" cell phenotype with decreased stiffness (Kinney, Saeed, & McDevitt, 2014; Lee et al., 2006). However, we found that prolonged silencing of ROCK1 in human iPSCs (6 days) resulted in cells that were actually two-fold stiffer than either the CDH1 KD

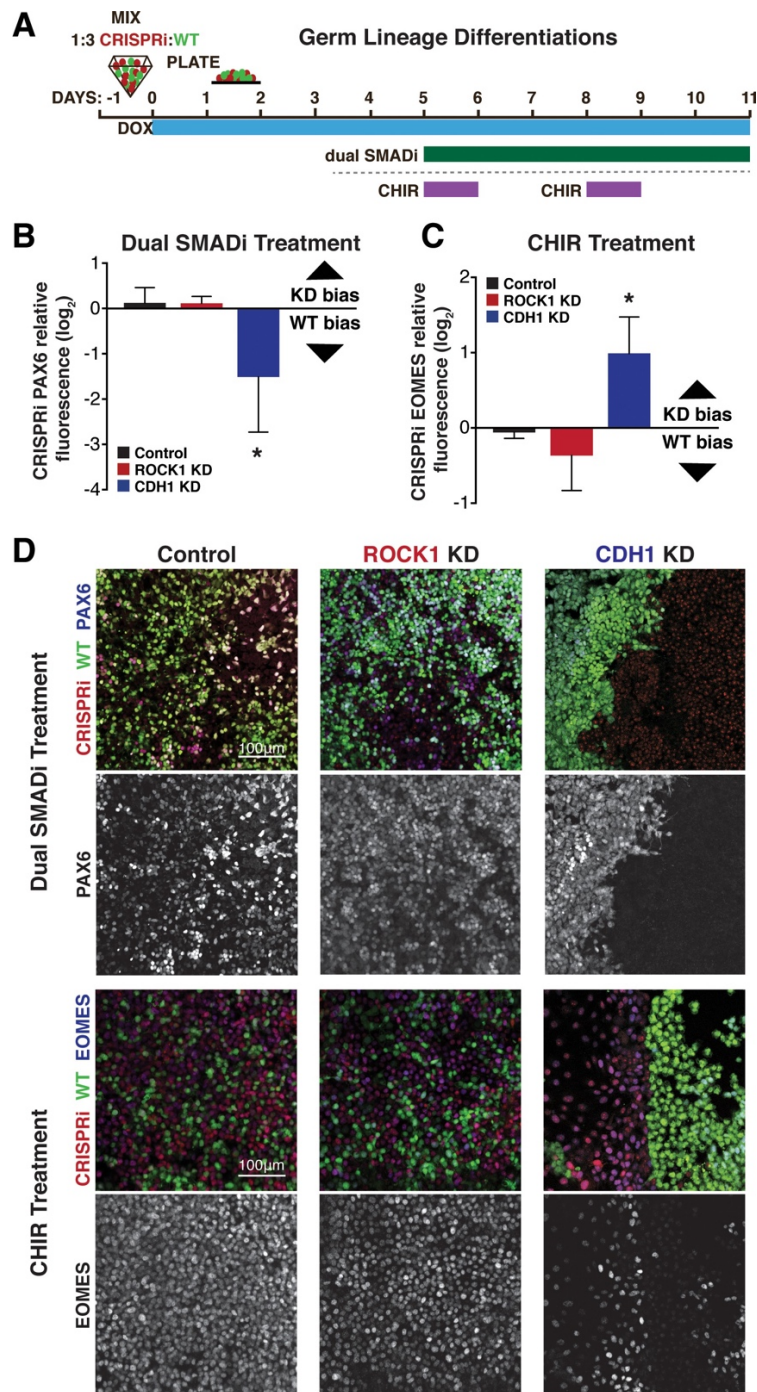


Figure 2.14: Germ lineage differentiations of mixed colonies. (A) Schematic of a 6 day ectoderm directed differentiation using dual SMAD inhibition or a complementary mesendoderm directed differentiation using CHIR treatment. (B) Quantification of PAX6 protein presence by immunofluorescence in WT and KD cells of mixed colonies after dual SMAD inhibition treatment. (n = 4 biological replicates with 25 images per biological replicate; * indicates significance; $p < 0.05$) (C) Quantification of EOMES protein presence by immunofluorescence in WT and KD cells of mixed colonies after CHIR treatment. (D) Representative images of mixed colony differentiations with either dual SMAD inhibition or CHIR treatment stained for PAX6 and EOMES, respectively.

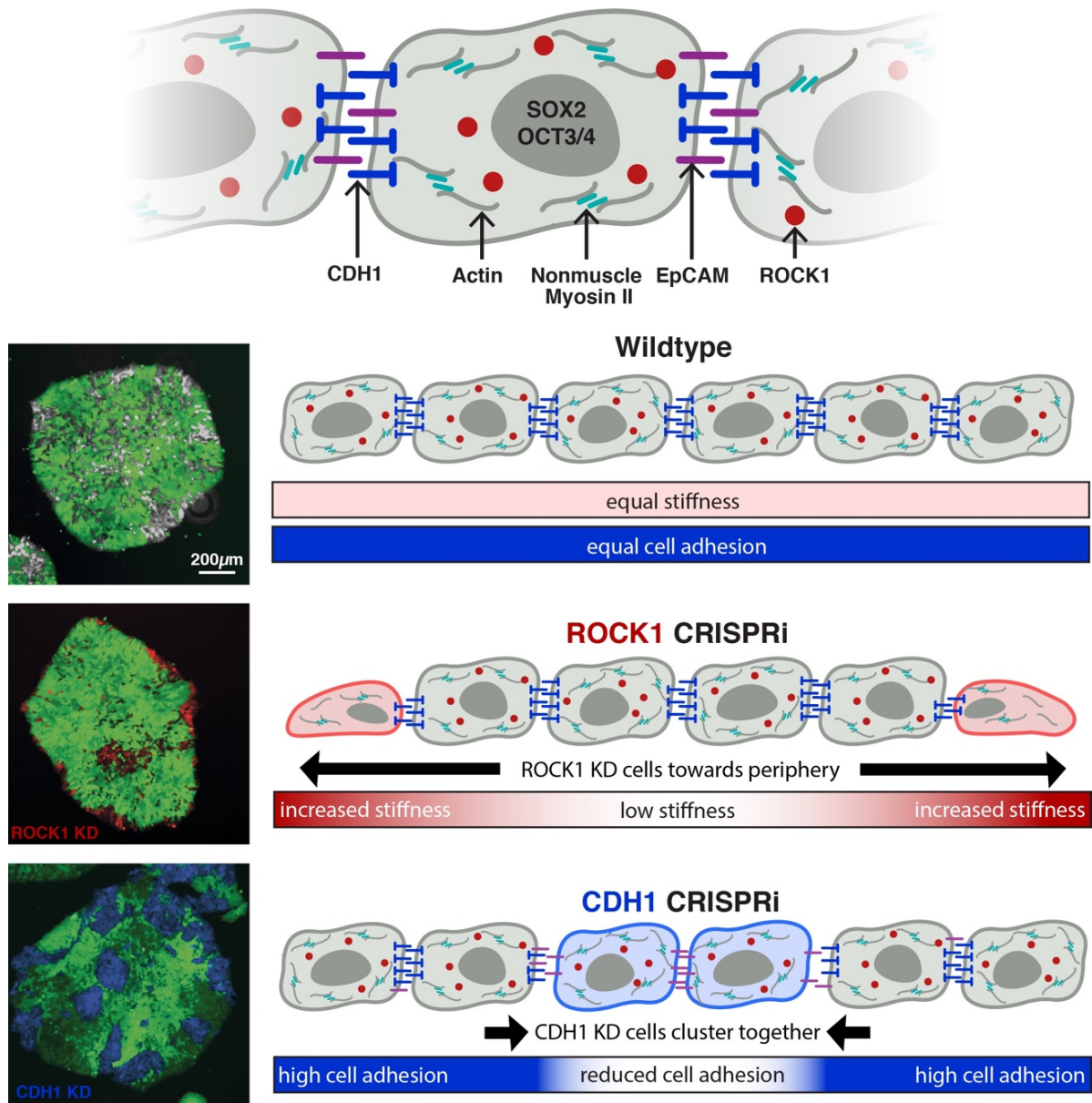


Figure 2.15: Inducible pattern emergence through the KD of molecules that affect human iPSC physical properties.(A) Schematic of working model of sub-population manipulation where controlled changes in cellular stiffness or cellular adhesion result in specific colony pattern formation. With mosaic KD, ROCK1 produces continuous radial separation of KD cells from WT, whereas CDH1 displays discrete islands of KD cells within the WT population.

cells or the control CRISPRi cells. The increased cortical stiffness of ROCK1 KD iPSCs could be due to the difference between the inhibition of an existing protein and KD of the gene. A small molecule inhibitor prevents the function of already existing proteins so that a small amount of functioning protein may escape the inhibitor's influence. In contrast, CRISPRi only needs to target the ROCK1 gene loci at two alleles to completely abolish protein translation, thus highlighting the strength of genetic perturbation. Additionally, ROCK inhibition is often used as a transient perturbation (24h), whereas long-term KD of ROCK1 (6 days) may induce compensatory effects within the cells that are responsible for the somewhat surprising results. Long-term ROCK1 KD compensation is a likely partial explanation why KD of ROCK1 prior to mixing resulted in radially partitioned populations, but post-mixing KD resulted in less segregated populations. The increased stacking of the ROCK1KD population with induction of knockdown after mixing with WT cells could also be a reflection of the difference between short term and long term ROCK1 KD and potential compensation effects. The observed 3D stacking may be caused by a compensatory contraction of WT cells as the ROCK1 KD cells lose the ability to control cytoskeletal contraction, causing 3D colony rounding. This effect would then be abrogated with prolonged ROCK1 KD as compensatory mechanisms within the ROCK1 KD cells take effect.

The emergence of autonomous patterning events and separate cell populations is often associated with differentiation, and in particular, CDH1 is involved in the control of morphogenesis in a range of species (C.A. Burdsal, C.H. Damsky, & R.A. Pedersen, 1993; D. Li et al., 2010). Historically, *in vitro* studies of mouse embryonic stem cells often describe CDH1 as a marker of stem cell pluripotency (L. Li, Bennett, & Wang, 2012; Soncin & Ward, 2011). However, while CDH1 is commonly expressed by pluripotent cells and CDH1 can replace OCT3/4 during fibroblast reprogramming to pluripotency (Redmer et al., 2011), CDH1 is not essential to maintain the pluripotent state (Larue et al., 1996; Soncin et al., 2009; Ying et al., 2008). Our results are consistent with the latter observations where CDH1 KD in human iPSCs did not disrupt the expression of pluripotency markers nor lead to a loss of epithelial phenotype, indicating that KD

of CDH1 alone was not sufficient to induce differentiation, but could result in rearrangement of colony structure. Furthermore, the observed maintenance of pluripotency and differentiation preferential to mesendoderm is consistent with recent studies revealing that changes in human CDH1 adhesions coordinate with *in vitro* human stem cell lineage decisions rather than pluripotency maintenance (Przybyla, Lakins, & Weaver, 2016). Changes in CDH1 influencing lineage fate decisions may explain the transient gene expression changes that we observed with the induction of KD in mixed colonies, where the loss of CDH1 potentially primes the cells to respond to a signal for differentiation, and without such a signal, the cells return to a ground state of pluripotency. Additionally, the large variety in gene regulation between biological replicates may be indicative of a transitional state of cells primed to differentiate that mosaic KD incites in the examined pluripotent stem cell colonies. Similar priming has been described in the context of cell-matrix adhesion where differentiation in response to TGF β signaling is primed by stiffness-dependent integrin signaling (Allen, Cooke, & Alliston, 2012); a comparable mechanism may explain the observed transient gene expression changes without loss of pluripotency in CDH1 KD human iPSCs.

The ability to manipulate distinct cell populations allows for robust modeling of human morphogenic events and, thus, an expanded understanding of human biology that can be exploited to develop physiologically realistic *in vitro* human tissue models. Cellular location within pluripotent colonies can be thought to parallel the effects seen in early developing blastocysts. A cell's location within the early embryo relays signals that dictate initial symmetry-breaking events, such as the decision to become trophectoderm instead of inner cell mass. Cells located within the center of an embryo maintain different adhesion contacts (Stephenson, Yamanaka, & Rossant, 2010) and are subjected to higher tension generated by neighboring cells (Samarage et al., 2015), which then feed back into lineage fate decisions. For example, the Hippo pathway is controlled by a cell's position within the early blastocyst, where the outer cell layer has the ability to polarize and sequester the signaling molecule angiomin away from adherens junctions, preventing the

phosphorylation and activation that would occur in an internal cell that maintained cell-cell contacts on all sides (Hirate et al., 2013). Additionally *in vitro* micro-patterned PSC colonies have been reported to display spatially dependent germ layer patterning upon differentiation (Etoc et al., 2016; Tewary et al., 2017; Warmflash, Sorre, Etoc, Siggia, & Brivanlou, 2014). The observed patterns in mosaic ROCK1 KD demonstrate that a targeted gene KD can affect the spatial location of a cell within a colony without perturbing its differentiation potential, and conversely the mosaic CDH1 KD demonstrates control of both lineage potential and spatial organization, offering two complementary methods to interrogate multicellular organization and morphogenic processes. Therefore, the system described here enables the potential to enhance and expand on these previous complementary studies by allowing for the manipulation of local multicellular neighborhoods through subpopulation organization and potential priming to certain lineage fates.

Complementary to changes in multicellular organization as a result of ROCK1 or CDH1 KD, there were compensatory changes in gene expression that occur in the WT population. In particular the WT population displayed significant changes in several genes associated with adhesion and lineage fate. For example, GATA4 is down regulated in WT cells in all mixed colonies. GATA4 is associated with mesendoderm lineages (Molkentin, Lin, Duncan, & Olson, 1997, p. 4; Zorn & Wells, 2009) and its down regulation may have effects on the WT cell's ability to properly differentiate to mesendoderm. However, there were no significant difference in differentiation potential quantified by EOMES expression when control or ROCK1 KD mixed colonies were directed toward a mesendodermal fate (Figure 2.14). This could potentially be due to the strength of the small molecule CHIR, a GSK3 inhibitor, in inducing the mesendoderm fate. Additionally, differences between the two populations may arise in longer differentiations, where the cells are allowed to mature beyond a progenitor stage. However, differences in maturation or cell type within a germ lineage may assist in the controlled co-emergence of multiple tissue types or even cell populations within a single tissue.

Overall, this study capitalized on the ability of CRISPRi to temporally perturb specific molecular regulators of physical cell properties, such as adhesion and tension, that resulted in differing multicellular patterns. Moreover, CRISPRi additionally offers the flexibility to target any gene of interest and timing of KD (Gordon et al., 2016; Mandegar et al., 2016), allowing for the creation of dynamic patterns through transient genetic KD that could be used to pre-pattern PSC colonies in various types of multicellular geometries before differentiation. Additionally, the ability to induce molecular asymmetry can also be applied to co-differentiation, where the temporal induction of specific heterotypic interactions, such as the presentation of a ligand or receptor, can give rise to the coordinated emergence of two (or more) cell types under the same culture conditions. In addition, mosaic induction of KD can be used to examine how signals propagate between cells, for example, interrogating how the networks between cells created by either mechanical (adhesions) or chemical gradients (gap junctions) affect lineage fate decisions. Furthermore, the predictable patterning events and potential for control over co-emergence that we establish in this study could aid the eventual control over morphogenic events in organoid systems. Organoids require coordinated heterotypic interactions in a 3D environment in order to self-organize (Bredenoord, Clevers, & Knoblich, 2017; Sasai, 2013); the ability to precisely predict and control the organization of multiple cell types in parallel would significantly improve the reproducibility and robustness of *in vitro* tissue modeling. Ultimately, this study identifies a novel strategy to direct the emergence of heterotypic cell populations to control multicellular organization in pluripotent stem cells, and subsequently facilitates the creation of robust models of morphogenesis necessary for the mechanistic study of human developmental tissue patterning and formation.

2.5 Conclusion

In conclusion, this first study provides an example of an *in vitro* system to interrogate cell-cell interactions and multicellular organization within human induced pluripotent stem cell colonies. We show that CRISPRi can be used to systematically control symmetry breaking events by inducing mosaic knockdown effectively providing a transition from homogeneity to heterogeneity. Therefore, mosaic patterning enables genetic interrogation of a range of complex emergent multicellular properties, which can facilitate better understanding of the molecular pathways that regulate symmetry-breaking during morphogenesis.

2.6 Bibliography

- Allen, J. L., Cooke, M. E., & Alliston, T. (2012). ECM stiffness primes the TGF β pathway to promote chondrocyte differentiation. *Molecular Biology of the Cell*, 23(18), 3731–3742. <https://doi.org/10.1091/mbc.E12-03-0172>
- Arboleda-Estudillo. (n.d.). Movement Directionality in Collective Migration of Germ Layer Progenitors. Retrieved January 11, 2016, from <http://www.sciencedirect.com/science/article/pii/S096098220902051X>
- Bredenoord, A. L., Clevers, H., & Knoblich, J. A. (2017). Human tissues in a dish: The research and ethical implications of organoid technology. *Science*, 355(6322), eaaf9414. <https://doi.org/10.1126/science.aaf9414>
- C.A. Burdsal, C.H. Damsky, & R.A. Pedersen. (1993, July 1). The role of E-cadherin and integrins in mesoderm differentiation and migration at the mammalian primitive streak | Development. Retrieved from Development website: <http://dev.biologists.org/content/118/3/829.abstract>
- Etoc, F., Metzger, J., Ruzo, A., Kirst, C., Yoney, A., Ozair, M. Z., ... Siggia, E. D. (2016). A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Developmental Cell*, 39(3), 302–315. <https://doi.org/10.1016/j.devcel.2016.09.016>
- Gordon, G. C., Korosh, T. C., Cameron, J. C., Markley, A. L., Begemann, M. B., & Pflieger, B. F. (2016). CRISPR interference as a titratable, trans-acting regulatory tool for metabolic engineering in the cyanobacterium *Synechococcus* sp. Strain PCC 7002. *Metabolic Engineering*, 38, 170–179. <https://doi.org/10.1016/j.ymben.2016.07.007>
- Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V. B., Akimoto, K., ... Sasaki, H. (2013). Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos. *Current Biology*, 23(13), 1181–1194. <https://doi.org/10.1016/j.cub.2013.05.014>

- Hookway, T. A., Butts, J. C., Lee, E., Tang, H., & McDevitt, T. C. (2015). Aggregate formation and suspension culture of human pluripotent stem cells and differentiated progeny. *Methods*. <https://doi.org/10.1016/j.ymeth.2015.11.027>
- Kinney, M. A., Saeed, R., & McDevitt, T. C. (2014). Mesenchymal morphogenesis of embryonic stem cells dynamically modulates the biophysical microtissue niche. *Scientific Reports*, *4*. <https://doi.org/10.1038/srep04290>
- Krieg, M., Arboleda-Estudillo, Y., Puech, P.-H., Käfer, J., Graner, F., Müller, D. J., & Heisenberg, C.-P. (2008). Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biology*, *10*(4), 429–436. <https://doi.org/10.1038/ncb1705>
- Larue, L., Antos, C., Butz, S., Huber, O., Delmas, V., Dominis, M., & Kemler, R. (1996). A role for cadherins in tissue formation. *Development*, *122*(10), 3185–3194.
- Lecuit, T., & Lenne, P.-F. (2007). Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature Reviews Molecular Cell Biology*, *8*(8), nrm2222. <https://doi.org/10.1038/nrm2222>
- Lee, J. S. H., Panorchan, P., Hale, C. M., Khatau, S. B., Kole, T. P., Tseng, Y., & Wirtz, D. (2006). Ballistic intracellular nanorheology reveals ROCK-hard cytoplasmic stiffening response to fluid flow. *Journal of Cell Science*, *119*(9), 1760–1768. <https://doi.org/10.1242/jcs.02899>
- Li, D., Zhou, J., Wang, L., Shin, M. E., Su, P., Lei, X., ... Wang, F. (2010). Integrated biochemical and mechanical signals regulate multifaceted human embryonic stem cell functions. *The Journal of Cell Biology*, *191*(3), 631–644. <https://doi.org/10.1083/jcb.201006094>
- Li, L., Bennett, S. A. L., & Wang, L. (2012). Role of E-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. *Cell Adhesion & Migration*, *6*(1), 59–73. <https://doi.org/10.4161/cam.19583>
- Ludwig, T. E., Bergendahl, V., Levenstein, M. E., Yu, J., Probasco, M. D., & Thomson, J. A. (2006). Feeder-independent culture of human embryonic stem cells. *Nature Methods*, *3*(8), 637–646. <https://doi.org/10.1038/nmeth902>

- Mandegar, M. A., Huebsch, N., Frolov, E. B., Shin, E., Truong, A., Olvera, M. P., ... Conklin, B. R. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell*, 18(4), 541–553. <https://doi.org/10.1016/j.stem.2016.01.022>
- McBeath, R., Pirone, D. M., Nelson, C. M., Bhadriraju, K., & Chen, C. S. (2004). Cell Shape, Cytoskeletal Tension, and RhoA Regulate Stem Cell Lineage Commitment. *Developmental Cell*, 6(4), 483–495. [https://doi.org/10.1016/S1534-5807\(04\)00075-9](https://doi.org/10.1016/S1534-5807(04)00075-9)
- Molkentin, J. D., Lin, Q., Duncan, S. A., & Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes & Development*, 11(8), 1061–1072. <https://doi.org/10.1101/gad.11.8.1061>
- Montero, J.-A., & Heisenberg, C.-P. (2004). Gastrulation dynamics: Cells move into focus. *Trends in Cell Biology*, 14(11), 620–627. <https://doi.org/10.1016/j.tcb.2004.09.008>
- Ohgushi, M., Minaguchi, M., & Sasai, Y. (2015). Rho-Signaling-Directed YAP/TAZ Activity Underlies the Long-Term Survival and Expansion of Human Embryonic Stem Cells. *Cell Stem Cell*, 17(4), 448–461. <https://doi.org/10.1016/j.stem.2015.07.009>
- Park, S., Kim, D., Jung, Y.-G., & Roh, S. (2015). Thiazovivin, a Rho kinase inhibitor, improves stemness maintenance of embryo-derived stem-like cells under chemically defined culture conditions in cattle. *Animal Reproduction Science*, 161, 47–57. <https://doi.org/10.1016/j.anireprosci.2015.08.003>
- Przybyla, L., Lakins, J. N., & Weaver, V. M. (2016). Tissue Mechanics Orchestrate Wnt-Dependent Human Embryonic Stem Cell Differentiation. *Cell Stem Cell*, 19(4), 462–475. <https://doi.org/10.1016/j.stem.2016.06.018>
- Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreira, A., Birchmeier, W., & Besser, D. (2011). E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Reports*, 12(7), 720–726. <https://doi.org/10.1038/embor.2011.88>

- Salbreux, G., Charras, G., & Paluch, E. (2012). Actin cortex mechanics and cellular morphogenesis. *Trends in Cell Biology*, 22(10), 536–545. <https://doi.org/10.1016/j.tcb.2012.07.001>
- Samarage, C. R., White, M. D., Álvarez, Y. D., Fierro-González, J. C., Henon, Y., Jesudason, E. C., ... Plachta, N. (2015). Cortical Tension Allocates the First Inner Cells of the Mammalian Embryo. *Developmental Cell*, 34(4), 435–447. <https://doi.org/10.1016/j.devcel.2015.07.004>
- Sasai, Y. (2013). Cytosystems dynamics in self-organization of tissue architecture. *Nature*, 493(7432), 318–326. <https://doi.org/10.1038/nature11859>
- Schäfer, G., Narasimha, M., Vogelsang, E., & Leptin, M. (2014). Cadherin switching during the formation and differentiation of the Drosophila mesoderm—Implications for epithelial-to-mesenchymal transitions. *Journal of Cell Science*, 127(Pt 7), 1511–1522. <https://doi.org/10.1242/jcs.139485>
- Soncin, F., Mohamet, L., Eckardt, D., Ritson, S., Eastham, A. M., Bobola, N., ... Ward, C. M. (2009). Abrogation of E-Cadherin-Mediated Cell–Cell Contact in Mouse Embryonic Stem Cells Results in Reversible LIF-Independent Self-Renewal. *STEM CELLS*, 27(9), 2069–2080. <https://doi.org/10.1002/stem.134>
- Soncin, F., & Ward, C. M. (2011). The Function of E-Cadherin in Stem Cell Pluripotency and Self-Renewal. *Genes*, 2(1), 229–259. <https://doi.org/10.3390/genes2010229>
- Stephenson, R. O., Yamanaka, Y., & Rossant, J. (2010). Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. *Development*, 137(20), 3383–3391. <https://doi.org/10.1242/dev.050195>
- Tewary, M., Ostblom, J., Prochazka, L., Zulueta-Coarasa, T., Shakiba, N., Fernandez-Gonzalez, R., & Zandstra, P. W. (2017). A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning. *Development*, dev.149658. <https://doi.org/10.1242/dev.149658>

- Ungrin, M. D., Joshi, C., Nica, A., Bauwens, C., & Zandstra, P. W. (2008). Reproducible, ultra high-throughput formation of multicellular organization from single cell suspension-derived human embryonic stem cell aggregates. *PloS One*, 3(2), e1565. <https://doi.org/10.1371/journal.pone.0001565>
- Walt, S. van der, Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., ... Yu, T. (2014). scikit-image: Image processing in Python. *PeerJ*, 2, e453. <https://doi.org/10.7717/peerj.453>
- Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D., & Brivanlou, A. H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature Methods*, 11(8), 847–854. <https://doi.org/10.1038/nmeth.3016>
- Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., ... Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature Biotechnology*, 25(6), 681–686. <https://doi.org/10.1038/nbt1310>
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., ... Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, 453(7194), 519–523. <https://doi.org/10.1038/nature06968>
- Zorn, A. M., & Wells, J. M. (2009). Vertebrate Endoderm Development and Organ Formation. *Annual Review of Cell and Developmental Biology*, 25(1), 221–251. <https://doi.org/10.1146/annurev.cellbio.042308.113344>

Chapter 3: Automated Design of Pluripotent Stem Cell Self-Organization

3.1 Introduction

During the early stages of embryonic development, patterned self-assembly of cells is essential for the organization of primitive germ layers, multicellular tissues, and complex organ systems (Montero and Heisenberg 2004). Similarly, human pluripotent stem cells (hPSCs) maintain the ability to self-organize, differentiate to all three germ layers, and generate 3D organoids that replicate primitive tissue structure and function (Sasai 2013; Warmflash et al. 2014; Bredenoord, Clevers, and Knoblich 2017); hence, hPSCs provide a robust and tractable system to observe, quantify, predict, and ultimately control collective cellular behaviors (Pir and Le Novère 2016). The ability to direct heterotypic cell self-organization and concurrently specify cell fate can enable the possibility of directing organogenesis via novel cell-intrinsic routes.

In this study, we paired CRISPR interference (CRISPRi) driven genetic perturbations of human induced pluripotent stem cells (hiPSCs) with computational modeling, machine learning, and mathematical optimization to facilitate a “closed loop” cycle of *in silico* hypothesis generation that could be experimentally validated *in vitro*. To predict multicellular pattern formation, we combined a multi-scale Cellular Potts model (Krieg et al. 2008; Ouchi et al. 2003; Graner and Glazier 1992; Pir and Le Novère 2016; Marée, Grieneisen, and Hogeweg 2007; Magno, Grieneisen, and Marée 2015) of behavior driven cell sorting with an automated machine learning and optimization procedure, referred to as “Multicellular Pattern Synthesis” (Briers et al. 2016; Bartocci et al. 2016), that consisted of four steps (Figure 3.1). First, we created a computational model of observed hiPSC self-organization that quantified collective stem cell dynamics and captured how targeted changes in the mechanical profiles of sub-populations of cells affected stem cell colony patterning. Second, a supervised machine learning classifier was trained to

quantify pattern similarity to the desired pattern using images from our computational model. Third, we employed mathematical optimization, specifically Particle Swarm Optimization (PSO), to simulate thousands of potential designs and identify specific experimental conditions that yielded unique patterns in *in silico* simulations. Finally, we tested the *in silico* predicted conditions with hiPSCs *in vitro* and obtained the desired multicellular patterns with similar frequency and quantitative characteristics, thereby validating the system. Furthermore, the patterning differentially impacted the subsequent cell fate commitment upon exposure to morphogen treatment (BMP4).

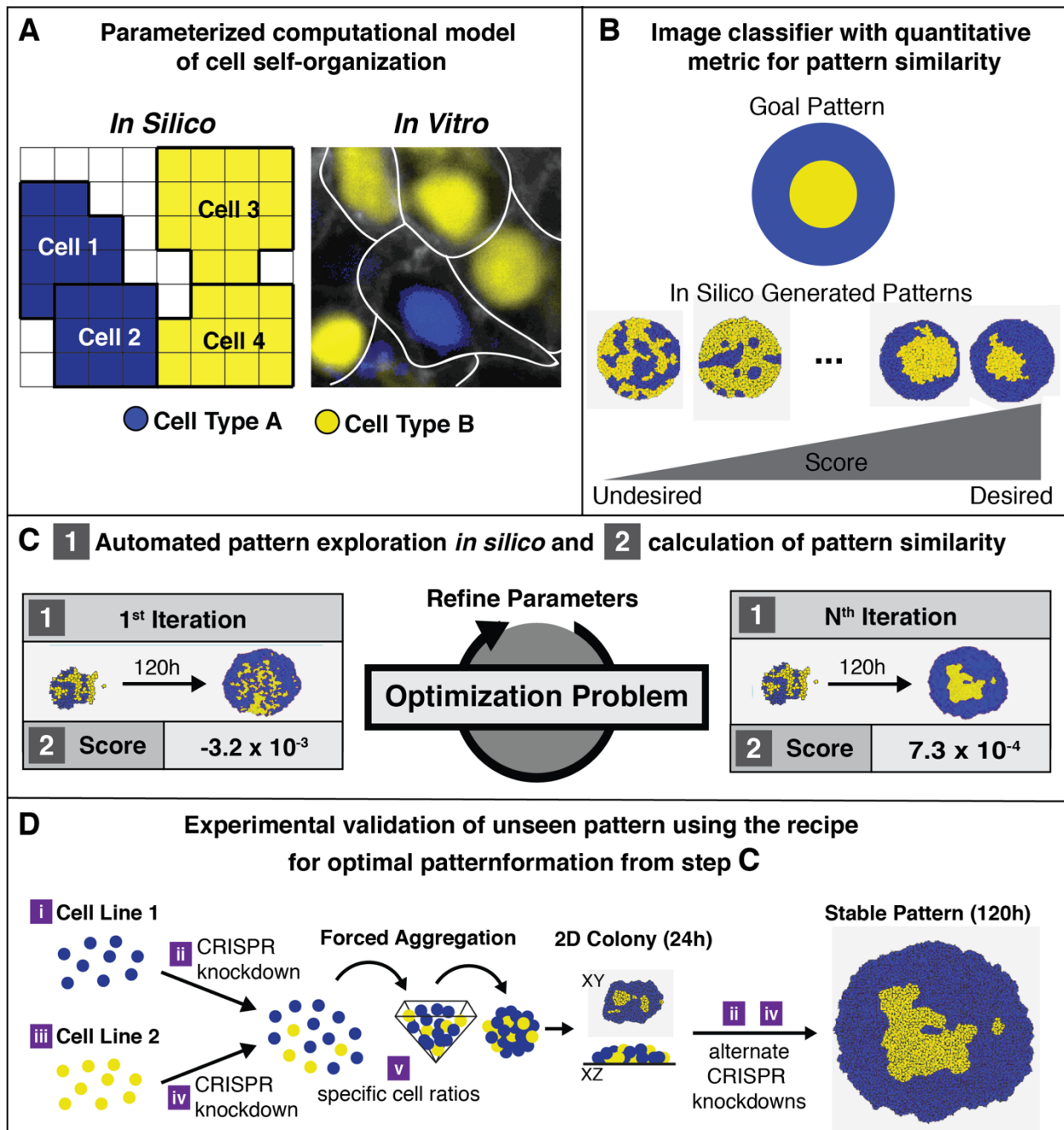


Figure 3.1: Overview of Synthesis of Spatial Patterns. Pattern Synthesis is a computational procedure used to predict the *in vitro* conditions needed to produce target spatial patterns. **A)** The first input to pattern synthesis is a parameterized computational model of mechanically driven spatial patterning in iPSC colonies. Five parameters of the computational model map directly to perturbations that can be made by experimentalists, and the output of the model was a series of images. **B)** The second input to pattern synthesis is a trained image classifier that produces a numerical score indicating the similarity of an image to the desired pattern class. In this scenario our desired pattern was a “Bullseye” pattern. **C)** Given the parameterized model and pattern classifier, Particle Swarm Optimization was used to explore parameter combinations, which map directly to *in vitro* perturbations, in order to identify the optimal conditions to produce the desired pattern *in silico*. **D)** Given the “recipe” of perturbations suggested by parameter optimization, we validate the control strategy is able to produce the desired pattern *in vitro*.

3.2 Materials and Methods

3.2.1 Cell Lines

All work with human iPSC lines was approved by the University of California, San Francisco Human Gamete, Embryo and Stem Cell Research (GESCR) Committee. Cell lines were derived from the parent line WTC (Coriell Cat. # GM25256) where the species of origin was confirmed by a LINE assay. All cell lines tested negative for mycoplasma using a MycoAlert Mycoplasma Detection Kit (Lonza). All human induced pluripotent stem cells (hiPSCs) were cultured at 37° C, seeded at a density of 12,000 cells per cm² in feeder-free media conditions on growth factor-reduced matrigel (BD Biosciences), and daily fed MTeSR™ medium (STEMCELL Technologies)(Ludwig et al., 2006). When hiPSC confluency reached 75%, cells were dissociated and singularized using Accutase (STEMCELL Technologies). Single cells were counted using a Invitrogen Countess Automated Cell Counter (Thermofisher Scientific), re-plated at previously described density, and in the first 24hrs after passaging, fed with MTeSR™ medium supplemented with the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10µM; Selleckchem) to promote survival (Park, Kim, Jung, & Roh, 2015; Watanabe et al., 2007).

3.2.2 Generation of CRISPRi knockdown iPSC lines

CRISPRi knockdown lines were previously generated as described in (Mandegar et al., 2016), where 20 base pair guides were designed using the Broad Institute sgRNA design website (Doench et al., 2016). 20 base pair sequences were cloned into the gRNA-CNKB vector using restriction enzyme BsmBI digestions, followed by ligation with T4 DNA ligase as described in (Mandegar et al., 2016). 200,000 cells of the CRISPRi-Gen1C or CRISPRi-Gen2 hiPSC lines from the Conklin Lab were nucleofected with individual gRNA vectors using the Human Stem Cell

Nucleofector Kit 1 solution with the Amaxa nucleofector 2b device (Lonza). Cells were then plated at increasing dilutions into 3 wells of a 6-well plate coated with growth factor-reduced matrigel (BD Biosciences) in MTeSR™ supplemented with Y-276932 (10µM) for 2 days. Then the nucleofected hiPSCs were treated with blasticidin (10µg/ml) for a selection period of 7 days. Surviving colonies for each gRNA were pooled and passaged in MTeSR™ with blasticidin (10µg/ml) and Y-27632 (10µM) for a single day then transitioned to MTeSR™ media only. After stable polyclonal populations of hiPSCs were established for each gRNA, cells were karyotyped by Cell Line Genetics (Libby et al., 2018)(Figure 3.2). Finally, knockdown efficiency was tested by the addition of doxycycline (2µM) to the culture media for 6 days and subsequent qPCR of mRNA levels of respective genes compared to time matched controls of the same line without CRISPRi induction.

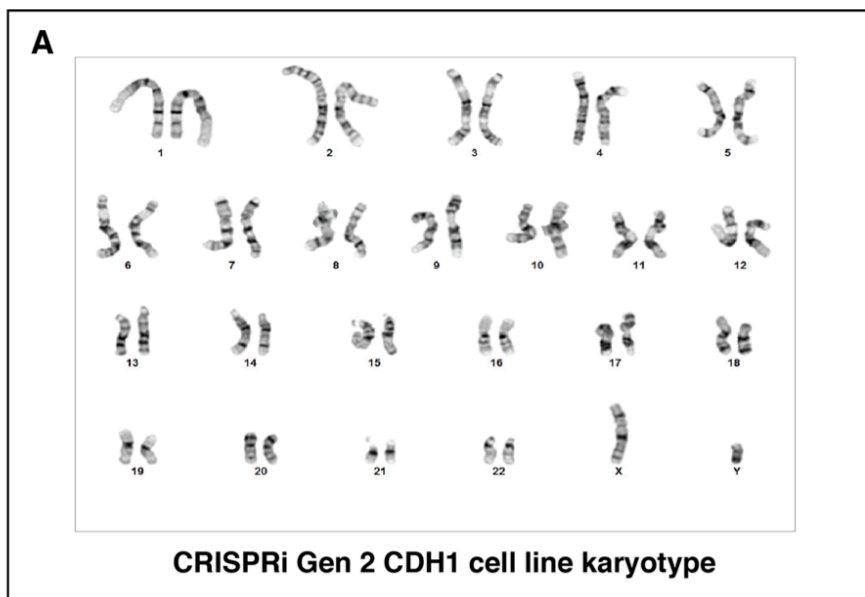


Figure 3.2: Karyotype of CRISPRi Gen2 CDH1 hiPS Cell line. (A) The Gen2 CDH1 hiPSC line was karyotypically normal.

3.2.3 Mixed Colony Generation

Mixed population hiPSC colonies were generated using forced aggregation via PDMS microwells in a 24-well tissue culture plate (~975 400X400µm wells per well)(Hookway, Butts, Lee, Tang, & McDevitt, 2016; Libby et al., 2018). hiPSCs were dissociated and singularized using Accutase (STEMCELL Technologies) and subsequently counted using an Invitrogen Countess Automated Cell Counter (Thermofisher Scientific). The proper ratios of cells to create 100 cell aggregates were then seeded into PDMS wells in MTeSR™ with Y-27632 (10µM), centrifuged at 200g for 5 minutes, and allowed to compact overnight (~18h). Aggregates were then washed out of the PDMS wells with fresh MTeSR™ and re-plated into a growth factor reduced matrigel (BD Biosciences) coated PerkinElmer CellCarrier™-96 plates at ~10/aggregates/cm² and fed daily with MTeSR™.

3.2.4 Immunofluorescence Staining and Imaging

Human iPSCs were fixed for 25 minutes with 4% paraformaldehyde (VWR) and subsequently washed 3 times with PBS. Fixed colonies were simultaneously blocked and permeabilized with a 1X PBS solution with 0.3% Triton X-100 (Sigma Aldrich) and 5% Normal Donkey Serum (Jackson ImmunoResearch) for 1 hour at room temperature. Samples were then incubated with primary antibodies overnight at 4°C in a 1X PBS solution with 1% Bovine serum albumin (Sigma Aldrich) and 0.3% Triton-X. Samples were washed 3 times and then incubated for 1 hour at room temperature with secondary antibodies. Primary antibodies used were: anti-OCT4 (SantaCruz 1:400), anti-SOX2 (AbCAM 1:400), and anti-Ecadherin (AbCAM 1:200). All secondary antibodies were used at 1:1000 and purchased from Life Technologies. Images were taken in one focal plane on the apical surface of hiPSC colonies.

Mixed colonies were imaged using a Zeiss Observer.Z1 (Zeiss) and an InCell Analyzer2000 (GE Healthcare) with a 10X objective, and confocal images were obtained using a Zeiss LSM880 Confocal w/ Airyscan (Zeiss) microscope with a 10X objective. Images were analyzed in ImageJ and in python using the skimage package (Walt et al., 2014).

3.2.5 Protein Quantification

Protein quantification for CDH1 KD was first quantified by immunofluorescence imaging of mixed colonies of WT-GFP hiPSCs and CDH1 KD colonies (Libby et al., 2018). Total fluorescence of CDH1 was measured by a python script that compared fluorescence of the CDH1 channel normalized to the amount of WT cells vs KD cells (determined by GFP fluorescence)(Fig. 3.3). This data was supplemented by Western blot data from the previously published KD of CDH1 and ROCK1 in (Libby et al., 2018).

3.2.6 mRNA quantification

The relative gene expression following CRISPRi knockdown was previously reported in (Libby et al., 2018) and used as a reference to establish knockdown timing curves used in our *in silico* simulations. As previously reported (Libby et al., 2018), total mRNA isolation from dissociated hiPSCs was performed using a RNeasy Mini Kit (QIAGEN) according to manufacturer's instructions and quantified with a Nanodrop 2000c Spectrometer (ThermoFisher). Obtained mRNA was then used to synthesize cDNA using an iScript cDNA Synthesis kit (BIORAD). A StepOnePlus Real-Time PCR system (Applied Biosciences) was used to quantify and detect gene expression by Fast SYBR Green Master Mix (ThermoFisher Scientific). Relative gene expression was determined by normalizing comparative threshold(Ct) values to the house keeping gene 18S rRNA. Gene expression was then displayed as a fold change comparison to

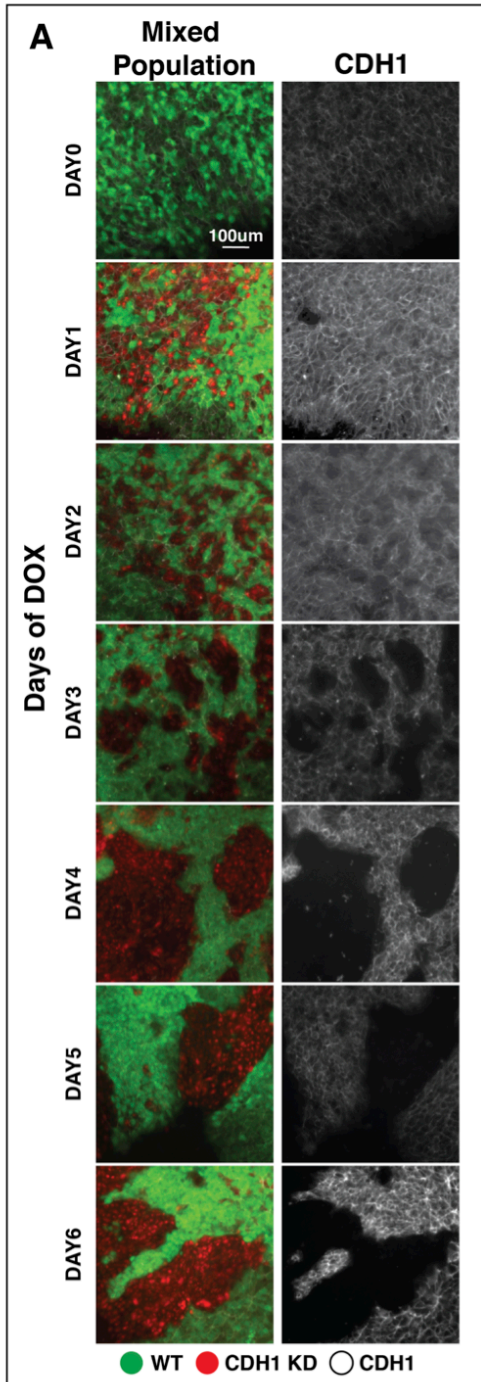


Figure 3.3: Knockdown of CDH1 quantified by immunofluorescence imaging (A) Mixed populations of wildtype (GFP+) and CDH1 KD cells (GFP-) were imaged daily and fluorescence of CDH1 levels was normalized to GFP fluorescence (n=10).

the day 0 control before the start of gene knockdown. The NCBI Primer-BLAST website was used to design the primers. Statistical analysis was conducted using a two-tailed unpaired t-test between any two groups ($p < 0.05$, $n = 3$).

3.2.7 Time Lapse Imaging

Mixed hiPSC colonies were imaged at the basal surface on optically clear PerkinElmer CellCarrier™-96 plates on an inverted AxioObserver Z1 (Zeiss) with an ORCA-Flash4.0 digital CMOS camera (Hamamatsu) with a 10X objective, where that single plane was used for parameter estimations. Using ZenPro software, colony locations were mapped and a single colony was imaged every 30 minutes over the course of 12 hours. Time lapse imaging occurred from hours 24-36 and from hours 96-108 after mixed colony plate down. The 12 hour series of images were then used to compare *in silico* to *in vitro* pattern formation and organization of cells. Additionally, mixed colonies of wildtype and CRISPRi-Gen1C without knockdown guides were imaged for 6 hours every 5 minutes with a 20X objective from hours 60-66 after plate down. These 6 hour image series were used to generate velocity values as previously described in section 3.2.13 (Velocity Characterization).

3.2.8 Comparison of *in vitro* and *in silico* Spatial Patterns

We used *in vitro* and *in silico* images to calculate the total number of cells in an image, the number of clusters, and the circularity of each cluster (Figure 3.4). Our work-flow for comparing patterns (Figure 3.5) involved splitting the images into single color RGB channels using the python module scikit-image (Walt et al., 2014).

For *in silico* images each channel represented a different cell type. After splitting the image into color channels we detected the number of islands in a colony. For *in silico* images, cells were

separated by a black border so we sequentially masked out the border, dilated the image, removed small objects, then removed small holes in the mask with scikit-image. Contiguous regions (8-connected) were considered clusters. We then overlaid a mask of individual cells onto each cluster using a logical AND comparison of the image masks to determine if the cell cluster met our criteria to be considered an island. Using only the cell clusters we considered islands, we then calculated the number of cells per island and the circularity of the islands using the formula $Circularity = 4\pi * Area/Perimeter^2$.

In contrast to the work-flow for simulation images, for *in vitro* images one channel represented all cell nuclei and the other channel represented cells stained for the protein CDH1, which delineated the CDH1 knockdown cells from the WT or ROCK1 knockdown cells. For the *in vitro* images, the CDH1 channel was thresholded and then dilated to create a CDH1+ cell mask followed by removal of small objects and holes to create a smooth segmentation. To generate the

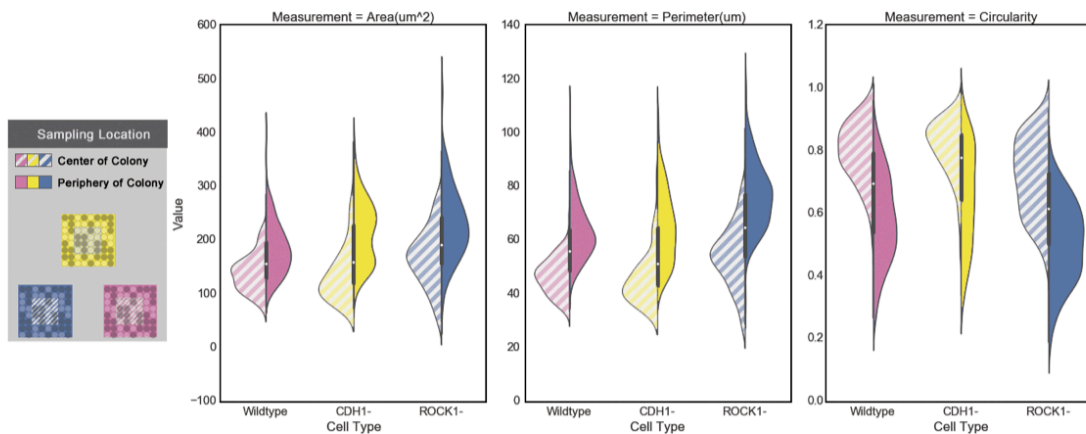


Figure 3.4: Characterizing *in vitro* morphological metrics. Violin Plots of single cell morphology (cell area, cell perimeter, cell circularity) for wildtype hPSCs, hPSC with CDH1 knocked down, and hPSC with ROCK1 knocked down in relation to the cell's position in the colony (n = 55 central, 55 edge per cell type). Periphery of the colony was defined as within 5 cell lengths from the edge of the colony. Cells have distinct distributions for both morphological measurements. The area and perimeter measurements with respect to each cell line were used to fit our computational model.

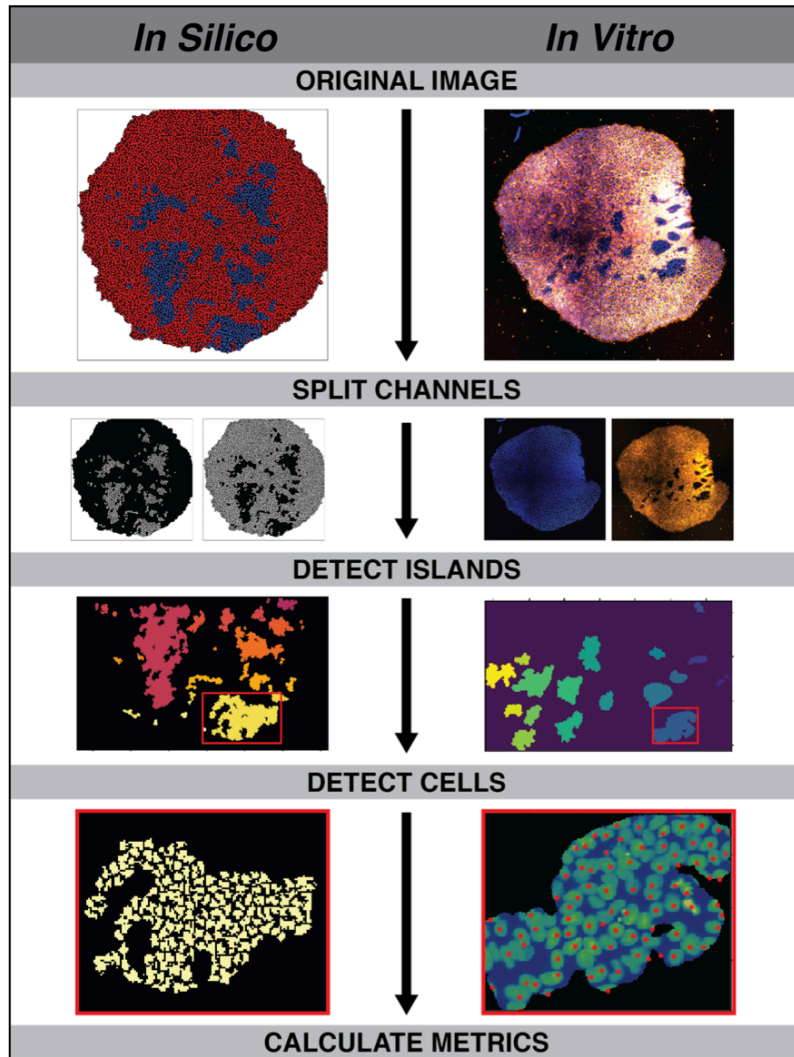


Figure 3.5: Segmentation workflow. *In silico* vs. *in vitro* image segmentation work flow to quantify and compare spatial patterns.

island masks, isolated CDH1 negative clusters were identified using the “label” function on the inverse of the CDH1+ mask. Individual cells were localized by detecting local maximum intensity in the DAPI channel images then the number of DAPI peaks per island were calculated using the logical AND of the island and CDH1 negative masks. Finally, we used the function "regionprops" to calculate the cluster area and perimeter for each island, which were then employed to calculate the island circularity with the above formula.

3.2.9 BMP4 differentiations

Successfully patterned hiPSC colonies were differentiated for 48h in MTeSR™ cell culture medium (STEMCELL Technologies) supplemented with BMP4 (R&D Systems) at a 50 $\mu\text{M}/\text{ml}$ concentration. The colonies were then fixed with 4% PFA for 25 min and subsequently analyzed.

3.2.10 Cellular Potts Model Environment

We modeled the mechanical properties of interacting human induced pluripotent stem cells (hiPSCs) with an extended Cellular Potts Model (CPM). In the model of mechanically driven self-organization in hiPSCs, cell—cell interaction mechanics were explained by four physical properties of cells. 1) cell-cell adhesion, 2) cortical tension, 3) conservation of volume, 4) and directionally persistent cell migration. Below, we describe how the extended CPM was used to recapitulate spatiotemporal patterns and predictively design de novo spatiotemporal behaviors.

We defined the environment of a CPM simulation S on a 2D square lattice domain $S \in \mathbb{Z}_+^2$. Each lattice site, $x = (m, n) \in S$, represented a coordinate location where $m \in \mathbb{Z}_+$ and $n \in \mathbb{Z}_+$ were the horizontal and vertical coordinates of each lattice site respectively. The spatial resolution of each lattice site was $1\mu\text{m}^2$ so that each square region of the grid is equal to 1 square micrometer.

To represent the location of hiPSCs, each lattice site x was assigned a value σ_x , conventionally called the *spin* or cell index (cell ID) of a site, from the set of cell indices $k \in K$ given $K = \{1, \dots, N(t)\}$ where $N(t)$ was the number of cells in the simulation at time t . Lattice sites that represent empty space where there is no hiPSC covering the lattice site were assigned a cell index of 0. In the CPM, a cell C_k was composed of multiple lattice sites where each lattice site represents a partial region of a cell or the surrounding media. A cell C_k was defined as the set of lattice sites with the same cell ID $C_k = \{x \in S: \sigma_x = k\}$. Since a single cell was composed of

multiple lattice sites, the CPM was able to capture fluctuations in a cell's shape with a granularity that is not possible with Type A cellular automaton or center-based models . Each cell was also assigned a cell type τ to notate the type of genetic perturbation (i.e. knockdown) of that cell which determined its intracellular and extracellular behaviors.

Next, we summarize two common metrics to describe cell morphology in a CPM simulation; cell area and cell membrane length. These metrics are important since their values in the model were directly measured from microscopy images. For a discussion of these metrics see .

Given that each lattice site had an area of $1\mu\text{m}$, the area of a cell at time t in the simulation was defined as the number of lattice sites encompassed by a cell:

$$a_{k,t} = |\{x \in S : \sigma_x = k\}|, \quad (\text{S1})$$

where $|\cdot|$ denoted the cardinality of a set. The time varying membrane length of a cell $p_{k,t}$, synonymously called the perimeter or surface length in other studies , was defined as the number of lattice interfaces bordering other cells or empty space:

$$l_{k,t} = 0.5 \times \sum_{\text{interfaces}\{x,x'\}} \delta(k, \sigma_x), \quad (\text{S2})$$

where x' represented any of the lattice sites adjacent to x , $(m \pm 1, n) \vee (m, n \pm 1)$ in 2D. The Kronecker symbol δ was defined by $\delta(u, v) = 1$ if $u = v$ and $\delta(u, v) = 0$ if $u \neq v$. An interface (x, x') was a shared border between lattice sites. To avoid counting adjacent lattice sites inside a cell, the CPM only summed interfaces between lattice sites with different cell ID's; when $\delta(\sigma_x, \sigma_{x'}) = 0$. Put simply, we were measuring the perimeter of each stem cell.

3.2.11 Cellular Potts Model Dynamics

The CPM uses a function called the **Hamiltonian** H to describe the energy (favorable behaviors) for any *configuration* of cells. Cell motility evolved by choosing a random lattice site x , a region of a cell-cell interface or a cell-media interface and attempted to copy it to a random neighboring lattice site x' . The Hamiltonian was defined as the sum of four constraints that represent four physical properties of simulated stem cells: 1) conservation of cell area, 2) locally polarized cell migration, 3) cell-cell adhesion, 4) and cell membrane length which commonly represents cortical tension. In the CPM, the goal was to minimize the Hamiltonian or minimize violations of the desired cellular behaviors. Therefore, each constraint calculated a decrease (reward) or increase (penalty) in the configuration energy due the collective properties of cells in the simulation.

When a change in a lattice site was proposed, this affects H . If the proposed change was accepted, the change in H was defined as ΔH . A proposed change for a cell's lattice site was accepted with the following probability:

$$\text{if } \Delta H < -Y, P(\sigma \rightarrow \sigma') = 1 \quad (\text{S3})$$

$$\text{otherwise, } P(\sigma \rightarrow \sigma') = e^{-(\Delta H + Y)/T}, \quad (\text{S4})$$

where the yield $Y = 0.1$ and the temperature $T = 10$. Simply, if the proposed change in local cell position resulted in less energy, then the change was accepted. If the proposed update would have resulted in greater energy (ΔH), then the change was only accepted with a very low probability. In this way, complex behaviors such as preferential cell-cell adhesions, cortical tension, and cell migration, are represented by a score and a weight, where the score represents a reward or penalty depending on the divergence of a cell from its target behavior, while the weight represents the relative importance of the respective cell behavior.

CPM Configuration Energy: The free energy for a configuration of cells was defined as the sum of four constraints: local cell-cell/cell-ECM adhesion, cell area conservation, cell membrane length, and locally polarized cell migration:

$$H = H_{adhesion} + H_{area} + H_{membraneLength} + H_{migration} \quad (S5)$$

For a configuration of cells, the free energy due to cell adhesion was

$$H_{adhesion} = \sum_{k \in K} J_{\tau(\sigma_x), \tau(\sigma_{x'})} (1 - \delta(\sigma_x, \sigma_{x'})), \quad (S6)$$

where $J_{\tau(\sigma_x), \tau(\sigma_{x'})}$ represented the cell adhesion strength between lattice sites σ_x and $\sigma_{x'}$ that was defined by their cell type $\tau(\sigma_{x'})$. $(1 - \delta_{\sigma_x, \sigma_{x'}})$ restricted these calculations to interfaces between cells instead of all lattice sites and improved the efficiency of the simulation. Although not explicit in our notation, the cell adhesion strength was a time-dependent function controlled by protein expression to mimic changes in cell behavior with inducible gene knockdown. The energy due to cells resisting changes from their resting area was defined as

$$H_{area} = \sum_{k \in K} \lambda_a (a_{k,t} - A_{k,t})^2, \quad (S7)$$

where $A_{k,t}$ represented the target area of a cell, $a_{k,t}$ represented the current area of a cell, and λ_a was the relative strength of area conservation term.

The cortical tension constraint was defined as:

$$H_{membraneLength} = \sum_{k \in K} \lambda_l (l_{k,t} - L_{k,t})^2, \quad (S8)$$

where $l_{k,t}$ represented the current membrane length of a cell at time t , $L_{k,t}$ was the target membrane length, and λ_l was the strength of the cortical tension constraint. As a proxy for increasing or decreasing the cell membrane length, the Equivalent Circular Perimeter (ECP) was used to set the membrane length for a cell given its current area. The ECP of a non-circular 2D object was defined as the perimeter of a circle with equivalent surface area as the non-circular object:

$$ECP(k) = 2 \sqrt{a_{k,t}\pi}.$$

The target membrane length was calculated using a membrane length proportionality constant:

$$L_{k,t} = r_k ECP(k), \quad (S9)$$

where r_k was the membrane proportionality constant. To find r_k the membrane length and area of cells were measured and divided by the ECP of the cell. This ratio of membrane length to ECP was equal to the membrane proportionality constant r_k . The ECP allowed us to calculate the membrane length of a cell of any area that would have a comparable shape to empirical measurements.

To capture directionally persistent cell migration, we modeled "polarized cell migration" as the tendency of cells to bias their movement in the same direction as their previous direction of movement as described in . Cells had a target direction \vec{t} based on previous movement where CPM updates in this direction were preferred (they decreased the energy in H). For each copy attempt $x \rightarrow x'$, the cell center was displaced in direction \vec{s} . The change in energy due to migration in this direction was defined as:

$$H_{migration} = -\mu a_k (\vec{t} \cdot \vec{s}), \quad (S10)$$

where μ was the strength of cell migration, $a_{k,t}$ was the cell area at time t , \vec{t} was a unit vector giving the target direction, and \vec{s} was a unit vector giving the current direction of a stem cell if the CPM update ($x \rightarrow x'$) was to be accepted. The function was multiplied by -1 since updates in the direction of \vec{t} have a dot product that approached +1 as the angle between \vec{t} and \vec{s} approached zero. Multiplying by -1 resulted in decreased configuration energy for cells moving in the same direction as the target direction vector.

For every MCS, the target direction at any time (\vec{t}_t) was updated continuously given the displacement of a cell's centroid $w\Delta O = \mathbf{O}_t - \mathbf{O}_{t-1}$ then transformed into a unit vector $\vec{s} =$

$\Delta\theta/|\Delta\theta|$. This target direction included how 'decay-time' D of the previous direction and the current cell displacements contributed to the current polarity of the cell:

$$\vec{t}_t = (1 - D)\vec{t}_{t-1} + D\vec{s}. \quad (\text{S11})$$

3.2.12 Physical Units and Other Cellular Phenomena of Cellular Potts Model

Cell division was symmetrical (the parent cell divided into 2 equally sized daughter cells), and the timing of cell division was asynchronous. This was achieved by assigning a uniformly distributed "division counter" d_c for each cell at $t = 0$ between 0 and the division time d_t . This counter was incremented at each time step of the simulation, and a cell would divide when $d_c = d_t$. d_c was then reset to 0 for both daughter cells. Cell division was assumed to be asynchronous amongst the population, and cell division times specific to each type of knockdown were incorporated into the model to provide an accurate depiction of population growth kinetics. Cell division times were calculated from *in vitro* doubling rates and modeled to be 18 hours for CDH1(-) cells, and 20 hours for all other cells.

3.2.13 Model Fitting to Empirical Data

In the main text we provide a brief explanation of the characterization experiments to fit our computational model. Here we describe the mathematical transformations, mapping functions, and model parameters associated with these characterization experiments. The parameters fit during this process are summarized in Table 3.1.

Morphology Characterization: Three types of colonies were characterized; purely wildtype, wildtype and CDH1 knockdown in a 1:1 ratio, and wildtype:ROCK1 knockdown in a 1:1 ratio. We measured the cell area, perimeter, and ECP at the center and periphery of colonies (Figure 3.4). The median cell area was used to set the target cell area ($A_{k,t}$) in our simulations (Figure 3.4,

Table 3.1). The median membrane proportionality constant ($r_k = \text{perimeter} / \text{ECP}$) was used to set the target membrane length $L_{k,t}$ in our simulations (Figure 3.4, Table 3.2).

Table 3.1 : List of design parameters that map to experimental perturbations. CL1 and CL2 are chosen from our library of mechanically tunable cell lines: CDH1-0, CDH1-70, CDH1-75, CDH1-90, ROCK1-20, Wildtype. The number following the cell line indicates the relative expression of the gene in comparison to WT. The knockdown times of CL1 and CL2 range from 120 hours before co-culture to 120 hours after co-culture in 24 hour increments. Finally, the abundance of CL2 cells in relation to CL1 can be vary from 5% to 95% of the colony in increments of 5%.

Design Parameter	Parameter Domain
Cell Line 1 (CL1)	{CDH1-0,CDH1-70,CDH1-75,CDH1-90,ROCK1-20, WT}
Cell Line 2 (CL2)	{CDH1-0,CDH1-70,CDH1-75,CDH1-90,ROCK1-20, WT}
Knockdown Time of CL1	[-120 hours, 120 hours]
Knockdown Time of CL2	[-120 hours, 120 hours]
Abundance of CL2	[5%, 95%]

Table 3.2: Model fitting parameters for the Space vs. Time, Protein Expression vs. Time, and the Protein Expression vs. Space characterizations. The partial knockdowns of CDH1 (Figure 3.6) used a response curve inferred to follow the same trend as the complete CDH1 knockdown (Figure 3.7F).*= since these knockdowns were different efficiencies of the CDH1 response curve the adhesion strength and cortical tension was inferred from the complete knockdown.

Morphology Characterization	Morphological Characterization	
Empirical Measurement	Model Parameter	Definition
Median cell area	A(k,t)	Target area of a cell.
Median cell membrane length;ratio of cell perimeter divided by Equivalent Circular Perimeter	r(k,t)	membrane length proportionality constant
Space vs Time (Velocity Characterization)	Velocity Characterization	
Empirical Measurement	Model Parameter	Definition
median velocity (0.29 um/minute)	MCS	monte carlo step size
Empirical Measurement	Model Parameter	Definition
median velocity (0.29 um/minute)	J(WT,WT)	adhesion strength
	u	strength of self-propulsion
	lambda_l	strength of cortical tension
	D	decay of velocity direction bias
Protein Expression vs Time (Temporal CRISPRi Knockdown Characterization)	Knockdown Characterization	
Empirical Measurement	Model Parameter	Definition
median protein expression after knockdown	K_m	repression coefficient
median protein expression after knockdown	n	hill coefficient

Protein Expression vs Space (Spatial Pattern Characterization)		Spatial Pattern Characterization	
Empirical Measurement		Model Parameter	Definition
spatial organization from knockdown of CDH1		$J(a,a)$	adhesion energy
spatial organization from knockdown of ROCK1		$\lambda_{a,l}$	strength of cortical tension

Morphology Characterization	
Empirical Measurement	Interpretation in the CPM
Median cell area	Lattice copy attempts to grow or shrink a cell to its target area are favorable
Median cell membrane length;ratio of cell perimeter divided by Equivalent Circular Perimeter	Lattice copy attempts to grow or shrink the cell perimeter to its target size are favorable. Decreasing the value $r(k,t)$ decreases the cell membrane length. Since there is less membrane to generate cell protrusions, cell migration also decreases.

Space vs Time (Velocity Characterization)	
Empirical Measurement	Interpretation in the CPM
median velocity (0.29 um/minute)	
Empirical Measurement	The time in hours for each lattice site of a membrane interface to make one copy attempt. This parameter relates to the maximum migration velocity
median velocity (0.29 um/minute)	Interpretation in the CPM adhesion energy per unit interface between wildtype stem cells. Weights how favorable cell migration is in comparison to maintaining the target cell area, membrane length, or maintaining cell adhesion contacts. regulates the ability of the cell membrane make protrusions and retraction. This affects the cell migration velocity The decay of the directional cell migration bias

Protein Expression vs Time (Temporal CRISPRi Knockdown Characterization)			
Empirical Measurement		Interpretation in the CPM	
median protein expression after knockdown		time (hours) when expression reaches half maximal expression	
median protein expression after knockdown		steepness of Hill Function	

Protein Expression vs Space (Spatial Pattern Characterization)	
Empirical Measurement	Interpretation in the CPM
spatial organization from knockdown of CDH1	steady state adhesion energy per unit interface between like stem cells after knockdown
spatial organization from knockdown of ROCK1	regulates the ability of the cell membrane make protrusions and retraction. This affects the cell migration velocity

Morphology Characterization							
Empirical Measurement (center/edge)	Wildtype	ROCK1-20	CDH1-0	CDH1-70	CDH1-75	CDH1-90	
Median cell area	137.40/ 177.17	169.94/ 223.42	121.13/2 23.42	121.13/ 223.42	121.13/ 223.42	121.13/ 223.42	121.13/ 223.42
Median cell membrane length;ratio of cell perimeter divided by Equivalent Circular Perimeter	1.12/ 1.32	1.17/ 1.41	1.10/ 1.23	1.10/ 1.23	1.10/ 1.23	1.10/ 1.23	1.10/ 1.23
Space vs Time (Velocity Characterization)							
Empirical Measurement	median velocity (0.29 um/minute)	(0.29 Global Parameter					
Empirical Measurement	median velocity (0.29 um/minute)	0.03					
		Wildtype	ROCK1-20	CDH1-0	CDH1-70	CDH1-75	CDH1-90
		-100	-	-	-	-	-
		6	6	6	6	6	6
		0.75	-	-	-	-	-
		0.003	0.003	0.003	0.003	0.003	0.003
Protein Expression vs Time (Temporal CRISPRi Knockdown Characterization)							
Empirical Measurement	Wildtype	ROCK1-20	CDH1-0	CDH1-70	CDH1-75	CDH1-90	
median protein expression after knockdown	-	46.78	48.28	48.28	48.28	48.28	48.28
median protein expression after knockdown	-	5	5	5	5	5	5
Protein Expression vs Space (Spatial Pattern Characterization)							
Empirical Measurement	Wildtype	ROCK1-20	CDH1-0	CDH1-70	CDH1-75	CDH1-90	
spatial organization knockdown of CDH1	from	-100	-95	-50	*-90	*-87.5	*-95
spatial organization knockdown of ROCK1	from	0.75	2	0.75	0.75	0.75	0.75

Velocity Characterization (Space vs. Time): We characterized space vs. time by measuring the velocity (change in distance over time) of wildtype cells in dense colonies. Mixed aggregates of 90% WT and 10% CRISPRi cells without a targeting guide were generated. With the addition of doxycycline (DOX) to the cell culture media, the CRISPRi no guide population expressed a cytoplasmic mCherry marker which allowed individual cells to be distinguished from the untagged

WT background (Figure 3.7B). 24 colonies were imaged for 6 hours at 5 minutes/image at 20X magnification creating a time series of 73 frames. Each frame was individually normalized and thresholded using non-local means. Cell migration tracks were generated by following matching contours between frames where matching contours share at least ten pixels overlap. We used watershed segmentation to separate adjacent cells. Instantaneous frame to frame velocity was calculated as

$$v_{inst} = [(x_{cm,2} - x_{cm,1})/\Delta t, (y_{cm,2} - y_{cm,1})/\Delta t], \quad (S12)$$

where $x_{cm,2}$ was the center of mass of each segmented cell body at the currently observed frame and $x_{cm,1}$ was the center of mass of each segmented cell body at the previous frame, and Δt was 5 minutes. Taking the average magnitude of the per-cell instantaneous velocity over 24 colonies gave a median velocity of 0.29 $\mu\text{m}/\text{minute}$.

We then ran parameter sweeps to fit model parameters that affect cell migration (Table 3.2):

- MCS - copy attempts per simulation hour.
- $J_{WT,WT}$ - adhesion energy or reward per micrometer of cell border between wildtype cells.
- μ - strength of self-propulsion
- λ_l - strength of cortical tension.

We chose the parameter combination where the simulation velocity distribution matched the empirical velocity distribution and remained a dense colony (0.34 $\mu\text{m}/\text{minute}$). It is important to note that we chose optimal model parameters using the distribution of cell velocities and not the median cell velocity. We ran 24 simulations to mimic the experimental setup of the *in vitro* characterization. Using the Mann Whitney U test, there was no significant difference in the distribution of cell velocities; p-value threshold of 0.05 and p-value for 24 *in silico* colonies was 0.051. After fitting the model to empirical data of cell morphology and velocity, we could recapitulate the cell morphology and collective cell migration of wildtype stem cell colonies without genetic modulation.

Temporal Knockdown Characterization (Protein Expression vs. Time) - CDH1: We characterized the time dependent knockdown of CDH1 expression which was responsible for changes in cell-cell adhesion (J). CDH1 was knocked down using CRISPRi, and the relative expression was measured for 6 consecutive days. The relative mRNA expression of CDH1 was quantified by quantitative PCR (n=3) and protein expression of CDH1 was measured by immuno-fluorescence microscopy (n=10), which displayed a 24 hour delay from the mRNA knockdown. The data was min-max normalized to a domain of [0,1] using the median expression for each day:

$$y(t_k)' = (y(t_k) - y_{min}) / (y_{max} - y_{min}), \quad (S13)$$

where t_k was the time since the knockdown, $y(t_k)$ was the expression at k hours after knockdown, y_{max} was the max expression from all days, and y_{min} was the minimum expression over all days. It is important to note that t_k is the time since knockdown and t is time since the initiation of co-culture experiments. Using least squares optimization (Python `scipy.optimize.curve_fit` function), the K_m (repression coefficient) and n (Hill coefficient) of the Hill Function for repression were fit to the normalized median expression to create a response function using least squares optimization:

$$F(t) = 1 / (1 + (K_m / t_k)^n), \quad (S14)$$

where K_m was the time half expression occurs, and n was the hill co-efficient. Functions were fit to the normalized median expression (per day) to create a time-dependent response function for CDH1 knocked down to 90%, 75%, 70%, and 2% of the original mRNA expression (Figure 3.7F, Figure 3.6). This range of knockdown efficiencies allowed us to computationally model how differing levels of CDH1 expression could impact spatial patterning. Given value for the parameters k_m and n (Table 3.2) we now had a continuous response function for the expression of CDH1 given a knockdown time that we could modulate. It is important to note that normalizing the relative expression to a domain of [0,1] allows us to stretch the response function to different parameter ranges in the spatial pattern characterization experiments.

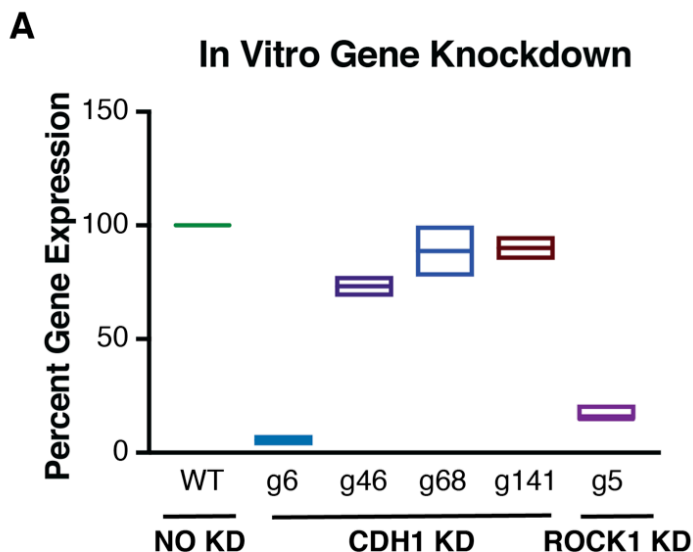


Figure 3.6: CRISPRi cell line mRNA knockdowns (A) Percent knockdown of either CDH1 or ROCK1 when compared to wildtype control populations on day six of DOX treatment (n=3).

Temporal Knockdown Characterization (Protein Expression vs. Time) - ROCK1: Using the same approach as the CDH1 knockdown, we created a Hill response function for ROCK1 knocked down to 20% mRNA expression. ROCK expression is represented by the strength of cortical tension (λ_t) parameter in our computational model. We assumed mRNA expression changed 24 hours ahead of protein expression, so we shifted the time axis forward by one day to account for the delay. The median expression for each day was min-max normalized to a domain of [0,1] (Equation S13). The K_m (repression coefficient) and n (Hill coefficient) of the hill function for repression were fit to the normalized median expression to create a response function using least squares optimization (Equation S14, Figure 3.7, Table 3.2). Due to the delay in protein knockdown compared to mRNA levels, the Hill functions were shifted by 24 hours to account for the delay in protein loss (Figure 3.7G), allowing us to model the average change in ROCK1 protein expression for individual cells over time.

Spatial Pattern Characterization (Protein Expression vs. Space): Given the previous characterization experiments, we were able to model cell proliferation, cell morphology, wildtype

cell migration, and temporal changes in the expression of CDH1 and ROCK1. However, the time-dependent modulation of protein express had to be mapped to changes in multi-cellular self-organization. To model the time-dependent modulation of cell-cell adhesion via CDH1 and cortical tension via ROCK1, fluorescent microscopy images were collected 96 hours after mixing either ROCK1 KD or CDH1 KD cells with wildtype hiPSCs. Then *in silico* parameter sweeps were run overlaying a range of parameters controlling the strength of adhesion or membrane stiffness. These two parameters rescaled their respective Hill Functions and produced a range of spatial patterns due to progressively weaker cell-cell adhesion or progressively stiffer cell membrane parameter values. We then conducted double-blind experiments to fix adhesion strength and membrane stiffness parameters which most closely matched *in vitro* spatial patterning for CDH1 and ROCK1 knockdowns respectively (Figure 3.7H-I).

$$F'(t) = y_{knockdown} + (y_{wildtype} - y_{knockdown}) * (1 + (K_m/t)^n), \quad (S15)$$

where $y_{knockdown}$ was the adhesion strength (J) or target membrane length (L_k, t) of knockdown cell lines in the model, and $y_{wildtype}$ was the adhesion strength (J) or target membrane length of wildtype cells in the model. In Equation 15, we scaled the normalized response function from $[0, 1]$ to the range of model parameters $[y_{knockdown}, y_{wildtype}]$. Given the characterization experiments of cell morphology, cell migration velocity, time-dependent modulation of cell mechanics, and the resulting spatial organization, the computational model was able to recapitulate the spatial patterning due to the CDH1 and ROCK1 knockdowns.

3.2.14 TSSL Scoring and Pattern Optimization

In order to automatically compare patterns produced by the model from different parameterizations and determine optimal parameter values, we needed a measure capable of

quantifying how close any given pattern was to the desired one. A very effective algorithm was proposed in for this purpose.

Quad-Tree Representation of an Image: Consider an RGB representation of an $m \times n$ image as the matrix A where the element $a_{ij} = \langle a_{ij}^{(r)}, a_{ij}^{(g)}, a_{ij}^{(b)} \rangle$ is the normalized RGB values for the pixel located on the i th row and j th column of the image. Thus,

$$0 \leq a_{ij}^{(c)} \leq 1 \text{ for } c \in \{r, g, b\}.$$

Given a matrix A , $A[i_s, i_e; j_s, j_e]$ was used to denote the submatrix created by selecting rows with indices from i_s to i_e and columns from j_s to j_e . Following, we represented the matrix A as a quad-tree. A quad-tree $Q = (\mathcal{V}, R)$ is a quaternary tree representation of matrix A where each vertex $v \in \mathcal{V}$ represents a submatrix of A and the relation $R \subset \mathcal{V} \times \mathcal{V}$ defines four children for each vertex that is not a leaf.

Figure 3.8Ai-ii demonstrates how a quad-tree is built from a matrix. In this figure, we label each edge in the quad-tree with the direction of the sub-matrix represented by the child: north west (NW), north east (NE), south west (SW), and south east (SE). In Figure 3.8Aii:

- v_0 represents the complete matrix A at quadrant level 1.
- v_1 represents the first quadrant of level 2 or $A[1, \lfloor m/2 \rfloor; 1, \lfloor n/2 \rfloor]$, where m is the total number of rows and n is the total number of columns in A .
- v_2 represents $A[\lfloor m/2 \rfloor + 1, m; 1, \lfloor n/2 \rfloor]$.
- v_3 represents $A[\lfloor m/2 \rfloor + 1, m; \lfloor n/2 \rfloor + 1, n]$.
- v_4 represents $A[1, \lfloor m/2 \rfloor; \lfloor n/2 \rfloor + 1, n]$.
- v_5 represents represents the first quadrant of level 3 or $A[1, \lfloor m/4 \rfloor; 1, \lfloor n/4 \rfloor]$,
- etc.

We used the procedure described in (Briers et al., 2016) to construct quad-trees, which is slightly different from (Bartocci et al., 2016). The assumption was made that A has a size of

$2^k \times 2^k$ so that each submatrix could be divided into four equal-sized partitions. Here, we relaxed this requirement by allowing non-equal submatrices to be children of a node. Furthermore, defined a leaf as a vertex of the quad-tree for which all the elements of a submatrix had the same values. While this approach works perfectly for the 32×32 network that is studied in that paper, it can be problematic for larger images since the number of vertices in a quad-tree grows exponentially as more levels are added to it. In this paper, we constructed quad-trees with a fixed depth of 5, regardless of the size and other characteristics of A .

The representation function $\mu^{(c)}(v): \mathcal{V} \rightarrow [0, b] \times [0, b]$ was defined for sub-matrix $A[i_s, i_e; j_s, j_e]$ represented by vertex $v \in \mathcal{V}$ of the quad-tree $Q = (\mathcal{V}, R)$ as follows :

$$\begin{aligned} \mu^{(c)}(v) &= (\mu_1^{(c)}, \mu_2^{(c)}) \\ \mu_1^{(c)} &= \frac{1}{(i_e - i_s + 1)(j_e - j_s + 1)} \sum_{i,j \in \{i_s, \dots, i_e\} \times \{j_s, \dots, j_e\}} a_{ij}^{(c)}, \\ \mu_2^{(c)} &= \frac{1}{(i_e - i_s + 1)(j_e - j_s + 1)} \sum_{i,j \in \{i_s, \dots, i_e\} \times \{j_s, \dots, j_e\}} (a_{ij}^{(c)} - \mu_1^{(c)})^2, \end{aligned} \quad (\text{S16})$$

where $c \in \{r, g, b\}$ was an RGB color. The function $\mu^{(c)}$ provided the mean value and variance for the concentration of RGB colors in a particular region of the space represented by the vertex v .

Quad-trees can be interpreted as multi resolution representation of images, as the nodes that appear in deeper levels provide statistical information for higher resolutions and nodes that appear on higher levels correspond to more global characteristics of an image.

Tree Spatial Superposition Logic: In, a formal logic, called tree spatial superposition logic (TSSL), was introduced. TSSL is capable of formally specifying global patterns in a network of locally interacting agents. The authors showed that this logic is sophisticated enough to describe complicated patterns such as Turing patterns in biochemical reaction-diffusion systems. In this paper, we used this logic to express various patterns that are studied here (Figure 3.9). First, we present a brief introduction to TSSL. The reader can refer to for a thorough explanation of this logic, definitions of syntax and semantics, and its properties.

A TSSL formula is recursively constructed using the following:

- Linear predicates over valuations for the representation function (Equation S16). For example: $\mu_1^{(r)} > 0.8$ or $\mu_1^{(b)} < 0.5$.
- Boolean operators, such as $\neg\phi$, $\phi_1 \wedge \phi_2$, and $\phi_1 \vee \phi_2$.
- Spatial operators: $\exists_B \circ \phi$, $\forall_B \circ \phi$, where B is a nonempty subset of the set of directions $\{NW, NE, SW, SE\}$.

The spatial operators $\exists_B \circ$ and $\forall_B \circ$ are read as *there exists in directions B next* and *for all directions B next*, respectively. $\exists_B \circ \phi$ is interpreted as follows: For at least one of the nodes located in the next level of the quad-tree labeled with one of the directions in B , ϕ must be satisfied. $\forall_B \circ \phi$ specifies that for all such nodes ϕ must be satisfied. We demonstrate how TSSL can be used to express spatial patterns through an example.

Consider a 4×4 pattern as illustrated in Figure 3.8A. This pattern can be expressed as the following TSSL formula φ . A portion of the quad-tree satisfying this formula is shown in Figure 3.8Aii.

$$\begin{aligned} \varphi = & \forall_{\{NE,SW,SE\}} \circ [\forall_{\{NW,SE\}} \circ \varphi_{\text{white}} \wedge \forall_{\{NE,SW\}} \circ \varphi_{\text{colored}}] \wedge \\ & \forall_{NW} \circ [\forall_{NW} \circ \varphi_{\text{white}} \wedge \forall_{\{NE,SW\}} \circ \varphi_{\text{blue}} \wedge \\ & \quad \forall_{SE} \circ (\forall_{\{NW,SE\}} \circ \varphi_{\text{white}} \wedge \forall_{\{NE,SW\}} \circ \varphi_{\text{blue}})], \end{aligned} \quad (\text{S17})$$

where

$$\begin{aligned} \varphi_{\text{white}} &= \mu_1^{(r)} = 1 \wedge \mu_1^{(g)} = 1 \wedge \mu_1^{(b)} = 1, \\ \varphi_{\text{colored}} &= \mu_1^{(r)} < 1 \vee \mu_1^{(g)} < 1 \vee \mu_1^{(b)} < 1, \\ \varphi_{\text{blue}} &= \mu_1^{(r)} = 0 \wedge \mu_1^{(g)} = 0 \wedge \mu_1^{(b)} \geq 0.5. \end{aligned}$$

TSSL formulas can be viewed as formal pattern descriptors or pattern classifiers. For instance, the formula of Equation S17 accepts a quad-tree derived from a checkerboard pattern and rejects any other quad-tree. Although TSSL is capable of describing complicated spatial behaviors in an image, it is difficult in general to write a formula that describes a complex pattern. In, the authors

proposed to use machine learning techniques in order to find such a formula from a given set of positive and negative examples.

Assume a set of positive images (\mathcal{Y}_+), illustrating a desirable pattern, and a set of negative images (\mathcal{Y}_-), in which the desirable pattern was not present, were available. We created a set \mathcal{L} from these images as:

$$\mathcal{L} = \{(Q_y, +) \mid y \in \mathcal{Y}_+\} \cup \{(Q_y, -) \mid y \in \mathcal{Y}_-\},$$

where Q_y was the quad-tree generated from image y . The set \mathcal{L} was separated into a learning set \mathcal{L}_L (used to train a classifier) and a testing set \mathcal{L}_T (used to test the classifier obtained from \mathcal{L}_L) such that $\mathcal{L} = \mathcal{L}_L \cup \mathcal{L}_T$. A rules-based learner called RIPPER was used to learn a set of classification rules from \mathcal{L}_L . Each of these rules was in the form:

$$R_i: C_i \Rightarrow Label_i,$$

where C_i was a Boolean formula over linear predicates over the representation values of the nodes of a quad-tree and $Label_i \in \{+, -\}$. We used the Weka workbench for implementing RIPPER. Each C_i was then translated into an equivalent TSSL formula Φ_i . Since the classification rules were interpreted as nested if-else statements, the TSSL formula equivalent to the entire set of classification rules corresponding to the positive class was written as:

$$\Phi_+ = \bigvee_{j \in R_+} (\Phi_j \wedge \bigwedge_{i=1, \dots, j-1} \neg \Phi_i), \quad (S18)$$

where R_+ was the set of indices of rules labeled positive.

Quantitative Robustness: A TSSL formula can be created for any desired spatial pattern by following the procedure described in the previous section. If this formula is evaluated as true for a given image, it means that the image contains the required pattern. On the other hand, a false evaluation of the formula means that the pattern does not exist. However, this qualitative evaluation of TSSL descriptors does not provide any information about how strongly an image demonstrates the required pattern.

In order to provide information about how strongly an image satisfies or violates the given property, TSSL was also equipped with a recursive quantitative semantics definition which assigned a real value to a TSSL formula ϕ with respect to vertex $v \in \mathcal{V}$ of quad-tree $Q = (\mathcal{V}, R)$; denoted by $\rho(\phi, v)$. The TSSL quantitative valuation was derived recursively as follows:

- $\rho(\mu_i^{(c)} \geq d, v) = \mu_i^{(c)}(v) - d.$
- $\rho(\mu_i^{(c)} \leq d, v) = d - \mu_i^{(c)}(v).$
- $\rho(\neg\phi, v) = -\rho(\phi, v).$
- $\rho(\phi_1 \wedge \phi_2, v) = \min(\rho(\phi_1, v), \rho(\phi_2, v)).$
- $\rho(\phi_1 \vee \phi_2, v) = \max(\rho(\phi_1, v), \rho(\phi_2, v)).$
- $\rho(\exists_B \circ \phi, v) = 0.25 \max_{b \in B}(\rho(\phi, v_b))$ where v_b was the child vertex of v with label b .
- $\rho(\forall_B \circ \phi, v) = 0.25 \min_{b \in B}(\rho(\phi, v_b))$ where v_b was the child vertex of v with label b .

It was proven in (Bartocci et al., 2016) that TSSL quantitative semantics are sound. In other words, a quad-tree Q satisfied a formula ϕ ($Q \models \phi$) if $\rho(\phi, v_0) > 0$ where v_0 was the root of Q , and Q violated ϕ ($Q \not\models \phi$) if $\rho(\phi, v_0) < 0$. Therefore, the problem of checking whether an image contains a pattern expressed as a TSSL formula was reduced to computing its quantitative valuation $\rho(\phi, v_0)$. Moreover, the absolute value of $\rho(\phi, v_0)$ was viewed as a measure of how strongly ϕ was satisfied (or violated) by Q . Hence, the quantitative valuation of a formula with respect to a quad-tree was called its *robustness*. This property is demonstrated in Figure 3.8.

Particle Swarm Optimization: Consider an agent-based model with a set of parameters $p \in \Omega \subset \mathbb{R}^{N_p}$, where Ω was the possible set of parameter ranges and N_p was the total number of parameters. For instance, in the model described in Section 3.2.10-13, we had $N_p = 5$ parameters with ranges specified in Table 3.1.

The output of the model was a sequence of T images where $A[t]$ was the image corresponding to time step $t \in \{0, 1, \dots, T\}$ and T was the total duration of simulation. Our goal was

to determine parameter values that result in emergence of a required pattern in the sequence of images derived from the model. Recall that we could specify the pattern using a TSSL formula Φ_{Pattern} . Moreover, each image $A[t]$ could be translated into a corresponding quad-tree $Q[t]$ with root $v_0[t]$. Therefore, for a fixed parameterization p , we could quantify the resulting sequence of images with $S(p)$ using the following equation:

$$S(p) = \max_{0 \leq t \leq T} \rho(\Phi_{\text{Pattern}}, v_0[t]), \quad (\text{S19})$$

where ρ was the TSSL robustness as described in the previous section. Note that since the model was stochastic in nature, $S(p)$ was a random variable and would have a different value every time a sample simulation was produced using the model with the parameters p . If $S(p) > 0$, there exists at least one image in that particular sequence for which the TSSL robustness was positive and the pattern was present. On the other hand, the pattern had not emerged in the sample simulation if $S(p) < 0$. We called $S(p)$ the *robustness* degree for parametrization p . Now, the problem became finding the parameterization p^* that maximized the score $S(p)$. Since $S(p)$ was a random variable, we choose to maximize its expected value:

$$p^* = \operatorname{argmax}_{p \in \Omega} \mathbb{E}(S(p)), \quad (\text{S20})$$

which means that we were looking for the parameterization p^* that *on average* produced patterns with highest possible robustness score. If we simulated the model n times from parameters p , the expected value could be approximated with the sample mean:

$$\mathbb{E}(S(p)) \approx \tilde{S}(p) = \frac{1}{n} \sum_{i=1}^n S_i(p), \quad (\text{S21})$$

where $S_i(p)$ was the robustness score for parameters p in the i th simulation. In general, a large sample is needed to achieve an accurate approximation. however, it was shown in that in practice, a relatively small n suffices for the purpose of optimization in Equation S20. In this paper, we computed the average robustness for three sample simulations in every case ($n = 3$).

Many optimization methods can be used to solve this optimization problem. Inspired by, we employed particle swarm optimization (PSO) to solve this problem. PSO is a heuristic solution to unconstrained optimization problems that is capable of solving problems with irregular search spaces, is easily distributable, and does not require the objective function to be differentiable.

The PSO algorithm worked as follows: The procedure began by randomly initializing a set of M particles with positions $z_i \in \Omega$ and velocities z'_i . The position of a particle was a candidate solution to Equation S20, and the velocity was a search direction from the current solution. Next, n simulations were produced and n sequences of quad-trees $Q[t](z_i)$ were created for each particle and the average robustness degree $\tilde{S}(z_i)$ was evaluated for each set of simulations represented by particle z_i . The position of the i th particle that had performed best so far was stored in the variable z_i^* , and the optimal value of z_i^* was denoted by z^* . After all particles had been evaluated, the positions and velocities were updated according to the following relations:

$$\begin{aligned} z'_i &\leftarrow Wz_i + \eta(r_p)(z_i^* - z_i) + \eta(r_g)(z^* - z_i) \\ z_i &\leftarrow z_i + z'_i, \end{aligned} \tag{S22}$$

where $\eta(r_i)$ was a random number uniformly distributed over $[0, r_i]$ and the parameters $W \in \mathbb{R}$, r_p, r_g are tuned by the user. This iterative process continued until a termination criterion was met.

If $\tilde{S}(p^*)$ was positive or negative but sufficiently close to zero, we had found the optimal parameterization of the model for the required pattern. This occurred for the Bullseye and Multi-Islands patterns. The optimal parameterization is shown in Figure 3.8Ci-ii. On the other hand, $\tilde{S}(p^*) \ll 0$ indicated that even for the best possible parameterization of the model, the required pattern did not emerge, meaning that the model was not capable of producing that pattern at all. This occurred for the Janus (Left-Right) pattern (Figure 3.8Ciii).

Figure 3.10A demonstrates two sample simulations, one for the Bullseye pattern and one for the Multi-Islands pattern. Figure 3.10B shows how the corresponding TSSL scores evolve over

time for each simulation. It is seen in this figure that the scores gradually improve until at some points the desired patterns are formed.

3.2.15 Quantification and Statistical Analysis

Mann-Whitney U-tests were used to compare *in vitro* and *in silico* experimental populations in Figure 3.7. Unpaired T-tests with Welch's correction were used to compare *in vitro* and *in silico* experimental populations in Figure 3.10. Error bars depicted in graphical representations signify 1 standard deviation unless otherwise specified in the figure legend. For each statistical test the number of replicates is described in the figure legend. Throughout the manuscript the symbol * signifies statistical significance at the 0.05 level unless otherwise specified.

3.2.16 Software

The mathematical model was implemented in Morpheus, v1.9.1. Model fitting of single-cell morphology, cell velocities, temporal knockdown characterizations, and spatial pattern characterizations were performed with custom Python code (modules: scipy, numpy, matplotlib, pandas, seaborn, scikit-image) (Jones, Oliphant, & Peterson, 2001). Image preprocessing, segmentation, and quantification of cell and colony morphology was performed with custom Python code (modules: numpy, scipy, scikit-image). Pattern classifications and specification mining was performed using the data mining software WEKA. Quantification of pattern similarity and pattern optimization were performed with TSSL (MATLAB + Java) which can be accessed from: <https://github.com/dmarcbriers/Multicellular-Pattern-Synthesis>.

3.2.17 Data Availability

All data is available in the main text or supplementary materials. A GitHub repository of analysis code can be found at: <https://github.com/dmarcbriers/Multicellular-Pattern-Synthesis>.

3.3 Results

3.3.1 Pattern Synthesis: *in silico* prediction and automated discovery of spatial behaviors

To observe multicellular pattern formation, we used a previously established hiPSC line with a DOX inducible CRISPRi system, allowing for temporal gene knockdown (KD) wherein mixed populations establish KD in only a portion of the colony, creating a symmetry breaking event and subsequent pattern formation (Libby et al. 2018; Mandegar et al. 2016). However, the generation of new patterns in a predictable manner requires the ability to test large numbers of experimental conditions that would require a massive amount of time and manual effort to comprehensively test the vast number of experimental parameters possible. For example, to experimentally explore the parameter space of a single gene knockdown where the following parameters are varied: knockdown timing (3 timing schemes tested), duration of experiment (5 durations tested), degree of gene knockdown (5 knockdown levels tested), and proportion of population that is knocked down (9 percentages tested), one would need to perform 675 total experiments. Given the biological variability we observe within our *in vitro* experiments (Figure 3.10), a power analysis suggests that a minimum of approximately 13 biological replicates would be necessary to detect significant differences between individual experiments (12.85 observations required, with significance assessed at $p < 0.05$, 80% probability of accepting the alternate hypothesis, corrected for multiple comparisons), yielding approximately 9,000 total conditions or roughly ninety four 96-well plates, where one well represents a single condition. Alternatively, a machine learning and optimization algorithm, such as “Pattern Synthesis” (Bartocci et al. 2016; Briers et al. 2016), can automatically and efficiently discover experimental

conditions and robustly predict the *de novo* self-organization of hiPSCs into desired target patterns.

Pattern Synthesis required two inputs: a model of hiPSC behavior, and images of the desired pattern (i.e. “goal”) outcomes. First, we developed a computational model of hiPSC colony organization as a result of a single gene KD (Figure 3.1A). Next, we generated images of desired and undesired spatial patterns to train a machine learning algorithm that establishes a pattern classifier with a quantitative metric of pattern similarity (Figure 3.1B) (Bartocci et al. 2016; Haghghi et al. 2015). Given these inputs, we formalized pattern discovery as an optimization problem where the objective was to maximize the similarity score of images from our computational model to our desired spatial pattern (Figure 3.1C). The variation between different simulations was based upon five categories of *in vitro* perturbations that could be readily created in hiPSC colonies (Figure 3.1D).

3.3.2 Data-Driven Computational Model of Human iPSC Self-Organization

It is challenging to both predict and control spatial patterning in human iPSCs since the design of multicellular systems rapidly increases in complexity when considering the dynamics of single cell mechanics and cell-cell interactions. These dynamics include, but are not limited to, temporal changes in interfacial tension associated proteins, cell type abundance, cell division, and cell migration velocities.

To capture the complex dynamic interactions involved in multicellular patterning, we developed a data-driven Cellular Potts Model (CPM) to predict spatial patterning in hiPSCs due to the time-dependent modulation of cell-cell adhesion and cortical tension. Using the CPM, we modeled an *in vitro* system consisting of two populations of iPSCs co-cultured for up to 120 hours. The model incorporates numerous biological parameters, such as cell area conservation, cell proliferation, cell-cell adhesion, intracellular cortical tension, cell migration, and time-dependent modulation of properties implicated in cellular organization as a result of mosaic KD. To connect

the *in silico* model to potential genetic targets for *in vitro* experimental manipulation, we focused on using CRISPRi KD, which provides precise temporal control over protein expression, of two molecules associated with regulating cellular mechanics and cell-cell interactions: E-cadherin (CDH1) and Rho associated coiled-coil containing protein kinase (ROCK1). CDH1 is a classical cadherin cell-cell adhesion molecule, whose modulation allows for changes in the adhesive interactions between neighboring cells, and ROCK1 is a protein kinase that regulates non-muscle myosin activity and indirectly modulates the actinomyosin cytoskeletal tension within and between cells. These two molecules contribute to feedback loops that regulate interfacial tension between cells within a tissue and facilitate the physical organization of multiple cell types making them ideal candidates that when knocked down alter the cellular organization within a pluripotent stem cell colony (Libby et al. 2018).

To fit the *in silico* model to an *in vitro* experimental training set, pairwise *in vitro* characterization experiments were performed to determine the relationship between space, time, and protein expression (Figure 3.7A) in wildtype (WT), CDH1 KD and ROCK1 KD hiPSCs. These relations were established by *in vitro* measurements of single cell morphological changes (Figure 3.4), migration velocity magnitudes (Figure 3.7B-D), protein expression changes (Figure 3.7E-G), and colony organization (Figure 3.7H-I) before and after mosaic KD of CDH1 and ROCK1 in hiPSC colonies. The purpose of these characterization experiments was two-fold: 1) to reduce the complex interactions into quantifiable relationships, 2) and to create a closed-loop mapping between *in vitro* perturbations and *in silico* simulation parameters.

To characterize cell morphology, brightfield images of wild-type (WT), CDH1(-), and ROCK1(-) cells were collected 6 days (144 hours) after gene knockdown. Single cell *in vitro* cell area and membrane length measurements (Figure 3.4) were acquired to set the target cell area and target cortical tension in the simulations, respectively (n= 110 per cell type). In the CPM, the weight associated with cortical tension constraint regulates how readily a cell can change its cell membrane length and relates to cell membrane stiffness. Due to differences in cell crowding in

the center versus the edge of colonies, cell morphology measurements were fixed given a cell's mechanical modulation and its radial position in the colony (n=55 central, 55 edge)(Table 3.2, Figure 3.4). Cell division was assumed to be asynchronous amongst the population, and cell division times specific to each type of knockdown were incorporated into the model to provide an accurate depiction of population growth kinetics.

The relationship between cells in space with respect to time was characterized by measuring the *in vitro* distribution of individual cell velocities, resulting in an empirical median velocity magnitude of $0.29 \pm 0.17 \mu\text{m}/\text{minute}$ (Figure 3.7B,D). The distribution of velocity magnitude values was then used to model collective cell migration as locally polarized motility where the direction of cell migration is influenced by the relative cell adhesion strength of neighboring cells (Czirók et al. 2013). Cell-cell adhesion and cortical tension parameters for WT cells were selected to mimic the *in vitro* velocity measurements, producing a comparable distribution with a median *in silico* velocity magnitude of $0.31 \pm 0.12 \mu\text{m}/\text{minute}$ (Figure 3.7C,D). Importantly, the *in silico* generated velocity distributions were not significantly different from the *in vitro* measured velocities (Mann-Whitney U test, $p = 0.0502$). An additional requirement that individual cells remained part of dense epithelial colonies without migrating from the exterior borders was also imposed to match the hiPSC phenotype observed *in vitro*. After fitting the model to empirical data of cell morphology and velocity, single cell morphology and collective cell migration of human iPSC colonies were accurately recapitulated.

To examine the relationship between protein expression time, CDH1 and ROCK1 were knocked down using CRISPRi, and the relative mRNA and protein expression was assessed for 6 consecutive days via qPCR, fluorescence microscopy, (Figure 3.7E, Figure 3.3) and Western blot analyses (Libby et al. 2018). Due to our previous observation of the phenotypic robustness of CDH1 knockdown in promoting cell self-organization(Libby et al. 2018), we designed several CRISPRi guide RNAs to target CDH1 producing different levels of transcriptional knockdown at

10%, 25%, 30%, and 98% compared to WT expression. A single guide RNA for ROCK1 knockdown was used to achieve 80% knockdown of WT expression levels (Figure 3.2). The data was normalized (min-max [0,1]) and Hill Functions were fit to the normalized median expression (per day) using least squares optimization to create a time-dependent response function for CDH1 knocked down to 90%, 75%, 70%, and 2% of the original mRNA expression (Figure 3.7F). This range of knockdown efficiency allowed us to computationally model how differing levels of CDH1 expression could impact spatial patterning. Using the same approach as the CDH1 knockdown, we created a Hill response function for ROCK1 knocked down to 20% mRNA expression. Because of a delay in protein knockdown compared to mRNA levels, the Hill functions were shifted by 24 hours to account for the delay in protein loss (Figure 3.7G), allowing us to model the average change in ROCK1 protein expression for individual cells over time.

Given the previous characterization experiments, we were able to model collective cell migration and temporal changes in cell mechanics. To model the spatial patterning due to the temporal modulation of cell-cell adhesion via CDH1 or cortical tension via ROCK1, either inducible ROCK1 knockdown or inducible CDH1 knockdown iPSCs were co-cultured with WT iPSCs, where knockdown of gene expression was induced upon mixing the two cell types. Then, images of the mixed populations were collected 96 hours after gene knockdown and co-culture. As previously reported (Libby et al. 2018), mixed colonies with a subpopulation of cells that had reduced CDH1 or ROCK1 expression produced distinct mosaic patterns due to reduced cell-cell adhesion and increased membrane stiffness properties respectively (Figure 3.7H-I (left)). *In silico*, parameter sweeps were run over a range of adhesion strength and membrane length values to explore the phenotypic space resulting from decreases in cell-cell adhesion and increases in membrane stiffness. Computationally varying the adhesion strength produced a variety of spatiotemporal patterns due to progressively weaker cell-cell adhesion or progressively stiffer cell membrane parameter values. Double-blind analysis of *in silico* and *in vitro* generated data sets was then conducted to identify parameters that yielded closely matching multicellular patterns (Figure 3.7H-

I (right)). Given the characterization experiments of cell morphology, cell migration speed, time-dependent modulation of cell mechanics, and the resulting spatial organization, the computational model was able to recapitulate the spatial patterning due to the CDH1 and ROCK1 knockdowns.

Overall, after incorporating *in vitro* measurements into our computational model, we accurately recapitulated hiPSC spatial patterns with the initial experimentally derived parameters in mixed colonies of WT and CDH1 KD cells or WT and ROCK1 KD cells (Libby et al. 2018).

3.3.3 Formulating Parameters for Design Automation

Given the success in matching the output of the computational model to experimental data, we then introduced five new design parameters to simulate *in vitro* experimental perturbations, allowing us to model exponentially more permutations of experimental design than would be feasible *in vitro*. The five design parameters were: i) the gene knockdown target of cell population 1, ii) the knockdown time for cell population 1, iii) the gene knockdown target of cell population 2, iv) the knockdown time for cell population 2, and v) the ratio of the distinct cell populations (Figure 3.1D, Table 3.1). These additional design parameters allowed us to convert trial-and-error based design into a mathematical optimization problem that could be computationally solved *in silico* without time-consuming and costly additional experiments. Although computational design frameworks for multicellular spatiotemporal patterning have been used in several previous studies (Tewary et al. 2017; Krieg et al. 2008; Marcon et al. 2016), they often propose underlying morphogenic mechanisms with limited perturbation potential *in vitro*. Thus, an *in silico* optimization framework that directly informs subsequent experimental designs is critical to survey the high-dimensional landscape of morphogenesis.

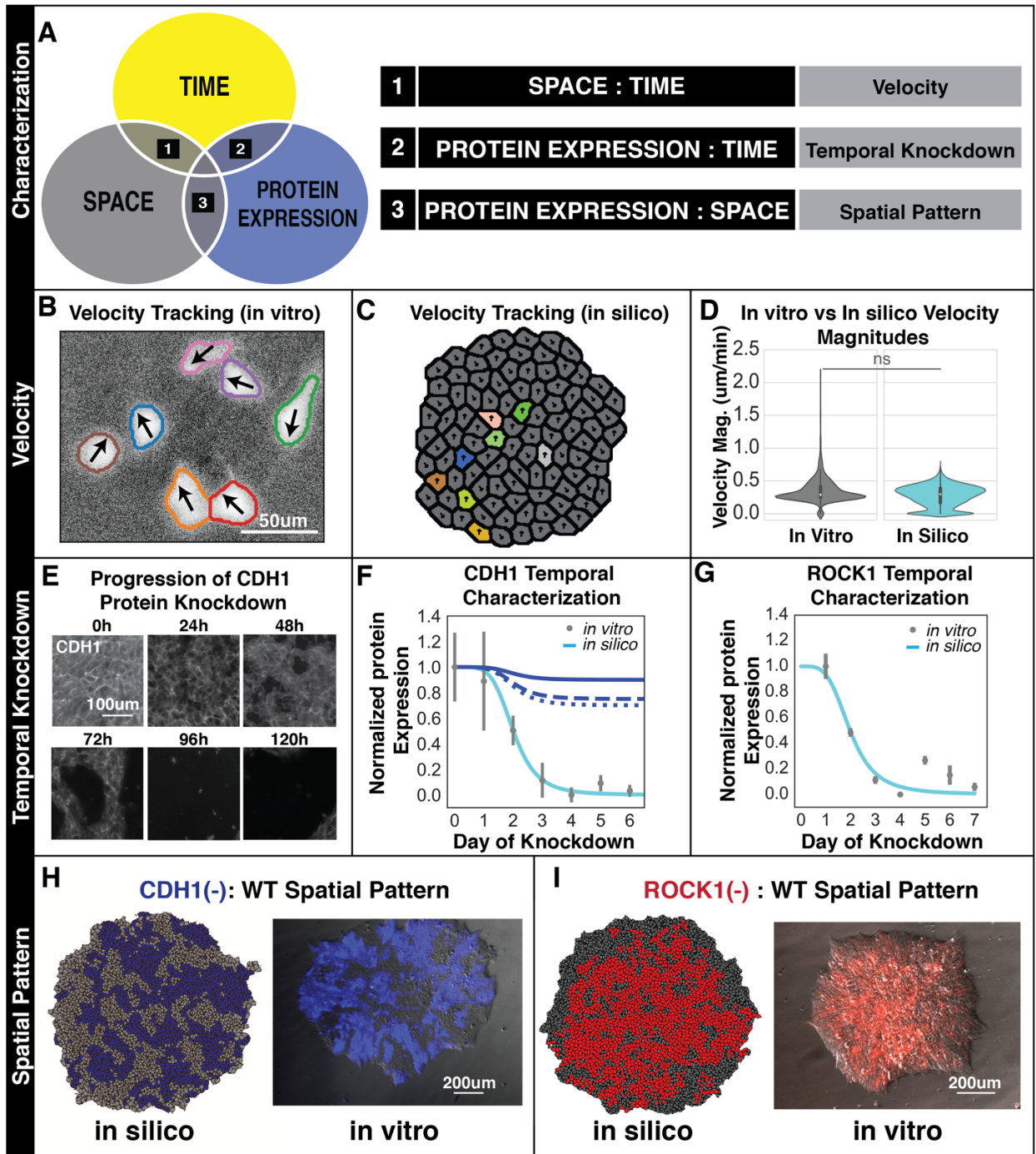


Figure 3.7: Pairwise experiments to characterize dynamic changes in spatiotemporal behaviors. A) We characterize cellular behaviors in a pairwise manner to reduce the complexity of possible interactions. Space, time, and protein expression are the minimally necessary properties to characterize and model spatiotemporal behaviors. Space-time relationships are captured with velocity characterizations, time-protein expression is captured characterizing the relative protein expression for several days after knockdown, and protein-space relationships are characterized by confocal microscopy imaging of spatial behavior due to cell mechanical perturbations. **B)** We performed paired *in vitro* and **C)** *in silico* experiments to match the velocity distributions of iPSCs. **D)** The grey violin plot represents the distribution of *in vitro* velocities, while the cyan violin plot represents the distribution of *in silico* velocities. Using the Mann-Whitney

U test, there was no statically significant difference in cell velocity (p value < 0.05). **E**) Representative images of DOX inducible modulation of protein expression. **F**) We used Hill Functions to mathematically model CDH1 knockdown over time from quantification of mRNA by qPCR ($n=3$) and then adding a 24h delay to account for protein production depicted by light blue lines. Grey circles represent the normalized median expression 0-6 days after CDH1 knockdown. Error bars represent 1 standard deviation from the mean. Dark blue line depicts Hill Function models of partial KD of CDH1. **G**) We use a Hill Function to model ROCK1 knocked down over time as previously described for CDH1 knockdown. ($n=3$) **H**) Paired *in vitro* and *in silico* images of spatial patterning 96 hours after CDH1 knockdown in a subpopulation of cells (blue). **I**) Paired *in vitro* and *in silico* images of spatial patterning 96 hours after ROCK1 knockdown in a subpopulation of cells (red). Given the previous characterizations, the relative strength of cell-cell adhesion and cortical tension can be tuned in the *in silico* simulations to recapitulate the spatiotemporal patterns observed *in vitro*.

3.3.4 Quantitative Pattern Classification

The second input to the Pattern Synthesis procedure was a supervised image classifier known as Tree Spatial Superposition Logic (TSSL) (Bartocci et al. 2016). TSSL uses a quadtree data structure to represent spatial relationships in an image at multiple levels of detail, where the highest level captures global aspects of an image, while the lower levels capture local spatial relationships. For example examining a checkerboard image with some variation (Fig. 3.8Ai), the TSSL would generate a unique quadtree (Fig. 3.8Aii) representing the levels of complexity within the image (Bartocci et al. 2016; Finkel and Bentley 1974; Jackins and Tanimoto 1983). A rule-based machine learning algorithm (RIPPER) (Cohen 1995) was employed to automatically learn a set of rules over the values of quadtree vertices specific to an *in silico* training set of 3,000 positive and 13,000 negative manually rendered images of cells precisely organized into target patterns, such that a quantitative score of pattern similarity could be assigned to any image from the associated quadtrees (Section 3.2.14) (Fig 3A.8iii). The magnitude of the similarity score, which can range from -1 to +1, indicates how strongly a simulation image matches (positive scores) or violates (negative scores) the target spatial behavior. Use of a TSSL robustness score replaces qualitative manual observation of simulation images with a quantitative score of pattern similarity.

Analogous to the checkerboard example, this algorithm can be applied to more complex images such as a target organizational pattern within the CPM (Fig. 3.8Aiv-v) where the generated

quadtree from the TSSL of each desired pattern is used to recognize and rank pattern similarity (Fig. 3.8Av). As a proof-of-principle, we first attempted a concentric ring (i.e. “Bullseye”) pattern, defined as one population of 50 or more connected cells completely surrounded by a second population (Fig. 3.8Ci). The annular Bullseye pattern was chosen because similar asymmetric cell organization occurs multiple times in human development, such as during early embryo compaction leading to the formation of the inner cell mass in the human blastocyst (Ziomek and Johnson 1980; Ducibella and Anderson 1975; Deglincerti et al. 2016). The second target was a Multi-Island pattern, consisting of at least three distinct clusters of 25 or more cells completely surrounded by a separate larger population (Fig. 3.8Cii). The Island pattern was chosen to demonstrate the reproducibility of previously observed segregation of cell populations *in vitro* (Libby et al. 2018) and to test whether this can be predictably controlled. To first demonstrate that the automated classifiers could reliably detect and distinguish between desired and undesired spatial patterns, the classifiers were tested using an *in silico* set of 1,000 positive and 5,000 negative images. The TSSL classifiers achieved a 98.2% classification accuracy for the Bullseye and 96.9% classification accuracy for the Multi-Islands pattern, meaning that the TSSL algorithm can properly recognize and score Bullseye or Island patterns nearly 100% of the time.

By quantifying how well images from an *in silico* multicellular arrangement matched images of our target organization, we enabled the optimization algorithm (described in the next section) to incrementally improve and learn a unique combination of design parameters that could give rise to a desired goal pattern.

3.3.5 Automated Discovery of Pattern Producing Conditions

The CPM allows simulation of more than 40,000 distinct parametric conditions and facilitating the study of emerging behaviors of hiPSCs much faster than *in vitro* experiments. Distributing the computation over 12 processors at 2.1 GHz on a server cluster, it only took approximately 5 minutes to simulate the evolution of one cell population over 120 hours. To

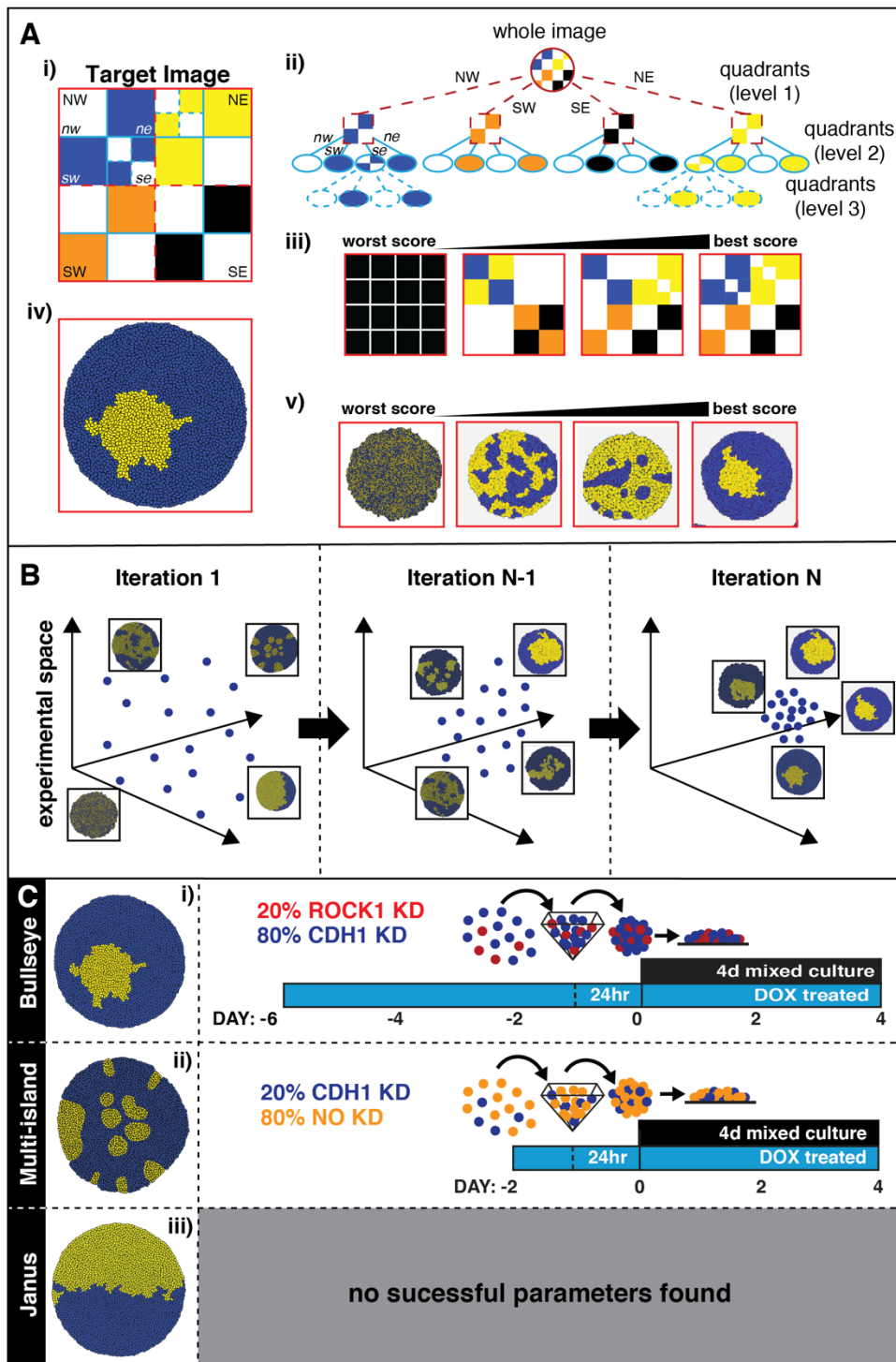


Figure 3.8: Quantitative Pattern Classifier with TSSL. A quadtree is used to represent an image at multiple levels of detail. **Ai-ii)** Shows a representative quad tree for an example checkerboard image. An image **Ai)** is subdivided into sequential quadrants until each quadrant is one singular color. This is then depicted as a tree **Aii)** where both the values and branches of the tree are specific to each image. **Aiii)** Given a quadtree representation of a target image, TSSL produces a numerical score corresponding to the similarity of an image to the desired target image. This score can then be used to rank images by similarity to the desired image. **Aiv)** An example image of a desired pattern generated in the CPM. **Av)** A pictorial

example of how the TSSL would be able to distinguish different CPM images and score them against the desired pattern. **B)** Schematic representation of a particle swarm algorithm depicted in a 3D search space where each particle represents an *in silico* simulation. With each iteration of the algorithm, PSO reduces the breadth of exploration in the experimental space and travels towards increasing TSSL scores, indicating that the optimization procedure has located *in silico* experiments that are generating patterns of increasing similarity to the goal pattern. **C)** Schematics of example target patterns given as classifiers in the machine learning pattern synthesis process and parameters produced by Pattern Synthesis that predict the creation of the desired patterns: **Ci)** Bullseye, **Cii)** Island, **Ciii)** Janus.

recapitulate this same experiment *in vitro*, 13 96-well plates would need to be cultured in parallel for 120 hours, demonstrating that *in silico* experimentation can accelerate parameter exploration more than a 1,000-fold. The simulation speed permitted examination of a wide range of different experimental conditions in a rapid and inexpensive manner, taking both the labor and reagent costs into account. However, due to the tens of thousands of experimental conditions to consider and the resulting months of computation for such a large number of simulations, it quickly became impractical to enumerate every possible set of conditions to identify parameter combinations that yielded the highest robustness score(s). Thus, to automate the discovery of conditions that yielded goal spatial patterns, we formulated the selection of experimental conditions as an optimization problem.

Using the TSSL-provided metric of image similarity, a Particle Swarm Optimization (Eberhart and Kennedy 1995) was employed to identify regions of the 5-dimensional design space, created by the available design parameters, with the highest probability of producing a target pattern (Bullseye or Multi-island)(Figure 3.8B). In brief, the Particle Swarm Optimization first explores the extremes of the 5-dimensional experimental space, where every extreme represents a set of experimental parameters that are run as an *in silico* experiment using the previously described CPM. Then the resulting patterns from this first set of *in silico* experiments are given scores. The algorithm then narrows its focus to the experimental space that produced experiments resulting in higher scores, doing this iteratively further selecting for the experimental space that is most likely to produce the highest TSSL score and therefore the patterns that most

closely resemble the goal pattern. A full explanation of the particle swarm algorithm can be found in the STAR Methods.

For any *in silico* simulation, where the previously described design parameters are varied to represent a different experimental condition, the Patterning Synthesis algorithm determined whether the generated pattern was successful (Figure 3.9A), and whether the similarity score improved over the simulated period of 120h by at least one order of magnitude, eventually reaching a steady state (Figure 3.9B). Analyzing the temporal dynamics of robustness scores provided insight into the exact time a pattern emerged *in silico*, and optimized design parameters for target patterns that closely resembled, but still resembled the desired spatial behavior. The final output of the particle swarm algorithm is a list of experimental parameters that are predicted to generate the desired pattern both in the *in silico* CPM and the *in vitro* stem cell culture system after 120 hours of mixed culture (Fig. 3.8C).

In addition to automating the design of *de novo* spatial patterns, we could also determine the feasibility of any spatial pattern given the tunable conditions of the system. Although it is impossible to exclude experimentally that a particular pattern can never be generated *in vitro* (it would require testing all possible conditions), *in silico* certain *de novo* patterns resulted in negative robustness scores (violating the pattern specification), indicating that the cell population, under the current perturbations available (knockdowns, mixing ratios, etc.) was unable to perfectly recapitulate the desired spatial behavior. For example, the algorithm was able to determine that a perfectly symmetrical “Janus” pattern (left-right)(Fig. 3.8Ciii) was not achievable with the primary experimental variables (i.e. timing of CDH1/ROCK1 knockdowns and the ratio of cell types co-cultured in an approximately 2D monolayer), indicating that additional parameters such as silencing of other genes may be necessary to yield such a pattern.

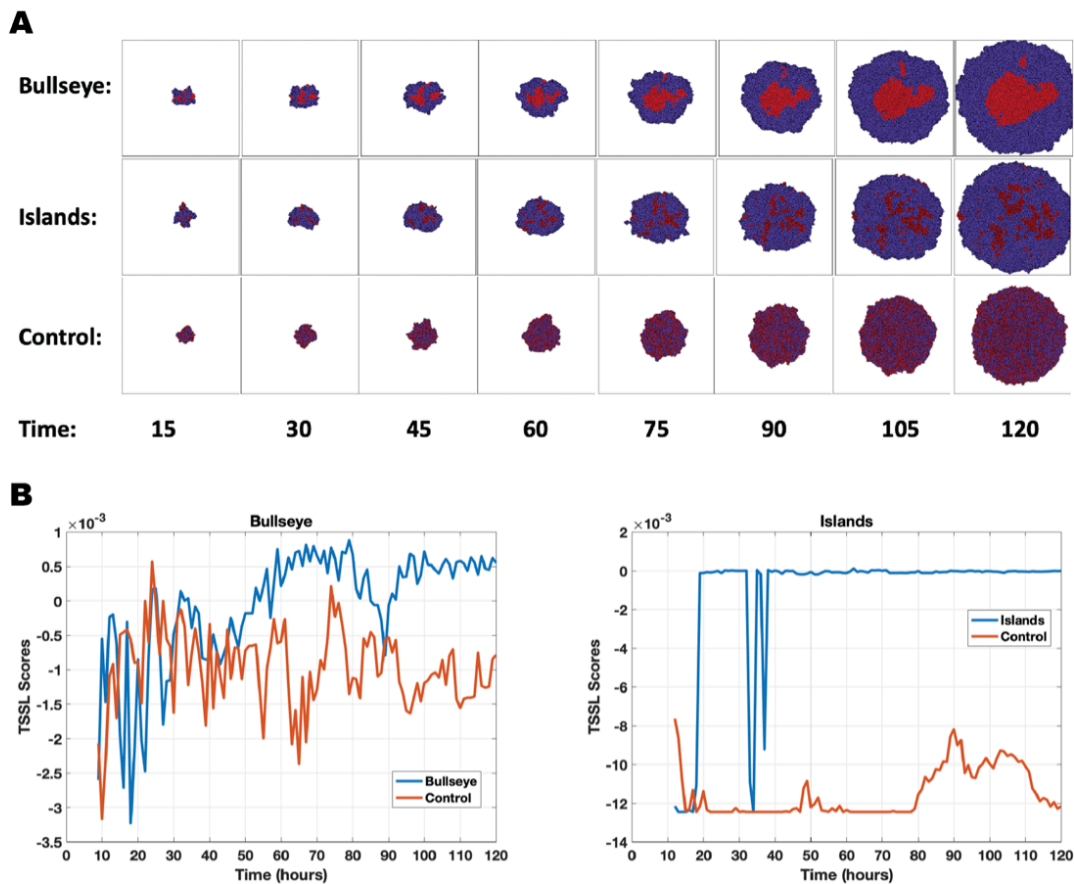


Figure 3.9: TSSL increase with pattern formation. (A) Example evolution of pattern in simulated hiPSC colonies. **(B)** Evolution of the TSSL robustness score over time for the sample simulations illustrated in A; producing the Bullseye pattern (left) and Multi-islands pattern (right).

3.3.6 *In Silico* Model Accurately Predicts *In Vitro* Experimental Validation

The Patterning Synthesis algorithm yielded different sets of instructions to produce either a Bullseye pattern or a Multi-Island pattern of hiPSCs. The Pattern Synthesis predicted that a mixture of 1:4 ROCK1 KD iPSCs to CDH1 KD iPSCs that were independently pretreated with DOX for 6 days prior to mixing and cultured together for 4 days would yield a Bullseye pattern (Fig. 3.8Ci) and that a mixture of WT cells with CDH1 KD at a ratio of 1:4 with DOX pretreatment of iPSCs for 48h prior to mixing would create the Multi-Island Pattern (Fig. 3.8Cii).

Based on these predictions, *in vitro* experiments were performed using the specified conditions, and the incidence of pattern formation was independently analyzed for *in silico* and *in vitro* results (Figure 3.10A-D). The experiments were performed with unrestricted colony growth (i.e. no patterned matrix restriction)(Tewary et al. 2017; Warmflash et al. 2014) to ensure that cellular organization within the colony was not driven by imposed boundary conditions. To account for colony size differences affecting the resulting patterns, only colony sizes within two standard deviations of the mean number of cells per colony were examined for pattern formation. We characterized the morphology of *in silico* and *in vitro* generated patterns by interrogating subpopulation cluster circularity, number of clusters, and cells per cluster within the colony (Figure 3.11). However, the *in vitro* experimental results were more variable and yielded a wider range of results, which may be due to biological variability in wet lab experimentation or subtle variations in cellular behavior that the *in silico* model does not take into account. Comparing the robustness scores generated for both the parallel *in silico* and *in vitro* experiments indicated that the optimal *in vitro* bullseye and islands patterns had larger robustness than their respective control images (at least an order of magnitude difference). The robustness scores are highly comparable only when they are calculated in the same setting also known as a domain; thus, a simulation vs. a simulation control is quite comparable whereas an *in silico* simulation vs. an *in vitro* experimental image will inherently differ to some extent. The *in silico* model and experimental optimization predicted that a Bullseye Pattern would be achieved ~12% of the time, which closely matched the *in vitro* frequency (~16%; Figure 3.10B). Similarly, a Multi-Island Pattern was predicted to occur 100% of the time by the model and was achieved in ~87% of the *in vitro* experiments (Figure 3.10D). Overall, these results demonstrate that *in silico* modeling accurately classified and predicted desired pattern formation achieved by hiPSC self-organization *in vitro*.

To determine how robust the predicted parameters were within the *in vitro* system, the proposed mixing ratios of the populations were incrementally varied by 10 percent (n=16 per condition) (Figure 3.10E, Figure 3.12). Robustness scores for each of the mixing ratios were

calculated (Figure 3.10Ei-ii) to compare how well each condition produced patterns similar to the respective target (Bullseye or Island). In Bullseye patterns, a 50 percent change in mixing ratio from the predicted parameters (80% CDH1KD : 20% ROCK1 KD) resulted in significant decrease in pattern robustness scores ($p < 0.05$) (Figure 3.10Ei). Despite an increase in the robustness scores for the Island patterns in the *in silico* experiments, there were no significant differences in the robustness scores calculated for the parallel *in vitro* experiments with varying population ratios (Figure 3.10Eii). Robustness scores produced by the TSSL algorithm *in vitro* were uniformly lower and had higher variance than the comparable *in silico* conditions (Figure 3.10Ei-ii), reflecting the greater difficulty in classifying natural images over the synthetic images generated by the Cellular Potts Model. Due to the domain change from *in silico* to *in vitro* images, the TSSL algorithm was less able to confidently recognize patterns and explain variability both within and across experiments, resulting in reduced discrimination between mixing ratios. Additionally, differences could be due to the fact that the CPM used is a 2D model that does not account for possible vertical movement within a hiPSC colony. However, since the primary goal of the TSSL was to enable *in silico* Pattern Optimization, the decreased classification power for *in vitro* images did not adversely impact the ability of Pattern Optimization to predict conditions that resulted in the desired target patterns.

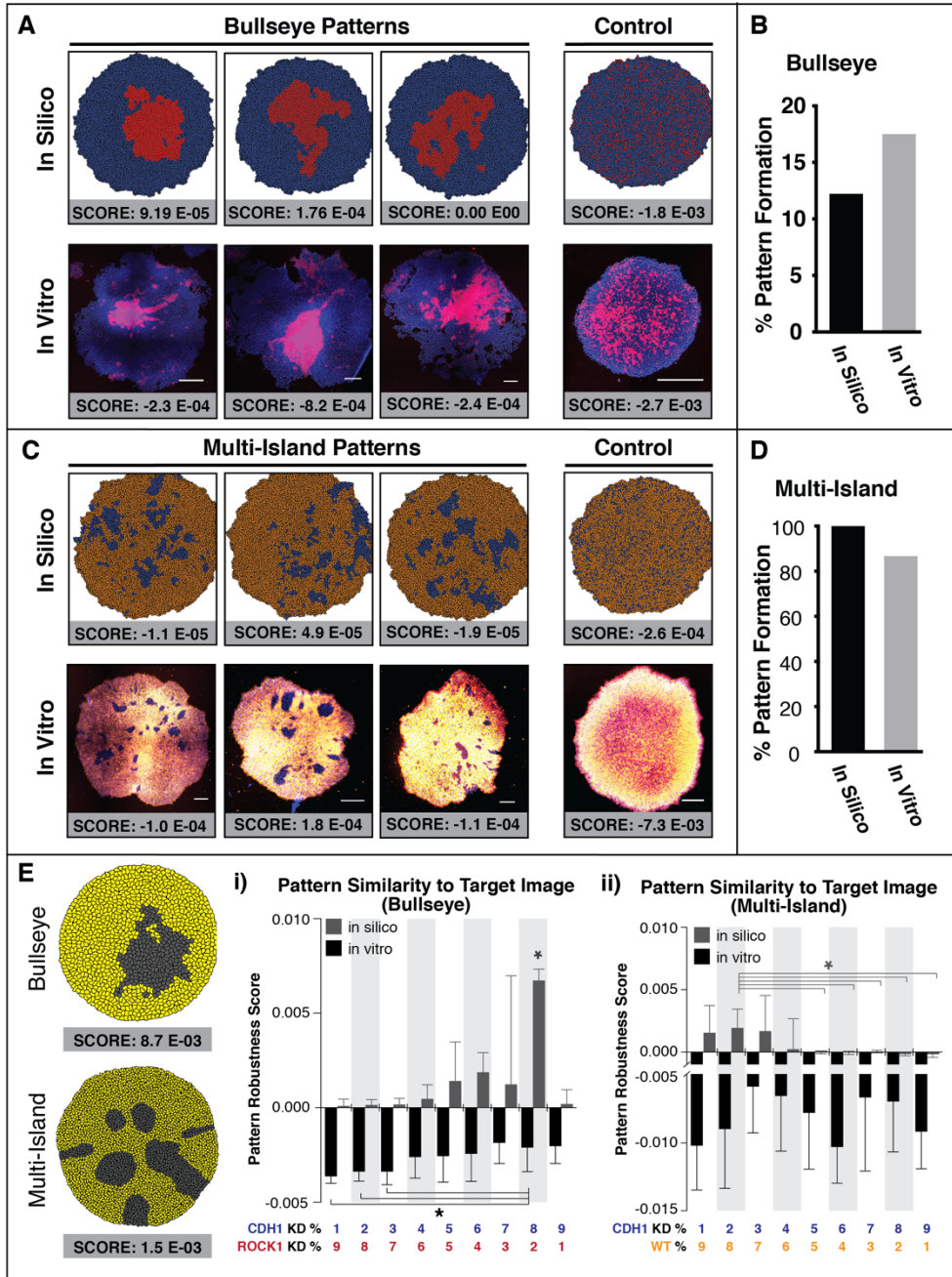


Figure 3.10: Computational synthesis of *de novo* spatial patterns and *in vitro* validation A-D) Comparisons of three simulations of patterns predicted *in silico* and the resulting patterns seen *in vitro* under the same experimental conditions (scale bars = 200 μ m). Pluripotent colonies stained for DAPI (blue) and CDH1 (red/orange) to distinguish populations by presence or absence of CDH1. TSSL robustness scores show how well a simulation matches our specification compared to parallel control *in silico* or *in vitro* experiment. Scores are only comparable if they are calculated in the same environment (simulation vs simulation but not simulation vs experimental image). Image classification in different environments is a well-known limitation in machine learning. **B,D)** Successful pattern creation rates, comparing in silico to in vitro results (Bullseye n = 286 colonies, Multi-Island n = 168 colonies). **E)** Proposed KD populations for the Bullseye Pattern and the Island Pattern were varied by 10% *in silico* and *in vitro* (n=10). Example target patterns (left) from the image set used to train the image classifier to identify and score Island and Bullseye patterns. i-ii) Robustness scores for the respective *in vitro* colonies as KD populations were varied by 10%

where an increase in robustness score indicates more similarity to the target pattern. The predicted parameters from the *in silico* optimization are highlighted in grey and the *in vitro* are in black. Significance is indicated by * with p values < 0.05 where n = 16 colonies per condition and error bars indicate standard deviations.

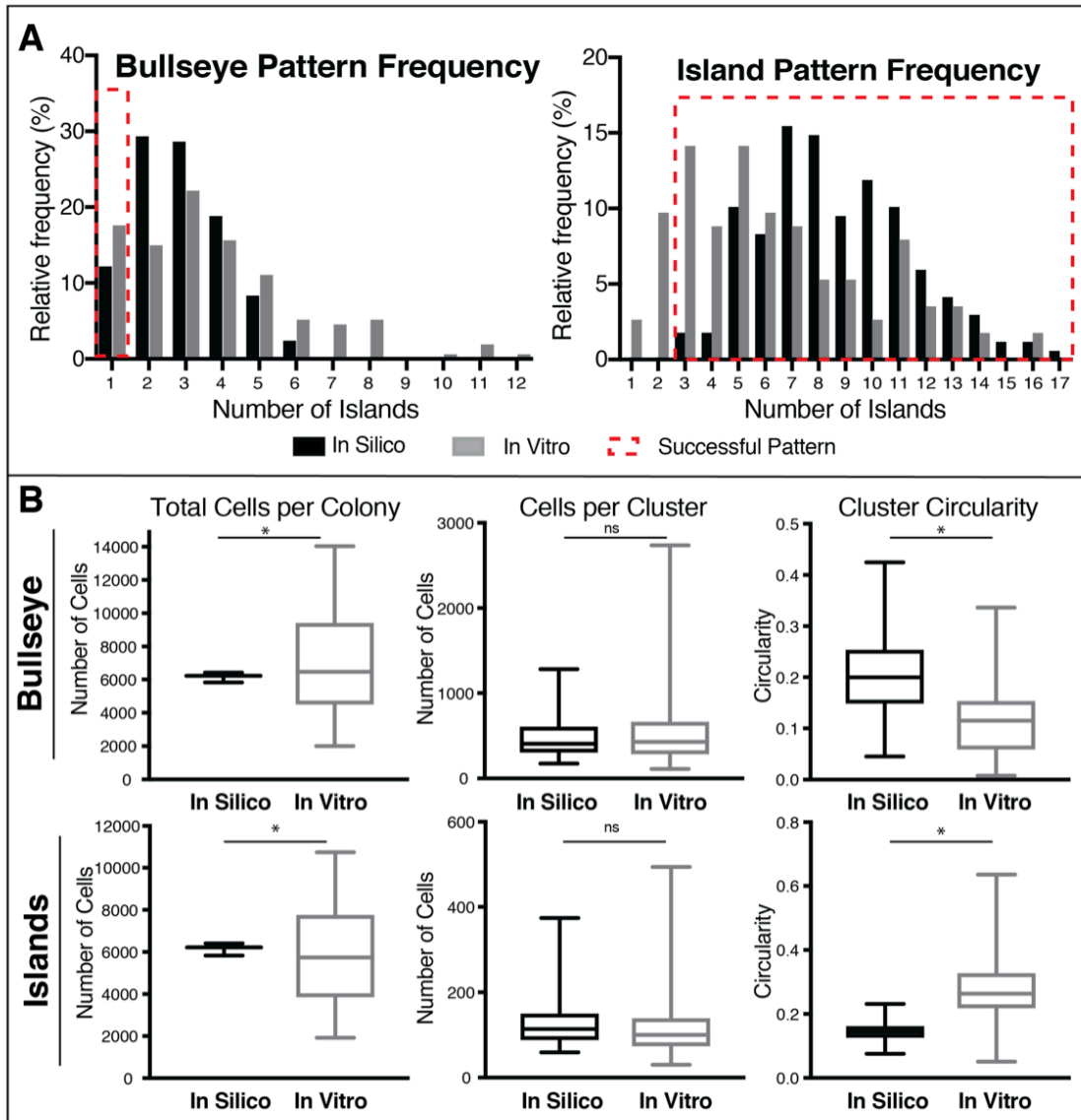


Figure 3.11: In silico vs. in vitro comparative metrics of generated target patterns. (A) Distribution of islands, where for the Multi-island patterns, an island is defined as a homotypic cluster of 25 or more cells, and for the Bullseye patterns, an island is defined as 50 or more cells. Successful Bullseye patterns will display one island, successful Multi-island patterns will display 3 or more islands. (Bullseyes n=148 colonies; Islands n=79) **(B)** Distribution of in silico and in vitro patterned colonies with regard to total cells per colony (Bullseye p = 0.0029, Island p = 0.0215), number of cells per cluster (Bullseye p = 0.9499, Island p = 0.9809), and cluster circularity (Bullseye p < 0.0001, Island p < 0.0001). In Bullseye patterns clusters consist of ROCK1 knockdown cells. In Multi-island patterns clusters consist of CDH1 knockdown cells. Significance is indicated by * (p < 0.05) using two tailed t-tests with Welch's correction. Error bars label the standard deviation of the populations.

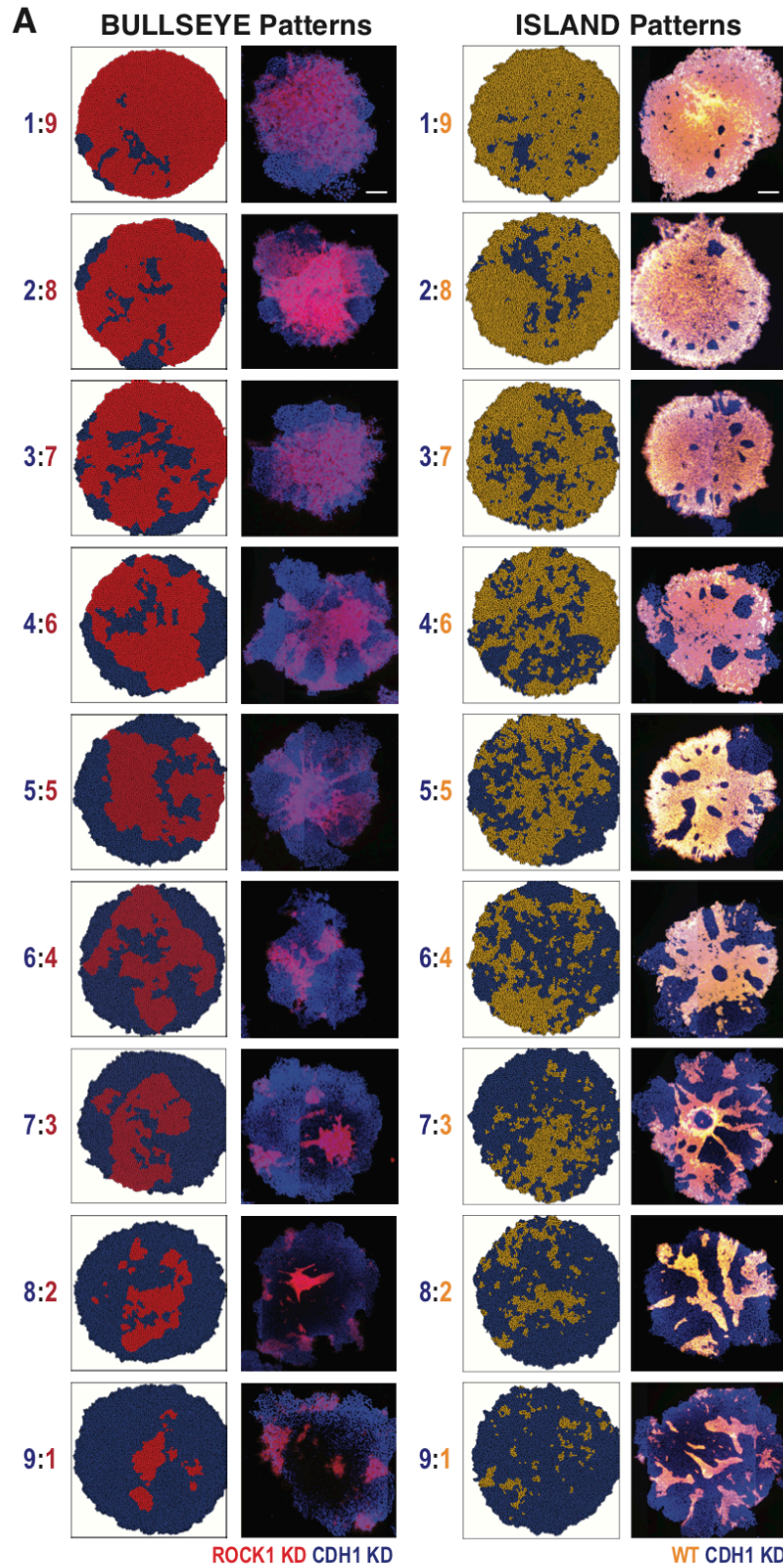


Figure 3.12: Titration of gene knockdown mixing ratios *in vitro* and *in silico*. (A) Example patterns generated in silico and in vitro when the mixing ratios of knockdown cell lines were varied by 10 percent increments (scale bars = 200 μ m).

3.3.7 Colony Organization Impacts Initial Patterns of iPSC Differentiation

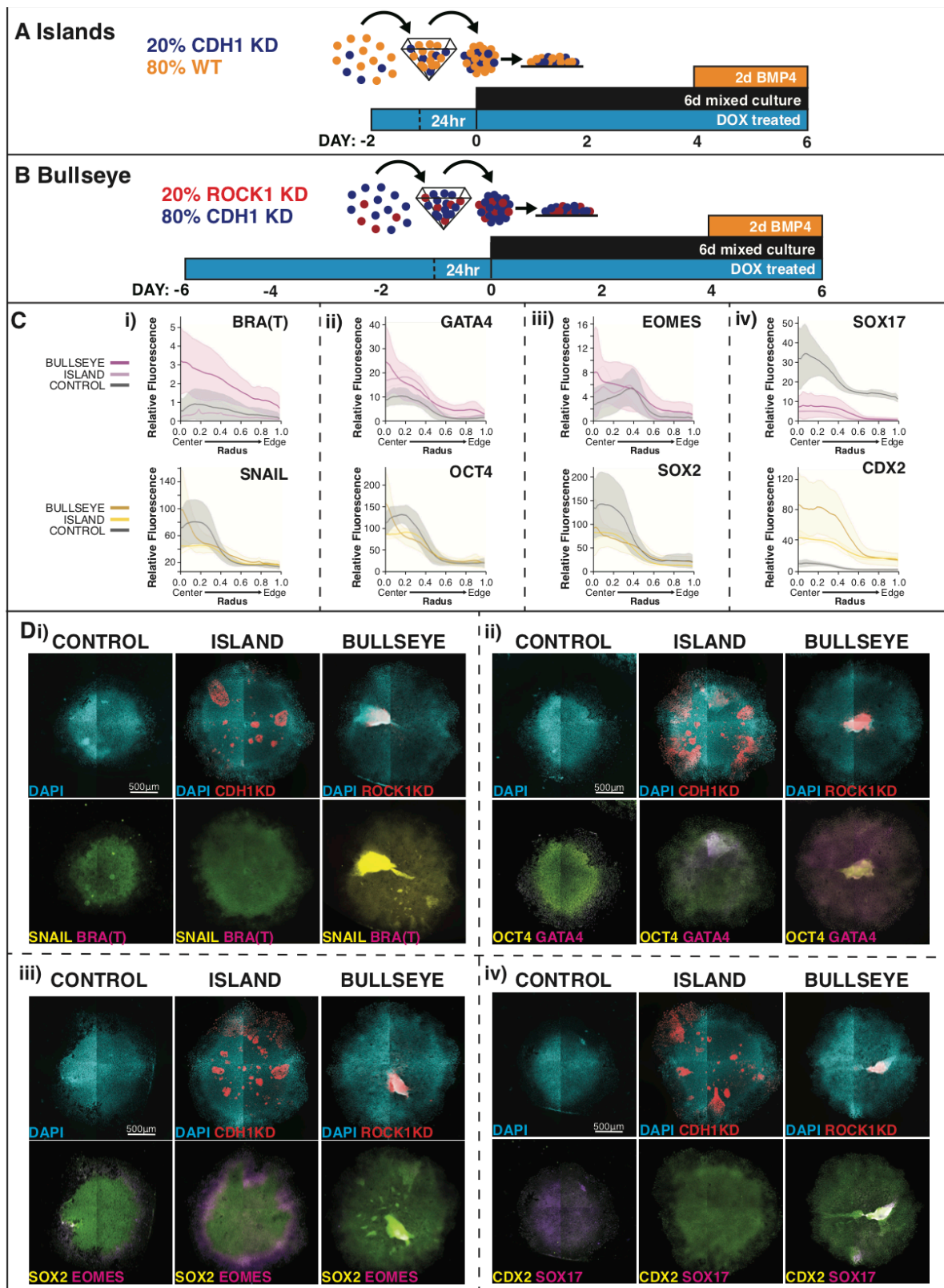
During human development, cell autonomous pattern formation is intimately coupled with cell fate decisions that lead to complex tissue structures. Therefore, we interrogated how the experimentally generated multicellular patterns affected subsequent hiPSC differentiation. We examined the initial cell fate commitment after treatment with BMP4 for 48 hrs (Figure 3.13A,B) with a panel of markers descriptive of different differentiation stages (Figure 3.13C). In brief, hiPSCs were marked by high OCT4 and SOX2 expression in the pluripotent state, then as differentiation proceeded, the first lineage fate decision was marked by upregulation of markers associated with the gastrulating primitive streak (Brachyury (BRA(T)), SNAIL). Cells then transitioned through a mesendodermal fate (EOMES) before displaying mesoderm (GATA4) or endoderm (SOX17) specific markers. The ectoderm lineage remained SOX2(+). Additionally CDX2 was used to mark both extra embryonic lineages and presumptive neural plate cells within the neuroectoderm. (Niwa et al. 2005, 2; Tewary et al. 2017; Wang et al. 2012; Warmflash et al. 2014). WT colonies displayed a radial differentiation pattern with central SOX2(+) OCT4(+) SNAIL(+) cells and a ring of EOMES(+) cells around the periphery indicating the beginning of mesendodermal specification (Figure 3.13C,D). The lack of robust BRA(T) expression was likely due to the transient nature of BRA(T) expression during mesendoderm induction so the time point examined (48h) in this experiment may have captured the tail end of expression. WT colonies displayed a slight increase in SOX17 at the center of the colonies, while GATA4 and CDX2 remained low throughout the colonies (Figure 3.13C,D). A similar radial pattern of cell differentiation was maintained in Island patterned colonies, although SOX17 expression was reduced and GATA4 and CDX2 expression increased (Figure 3.13C,D). The Bullseye patterned colonies displayed a slight increase in BRA(T) expression at the center of the colonies overlapping with the central island that defines a bullseye pattern. Additionally, for the Bullseye patterns GATA4 expression was increased across the entire colony, the radial ring of EOMES was expanded to the entire colony, and high levels of SOX2, OCT4, and SNAIL were displayed in the

center of the colony. These results suggest that the central region of Bullseye colonies underwent lineage transitions through the mesendodermal fate to the mesoderm lineage and displayed an expansion of the CDX2 positive cells. The Bullseye EOMES expression pattern was distinctly different from the control and island patterned colonies that formed a ring of EOMES expression, indicating a positional change in fate acquisition dictated by the establishment of the Bullseye pattern. Thus, the genetic manipulations used to control multicellular organization of human PSCs also influenced the initial differentiated phenotypes of the patterned colonies.

3.4 Discussion

Cell-intrinsic patterning of multicellular stem and progenitor populations is a critical feature of morphogenic events that occur throughout early development (Montero and Heisenberg 2004; Sasai 2013; Ducibella and Anderson 1975; Deglincerti et al. 2016). Thus, systems in which multicellular organization can be robustly controlled and perturbed will help to elucidate key mechanisms in development and symmetry breaking events. Currently, the study of symmetry breaking events often involves the manipulation of cell extrinsic factors, for example, varying morphogen gradients (Demers et al. 2016; Geun Chung et al. 2005), changes in substrate patterning (Hsiao et al. 2009; Théry et al. 2006) and/or the creation of restrictive boundary conditions (Warmflash et al. 2014; Tewary et al. 2017; Théry 2010). In contrast, attempts to influence patterning events using synthetic biology approaches often rely on implementation of an artificial circuit that uses reaction diffusion gradients to establish multicellular patterns (Greber and Fussenegger 2010; Sekine, Shibata, and Ebisuya 2018; Sohka et al. 2009; Toda et al. 2018).

In this study, we demonstrate the induction of active multicellular organization through controlled perturbation of intrinsic cell mechanisms without imposing exogenous boundary conditions. We developed a computational model capable of predicting empirically testable experimental perturbations (gene KD, mixing ratios, etc.) to generate desired multicellular spatial patterns in hiPSC colonies. Using agent-based model predictions of spatiotemporal pattern



is a WT cell colony. **C)** Line graphs denote relative radial fluorescence of lineage markers from the center $r=0$ to the edge of the colony $r=1$ where all colonies were normalized to depict a radius from 0 to 1. The dark line indicates the mean fluorescence across 10 colonies and the matching light color fill represents the standard deviation. The four quadrants indicate: i) SNAIL and BRA(T) ii) OCT4 and GATA4 iii) SOX2 and EOMES iv) CDX2 and SOX17. **D)** Tiled fluorescence images of pure population control differentiations followed by Island and Bullseye pattern differentiations where the scale is consistent across images ($500\mu\text{m}$ for the depicted scale bar in the top-left image of each quadrant). Each quadrant represents a different set of lineage staining where the top row of the quadrant depicts the DAPI stain (cyan) and the KD population specific to the indicated pattern (red) and the bottom row of the quadrant indicates: i) SNAIL and BRA(T) ii) OCT4 and GATA4 iii) SOX2 and EOMES iv) CDX2 and SOX17 depicted in yellow and magenta respectively.

formation, we were able to predict and achieve new patterns *in silico* and *in vitro* without using extrinsic patterning methods (i.e. hydrogels, micropatterning). Optimized design parameters achieved desired organization of cells within a colony and promoted the expected initial mesendoderm or ectoderm lineage fates. Ultimately, these results demonstrate that machine learning and mathematical optimization enable predictive and controlled spatial self-organization of heterogeneous populations of pluripotent cells, which is a critical initial step for hiPSC self-assembly prior to lineage commitment and subsequent organoid and tissue formation.

Previous attempts to pair computational models with experimental morphogenic systems have been largely observational and rarely demonstrate the ability to design phenotypes *in silico* that can be recapitulated *in vitro*. In this study, both the *in silico* and *in vitro* aspects can be adapted to additional parameters, truly taking advantage of machine learning and optimization to generate desired multicellular patterns. With respect to extending *in vitro* perturbations, CRISPR technology can be adapted to repress or activate any accessible genes related to cell patterning and organogenesis. As additional biological parameters are considered, we can quantitatively characterize the effect on cell patterning, and the *in silico* model can be refined to take those factors into account (Briers et al. 2016; E. White et al. 2015), enabling interrogation beyond cell mechanics and into other realms of cell-cell communication such as paracrine signaling gradients and gap junction connectivity to allow for more robust pattern formation than that described by only manipulating cellular mechanics (White et al. 2013; Glen, McDevitt, and Kemp 2018).

Ultimately, the combination of agent-based modeling, machine learning, and directed symmetry breaking provides a novel route to engineer complex multicellular tissue structures that go far beyond simple observation of pattern formation, and facilitate targeted mechanistic studies that address fundamental principles of development and morphogenesis leading to robust practices for complex *in vitro* tissue formation.

3.5 Conclusion

In conclusion, this second study provides an example of combined genetic engineering with computational modeling, machine learning, and mathematical pattern optimization to create a data-driven approach to control hPSC self-organization by knockdown of genes previously shown to affect stem cell colony organization. These results demonstrate that morphogenic dynamics can be accurately predicted through model-driven exploration of hPSC behaviors via machine learning, thereby enabling spatial control of multicellular patterning to engineer human organoids and tissues.

3.6 Bibliography

- Bartocci, E., E. Aydin Gol, I. Haghghi, and C. Belta. 2016. "A Formal Methods Approach to Pattern Recognition and Synthesis in Reaction Diffusion Networks." *IEEE Transactions on Control of Network Systems* PP (99): 1–1. <https://doi.org/10.1109/TCNS.2016.2609138>.
- Bredenoord, Annelien L., Hans Clevers, and Juergen A. Knoblich. 2017. "Human Tissues in a Dish: The Research and Ethical Implications of Organoid Technology." *Science* 355 (6322): eaaf9414. <https://doi.org/10.1126/science.aaf9414>.
- Briers, D., I. Haghghi, D. White, M. L. Kemp, and C. Belta. 2016. "Pattern Synthesis in a 3D Agent-Based Model of Stem Cell Differentiation." In *2016 IEEE 55th Conference on Decision and Control (CDC)*, 4202–7. <https://doi.org/10.1109/CDC.2016.7798907>.
- Cohen, William W. 1995. "Fast Effective Rule Induction." In *Machine Learning Proceedings 1995*, 115–123. Elsevier.
- Czirók, András, Katalin Varga, Előd Méhes, and András Szabó. 2013. "Collective Cell Streams in Epithelial Monolayers Depend on Cell Adhesion." *New Journal of Physics* 15 (7): 075006. <https://doi.org/10.1088/1367-2630/15/7/075006>.
- Deglincerti, Alessia, Gist F. Croft, Lauren N. Pietila, Magdalena Zernicka-Goetz, Eric D. Siggia, and Ali H. Brivanlou. 2016. "Self-Organization of the *in Vitro* Attached Human Embryo." *Nature* 533 (7602): 251–54. <https://doi.org/10.1038/nature17948>.
- Demers, Christopher J., Prabakaran Soundararajan, Phaneendra Chennampally, Gregory A. Cox, James Briscoe, Scott D. Collins, and Rosemary L. Smith. 2016. "Development-on-Chip: In Vitro Neural Tube Patterning with a Microfluidic Device." *Development* 143 (11): 1884–92. <https://doi.org/10.1242/dev.126847>.
- Ducibella, Thomas, and Everett Anderson. 1975. "Cell Shape and Membrane Changes in the Eight-Cell Mouse Embryo: Prerequisites for Morphogenesis of the Blastocyst." *Developmental Biology* 47 (1): 45–58. [https://doi.org/10.1016/0012-1606\(75\)90262-6](https://doi.org/10.1016/0012-1606(75)90262-6).

- Eberhart, R., and J. Kennedy. 1995. "A New Optimizer Using Particle Swarm Theory." In , *Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995. MHS '95*, 39–43. <https://doi.org/10.1109/MHS.1995.494215>.
- E. White, Douglas, Jonathan B. Sylvester, Thomas J. Levario, Hang Lu, J. Todd Strelman, Todd C. McDevitt, and Melissa L. Kemp. 2015. "Quantitative Multivariate Analysis of Dynamic Multicellular Morphogenic Trajectories." *Integrative Biology* 7 (7): 825–33. <https://doi.org/10.1039/C5IB00072F>.
- Finkel, R. A., and J. L. Bentley. 1974. "Quad Trees a Data Structure for Retrieval on Composite Keys." *Acta Informatica* 4 (1): 1–9. <https://doi.org/10.1007/BF00288933>.
- Geun Chung, Bong, Lisa A. Flanagan, Seog Woo Rhee, Philip H. Schwartz, Abraham P. Lee, Edwin S. Monuki, and Noo Li Jeon. 2005. "Human Neural Stem Cell Growth and Differentiation in a Gradient-Generating Microfluidic Device." *Lab on a Chip* 5 (4): 401–6. <https://doi.org/10.1039/B417651K>.
- Glen, Chad M., Todd C. McDevitt, and Melissa L. Kemp. 2018. "Dynamic Intercellular Transport Modulates the Spatial Patterning of Differentiation during Early Neural Commitment." *Nature Communications* 9 (1): 4111. <https://doi.org/10.1038/s41467-018-06693-1>.
- Graner, François, and James A. Glazier. 1992. "Simulation of Biological Cell Sorting Using a Two-Dimensional Extended Potts Model." *Physical Review Letters* 69 (13): 2013–16. <https://doi.org/10.1103/PhysRevLett.69.2013>.
- Greber, David, and Martin Fussenegger. 2010. "An Engineered Mammalian Band-Pass Network." *Nucleic Acids Research* 38 (18): e174–e174. <https://doi.org/10.1093/nar/gkq671>.
- Haghighi, Iman, Austin Jones, Zhaodan Kong, Ezio Bartocci, Radu Gros, and Calin Belta. 2015. "SpaTeL: A Novel Spatial-Temporal Logic and Its Applications to Networked Systems." In *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, 189–198. HSCC '15. New York, NY, USA: ACM. <https://doi.org/10.1145/2728606.2728633>.

- Hsiao, Sonny C., Betty J. Shum, Hiroaki Onoe, Erik S. Douglas, Zev J. Gartner, Richard A. Mathies, Carolyn R. Bertozzi, and Matthew B. Francis. 2009. "Direct Cell Surface Modification with DNA for the Capture of Primary Cells and the Investigation of Myotube Formation on Defined Patterns." *Langmuir* 25 (12): 6985–91. <https://doi.org/10.1021/la900150n>.
- Jackins, C. L., and S. L. Tanimoto. 1983. "Quad-Trees, Oct-Trees, and K-Trees: A Generalized Approach to Recursive Decomposition of Euclidean Space." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-5 (5): 533–39. <https://doi.org/10.1109/TPAMI.1983.4767433>.
- Krieg, M., Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. J. Müller, and C.-P. Heisenberg. 2008. "Tensile Forces Govern Germ-Layer Organization in Zebrafish." *Nature Cell Biology* 10 (4): 429–36. <https://doi.org/10.1038/ncb1705>.
- Libby, Ashley RG, David A Joy, Po-Lin So, Mohammad A Mandegar, Jonathon M Muncie, Federico N Mendoza-Camacho, Valerie M Weaver, Bruce R Conklin, and Todd C McDevitt. 2018. "Spatiotemporal Mosaic Self-Patterning of Pluripotent Stem Cells Using CRISPR Interference." *ELife* 7 (October): e36045. <https://doi.org/10.7554/eLife.36045>.
- Magno, Ramiro, Verônica A. Grieneisen, and Athanasius FM Marée. 2015. "The Biophysical Nature of Cells: Potential Cell Behaviours Revealed by Analytical and Computational Studies of Cell Surface Mechanics." *BMC Biophysics* 8 (May): 8. <https://doi.org/10.1186/s13628-015-0022-x>.
- Mandegar, Mohammad A., Nathaniel Huebsch, Ekaterina B. Frolov, Edward Shin, Annie Truong, Michael P. Olvera, Amanda H. Chan, et al. 2016. "CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs." *Cell Stem Cell* 18 (4): 541–53. <https://doi.org/10.1016/j.stem.2016.01.022>.
- Marcon, Luciano, Xavier Diego, James Sharpe, and Patrick Müller. 2016. "High-Throughput Mathematical Analysis Identifies Turing Networks for Patterning with Equally Diffusing

- Signals.” Edited by Naama Barkai. *ELife* 5 (April): e14022.
<https://doi.org/10.7554/eLife.14022>.
- Marée, Athanasius F. M., Verônica A. Grieneisen, and Paulien Hogeweg. 2007. “The Cellular Potts Model and Biophysical Properties of Cells, Tissues and Morphogenesis.” In *Single-Cell-Based Models in Biology and Medicine*, edited by Alexander R. A. Anderson, Mark A. J. Chaplain, and Katarzyna A. Rejniak, 107–36. Basel: Birkhäuser Basel.
https://doi.org/10.1007/978-3-7643-8123-3_5.
- Montero, Juan-Antonio, and Carl-Philipp Heisenberg. 2004. “Gastrulation Dynamics: Cells Move into Focus.” *Trends in Cell Biology* 14 (11): 620–27.
<https://doi.org/10.1016/j.tcb.2004.09.008>.
- Niwa, Hitoshi, Yayoi Toyooka, Daisuke Shimosato, Dan Strumpf, Kadue Takahashi, Rika Yagi, and Janet Rossant. 2005. “Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation.” *Cell* 123 (5): 917–29.
<https://doi.org/10.1016/j.cell.2005.08.040>.
- Ouchi, Noriyuki Bob, James A. Glazier, Jean-Paul Rieu, Arpita Upadhyaya, and Yasuji Sawada. 2003. “Improving the Realism of the Cellular Potts Model in Simulations of Biological Cells.” *Physica A: Statistical Mechanics and Its Applications* 329 (3): 451–58.
[https://doi.org/10.1016/S0378-4371\(03\)00574-0](https://doi.org/10.1016/S0378-4371(03)00574-0).
- Pir, Pınar, and Nicolas Le Novère. 2016. “Mathematical Models of Pluripotent Stem Cells: At the Dawn of Predictive Regenerative Medicine.” In *Systems Medicine*, edited by Ulf Schmitz and Olaf Wolkenhauer, 331–50. New York, NY: Springer New York.
https://doi.org/10.1007/978-1-4939-3283-2_15.
- Sasai, Yoshiki. 2013. “Cytosystems Dynamics in Self-Organization of Tissue Architecture.” *Nature* 493 (7432): 318–26. <https://doi.org/10.1038/nature11859>.

- Sekine, Ryoji, Tatsuo Shibata, and Miki Ebisuya. 2018. "Synthetic Mammalian Pattern Formation Driven by Differential Diffusivity of Nodal and Lefty." *Nature Communications* 9 (1): 5456. <https://doi.org/10.1038/s41467-018-07847-x>.
- Sohka, Takayuki, Richard A. Heins, Ryan M. Phelan, Jennifer M. Greisler, Craig A. Townsend, and Marc Ostermeier. 2009. "An Externally Tunable Bacterial Band-Pass Filter." *Proceedings of the National Academy of Sciences* 106 (25): 10135–40. <https://doi.org/10.1073/pnas.0901246106>.
- Tewary, Mukul, Joel Ostblom, Laura Prochazka, Teresa Zulueta-Coarasa, Nika Shakiba, Rodrigo Fernandez-Gonzalez, and Peter W. Zandstra. 2017. "A Stepwise Model of Reaction-Diffusion and Positional-Information Governs Self-Organized Human Peri-Gastrulation-like Patterning." *Development*, January, dev.149658. <https://doi.org/10.1242/dev.149658>.
- Théry, Manuel. 2010. "Micropatterning as a Tool to Decipher Cell Morphogenesis and Functions." *J Cell Sci* 123 (24): 4201–13. <https://doi.org/10.1242/jcs.075150>.
- Théry, Manuel, Victor Racine, Matthieu Piel, Anne Pépin, Ariane Dimitrov, Yong Chen, Jean-Baptiste Sibarita, and Michel Bornens. 2006. "Anisotropy of Cell Adhesive Microenvironment Governs Cell Internal Organization and Orientation of Polarity." *Proceedings of the National Academy of Sciences* 103 (52): 19771–76. <https://doi.org/10.1073/pnas.0609267103>.
- Toda, Satoshi, Lucas R. Blauch, Sindy K. Y. Tang, Leonardo Morsut, and Wendell A. Lim. 2018. "Programming Self-Organizing Multicellular Structures with Synthetic Cell-Cell Signaling." *Science*, May, eaat0271. <https://doi.org/10.1126/science.aat0271>.
- Wang, Zheng, Efrat Oron, Brynna Nelson, Spiro Razis, and Natalia Ivanova. 2012. "Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells." *Cell Stem Cell* 10 (4): 440–54. <https://doi.org/10.1016/j.stem.2012.02.016>.

- Warmflash, Aryeh, Benoit Sorre, Fred Etoc, Eric D. Siggia, and Ali H. Brivanlou. 2014. "A Method to Recapitulate Early Embryonic Spatial Patterning in Human Embryonic Stem Cells." *Nature Methods* 11 (8): 847–54. <https://doi.org/10.1038/nmeth.3016>.
- White, Douglas E., Melissa A. Kinney, Todd C. McDevitt, and Melissa L. Kemp. 2013. "Spatial Pattern Dynamics of 3D Stem Cell Loss of Pluripotency via Rules-Based Computational Modeling." *PLOS Computational Biology* 9 (3): e1002952. <https://doi.org/10.1371/journal.pcbi.1002952>.
- Ziomek, C. A., and M. H. Johnson. 1980. "Cell Surface Interaction Induces Polarization of Mouse 8-Cell Blastomeres at Compaction." *Cell* 21 (3): 935–42. [https://doi.org/10.1016/0092-8674\(80\)90457-2](https://doi.org/10.1016/0092-8674(80)90457-2).

Chapter 4: Loss of E-cadherin promotes extraembryonic fates accompanying multilineage differentiation

4.1 Introduction

The previous chapters discuss the importance of multicellular pattern formation in early embryonic development and how human induced pluripotent stem cells (iPSCs) can be used to model symmetry breaking and subsequent organizational events. However, the previous studies do not take into account how the additional degrees of freedom resulting from a three dimensional space contributes to the morphogenesis and emergence of form within the early embryo. Therefore, a three dimensional human iPSC culture system can be used to explore the influence of cell-cell adhesions on the multi-dimensional organization and patterning of the human embryo.

The human embryo transitions through several key three dimensional organizational states. First, the embryo generates cystic cavities and apolar groups of cells in the process of compaction forming the trophectoderm and inner cell mass (Campbell et al. 1995; Ziomek and Johnson 1980; Fleming 1987). The embryo then undergoes zona pellucida hatching, which involves epithelial sheet protrusions and extensions as the early blastocyst degrades a section of the surrounding zona pellucida and propels itself out of the now broken encasing. Similarly, gastrulation involves the delamination and invasion of the cells that comprise the primitive streak into the space between the epiblast and developing yolk sac, creating an invagination (Tada, Concha, and Heisenberg 2002; Montero and Heisenberg 2004; Sun and Stathopoulos 2018). Therefore, to interrogate how symmetry breaking regulates morphogenic events like compaction and gastrulation in a human context, the development of 3D human iPSC platforms that can both recapitulate lineage population emergence as well as multicellular organization are essential.

In this study, we interrogate symmetry breaking events in 3D culture. Specifically, we examine how heterogenous changes in adhesion through mosaic knockdown of the cell-cell

adhesion molecule E-cadherin (CDH1) regulate both 3D aggregate organization and emergent lineages. We genetically induce mosaic loss of CDH1 by CRISPRi in 50 cell aggregates encapsulated in alginate hydrogel, employing hydrogel encapsulation to mimic the physically confined environment of the zona pellucida. We show that 50 cell aggregates in suspension culture undergo population emergence of the three germ lineages similar to pre-implantation embryos. Additionally, we show that the combination of CDH1 knockdown and encapsulation lead to specific emergence of extraembryonic populations in aggregates after 5 days of culture, highlighting potential regulation of trophoblast development triggered by microenvironment.

4.2 Materials and Methods

4.2.1 Human Induced Pluripotent Stem Cell Culture

All work with human induced pluripotent stem cells (iPSCs) was approved by the University of California, San Francisco Human Gamete, Embryo, and Stem Cell Research (GESCR) Committee. Human iPSC lines were derived from the WTC11 line (Coriell Cat. #GM25256), the WTB line (Conklin Lab)(Miyaoaka et al. 2014), and the Allen Institute WTC11-LaminB cell line (AICS-0013 cl.210). All cell lines were karyotyped by Cell Line Genetics and reported to be karyotypically normal. Additionally, all cell lines tested negative for mycoplasma using a MycoAlert Mycoplasma Detection Kit (Lonza).

Human iPSCs were cultured on growth factor reduced Matrigel (Corning Life Sciences) and fed daily with mTeSRTM-1 medium (STEMCELL Technologies)(Ludwig et al. 2006). Cells were passaged by dissociation with Accutase (STEM CELL Technologies) and re-seeded in mTeSRTM-1 medium supplemented with the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10 μ M; Selleckchem)(Park et al. 2015) at a seeding density of 12,000 cell per cm².

4.2.2 Encapsulated Mixed Aggregate Generation

Encapsulation was performed using previously reported protocols (Wilson et al. 2014) where encapsulated cell aggregates of ~50 cells were created using 400 X 400 μm PDMS microwell inserts in 24-well plates (~975 microwells per well) similar to previously published protocols (Hookway et al. 2015; Ungrin et al. 2008). Dissociated iPSC cultures were resuspended in mTeSRTM-1 supplemented with Y-27632(10 μM), mixed at proper ratios and concentration (50 cells/well), added to microwells, and centrifuged (200 rcf). After 18 hours of formation, 50 cell aggregates were transferred into 1.5% ultrapure MVM alginate (Pronova) mixed with Laminin from Engelbreth-Holm-Swarm murine sarcoma (6 $\mu\text{g}/\text{mL}$; Sigma Aldrich) at a concentration of 16,000 aggregates/mL alginate. Alginate solution was prepared by mixing the appropriate amount of alginate to generate a 1.5% solution into calcium-free DMEM (Gibco) and sterilized by autoclave. Beads encapsulating single aggregates were generated using an electrostatic bead generator (Nisco). A 400 μm nozzle and syringe pump (flow of 6mL/hour) was used to extrude alginate solution with aggregates and dropped into a 100 mM calcium chloride (EMD) bath to trigger hardening of the alginate into a gel. Encapsulated aggregates were then washed 3X with DPBS containing calcium and magnesium (ThermoFisher Scientific) and once with mTeSRTM-1 medium (STEMCELL Technologies). Encapsulated aggregates were then allowed to recover for 24 hours in mTeSRTM-1 medium (STEMCELL Technologies) in rotary suspension. Then aggregates were fed daily with mTeSRTM-1 medium supplemented with doxycycline (DOX)(2 μM) to induce CDH1 knockdown.

4.2.3 Un-encapsulated Mixed Aggregate Generation

Unencapsulated 50 cell aggregates were created by dissociation of human iPSCs with Accutase (STEM CELL Technologies) and reseeded in mTeSRTM-1 medium (STEMCELL

Technologies) supplemented with Y-276932 (10 μ M; Selleckchem) into 96-well non-adherent round bottom plates, where a total of 50 cells were seeded per well. After 18 hours of aggregate formation, Y-276932 was removed from the media and DOX(2 μ M) was supplemented into the mTeSRTM-1 medium to induce CDH1 knockdown. Aggregates were then fed daily with mTeSRTM-1 medium supplemented with DOX (2 μ M).

4.2.4 Immunofluorescence Staining and Imaging

Aggregates were unencapsulated by washing 3X with a sodium citrate solution (55mM, Sigma), fixed with 4% paraformaldehyde (VWR) for 40 minutes, and then washed three times with PBS. Aggregates to be used for histology were embedded in HistoGel Specimen Processing Gel (Thermo Fisher) prior to paraffin processing. Paraffin embedded samples were sectioned in 5 μ m sections, baked for 1 hour at 60°C, and subsequently stained for H&E. For immunofluorescent staining, epitope retrieval was performed by submersing slides in Citrate Buffer pH 6.0 (Vector Laboratories) in a 95°C water bath for 35min. Samples were permeabilized in 0.2% Triton X-100 (Sigma-Aldrich) for 5min, blocked in 1.5% normal donkey serum (Jackson Immunoresearch) for 1hour, and probed with primary and secondary antibodies against SOX2, PAX6, T, NES, TUBB3, and CDH2 (Table S3). Coverslips were mounted with anti-fade mounting medium (ProlongGold, Life Technologies) and samples were imaged on a Zeiss Axio Observer Z1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 camera.

4.2.5 Whole Mount Lightsheet Imaging

4% paraformaldehyde-fixed paraffin-embedded samples (see “*Histology, Immunocytochemistry, and Imaging*”) were permeabilized with 0.3% Triton X-100 (Sigma-Aldrich) for 5min, blocked in 5% normal donkey serum (Jackson Immunoresearch) for 1 hour,

and probed with primary and secondary antibodies against PAX6 and T (TableS3) for 2 hours. Samples were then embedded in 1.5% low melt agarose (BioReagent) and drawn up into ~1mm imaging capillaries and subsequently imaged on the Zeiss Z.1 Light sheet Microscope equipped with a PCO.edge SCMOS camera.

4.2.6 Single Cell RNA Sequencing Sample and Library Preparation

Multiple organoid samples were combined and processed together using the MULTI-Seq technology (McGinnis et al. 2019). Organoids were singularized using Accutase (STEMCELL Technologies) and washed with cold PBS. Cells were resuspended in PBS with lipid-modified Anchor and Barcode oligonucleotides (gift from Zev Gartner) and incubated on ice for 5 min. A co-Anchor oligo was then added in order to stabilize membrane retention of the barcodes incubated for an additional 5 min on ice. Excess lipid-modified oligos were quenched with 1% BSA in PBS, washed with cold 1% BSA solution, and counted using a Countess II FL (Life Technologies). Single cell GEMs and subsequent libraries were then prepared using the 10X Genomics Single Cell V2 protocol with an additional anchor specific primer during cDNA amplification to enrich barcode sequences.(McGinnis et al. 2019) Short barcode sequences (approx. 65-100bp determined by Bioanalyzer) were purified from cDNA libraries with two sequential SPRI bead cleanups. Barcode library preparation was performed according to the KAPA HiFi Hotstart (Kapa Biosystems) protocol to functionalize with the P5 sequencing adapter and library-specific RPIX barcode. Purified ~173bp barcode fragments were isolated with another SPRI bead cleanup and validation by Bioanalyzer.

The sample library was sequenced on an Illumina NovaSeq yielding an average of 41,112 reads per cell and 6,444 cells. The MULTI-Seq barcode library was sequenced on an Illumina NextSeq yielding an average of 9,882 reads per barcode and enabling sample assignment for

4,681 of 6,124 unique UMIs detected (76.4% recovery), using the demultiplexing code provided by the MULTI-Seq protocol (McGinnis et al. 2019).

4.2.9 Genome Annotation, RNA-seq Read Mapping, and Estimation of Gene and Isoform Expression

The sample library was aligned to the human GRCh38 reference genome using Cell Ranger v1.2.0 (10x Genomics). Gene expression levels were assessed using the Seurat v3.0.0 analysis pipeline (Butler et al. 2018). First cells were removed with fewer than 200 detected genes, fewer than 1,000 total detected transcripts, or which had greater than 10% mitochondrial gene expression. Next, expression levels were log normalized, and the top 2,000 variable genes calculated using the VST algorithm. The top 20 principal components were used to group cells into 12 clusters using a resolution of 0.4. Finally, top markers were detected for each cluster by detecting the top differentially expressed genes between one cluster and the remaining data set, where at least 25% of cells in the cluster expressed the gene and the gene was expressed at least 0.25 \log_2 fold-change different from the remaining population. Clusters and gene expression were visualized on a two dimensional UMAP projection of the first 20 principal components.

4.2.8 Data Acquisition, Processing, and Merging

Human single cell RNA sequencing data sets from previously published papers (Yan et al. 2013; Petropoulos et al. 2016; Blakeley et al. 2015) were downloaded from Gene Expression Omnibus (GEO66507, GSE36552, E-MTAB-3929). FastQ data sets were aligned to the human GRCh38 reference genome using STAR aligner (Dobin et al. 2013) and then a counts matrix was generated using the featureCounts software (Liao, Smyth, and Shi 2014). For the entire data set, all counts matrices were concatenated into one matrix. Each matrix was then read as a Seurat object, which could then be combined with other data sets and analyzed using the Seurat v3.0.0

analysis pipeline (Butler et al. 2018). Batch correction was performed using the Harmony algorithm (Korsunsky et al. 2018).

4.2.9 Cluster Analysis

To assign cluster identity, the top markers for each cluster were tested for GO term enrichment using the biological process “enrichGO” function in the R package “clusterProfiler” v3.12. (Yu et al. 2012) In addition, differentiation and lineage specification in each cluster was assessed by examining expression level of panels of pluripotency, mesendoderm, endoderm, ectoderm, mesoderm, and trophectoderm markers. Computational trajectory analysis of transitions between cell states was performed using Monocle 3 (Trapnell et al. 2014).

4.3 Results

4.3.1 Loss of CDH1 Promotes Protrusion Morphology when Encapsulated

To better mimic the scale and environment of the blastocyst as well as prevent aggregate fusion, 50 cell human iPSC aggregates were encapsulated in 1.5% alginate mixed with laminin and cultured for 6 days. By combining a physical barrier (unfunctionalized alginate) and a basement ECM molecule (laminin), alginate encapsulation served as a proxy for zona pellucida encapsulation (Figure 4.1A). To interrogate population emergence similar to loss of CDH1 during gastrulation, CDH1 loss was induced in either 0%, 25% or 100% of the aggregate population. At the start of CDH1 knockdown, CDH1(+), CDH1(-), and mixed aggregates displayed polarized behavior, creating a single layer of cells surrounding a cystic cavity at the center of each aggregate (Figure 4.1B). Over time, a bilayer of cells assembled at the growing aggregate surface while maintaining the central cyst, indicating that the polar behavior of human iPSCs in monolayer culture is initially maintained in suspension culture at this scale.

At day 3 of CDH1 knockdown, all aggregates displayed bilayers around a cystic cavity. Mixed aggregates displayed organization similar to those reported in the monolayer culture discussed in Chapter 2 where the CDH1(+) and CDH1(-) cells sorted from one another. Furthermore, the CDH1(-) cells clustered within the outer layer of the cystic aggregates, suggesting preference for edge of the aggregate (Figure 1C). Interestingly, while the CDH1(+) aggregates and the mixed aggregates remained spherical in shape, the CDH1(-) aggregates began to display protrusions where a group of cells collectively extended from the aggregate main body (Figure 1B,D). These extensions displayed similar morphologies to those present during zona pellucida hatching where extensions remained epithelial with no observed single cells migrating away from the aggregate into the alginate encapsulation.

4.3.2 Unencapsulated and encapsulated CDH1(-) aggregates display differing morphologies

To interrogate whether the observed extensions were a result of CDH1 knockdown or a response to encapsulation itself, 50 cell human iPSC aggregates were cultured without encapsulation over 5 days (Figure 4.2A,B). Similar to encapsulated aggregates, cystic cavities formed in unencapsulated aggregates by day 1 and persisted through culture. By day 3 in mixed aggregates, CDH1(+) cells had segregated from CDH1(-) cells, mirroring the behavior seen in encapsulated mixed aggregates (Figure 4.2C). At day 5 of knockdown CDH1(-) aggregates displayed morphologies where an outer layer of cells peeled off of an inner epithelial core displaying a cystic cavity (Figure 4.2D). The entire aggregate lost OCT4 expression while outer layer displayed EOMES expression and the inner layer displayed SOX2 expression, indicating a loss of pluripotency and subsequent bifurcation of cell populations into the mesendoderm and ectoderm lineages (Figure 4.2D). This suggests that the complete loss of CDH1 accelerated differentiation of human iPSCs while still allowing for emergence of multiple lineages.

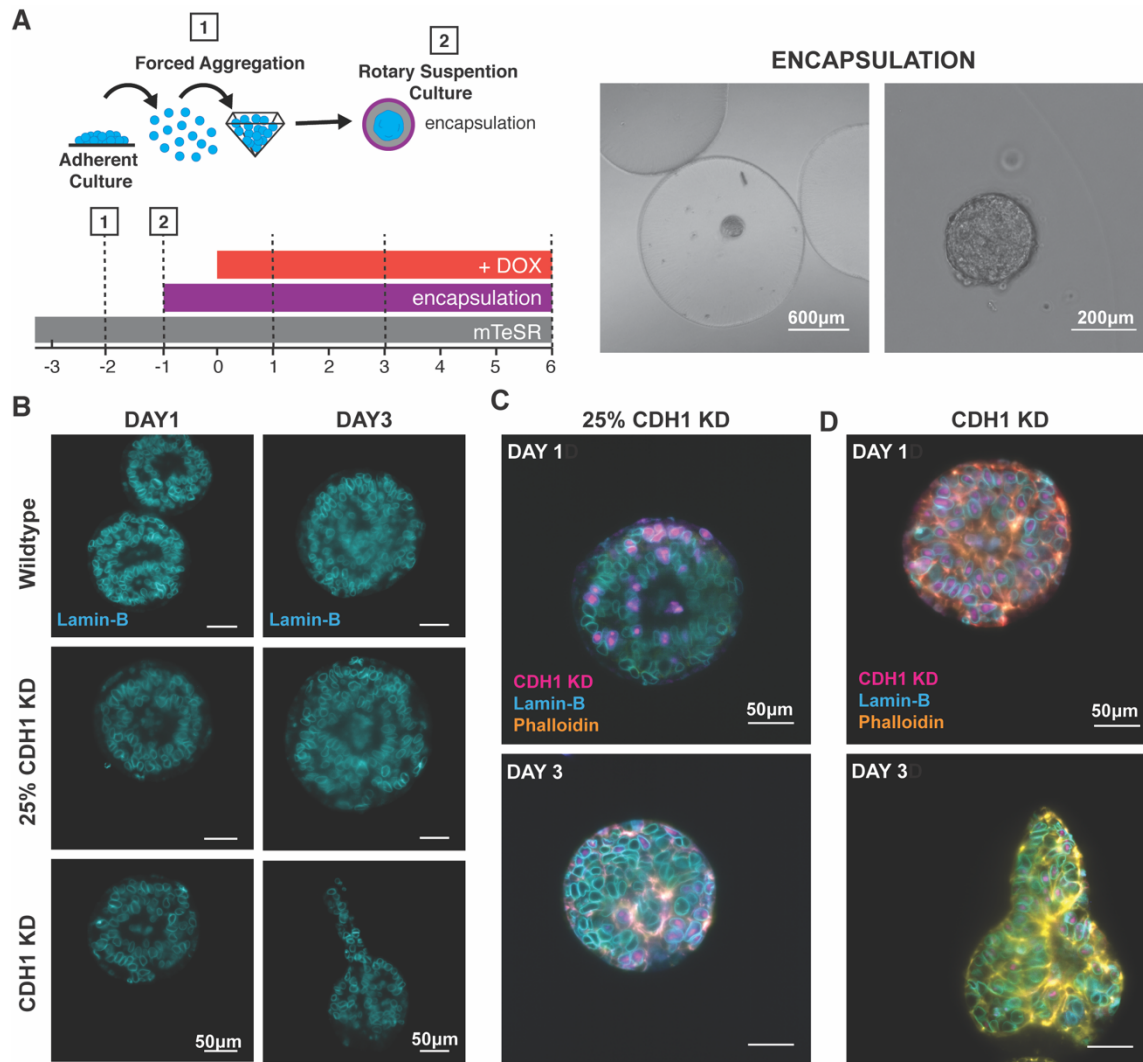


Figure 4.1: Encapsulated human iPSC 50 cell aggregate morphologic changes. (A) A schematic of the experimental set up of encapsulation (left) and example images of encapsulated 50 cell human iPSC aggregates in alginate gel 20 minutes post encapsulation. **(B)** Optical sections of aggregates demonstrating evolution of morphologies on day 1 and day 3. **(C)** Optical section of mixed aggregates with 25% CDH1 knockdown. **(D)** Optical sections of aggregates lacking CDH1.

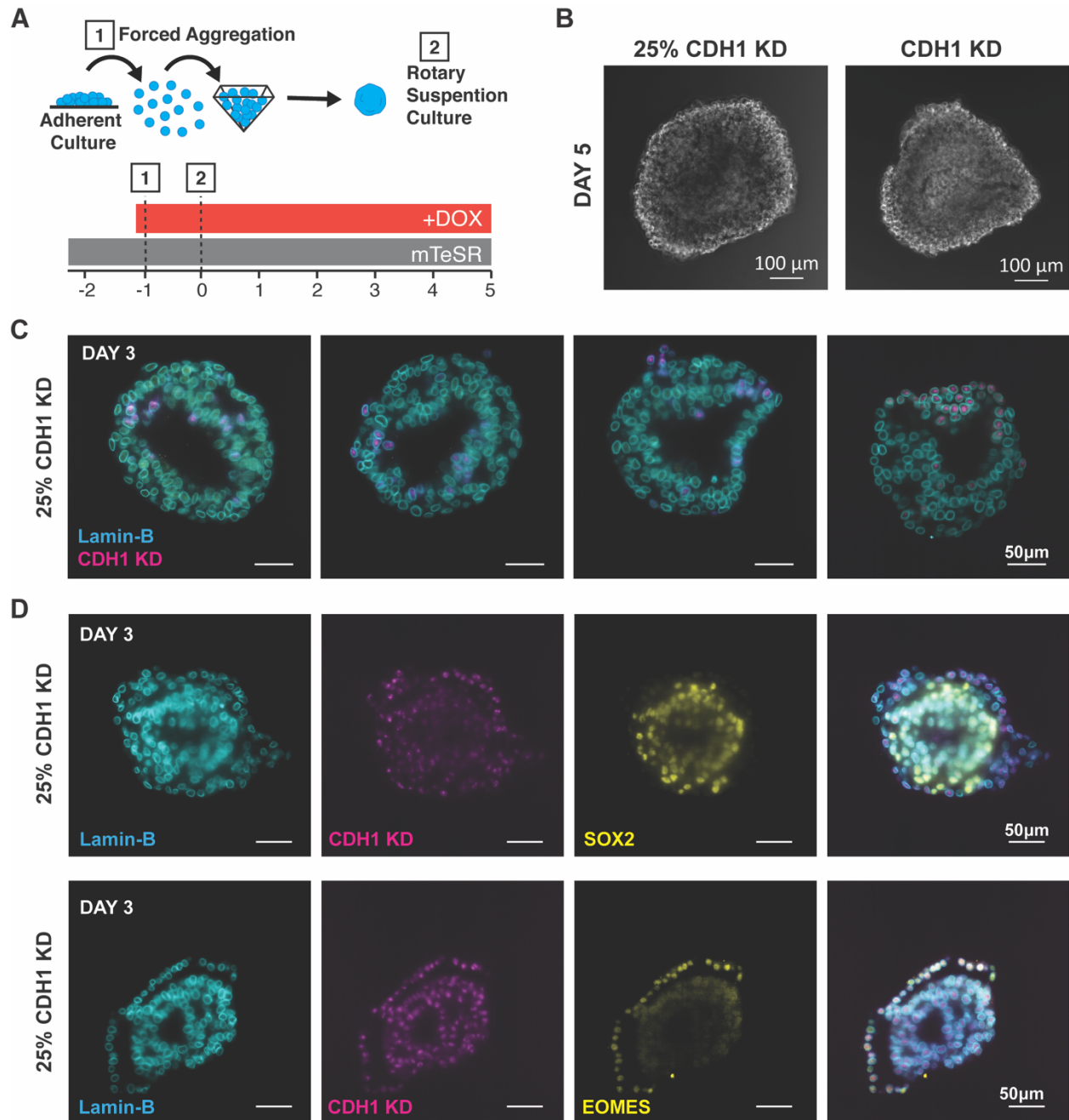


Figure 4.2: Unencapsulated human iPSC 50 cell aggregate morphologies. (A) A schematic of the experimental set up. (B) Example images of non-encapsulated 50 cell human iPSC aggregates in 96-well round bottom plates. (C) Optical sections of mixed aggregates with 25% CDH1 knockdown demonstrating maintenance of cysts and segregation of CDH1(-) cells by day 3. (D) Optical sections of aggregates lacking CDH1 demonstrating changes in morphology and lineage fate spatial segregation.

4.3.3 Emergence of Extraembryonic-like Cell Fates in Encapsulated Aggregates

With the emergence of multiple morphologies and populations in the aggregates, we interrogated the transcriptional expression between encapsulated and unencapsulated aggregates by single cell RNA sequencing, which allowed for identification of the diversity of cellular populations within the differentiating aggregates. The transcriptomes of encapsulated and unencapsulated aggregates at day 1, 3, and 6 after CDH1 knockdown were examined across CDH1(+), CDH1(-), and mixed aggregates (Figure 4.3A,B). The resulting data set of single cell transcriptomes clustered into 12 clusters representing 8 cell states by lineage markers. At day 1, all aggregate types overlapped in a cluster marked by high expression of pluripotency genes (OCT4, SOX2, NANOG) (Figure 4.3C). By day 3 and 5 aggregates transitioned through multiple lineage states eventually clustering into populations expressing markers of mesendoderm, ectoderm, endoderm, and mesoderm (Figure 4.3C). Interestingly, clusters 8 and 11 displayed markers of the extraembryonic trophoderm (CDX2, GATA3, CDH3, HAND1) (Figure 3C). Gene Ontology analysis of these two clusters revealed gene networks associated with placental development and reproductive system development (Figure 4.3D). Furthermore, these clusters were largely populated by CDH1(-) cells that had undergone encapsulation (Figure 4.3E) and high in expression of pathways associated with trophoderm development such as Hippo signaling and noncanonical Wnt signaling (Nishioka et al. 2009; Hirate et al. 2013; Gueth-Hallonet and Maro 1992; Tada, Concha, and Heisenberg 2002) (Figure 4.3F).

To computationally investigate whether this extraembryonic-like population was diverging from the other three germ lineages early on in the spontaneous aggregate differentiation, cellular trajectories indicating transitions between populations were examined. Cellular trajectory reconstruction revealed that the differentiating population bifurcated after day 1, with one branch transitioning through primitive streak and then mesoderm lineage states, while a separate trajectory went toward an extra-embryonic lineage state (Figure 4.4A). The mesendoderm, endoderm, and ectoderm germ lineages clustered separately from the lineage trajectories of

pluripotency to mesoderm or to extraembryonic (Figure 4.4A), indicating that the transition from pluripotency occurred on a faster time scale that was not captured by sampling at day 1,3, and 5, preventing computational reconstruction of the complete lineage trajectory from pluripotency. Overall, the early emergence of the bifurcation between the germ lineages specific to the epiblast (mesoderm, endoderm, ectoderm) and the extraembryonic-like lineage in pseudo-time resembles the population emergence at compaction of the embryo (Figure 4.4B), suggesting that this system can be used to model early developmental lineage bifurcations.

4.3.4 Human iPSC Aggregates Have Similar Transcriptomes to Pre-implantation Human Embryos

To interrogate whether the emergence of lineages seen on our encapsulated and unencapsulated cell aggregates resembled that of the early human embryo (Figure 4.5A), our *in vitro* data set was clustered with three previously published single cell transcriptome data sets from preimplantation embryos (Petropoulos et al. 2016; Yan et al. 2013; Blakeley et al. 2015) (Figure 4.5). The preimplantation blastocysts clusters overlapped with *in vitro* aggregate cell transcriptomes, clustering together in the Yan et al. data set (Figure 4.5B) and in similar multidimensional space in the Petropoulos et. al. and Blakeley et al. data sets (Figure 4.5C,D). This indicates that although not perfectly replicating the identity of early human embryos, the *in vitro* culture system demonstrates similar transcriptomics. Interestingly, the extraembryonic-like cells from the encapsulated aggregates, marked by high CDX2 and HAND1 expression, clustered closely to the preimplantation embryos (Figure 4.5B,C,D). Overall, the similarities in transcriptome suggest that encapsulation paired with CDH1 knockdown enhances activity of gene regulatory networks that mimic processes in preimplantation embryos.

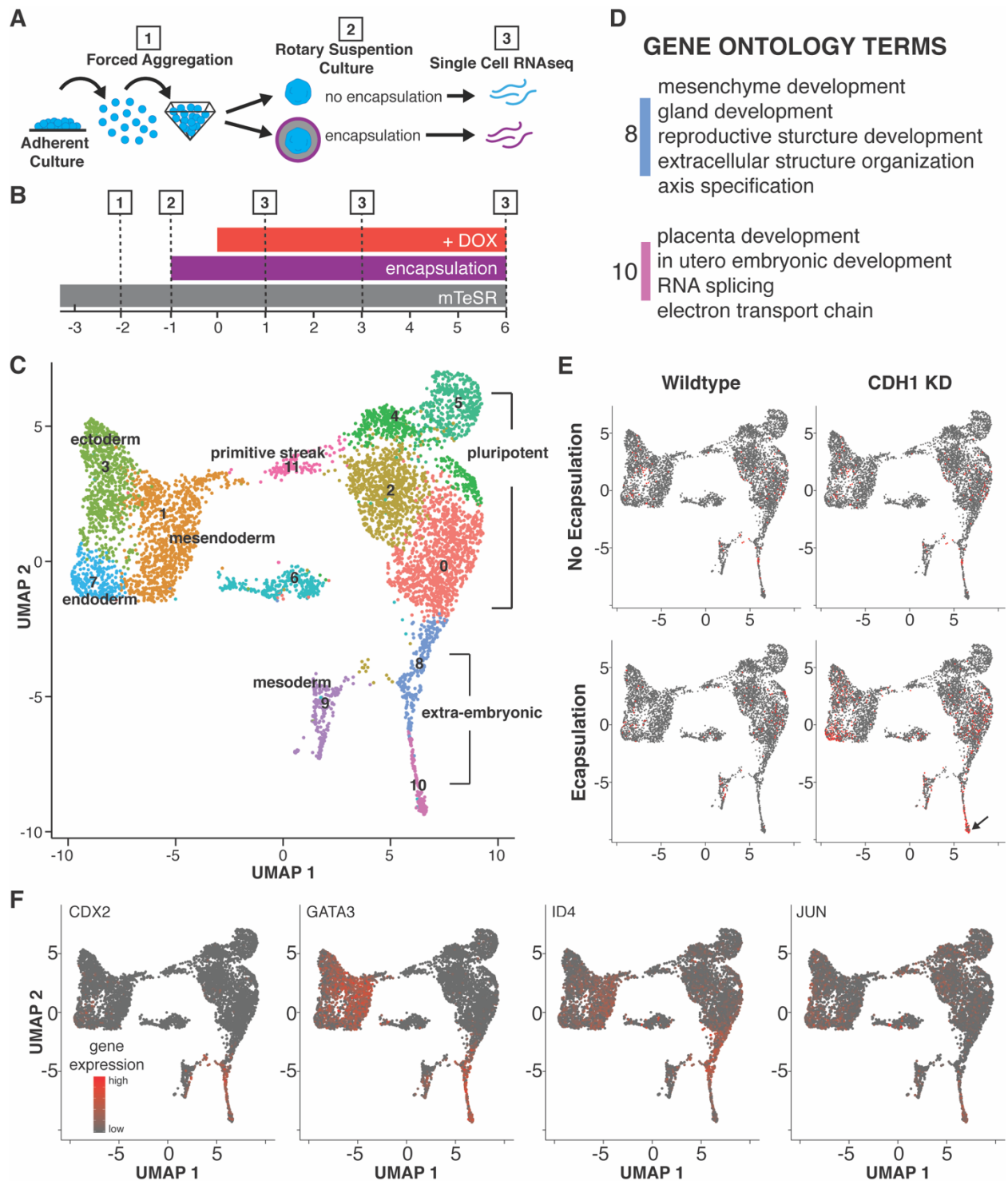


Figure 4.3: Single cell sequencing analysis of encapsulated and unencapsulated aggregates. (A,B) A schematic of the experimental set up. **(C)** UMAP demonstrating 12 clusters of cell populations at day 1,3, and 6. **(D)** Gene ontology terms for extraembryonic-like clusters 8 and 10 ($p < 0.05$). **(E)** Distribution of encapsulated and CDH1 knockdown cells within the UMAP clusters of all samples. **(F)** Distribution of genes associated with trophoblast.

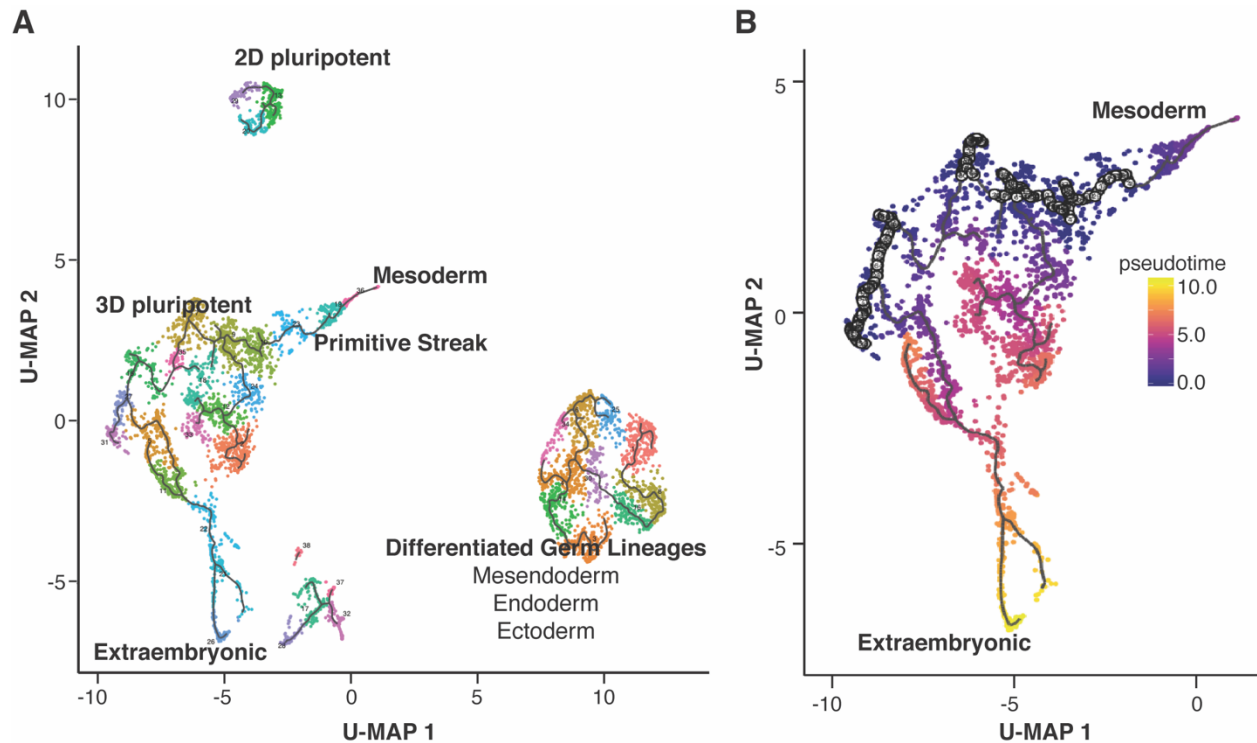


Figure 4.4: Lineage trajectories of 50 cell aggregates. (A) Lineage trajectory reconstruction using Monocle demonstrating Mesoderm and Extraembryonic cell fates on separate lineage branches. **(B)** Pseudotime reconstruction of the mesoderm and extraembryonic lineages.

4.4 Discussion

The emergence and coincident organization of multiple lineages in the developing embryo is essential to the functional structure of developmental and reproductive tissues and additionally establishes the body plan necessary for proper development. However, precisely how this process is robustly regulated with the limited signaling molecules and morphogens available to the developing embryo is not fully understood. In this study, we demonstrate that suspension culture of hydrogel encapsulated 50 cell human iPSC aggregates both transcriptionally and morphologically replicates aspects of early human development and lineage emergence without

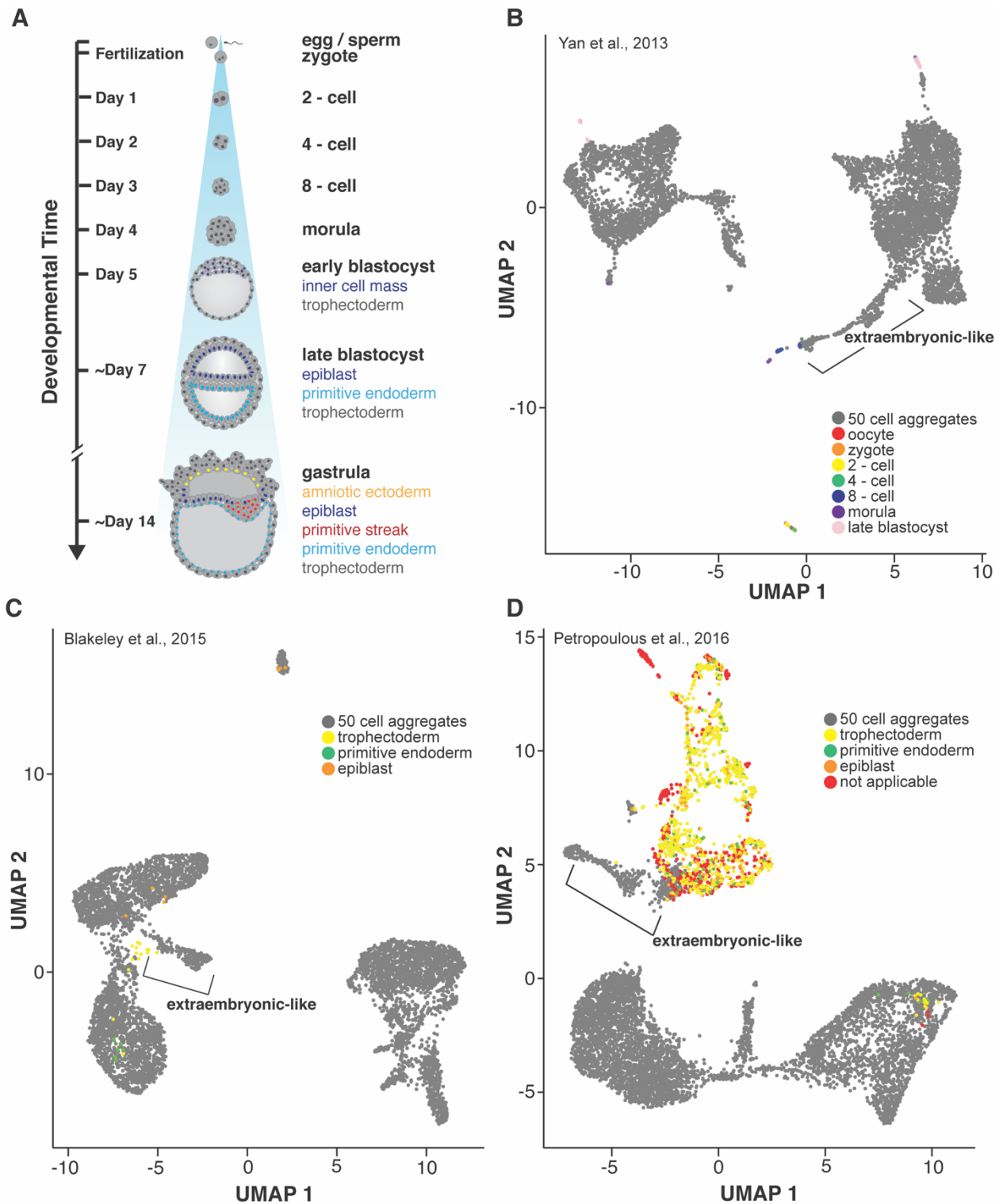


Figure 4.5: Comparison of *in vitro* aggregate transcriptome with preimplantation human embryos. (A) Schematic of early human embryonic development. (B-C) UMAPs displaying 50 cell encapsulated and non-encapsulated aggregate transcriptomes (gray) with three single-cell sequencing data sets (Yan et al., Blakeley et al., and Petropoulos et al., respectively) where the extraembryonic-like population is labeled with brackets.

addition of exogenous morphogens. Additionally, CDH1 knockdown within encapsulated aggregates enabled the emergence of a previously unreported extraembryonic-like population reminiscent of trophoblasts. These results demonstrate, first that human iPSCs inherently can establish signaling networks necessary to trigger symmetry breaking, and second that the changes in adhesion in the developing embryo may directly control lineage decisions, rather than simply emerging as a result of fate commitment. Furthermore, this study suggests that microenvironment plays a role in the coordination of lineage fate decisions, highlighting a possible regulatory mechanism employed by the early embryo to robustly control population emergence with a limited milieu of signaling molecules.

Currently the study of symmetry breaking events *in vitro* often rely on the exogenous presence of morphogens to trigger emergence of patterning and multiple cell types in monolayer or suspension culture (Lancaster and Knoblich 2014; Beccari et al. 2018; Brink et al. 2014, 3; Simunovic et al. 2019; Warmflash et al. 2014; Tewary et al. 2017) where a single morphogen regime gives rise to symmetry breaking events that derive multiple cell types. However, the emergence of multiple germ lineages in this study suggests that only the transition from monolayer culture to suspension culture is sufficient to enable the emergence of multiple lineages. As the embryo develops it goes through regimented changes in polarity that dictate developmental transitions. For example, compaction relies on the polarization of the trophoblasts to allow for cavity formation (Hirate et al. 2013; Gueth-Hallonet and Maro 1992), while gastrulation relies on the loss of polarity as invaginating cells begin their migration across the embryo (Tada, Concha, and Heisenberg 2002; Montero and Heisenberg 2004). Therefore, it is possible that the forced loss of adhesions in the CDH1 knockdown aggregates are triggering a signaling network that regulates polarity dependent lineage transitions.

Additionally, it has been previously reported that the mechanical micro-environment has the potential to regulate lineage fate emergence in pluripotent stems cells suggesting that micro-environment may contribute to embryonic development as well (Przybyla, Lakins, and Weaver

2016; Krieg et al. 2008). The emergence of a trophoblast-like population primarily in encapsulated CDH1 knockdown aggregates suggests that the physical micro-environment provided by encapsulation provides necessary cues for extraembryonic fate specification. Interestingly, trophoblast development is primarily regulated by Hippo signaling, a mechanically responsive signaling pathway (Nishioka et al. 2009; Janmey et al. 2013). Thus, it is possible that the enhanced emergence of trophoblast-like cells with encapsulation is a result of a mechanically triggered upregulation in Hippo signaling. Ultimately, the combination of suspension culture and encapsulation enables better replication of embryonic organization *in vitro*, highlighting additional mechanisms by which embryos maintain precise regulation over their development.

4.5 Conclusion

In conclusion, this third study provides an example of how combining bioengineering platforms such as encapsulation enables interrogation and discovery of regulators involved in the developmental processes underlying 3D morphogenesis and lineage emergence in the early embryo. These results demonstrate the combinatorial effect of adhesion regulation and microenvironment impacts lineage emergence and population morphogenesis in human iPSCs, potentially reflecting mechanisms within the early human embryo to robustly regulate development.

4.6 Bibliography

- Beccari, Leonardo, Naomi Moris, Mehmet Girgin, David A. Turner, Peter Baillie-Johnson, Anne-Catherine Cossy, Matthias P. Lutolf, Denis Duboule, and Alfonso Martinez Arias. 2018. "Multi-Axial Self-Organization Properties of Mouse Embryonic Stem Cells into Gastruloids." *Nature*, October, 1. <https://doi.org/10.1038/s41586-018-0578-0>.
- Blakeley, Paul, Norah M. E. Fogarty, Ignacio del Valle, Sissy E. Wamaita, Tim Xiaoming Hu, Kay Elder, Philip Snell, Leila Christie, Paul Robson, and Kathy K. Niakan. 2015. "Defining the Three Cell Lineages of the Human Blastocyst by Single-Cell RNA-Seq." *Development* 142 (18): 3151–65. <https://doi.org/10.1242/dev.123547>.
- Brink, Susanne C. van den, Peter Baillie-Johnson, Tina Balayo, Anna-Katerina Hadjantonakis, Sonja Nowotschin, David A. Turner, and Alfonso Martinez Arias. 2014. "Symmetry Breaking, Germ Layer Specification and Axial Organisation in Aggregates of Mouse Embryonic Stem Cells." *Development* 141 (22): 4231–42. <https://doi.org/10.1242/dev.113001>.
- Butler, Andrew, Paul Hoffman, Peter Smibert, Efthymia Papalexi, and Rahul Satija. 2018. "Integrating Single-Cell Transcriptomic Data across Different Conditions, Technologies, and Species." *Nature Biotechnology* 36 (5): 411–20. <https://doi.org/10.1038/nbt.4096>.
- Campbell, S., H. R. Swann, M. W. Seif, S. J. Kimber, and J. D. Aplin. 1995. "Integrins and Adhesion Molecules: Cell Adhesion Molecules on the Oocyte and Preimplantation Human Embryo." *Human Reproduction* 10 (6): 1571–78. <https://doi.org/10.1093/HUMREP/10.6.1571>.
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.

- Fleming, Tom P. 1987. "A Quantitative Analysis of Cell Allocation to Trophectoderm and Inner Cell Mass in the Mouse Blastocyst." *Developmental Biology* 119 (2): 520–31. [https://doi.org/10.1016/0012-1606\(87\)90055-8](https://doi.org/10.1016/0012-1606(87)90055-8).
- Gueth-Hallonet, Catherine, and Bernard Maro. 1992. "Cell Polarity and Cell Diversification during Early Mouse Embryogenesis." *Trends in Genetics* 8 (8): 274–79. [https://doi.org/10.1016/0168-9525\(92\)90253-Z](https://doi.org/10.1016/0168-9525(92)90253-Z).
- Hirate, Yoshikazu, Shino Hirahara, Ken-ichi Inoue, Atsushi Suzuki, Vernadeth B. Alarcon, Kazunori Akimoto, Takaaki Hirai, et al. 2013. "Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos." *Current Biology* 23 (13): 1181–94. <https://doi.org/10.1016/j.cub.2013.05.014>.
- Hookway, Tracy A., Jessica C. Butts, Emily Lee, Hengli Tang, and Todd C. McDevitt. 2015. "Aggregate Formation and Suspension Culture of Human Pluripotent Stem Cells and Differentiated Progeny." *Methods*. <https://doi.org/10.1016/j.ymeth.2015.11.027>.
- Janmey, Paul A., Rebecca G. Wells, Richard K. Assoian, and Christopher A. McCulloch. 2013. "From Tissue Mechanics to Transcription Factors." *Differentiation, Mechanotransduction*, 86 (3): 112–20. <https://doi.org/10.1016/j.diff.2013.07.004>.
- Korsunsky, Ilya, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po-Ru Loh, and Soumya Raychaudhuri. 2018. "Fast, Sensitive, and Accurate Integration of Single Cell Data with Harmony." *BioRxiv*, November, 461954. <https://doi.org/10.1101/461954>.
- Krieg, M., Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. J. Müller, and C.-P. Heisenberg. 2008. "Tensile Forces Govern Germ-Layer Organization in Zebrafish." *Nature Cell Biology* 10 (4): 429–36. <https://doi.org/10.1038/ncb1705>.
- Lancaster, Madeline A., and Juergen A. Knoblich. 2014. "Organogenesis in a Dish: Modeling Development and Disease Using Organoid Technologies." *Science* 345 (6194): 1247125. <https://doi.org/10.1126/science.1247125>.

- Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. "FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30 (7): 923–30. <https://doi.org/10.1093/bioinformatics/btt656>.
- Ludwig, Tenneille E., Veit Bergendahl, Mark E. Levenstein, Junying Yu, Mitchell D. Probasco, and James A. Thomson. 2006. "Feeder-Independent Culture of Human Embryonic Stem Cells." *Nature Methods* 3 (8): 637–46. <https://doi.org/10.1038/nmeth902>.
- McGinnis, Christopher S., David M. Patterson, Juliane Winkler, Daniel N. Conrad, Marco Y. Hein, Vasudha Srivastava, Jennifer L. Hu, et al. 2019. "MULTI-Seq: Sample Multiplexing for Single-Cell RNA Sequencing Using Lipid-Tagged Indices." *Nature Methods* 16 (7): 619–26. <https://doi.org/10.1038/s41592-019-0433-8>.
- Miyaoka, Yuichiro, Amanda H. Chan, Luke M. Judge, Jennie Yoo, Miller Huang, Trieu D. Nguyen, Paweena P. Lizarraga, Po-Lin So, and Bruce R. Conklin. 2014. "Isolation of Single-Base Genome-Edited Human IPS Cells without Antibiotic Selection." *Nature Methods* 11 (3): 291–93. <https://doi.org/10.1038/nmeth.2840>.
- Montero, Juan-Antonio, and Carl-Philipp Heisenberg. 2004. "Gastrulation Dynamics: Cells Move into Focus." *Trends in Cell Biology* 14 (11): 620–27. <https://doi.org/10.1016/j.tcb.2004.09.008>.
- Nishioka, Noriyuki, Ken-ichi Inoue, Kenjiro Adachi, Hiroshi Kiyonari, Mitsunori Ota, Amy Ralston, Norikazu Yabuta, et al. 2009. "The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass." *Developmental Cell* 16 (3): 398–410. <https://doi.org/10.1016/j.devcel.2009.02.003>.
- Park, Sangkyu, Daehwan Kim, Yeon-Gil Jung, and Sangho Roh. 2015. "Thiazovivin, a Rho Kinase Inhibitor, Improves Stemness Maintenance of Embryo-Derived Stem-like Cells under Chemically Defined Culture Conditions in Cattle." *Animal Reproduction Science* 161 (October): 47–57. <https://doi.org/10.1016/j.anireprosci.2015.08.003>.

- Petropoulos, Sophie, Daniel Edsgård, Björn Reinius, Qiaolin Deng, Sarita Pauliina Panula, Simone Codeluppi, Alvaro Plaza Reyes, Sten Linnarsson, Rickard Sandberg, and Fredrik Lanner. 2016. "Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos." *Cell* 165 (4): 1012–26. <https://doi.org/10.1016/j.cell.2016.03.023>.
- Przybyla, Laralynne, Johnathon N. Lakins, and Valerie M. Weaver. 2016. "Tissue Mechanics Orchestrate Wnt-Dependent Human Embryonic Stem Cell Differentiation." *Cell Stem Cell* 19 (4): 462–75. <https://doi.org/10.1016/j.stem.2016.06.018>.
- Simunovic, Mijo, Jakob J. Metzger, Fred Etoc, Anna Yoney, Albert Ruzo, Iain Martyn, Gist Croft, Dong Shin You, Ali H. Brivanlou, and Eric D. Siggia. 2019. "A 3D Model of a Human Epiblast Reveals BMP4-Driven Symmetry Breaking." *Nature Cell Biology* 21 (7): 900–910. <https://doi.org/10.1038/s41556-019-0349-7>.
- Sun, Jingjing, and Angelike Stathopoulos. 2018. "FGF Controls Epithelial-Mesenchymal Transitions during Gastrulation by Regulating Cell Division and Apicobasal Polarity." *Development* 145 (19): dev161927. <https://doi.org/10.1242/dev.161927>.
- Tada, Masazumi, Miguel L Concha, and Carl-Philipp Heisenberg. 2002. "Non-Canonical Wnt Signalling and Regulation of Gastrulation Movements." *Seminars in Cell & Developmental Biology* 13 (3): 251–60. [https://doi.org/10.1016/S1084-9521\(02\)00052-6](https://doi.org/10.1016/S1084-9521(02)00052-6).
- Tewary, Mukul, Joel Ostblom, Laura Prochazka, Teresa Zulueta-Coarasa, Nika Shakiba, Rodrigo Fernandez-Gonzalez, and Peter W. Zandstra. 2017. "A Stepwise Model of Reaction-Diffusion and Positional-Information Governs Self-Organized Human Peri-Gastrulation-like Patterning." *Development*, January, dev.149658. <https://doi.org/10.1242/dev.149658>.
- Trapnell, Cole, Davide Cacchiarelli, Jonna Grimsby, Prapti Pokharel, Shuqiang Li, Michael Morse, Niall J. Lennon, Kenneth J. Livak, Tarjei S. Mikkelsen, and John L. Rinn. 2014. "The Dynamics and Regulators of Cell Fate Decisions Are Revealed by Pseudotemporal

- Ordering of Single Cells.” *Nature Biotechnology* 32 (4): 381–86.
<https://doi.org/10.1038/nbt.2859>.
- Ungrin, Mark D., Chirag Joshi, Andra Nica, Céline Bauwens, and Peter W. Zandstra. 2008. “Reproducible, Ultra High-Throughput Formation of Multicellular Organization from Single Cell Suspension-Derived Human Embryonic Stem Cell Aggregates.” *PloS One* 3 (2): e1565. <https://doi.org/10.1371/journal.pone.0001565>.
- Warmflash, Aryeh, Benoit Sorre, Fred Etoc, Eric D. Siggia, and Ali H. Brivanlou. 2014. “A Method to Recapitulate Early Embryonic Spatial Patterning in Human Embryonic Stem Cells.” *Nature Methods* 11 (8): 847–54. <https://doi.org/10.1038/nmeth.3016>.
- Wilson, Jenna L., Mohamad Ali Najia, Rabbia Saeed, and Todd C. McDevitt. 2014. “Alginate Encapsulation Parameters Influence the Differentiation of Microencapsulated Embryonic Stem Cell Aggregates.” *Biotechnology and Bioengineering* 111 (3): 618–31. <https://doi.org/10.1002/bit.25121>.
- Yan, Liying, Mingyu Yang, Hongshan Guo, Lu Yang, Jun Wu, Rong Li, Ping Liu, et al. 2013. “Single-Cell RNA-Seq Profiling of Human Preimplantation Embryos and Embryonic Stem Cells.” *Nature Structural & Molecular Biology* 20 (9): 1131–39. <https://doi.org/10.1038/nsmb.2660>.
- Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. “ClusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters.” *OMICS: A Journal of Integrative Biology* 16 (5): 284–87. <https://doi.org/10.1089/omi.2011.0118>.
- Ziomek, C. A., and M. H. Johnson. 1980. “Cell Surface Interaction Induces Polarization of Mouse 8-Cell Blastomeres at Compaction.” *Cell* 21 (3): 935–42. [https://doi.org/10.1016/0092-8674\(80\)90457-2](https://doi.org/10.1016/0092-8674(80)90457-2).

Chapter 5: Axial Elongation of Caudalized Human PSC Organoids

Mimics Neural Tube Development

5.1 Introduction

During vertebrate embryogenesis post gastrulation, repeated symmetry breaking events shape the embryo as more and more complex structures containing a wide variety of cell types. In particular a key developmental symmetry breaking event beginning at gastrulation is the formation of the anterior-posterior axis and subsequent elongation of the embryo. This patterning event segments the main body axis and specifies the tissues that will eventually contribute to the spinal cord (Steventon et al. 2016; Wilson, Olivera-Martinez, and Storey 2009; Schiffmann 2006; Yamaguchi 2001). The spine is a crucial structure that both enables physical support as well as protection of essential neural projections that connect the body to the brain where dysregulation of the key processes leading to its formation results in several congenital abnormalities (Kaplan, Spivak, and Bendo 2005). Although this process has long been studied in model organisms such as chick and amphibians, it is complicated to study the multi-cellular interactions that drive this axial extension at high spatio-temporal resolution in mammalian embryos because it occurs post-implantation (Viebahn 1999; Beddington and Robertson 1999).

In recent years, protocols for human pluripotent stem cell organoids capable of recapitulating structures reminiscent of those in developing tissues have been largely developed (Bredenoord, Clevers, and Knoblich 2017; Warmflash et al. 2014; Lancaster and Knoblich 2014). However, to date, many of the discovered organoid systems are limited to radial or isotropic tissues, with the exception being 3D gastruloid cultures (Beccari et al. 2018) where all three germ lineages contribute to the signaling and self-organization in mouse embryonic stem cells to generate axial extension and now the directed protrusions discussed in Chapter 4. Here we report the development of a human induced pluripotent stem cell (hiPSC) organoid model of axial

extension in a neuronal directed differentiation that recapitulates many of the anisotropic morphologic features and gene expression profiles of *in vivo* axial elongation, enabling *in vitro* study of the multicellular interactions that pattern the human spinal cord.

5.2 Materials and Methods

5.2.1 Human Induced Pluripotent Stem Cell Line Generation and Culture

All work with human induced pluripotent stem cells (iPSCs) was approved by the University of California, San Francisco Human Gamete, Embryo, and Stem Cell Research (GESCR) Committee. Human iPSC lines were derived from the WTC11 line (Coriell Cat. #GM25256), the WTB line (Conklin Lab)(Miyaoaka et al. 2014), and the Allen Institute WTC11-LaminB cell line (AICS-0013 cl.210). All cell lines were karyotyped by Cell Line Genetics and reported to be karyotypically normal. Additionally, all cell lines tested negative for mycoplasma using a MycoAlert Mycoplasma Detection Kit (Lonza).

Human iPSCs were cultured on growth factor reduced Matrigel (Corning Life Sciences) and fed daily with mTeSRTM-1 medium (STEMCELL Technologies)(Ludwig et al. 2006). Cells were passaged by dissociation with Accutase (STEM CELL Technologies) and re-seeded in mTeSRTM-1 medium supplemented with the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10 μ M; Selleckchem)(Park et al. 2015) at a seeding density of 12,000 cell per cm².

The generation of the Chordin and Noggin CRISPRi lines first involved TALEN mediated insertion of the CRISPRi cassette pAAVS1-NDi-CRISPRi (Gen1) (Addgene) to the AAVS1 locus of the Allen Institute WTC11-LaminB (AICS-0013 cl.210) cell line. Following antibiotic selection of clones that received the CRISPRi cassette, CRISPRi gRNAs were generated targeting Noggin and Chordin (Table 5.1) using the Broad Institute GPP Web Portal and cloned into the gRNA-CKB (Addgene) following the previously described protocol(Mandegar et al. 2016). Guide RNA

vectors were nucleofected into the LaminB CRISPRi iPSC line using a P3 Primary Cell 96-well Nucleofector™ Kit (Lonza) and the 4D Nucleofector X Unit (Lonza) following manufacturer instructions. Nucleofected cells were allowed to recover in mTeSR™-1 medium supplemented with Y-276932 (10 µM) and then underwent antibiotic selection with blasticidin (ThermoFisher Scientific; 10 µg/ml) following the previously published protocol (Libby et al. 2018; Mandegar et al. 2016). Knockdown efficiency was evaluated by addition of doxycycline to the daily feeding media over the course of 5 days, collection of mRNA, and subsequent quantification of gene expression by qPCR.

Table 5.1: CRISPRi guides.

Gene Target (Symbol)	Guide sequence
Noggin (NOG)	CTCCTCTCCCGGGTCTACTG
Chordin (CHRD)	AAGGAGCCGCTGCCCGTTCCG

5.2.2 Organoid Differentiation

Organoid differentiations are a modified protocol of previously a published spinal cord interneuron differentiation protocol(Butts et al. 2017). Human iPSCs were seeded at 125000 cells/cm² in mTeSR™-1 medium supplemented with the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10 µM; Selleckchem) and small molecule GSK inhibitor CHIR99021 (2µM, 4µM, or 6µM; Selleckchem). Two days later, cells were singularized with Accutase (STEMCELL Technologies), counted using a Countess II FL (Life Technologies), and seeded into 800µm X 800µm PDMS microwell inserts in a 24 well plate (~270 wells/insert)(Hookway et al. 2015). After ~18 hours, condensed aggregates were transferred to rotary culture in mTeSR™-1 medium supplemented with Y-276932 (10 µM; Selleckchem), CHIR99021 (2µM, 4µM, or 6µM; Selleckchem), ALK5 small molecule inhibitor SB431542 (10µM, Selleckchem), and small molecule BMP inhibitor LDN193189 (0.2µM, Selleckchem). Organoids were fed every other day for up to 17 days. Y-276932 was removed from the media at day 3. At

day 5 organoids are transferred to Neural Induction Media (DMEM F:12 (Corning), N2 supplement (Life Technologies), L-Glutamine (VWR), 2 μ g/ml heparin (Sigma Aldrich), non-essential amino acids (Mediatech INC), penicillin-streptomycin (VWR), supplemented with fresh 0.4 μ g/ml ascorbic acid (Sigma Aldrich) and 10ng/ml brain derived neurotrophin factor (BDNF, R&D Systems)) supplemented with CHIR99021 (2 μ M, 4 μ M, or 6 μ M; Selleckchem), SB431542 (10 μ M, Selleckchem), and LDN193189 (0.2 μ M, Selleckchem). At day 7 organoids were resuspended in Neural Induction Media supplemented with retinoic acid (10nM, Sigma Aldrich), purmorphomine (300nM, EMD Millipore) and N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester (DAPT D5942, 1 μ M, Sigma-Aldrich).

5.2.3 Organoid Elongation Imaging and Quantification

Day 5 organoids from a single microwell insert were individually transferred using wide bore pipette tips into the center 60 wells of an uncoated ultra-low attachment 96-well plate (Corning), seeding exactly one aggregate per well, with the remaining organoids maintained in rotary culture through day 7. Using an inverted Axio Observer Z1 (Zeiss) microscope with incubation (Zeiss Heating Unit XL S, maintained at 37 °C, 5% CO₂), all 60 wells were imaged using an AxioCam MRm (Zeiss) digital CMOS camera at 5x magnification (NA 0.16, 2.6 μ m x 2.6 μ m per pixel). Each well was imaged in TL Brightfield every 20 minutes for 48 hours giving a total of 145 frames. At the end of imaging (day 7), 31 aggregates from the parallel rotary culture were imaged at 5x to generate a comparison image set.

To segment the organoids, all well images were first aligned by fitting a truncated quadratic curve to the average image intensity, then solving for the peak of maximum intensity, which was assumed to be the well center. Next, the average lighting inhomogeneity was calculated as the pixel-wise median of all 60 aligned well images, which was then subtracted from the individual aligned frames. After background correction, individual aggregates were isolated by finding

objects brighter than 0.83% of maximum intensity, but less than 3.0% of maximum intensity, with object size greater than 2,000 pixels, eccentricity greater than 0.1, and solidity greater than 40%. A bounding box 2 mm x 2 mm around the center of each of these objects was calculated and all frames of the time series cropped to this bounding box to reduce memory usage. To detect the region of maximum motion in the time series, the difference image between each pair of sequential images was calculated, and then the pixelwise standard deviation was calculated over all difference images in a given region. This standard deviation image was then thresholded at between 0.01 and 0.03 (AU) depending on the remaining lighting inhomogeneity in the image, producing a ring-shaped mask around the periphery of each aggregate. Finally, using the interior of the mask as the aggregate seed and the exterior as the background seed for the first frame, aggregates were segmented using anisotropic diffusion (Grady 2006), evolving the foreground and background seeds using the contour calculated from the previous frame for subsequent segmentations. Segmentation, labeling, and metrology were all performed using the python package `sckit-image`. (Walt et al. 2014)

Segmentations were manually inspected for accuracy, with 45 of 60 determined as having no or only minor flaws, with the remaining 15 excluded from automated analysis. Using the high quality segmentations only, each aggregate time series was then analyzed to examine geometry change over time. For each contour at each time point, we calculated contour area, contour perimeter, minimum, maximum and mean distance from contour center of mass to the perimeter. As additional non-dimensional measures of shape, we calculated the ratio of maximum to minimum radius and aggregate circularity. Aggregates were also manually classified as “extending”, “partially extending”, or “non-extending” by examining each video. Aggregates assigned to “extending” exhibited at least one, and at most two large protuberances that extended at least 100 μm from the main body. Partially extending aggregates exhibited at least one, and often many protuberances, all of which failed to extend robustly past the 100 μm demarcation.

Non-extending aggregates were any aggregates that failed to generate any extensions over the observation period.

5.2.4 Real Time Quantitative Polymerase Chain Reaction

Total RNA was isolated from human iPSCs using an RNAeasy Mini Kit (QIAGEN) according to manufacturer's instructions. Subsequently, cDNA was generated using an iScript cDNA Synthesis kit (BIORAD) and the reaction was run on a SimpliAmp thermal cycler (Life Technologies). Quantitative PCR using Fast SYBR Green Master Mix (ThermoFisher Scientific) was run on a StepOnePlus Real-Time PCR system (Applied Biosciences). Relative gene expression was determined by normalizing to the housekeeping gene 18S rRNA, using the comparative threshold (C_T) method. Gene expression was displayed as fold change of each sample (Noggin CRISPRi or Chordin CRISPRi) versus the time matched wildtype control with no knockdown. The primers were obtained from the Harvard Primer bank or designed using the NCBI Primer-BLAST website (Table 5.2).

Table 5.2: qPCR primers. Forward and reverse primers used in quantitative PCR experiments.

Gene Name (Symbol)	Forward Primer	Reverse Primer
RNA, 18S Ribosomal 5 (RNA18S5, 18S)	CTTCCACAGGAGGCCTACA	CTTCGGCCCACACCCTTAAT
Snail (SNAI1)	TAGCGAGTGGTTCTTCTGCG	GCCAGTCCAAGTCTTCTCGG
Brachyury (T)	TTTCCAGATGGTGAGAGCCG	CCGATGCCTCAACTCTCCAG
SRY-box 2 (SOX2)	TCAGGAGTTGTCAAGGCAGAGAAG	GCCGCCGCCGATGATTTGTTATTAT
NK1 Homeobox 2 (NKX1.2)	CCCTCCCACCACAAGATTTCT	GACCTCCGCCAAACTTTTCT
T-Box 6 (TBX6)	CATCCACGAGAATTGTACCCG	AGCAATCCAGTTTAGGGGTGT
RNA Binding Fox-1 Homolog 3 (RFX3, NEUN)	ACGATCGTAGAGGGACGGAA	AATTCAGGCCCGTAGACTGC
Paired Box 6 (PAX6)	GAGCGAGCGGTGCATTTG	TCAGATTCCTATGCTGATTGGTGAT
Motor Neuron And Pancreas Homeobox 1 (MNX1, HB9)	TCTCTTAACGGGAAGGGGCA	CTAATTCAGGGCGCTCTCGG
LIM Homeobox 2 (LHX2)	AAGTTCAGGCGCAACCTCTT	AAGACGGACGTCACAGTTGG
LIM Homeobox 5 (LHX5)	GTGCAAAGACGACTACCTGAG	CGGTCCGTACAGGATGACAC

Gene Name (Symbol)	Forward Primer	Reverse Primer
Engrailed Homeobox 1 (EN1)	CGCCCAGTTTTCGTTTTCGTT	GCAGAACAGACAGACCGACA
Visual System Homeobox 2 (VSX2, CHX10)	CGGCGACACAGGACAATCTT	CCTGTATCCTGTCTTCCGGC
Oligodendrocyte Transcription Factor 2 (OLIG2)	CGCATCCAGATTTTCGGGTC	AAAAGGTCATCGGGCTCTGG
Hes Family BHLH Transcription Factor 1 (HES1)	TCAACACGACACCGGATAAAC	GCCGCGAGCTATCTTTCTTCA
Chordin (CHRD) Noggin (NOG)	TATGCCTTGGACGAGACGTG GCTGCGGAGGAAGTTACAGA	ATGTTCTTGCAGCTGACCCT ACGAGCGCTTACTGAAGCAG

5.2.5 Histology, Immunocytochemistry and Imaging

Aggregates were fixed with 4% paraformaldehyde (VWR) for 40 minutes, washed three times with PBS. Aggregates to be used for histology were embedded in HistoGel Specimen Processing Gel (Thermo Fisher) prior to paraffin processing. Paraffin embedded samples were sectioned in 5µm sections, baked for 1 hour at 60°C, and subsequently stained for H&E. For immunofluorescent staining, epitope retrieval was performed by submersing slides in Citrate Buffer pH 6.0 (Vector Laboratories) in a 95°C water bath for 35min. Samples were permeabilized in 0.2% Triton X-100 (Sigma-Aldrich) for 5min, blocked in 1.5% normal donkey serum (Jackson Immunoresearch) for 1hour, and probed with primary and secondary antibodies against SOX2, PAX6, T, NES, TUBB3, and CDH2 (Table 5.2). Coverslips were mounted with anti-fade mounting medium (ProlongGold, Life Technologies) and samples were imaged on a Zeiss Axio Observer Z1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 camera.

Table 5.3: Antibodies. Antibodies and their dilutions used in experiments.

Gene Target (Symbol)	Species	Company (cat. #)	Dilution
SRY-box 2 (SOX2)	mouse	Abcam (ab79351)	1:200
Paired Box 6 (PAX6)	rabbit	ThermoFisher Scientific (42-6600)	1:200
Brachyury (T)	goat	ThermoFisher Scientific (PA5-46984)	1:400
Nestin (NES)	mouse	Santa Cruz (SC-23927)	1:400
Hoescht DNA stain	NA	ThermoFisher Scientific (62249)	1:10000
Tubulin Beta 3 Class III (TUBB3)	rabbit	Biolegend (802001)	1:600
N-cadherin (CDH2)	rabbit	Abcam (ab76057)	1:400

5.2.6 Whole Mount Lightsheet Imaging

4% paraformaldehyde-fixed paraffin-embedded samples (see “*Histology, Immunocytochemistry, and Imaging*”) were permeabilized with 0.3% Triton X-100 (Sigma-Aldrich) for 5min, blocked in 5% normal donkey serum (Jackson ImmunoResearch) for 1 hour, and probed with primary and secondary antibodies against PAX6 and T (TableS3) for 2 hours. Samples were then embedded in 1.5% low melt agarose (BioReagent) and drawn up into ~1mm imaging capillaries and subsequently imaged on the Zeiss Z.1 Light sheet Microscope equipped with a PCO.edge SCMOS camera.

5.2.7 Single Cell RNA Sequencing Sample and Library Preparation

Multiple organoid samples were combined and processed together using the MULTI-Seq technology (McGinnis et al. 2019). Organoids were singularized using Accutase (STEMCELL Technologies) and washed with cold PBS. Cells were resuspended in PBS with lipid-modified Anchor and Barcode oligonucleotides (gift from Zev Gartner) and incubated on ice for 5 min. A co-Anchor oligo was then added in order to stabilize membrane retention of the barcodes incubated for an additional 5 min on ice. Excess lipid-modified oligos were quenched with 1% BSA in PBS, washed with cold 1% BSA solution, and counted using a Countess II FL (Life

Technologies). Single cell GEMs and subsequent libraries were then prepared using the 10X Genomics Single Cell V2 protocol with an additional anchor specific primer during cDNA amplification to enrich barcode sequences.(McGinnis et al. 2019) Short barcode sequences (approx. 65-100bp determined by Bioanalyzer) were purified from cDNA libraries with two sequential SPRI bead cleanups. Barcode library preparation was performed according to the KAPA HiFi Hotstart (Kapa Biosystems) protocol to functionalize with the P5 sequencing adapter and library-specific RPIX barcode. Purified ~173bp barcode fragments were isolated with another SPRI bead cleanup and validation by Bioanalyzer.

The sample library was sequenced on an Illumina NovaSeq yielding an average of 41,112 reads per cell and 6,444 cells. The MULTI-Seq barcode library was sequenced on an Illumina NextSeq yielding an average of 9,882 reads per barcode and enabling sample assignment for 4,681 of 6,124 unique UMIs detected (76.4% recovery), using the demultiplexing code provided by the MULTI-Seq protocol (McGinnis et al. 2019).

5.2.8 Genome Annotation, RNA-seq read mapping, and Estimation of Gene and Isoform Expression

The sample library was aligned to the human GRCh38 reference genome using Cell Ranger v1.2.0 (10x Genomics). Gene expression levels were assessed using the Seurat v3.0.0 analysis pipeline (Butler et al. 2018). First cells were removed with fewer than 200 detected genes, fewer than 1,000 total detected transcripts, or which had greater than 10% mitochondrial gene expression. Next, expression levels were log normalized, and the top 2,000 variable genes calculated using the VST algorithm. The top 20 principal components were used to group cells into 6 clusters using a resolution of 0.2. Finally, top markers were detected for each cluster by detecting the top differentially expressed genes between one cluster and the remaining data set,

where at least 25% of cells in the cluster expressed the gene and the gene was expressed at least 0.25 log₂ fold-change different from the remaining population. Clusters and gene expression were visualized on a two dimensional UMAP projection of the first 20 principal components.

5.2.9 Cluster Analysis

To assign cluster identity, the top markers for each cluster were tested for GO term enrichment using the biological process “enrichGO” function in the R package “clusterProfiler” v3.12. (Yu et al. 2012) In addition, differentiation maturity in each cluster was assessed by examining expression level of panels of early neuroectoderm markers, proliferation markers, markers of neuron fate commitment, and markers of cell types present in neural tube formation and axial extension. Finally, to assess rostral caudal position of each, panels of HOX genes, as well as marker genes of Wnt and FGF signaling were examined to assign rough position of each cluster along the head-tail axis (Martin 2016; Carpenter 2002).

5.2.10 RNAScope

In situ hybridization for HOXB1, HOXC6, HOXB9 (probe information in Table 3.3) was performed on sections of 4% paraformaldehyde-fixed paraffin-embedded samples (see “*Histology, Immunocytochemistry, and Imaging*”) using the RNAScope Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics) and following the protocol outlined in User Manual 323100-USM. Sections were imaged on a Zeiss Axio Observer Z1 inverted microscope equipped with a Hamamatsu ORCA-Flash 4.0 camera.

Table 5.4: RNAScope Probes.

Gene Target (Symbol)	channel
HOXB1	C2
HOXC6	C1
HOXB9	C3

5.2.12 Statistical Analysis

Each experiment was performed with at least three biological replicates. Multiple comparisons were used to compare multiple groups followed by unpaired T-tests between two groups subject to a post-hoc Bonferroni correction. In gene expression analysis, three replicates were used for each condition, and all gene expression was normalized to control wildtype populations. Significance was specified as $P\text{-values} < 0.05$. P-values are specified in figure legends.

5.3 Results

5.3.1 Emergence of Axial Extensions from Neuronal Organoids

The benefit studying axial elongation in an organoid model is that it recapitulates developing tissues by generating organized structures in 3D. We observed that when a previously described neuronal differentiation protocol (Butts et al. 2017) was translated to a 3D orbital shaker tissue culture platform (Carpenedo, Sargent, and McDevitt 2007; Hookway et al. 2015), anisotropic extensions emerged from the main aggregate body after 5 days in culture (Figure 5.1A-C). These extensions persisted across aggregates of multiple sizes (500, 3000, 10000 cells at seeding). However, variation in the degree of extension was observed, with robustly elongating organoids in some wells and nearly spherical organoids in others. To determine if culture density was affecting the efficiency of extension, 3000-cell organoids were cultured at densities of 500, 1000, or 2000 organoids per well in suspension (Figure 5.2E). As organoid density increased, extensions became less pronounced, less frequent, and thinner; at the highest densities, extensions disappeared by day 15. At low culture densities, aggregate extension was robust. Organoids remained elongated through 15 days (Figure 5.1C), suggesting that density-mediated signaling, such as paracrine effects or nutrient deprivation, play a role in inducing and maintaining

extensions. In extending organoids, histology revealed internal elongated epithelial compartments with multiple sections devoid of cells (Figure 5.1E,F), To ensure that elongation did not result from the fusion of multiple aggregates, we transferred individual organoids at day 5 into the wells of low-adherence 96-well plates and imaged them over the course of 2 days. Extending organoids doubled the ratio of their longest axis relative to their shortest axis (Figure 5.1E, Figure 5.2A-C), demonstrating that elongation occurs through an aggregate-autonomous symmetry-breaking process reminiscent of *in vivo* axial elongation, and not through aggregate fusion. By day 6, organoids displayed markers of neural differentiation such as SOX2, PAX6, and Nestin (Figure 5.1G,H). The extending organoids also contained brachyury- (T) positive cells (Figure 5.1G,H) that often co-localized with SOX2 (Figure 5.1G, white arrows). The presence of SOX2+T+ cells suggests the emergence of neuro-mesodermal progenitors (NMPs)(Gouti et al. 2014) a population which, *in vivo*, contributes to the closing neural tube and to paraxial mesoderm (Henrique et al. 2015). Furthermore, by day 3 of culture, two patterns of T expression emerged: radial in spherical organoids versus concentrated in a pocket of cells in extending organoids (Figure S1D). Gene expression analysis by qPCR showed that both types of organoids expressed high levels of SOX2, T, migratory genes such as SNA1, and the NMP transcription factor NKX1.2 at day 10 of differentiation (Figure 5.3A). By day 17, expression of these genes had persisted or increased in elongated organoids, whereas it had disappeared in spherical organoids (Figure 5.3A), suggesting that the lack of extension was due to the absence of an NMP population.

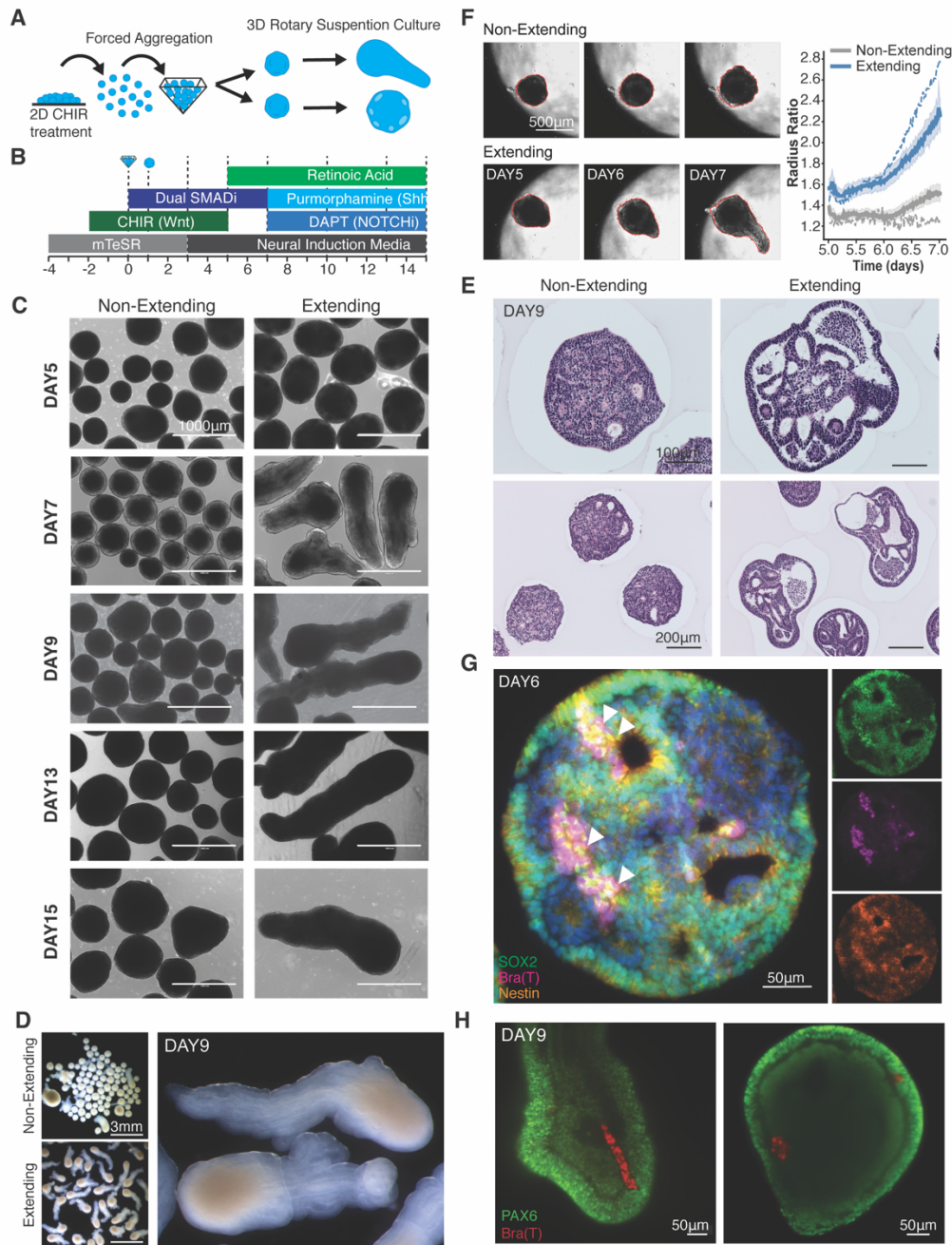


Figure 5.1: Emergence of axial extensions from neuronal organoids. (A-B) Schematic of 3D culture experimental set up and differentiation protocol timeline. (C) Brightfield images of extending and non-extending aggregates over the course of 15 days. (D) Stereoscope images of extending and non-extending aggregates at day 9 of differentiation. (E) (LEFT) Frames from video time-course tracking aggregate extensions in a static 96 well plate culture. (RIGHT) The graph depicts radius ratio where the solid lines represent the mean radius ratio of extending and non-extending aggregates with 95 percent confidence interval depicted by the surrounding light color ($n = 24$ non-elongating, $n = 36$ elongating). The dotted lines are the two pictured aggregates on the left. (F) Immunofluorescence imaged section of extending aggregate at day 6 of differentiation. White arrows indicate SOX2(+)T(+) cells. (G) Immunofluorescence optical sections of extending aggregate at day 9 of differentiations displaying both PAX6 staining and persistence of T(+) cells.

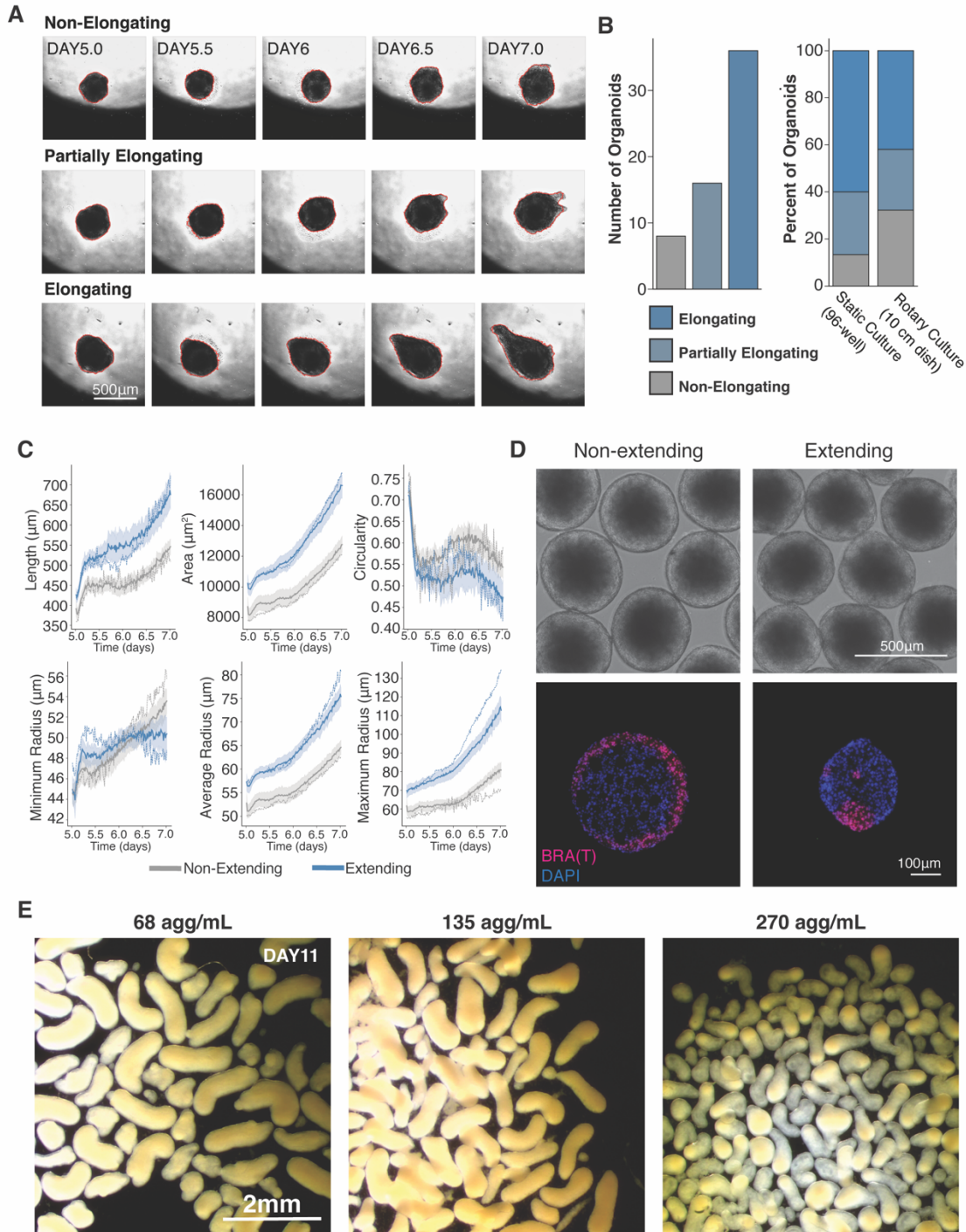


Figure 5.2: Aggregate extension phenotypes. (A) Frames from supplementary movie depicting time-course of extension from days 5-7 of differentiation. **(B)** Quantification of organoid extensions by number of extension and percent of the entire sample ($n = 8$ non-extending, $n = 16$ partial extending, $n = 36$ extending organoids). **(C)** Quantification of length, area, circularity, and radius of extending and non-extending aggregates. Dark solid line represents the mean value with 95 percent confidence interval depicted by the surrounding light color ($n = 24$ non-elongating, $n = 36$ elongating). **(D)** Immunofluorescence staining of brachyury (T) on day 3 of the differentiation in extending and non-extending populations as compared to corresponding brightfield images. **(E)** Stereoscope images of extending aggregates at day 11 of differentiation when cultured at different densities in rotary culture.

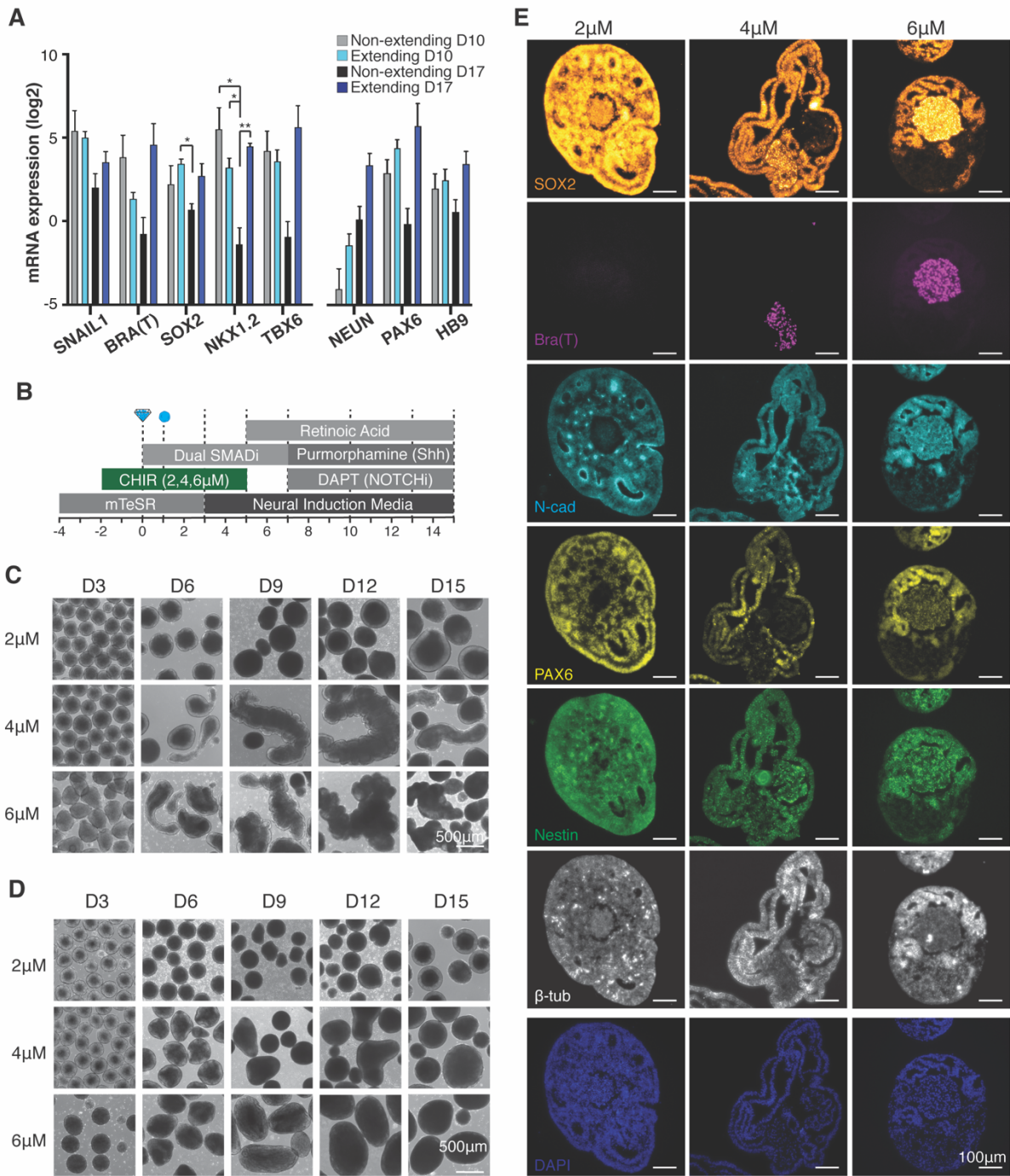


Figure 5.3: Wnt mediated increase in extensions and SOX2(+)T(+) progenitor populations. (A) QPCR of mRNA expression from extending and non-extending at day 10 and day 17 of differentiation (n=3). Error bars mark standard differentiations of 3 biological replicates. Significance marked by * ($p < 0.001$) and ** ($p < 0.00001$). **(B)** Schematic experimental set up and differentiation protocol timeline with increased CHIR dosing. **(C)** Brightfield images of differentiation time-course conducted in the WTC hiPS cell line. **(D)** Brightfield images of differentiation time-course conducted in the WTB hiPS cell line. **(E)** Immunofluorescence of paraffin sectioned aggregates at day 10 of differentiation at 2 μM, 4 μM, and 6 μM CHIR concentration.

5.3.2 Wnt Mediated Increase in Extensions and SOX2(+)T(+) Progenitor Populations

The caudalization of the spinal cord (Yamaguchi 2001) and the creation of NMPs (Wilson, Olivera-Martinez, and Storey 2009; Gouti et al. 2014) depend on Wnt signaling. We therefore examined whether increasing Wnt signaling would increase organoid elongation. Providing the small-molecule Wnt agonist CHIR99021 (CHIR) in increasing doses (2 μ M, 4 μ M, and 6 μ M; Figure 5.3B) increased the proportion of extending organoids. Organoids derived from an hiPSC cell line (WTB) which had previously failed to elongate at lower CHIR doses showed a similar dose-dependent increase in elongation (Figure 5.3C,D), though not to the same extent as WTC-derived organoids. Hence, inherent differences between hiPSC lines can influence their capacity for axial extension. At day 10 of WTC differentiation, a number of extending organoids from the 4 μ M- and 6 μ M-CHIR cultures harbored cells co-expressing SOX2 and T, usually at their centers (Figure 5.3E). These organoids expressed SOX2 and PAX6 at lower levels than organoids exposed to 2 μ M CHIR, despite maintaining comparable expression of the PAX6 target Nestin, indicating a loss of progenitor states in the 4 μ M- and 6 μ M-CHIR aggregates. Additionally, 4 μ M and 6 μ M-CHIR organoids displayed high N-cadherin and β 3-Tubulin expression, indicating the emergence of more mature populations of neurons (Figure 5.3E).

5.3.3 Regionalized HOX Gene Expression in Extending Aggregates

To better understand the cell diversity within extending aggregates and the differences between aggregates that do and do not extend, we performed single-cell transcriptomics at day 10 of differentiation. We found six clusters containing equivalent amounts of cells from extending and non-extending aggregates, with the exception of cluster 2, which contained a majority of cells from non-extending aggregates (Figure 5.4A, Figure 5.5A). All clusters displayed neural markers such as SOX2, N-cadherin, and Nestin (Figure 5.4C, Figure 5.5B,C). Clusters 0 and 1 were high in genes associated with early neural progenitors such as SOX2, CDX2, and PAX6, as well as

FGF8. They also contained a small number of cells expressing NKX1.2, a marker of NMPs (Figure 5.4C), which was not detected in clusters 2-5. However, MIXL1 and MEOX1 were both highly expressed in clusters 2-4, suggesting that these clusters contain mesodermal cells (Figure 5.4B, Figure 5.5B,C). The presence of markers associated with the caudalization of the extending neural tube, such as FGF8 and CYP1B1 (Figure 5.5C), prompted us to examine the expression profile of HOX genes (Figure 5.4D). Both extending and non-extending aggregates expressed genes associated with the hindbrain, and with the brachial and thoracic regions of the neural tube, and only a minor increase in HOX gene expression was seen in extending relative to non-extending aggregates (average 0.4 log₂ fold change). To assess positional information, we monitored HOX gene transcripts marking the hindbrain (HOXB1), the brachial (HOXC6), and the thoracic (HOXB9) regions of the neural tube via RNAscope in day 7 and day 10 aggregates. While non-extending aggregates expressed both hindbrain and brachial HOX genes, the expression was distributed radially throughout the spherical aggregates. In contrast, the extending aggregates displayed regionalized expression of hindbrain, brachial and thoracic HOX genes along the extension axis, as seen in the neural tube of developing vertebrates. Specifically, we observed a high concentration of cells expressing HOXB1 (hindbrain) in the central mass of the aggregate and HOXC6 (brachial) and HOXB9 (thoracic) in the extensions (Figure 5.4F). Interestingly, HOXC6 and HOXB9 often overlapped within cells in the extending aggregates at day 7 (Figure 5.4F), suggesting that while these regions extend away from the main hindbrain aggregate mass, the cells have yet to segregate as specific regions of the spinal cord. At day 10 the HOX gene expression was reduced in both non-extending and extending aggregates, but in extending aggregates, HOXB9 continued to localize to the extending region. Although nearly all non-extending aggregates displayed radial HOX gene expression, regionalized HOXB9 expression was detected in one spherical aggregate (Figure 5.4G), indicating that aggregate extension is not a requirement for sub-populations to organize into specific HOX domains.

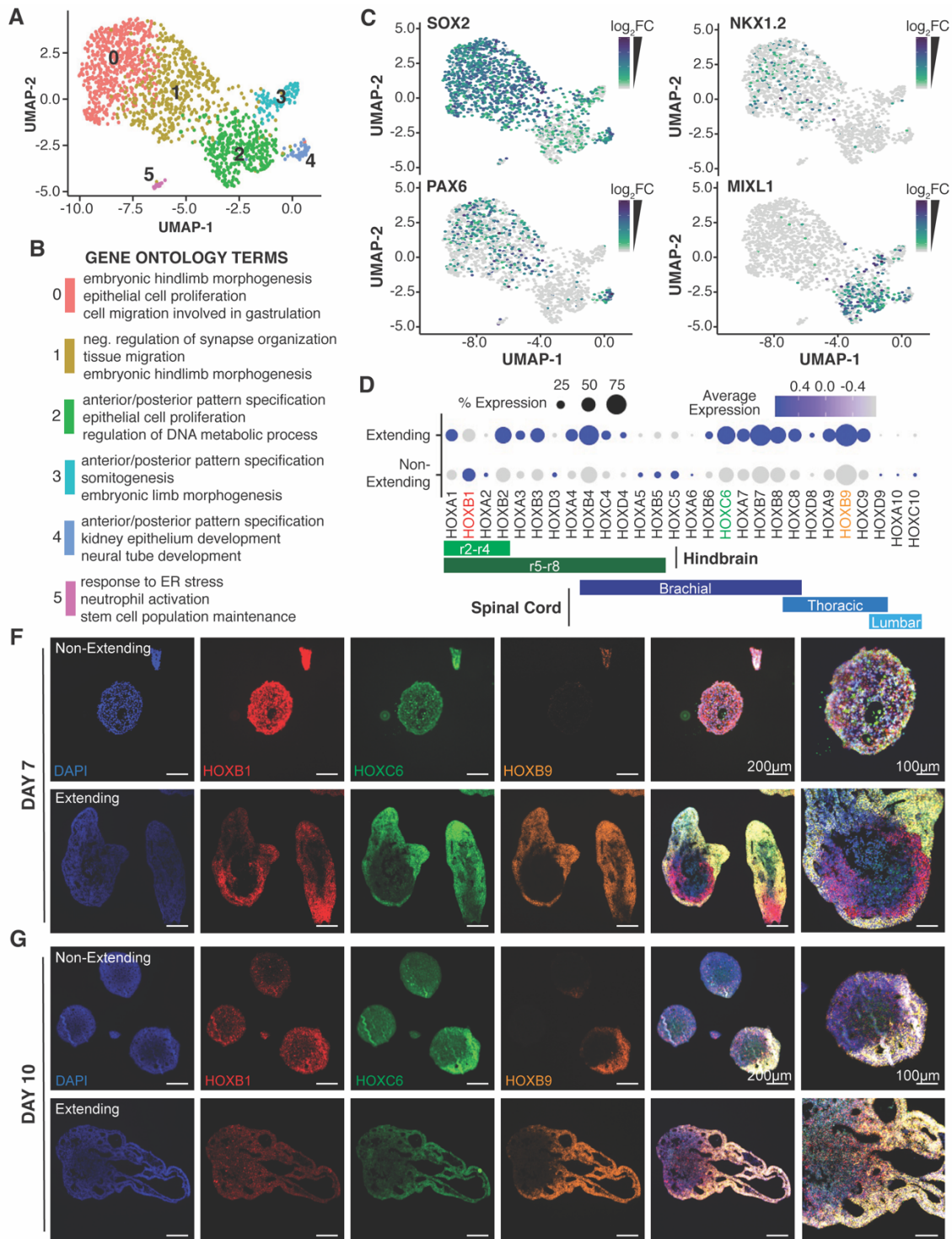


Figure 5.4: Progenitor and HOX gene expression in extending and non-extending aggregates. (A) UMAP of single cell RNA sequencing of extending and non-extending aggregates at day 10 of differentiation. **(B)** Gene Ontology terms assigned to RNA sequencing clusters. **(C)** UMAPs for SOX2, PAX6, NKX1.2, MIXL1 expression in extending and non-extending populations. Heatmaps depict a normalized increase in log₂ foldchange from min expression to max expression of the respective gene. **(D)** Detected HOX gene expression across elongating and non-elongating aggregates. Schematic depicting HOX gene regions of the spinal cord. **(F-G)** RNAscope based immunofluorescence imaging of HOX genes in extending and non-extending aggregates at days 7 and 10 of differentiation.

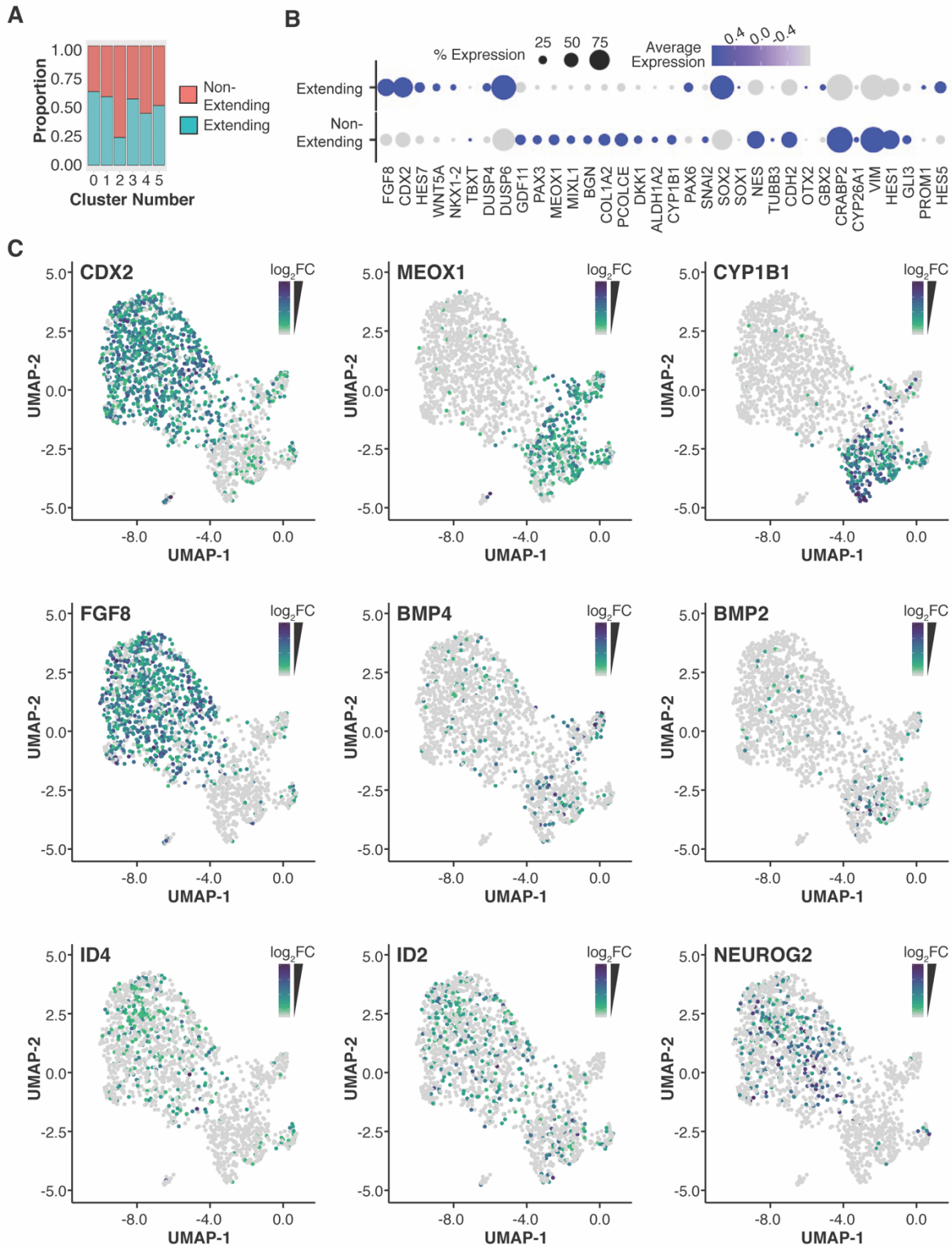


Figure 5.5: Gene expression analysis of extending and non-extending aggregates by single cell RNA sequencing. A) Proportion of extending and non-extending aggregates in each cluster of the generated UMAP. B) Dot plot of genes relevant to axial extension. C) UMAP plots of gene expression for a subset of genes involved in axial patterning and extension.

5.3.4 Manipulation of Endogenous BMP Signaling by CRISPR Interference

Because the extensions mimicked aspects of the developing neural tube, we asked how they would respond to disruptions in signaling pathways known to control axial extension and patterning. The BMP pathway is critical to establishing and maintaining both the anterior-posterior Wnt and FGF gradients that drive axial elongation and the BMP-SHH dorsal-ventral gradients that organize the developing spinal cord (Yamaguchi 2001; McMahon et al. 1998; Bel-Vialar, Itasaki, and Krumlauf 2002; Corral and Storey 2004). We assessed the role of BMP signaling in our organoid system by knocking down its regulators Noggin and Chordin. We introduced RNA guides targeting Noggin and Chordin into a Lamin-B labeled hiPSC line (Allen Institute) harboring a doxycycline (DOX) inducible CRISPR interference system (CRISPRi) (Libby et al. 2018; Mandegar et al. 2016) (Figure 5.6A). Noggin and Chordin gene expression decreased 10 fold 5 days after the addition of DOX (Figure 5.6B). In mouse development, Noggin is expressed along the developing neural tube and inhibits dorsally expressed BMP. Noggin deletion leads to increased BMP signaling at the posterior axis, failure of neural tube closure, and elongation of the developing tail (McMahon et al. 1998). Chordin is a BMP shuttling molecule that concentrates BMP at the ventral side of the neural tube where mouse knock-out models exhibit a shortened body axis and underdeveloped anterior spines (Bachiller et al. 2003). Similarly, knocking down Noggin led to organoids with long extensions that surpassed the length of the wildtype aggregate extensions, while knocking down Chordin led to aggregates with minimal to no extension (Figure 5.6C). Both the Noggin and Chordin knockdowns increased expression of dorsal neural progenitor fates (Figure 5.6E), suggesting that reduction of BMP inhibitors led to increased dorsalization of neural progenitors within the organoid, likely through increased endogenous exposure to BMP signaling (Figure 5.6D). Overall, modulating BMP signaling in the extending aggregates induced phenotypic and cell fate specification events analogous to those occurring *in vivo* during neural tube development.

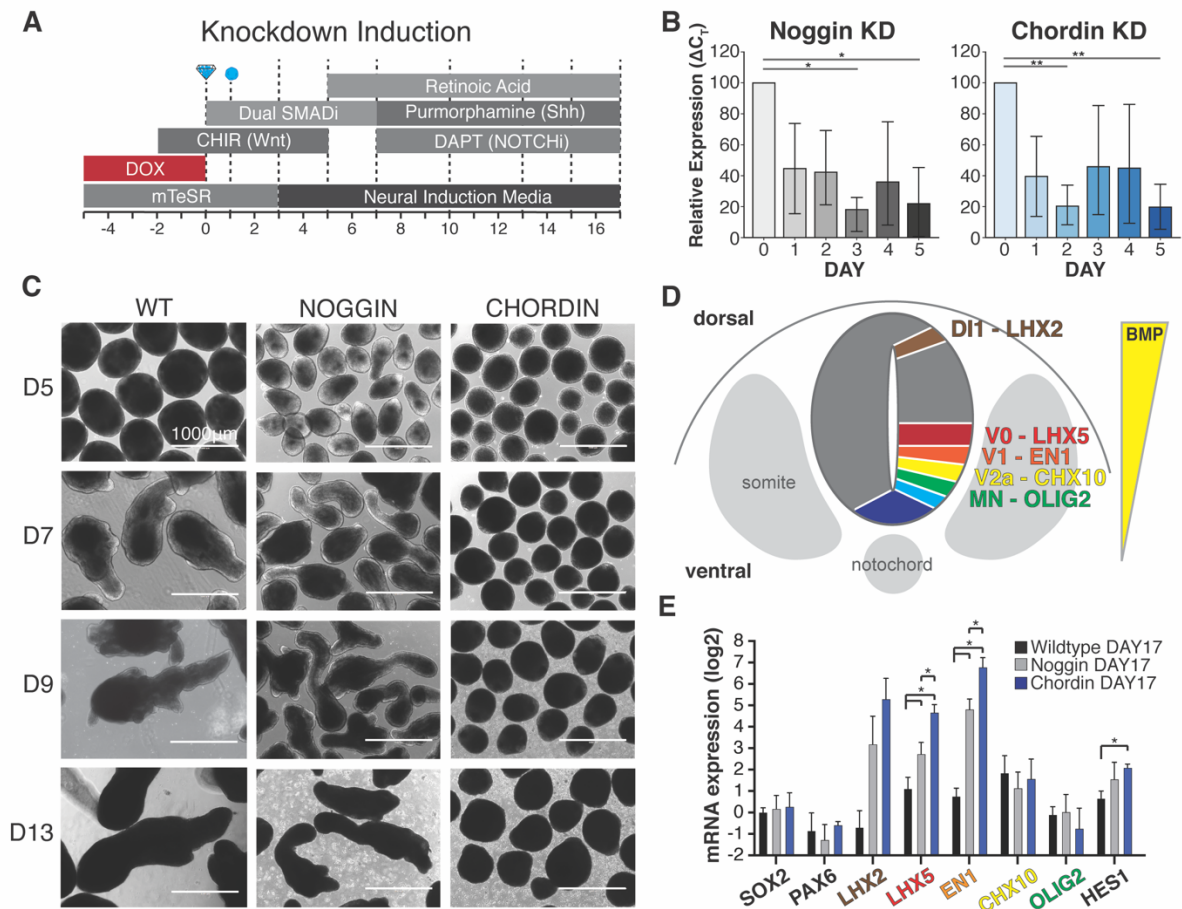


Figure 5.6: Manipulation of endogenous BMP signaling by CRISPR interference. (A) Schematic experimental set up and differentiation protocol timeline with knockdown induction. **(B)** Quantification of Noggin and Chordin knockdown efficiency by reduction in mRNA expression levels measured by qPCR. **(C)** Brightfield images of differentiation time course of knockdown iPSC lines. **(D)** Schematic depicting BMP signaling and progenitor specification within the developing neural tube. **(E)** Quantification of mRNA expression of progenitor specific transcription factors by qPCR at day 17 of differentiation (n=3). Error bars depict standard deviations of 3 biological replicates. Significance depicted by * (P<0.05).

5.4 Discussion

In this study, we demonstrate a novel organoid model of axial extension and neural tube-like formation. By combining 3D culture with early Wnt modulation of a neuronal differentiation protocol allows for the multi-population emergence of caudal spinal cord phenotypes, including NMPs, which lead to axial extension of aggregates. Furthermore, these extending aggregates maintain the ability to respond to perturbations in the BMP signaling pathway in manners

consistent with *in vivo* vertebrate biology. By recapitulating many of the morphologic and cell fate decision events associated with axial elongation, the neural extension organoids provide a nuanced multicellular tissue-like model of the neural tube that enables the study of the underlying gene regulatory networks in conjunction with morphological events and structural outcomes that enable the development of the human nervous system.

5.5 Conclusion

In conclusion, this fourth study provides an example of stem cell engineering paired with concepts in developmental biology to generate a unique organoid platform that displays emergent lineages and morphologies inherent to the developing neural tube. These results then demonstrate that this emergent behavior is dependent on early Wnt signaling to trigger a lineage fate decision that generates multipotent neuromesodermal progenitors, enabling future studies interrogating mechanisms behind progenitor patterning and differentiation that influence morphogenic phenotypes such as axial elongation.

5.6 Bibliography

- Bachiller, Daniel, John Klingensmith, Natalya Shneyder, Uyen Tran, Ryan Anderson, Janet Rossant, and E. M. De Robertis. 2003. "The Role of Chordin/Bmp Signals in Mammalian Pharyngeal Development and DiGeorge Syndrome." *Development* 130 (15): 3567–78. <https://doi.org/10.1242/dev.00581>.
- Beccari, Leonardo, Naomi Moris, Mehmet Girgin, David A. Turner, Peter Baillie-Johnson, Anne-Catherine Cossy, Matthias P. Lutolf, Denis Duboule, and Alfonso Martinez Arias. 2018. "Multi-Axial Self-Organization Properties of Mouse Embryonic Stem Cells into Gastruloids." *Nature*, October, 1. <https://doi.org/10.1038/s41586-018-0578-0>.
- Beddington, Rosa S. P., and Elizabeth J. Robertson. 1999. "Axis Development and Early Asymmetry in Mammals." *Cell* 96 (2): 195–209. [https://doi.org/10.1016/S0092-8674\(00\)80560-7](https://doi.org/10.1016/S0092-8674(00)80560-7).
- Bel-Vialar, Sophie, Nobue Itasaki, and Robb Krumlauf. 2002. "Initiating Hox Gene Expression: In the Early Chick Neural Tube Differential Sensitivity to FGF and RA Signaling Subdivides the HoxB Genes in Two Distinct Groups." *Development* 129 (22): 5103–15.
- Bredenoord, Annelien L., Hans Clevers, and Juergen A. Knoblich. 2017. "Human Tissues in a Dish: The Research and Ethical Implications of Organoid Technology." *Science* 355 (6322): eaaf9414. <https://doi.org/10.1126/science.aaf9414>.
- Butler, Andrew, Paul Hoffman, Peter Smibert, Efthymia Papalexi, and Rahul Satija. 2018. "Integrating Single-Cell Transcriptomic Data across Different Conditions, Technologies, and Species." *Nature Biotechnology* 36 (5): 411–20. <https://doi.org/10.1038/nbt.4096>.
- Butts, Jessica C., Dylan A. McCreedy, Jorge Alexis Martinez-Vargas, Frederico N. Mendoza-Camacho, Tracy A. Hookway, Casey A. Gifford, Praveen Taneja, Linda Noble-Haeusslein, and Todd C. McDevitt. 2017. "Differentiation of V2a Interneurons from Human Pluripotent

- Stem Cells." *Proceedings of the National Academy of Sciences* 114 (19): 4969–74. <https://doi.org/10.1073/pnas.1608254114>.
- Carpenido, Richard L., Carolyn Y. Sargent, and Todd C. McDevitt. 2007. "Rotary Suspension Culture Enhances the Efficiency, Yield, and Homogeneity of Embryoid Body Differentiation." *STEM CELLS* 25 (9): 2224–34. <https://doi.org/10.1634/stemcells.2006-0523>.
- Carpenter, Ellen M. 2002. "Hox Genes and Spinal Cord Development." *Developmental Neuroscience* 24 (1): 24–34. <https://doi.org/10.1159/000064943>.
- Corral, Ruth Diez del, and Kate G. Storey. 2004. "Opposing FGF and Retinoid Pathways: A Signalling Switch That Controls Differentiation and Patterning Onset in the Extending Vertebrate Body Axis." *BioEssays* 26 (8): 857–69. <https://doi.org/10.1002/bies.20080>.
- Gouti, Mina, Anestis Tsakiridis, Filip J. Wymeersch, Yali Huang, Jens Kleinjung, Valerie Wilson, and James Briscoe. 2014. "In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity." *PLOS Biology* 12 (8): e1001937. <https://doi.org/10.1371/journal.pbio.1001937>.
- Grady, L. 2006. "Random Walks for Image Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November. <https://doi.org/10.1109/TPAMI.2006.233>.
- Henrique, Domingos, Elsa Abranches, Laure Verrier, and Kate G. Storey. 2015. "Neuromesodermal Progenitors and the Making of the Spinal Cord." *Development* 142 (17): 2864–75. <https://doi.org/10.1242/dev.119768>.
- Hookway, Tracy A., Jessica C. Butts, Emily Lee, Hengli Tang, and Todd C. McDevitt. 2015. "Aggregate Formation and Suspension Culture of Human Pluripotent Stem Cells and Differentiated Progeny." *Methods*. <https://doi.org/10.1016/j.ymeth.2015.11.027>.

- Kaplan, Kevin M., Jeffrey M. Spivak, and John A. Bendo. 2005. "Embryology of the Spine and Associated Congenital Abnormalities." *The Spine Journal* 5 (5): 564–76. <https://doi.org/10.1016/j.spinee.2004.10.044>.
- Lancaster, Madeline A., and Juergen A. Knoblich. 2014. "Organogenesis in a Dish: Modeling Development and Disease Using Organoid Technologies." *Science* 345 (6194): 1247125. <https://doi.org/10.1126/science.1247125>.
- Libby, Ashley RG, David A Joy, Po-Lin So, Mohammad A Mandegar, Jonathon M Muncie, Federico N Mendoza-Camacho, Valerie M Weaver, Bruce R Conklin, and Todd C McDevitt. 2018. "Spatiotemporal Mosaic Self-Patterning of Pluripotent Stem Cells Using CRISPR Interference." *ELife* 7 (October): e36045. <https://doi.org/10.7554/eLife.36045>.
- Ludwig, Tenneille E., Veit Bergendahl, Mark E. Levenstein, Junying Yu, Mitchell D. Probasco, and James A. Thomson. 2006. "Feeder-Independent Culture of Human Embryonic Stem Cells." *Nature Methods* 3 (8): 637–46. <https://doi.org/10.1038/nmeth902>.
- Mandegar, Mohammad A., Nathaniel Huebsch, Ekaterina B. Frolov, Edward Shin, Annie Truong, Michael P. Olvera, Amanda H. Chan, et al. 2016. "CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs." *Cell Stem Cell* 18 (4): 541–53. <https://doi.org/10.1016/j.stem.2016.01.022>.
- Martin, Benjamin L. 2016. "Factors That Coordinate Mesoderm Specification from Neuromesodermal Progenitors with Segmentation during Vertebrate Axial Extension." *Seminars in Cell & Developmental Biology*, Bone development and disease, 49 (January): 59–67. <https://doi.org/10.1016/j.semcdb.2015.11.014>.
- McGinnis, Christopher S., David M. Patterson, Juliane Winkler, Daniel N. Conrad, Marco Y. Hein, Vasudha Srivastava, Jennifer L. Hu, et al. 2019. "MULTI-Seq: Sample Multiplexing for Single-Cell RNA Sequencing Using Lipid-Tagged Indices." *Nature Methods* 16 (7): 619–26. <https://doi.org/10.1038/s41592-019-0433-8>.

- McMahon, Jill A., Shinji Takada, Lyle B. Zimmerman, Chen-Ming Fan, Richard M. Harland, and Andrew P. McMahon. 1998. "Noggin-Mediated Antagonism of BMP Signaling Is Required for Growth and Patterning of the Neural Tube and Somite." *Genes & Development* 12 (10): 1438–52. <https://doi.org/10.1101/gad.12.10.1438>.
- Miyaoka, Yuichiro, Amanda H. Chan, Luke M. Judge, Jennie Yoo, Miller Huang, Trieu D. Nguyen, Paweena P. Lizarraga, Po-Lin So, and Bruce R. Conklin. 2014. "Isolation of Single-Base Genome-Edited Human IPS Cells without Antibiotic Selection." *Nature Methods* 11 (3): 291–93. <https://doi.org/10.1038/nmeth.2840>.
- Park, Sangkyu, Daehwan Kim, Yeon-Gil Jung, and Sangho Roh. 2015. "Thiazovivin, a Rho Kinase Inhibitor, Improves Stemness Maintenance of Embryo-Derived Stem-like Cells under Chemically Defined Culture Conditions in Cattle." *Animal Reproduction Science* 161 (October): 47–57. <https://doi.org/10.1016/j.anireprosci.2015.08.003>.
- Schiffmann, Yoram. 2006. "Symmetry Breaking and Convergent Extension in Early Chordate Development." *Progress in Biophysics and Molecular Biology* 92 (2): 209–31. <https://doi.org/10.1016/j.pbiomolbio.2005.10.002>.
- Steventon, Ben, Fernando Duarte, Ronan Lagadec, Sylvie Mazan, Jean-François Nicolas, and Estelle Hirsinger. 2016. "Species-Specific Contribution of Volumetric Growth and Tissue Convergence to Posterior Body Elongation in Vertebrates." *Development* 143 (10): 1732–41. <https://doi.org/10.1242/dev.126375>.
- Viebahn, Christoph. 1999. "3 The Anterior Margin of the Mammalian Gastrula: Comparative and Phylogenetic Aspects of Its Role in Axis Formation and Head Induction." In *Current Topics in Developmental Biology*, edited by Roger A. Pedersen and Gerald P. Schatten, 46:63–103. Academic Press. [https://doi.org/10.1016/S0070-2153\(08\)60326-9](https://doi.org/10.1016/S0070-2153(08)60326-9).
- Walt, Stéfan van der, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. 2014. "Scikit-Image: Image Processing in Python." *PeerJ* 2 (June): e453. <https://doi.org/10.7717/peerj.453>.

- Warmflash, Aryeh, Benoit Sorre, Fred Etoc, Eric D. Siggia, and Ali H. Brivanlou. 2014. "A Method to Recapitulate Early Embryonic Spatial Patterning in Human Embryonic Stem Cells." *Nature Methods* 11 (8): 847–54. <https://doi.org/10.1038/nmeth.3016>.
- Wilson, Valerie, Isabel Olivera-Martinez, and Kate G. Storey. 2009. "Stem Cells, Signals and Vertebrate Body Axis Extension." *Development* 136 (10): 1591–1604. <https://doi.org/10.1242/dev.021246>.
- Yamaguchi, Terry P. 2001. "Heads or Tails: Wnts and Anterior–Posterior Patterning." *Current Biology* 11 (17): R713–24. [https://doi.org/10.1016/S0960-9822\(01\)00417-1](https://doi.org/10.1016/S0960-9822(01)00417-1).
- Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. "ClusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters." *OMICS: A Journal of Integrative Biology* 16 (5): 284–87. <https://doi.org/10.1089/omi.2011.0118>.

Chapter 6: Future Considerations

The studies presented in this dissertation develop methods to interrogate the emergence of heterogeneous populations and then use these methods to investigate the roles of cell-cell adhesion on lineage fate decisions within human induced pluripotent stem cells (iPSCs). In Chapter 2, a mosaic knockdown system was described in human iPSCs. Genes associated with cellular mechanics were knocked down in subpopulations of human iPSC colonies where mosaic knockdown triggered colonies rearrangements into gene knockdown specific patterns without exit of pluripotency. Furthermore, when allowed to differentiate, mixed populations displayed preferential differentiation capacity that was only present in mosaic knockdowns, indicating that a symmetry breaking event itself influenced subsequent lineage fate decisions. In Chapter 3, a computational model of the mosaic knockdown system in human iPSCs was developed. This extended Cellular Potts Model used empirical data from *in vitro* experiments to derive model parameters that recapitulate *in vitro* behavior. Then, using a pattern classification algorithm (TSSL) and a parameter optimization particle swarm algorithm, experimental outcomes for desired *de novo* patterns were tested *in silico*. Once the optimized experimental design was discovered, it was reciprocally tested in *in vitro* culture, successfully generating the desired pattern at the predicted rates. This study demonstrated a unique example of paired computational modeling and machine learning prediction that can be applied to a tractable *in vitro* system, where both the model and the biological phenomenon can be directly related and measured. In Chapter 4, the mosaic knockdown system was translated to a 3D culture platform where 50 cell aggregate culture was facilitated by hydrogel encapsulation. Over 7 days of culture, the 50 cell human iPSC aggregates displayed the emergence of all three germ lineages and morphologies specific to each lineage, effectively enabling a gastruloid model. Interestingly, the transcriptional expression of the gastruloids by single cell RNA sequencing clustered with preimplantation human embryos,

highlighting the biological relevance of the gastruloid model. Furthermore, the specific micro-environment of encapsulation facilitated the emergence of extraembryonic lineages outside of the classical first three germ layers. In Chapter 5, later stages in development were modeled through the generation of neural tube organoids. These neural tube organoids displayed axial anisotropic extension in response to Wnt signaling, a pool of neuromesodermal progenitors, and regionalized HOX gene expression. Furthermore, the organoids responded to genetic knockdown of BMP pathway inhibitors in a similar manner to genetic knockout mice, both in morphology of extensions and progenitor lineage fate decisions. Together, these four studies capitalized on principles in genome engineering, tissue engineering, and developmental biology to interrogate mechanisms that drive the emergence of heterogeneous populations and how symmetry breaking influences morphogenesis in multiple types of culture and at multiple stages of human development. While these studies were able to establish engineered control over symmetry breaking events and multiple organoid systems, they also bring up additional questions that can be further explored in future studies.

6.1 CRISPRi Colony Pattern Formation Screens

The classical study of early development has been focused largely on embryology and genetic knockouts that cause congenital mutations, which can be limiting due to the time constraints, technical barriers, and resources needed to generate genetically manipulated organisms particularly in vertebrate species (Adams et al. 2013). However, the advent of pluripotent stem cell culture and complex organoid systems has enabled the study of many developmental processes in human specific models that offer a multitude of tangible genetic manipulations (Bredenoord, Clevers, and Knoblich 2017; Lancaster and Knoblich 2014). In particular, CRISPR interference (CRISPRi) technology offers a wide range of future directions to the described projects in this dissertation as the ease of generating guide RNAs for various targets facilitates an investigation of a multitude of genes *in vitro* that have been difficult to interrogate

due to embryonic lethality (Clamons and Murray 2017; Boettcher and McManus 2015; Larson et al. 2013). Moreover, due to the seemingly unlimited number of experiments that CRISPRi might facilitate, it would be beneficial to constrain the number of key gene targets through a systematic screen of many genes at once.

Classically screens look for easily distinguishable read outs, such as cell death or expression of a particular gene (Bassik et al. 2013; Kampmann, Bassik, and Weissman 2014), that enable high throughput scanning of multiple targets. However, the complicated nature of phenotypes and multi-gene interactions involved in developmental morphogenesis as well as the time needed to allow morphogenic events to occur has limited the use of screens in developmental biology to invertebrates such as *Drosophila* and *C.elegans* (Ahnn and Fire 1994). Therefore, pairing the 2D platform for pattern emergence described in Chapter 2 with pooled guide CRISPRi screening technology provides a high throughput manner in which 2D patterning events can be characterized as a proxy to the more complicated morphologies seen in the embryo. A pooled screen will allow for systematic review of many gene knockdowns and analysis of cell segregation would highlight genes necessary for morphogenic organization events. Furthermore, this would allow for genetic interrogation of early embryonic patterning in a human context. A patterning screen could then be further expanded to encompass lineage fate decisions by including live reporters for lineage specific markers, enabling interrogation of how position within a 2D tissue influences lineage and subsequent patterning. Overall, the expansion of CRISPRi targets in the form of a screen would allow for high throughput interrogation of a multitude of genetic regulators of multicellular patterning, enabling the establishment of the main regulators of morphogenesis within the embryo.

6.2 Expanding CRISPRi Perturbations in the Cellular Potts Model

The computational model described in Chapter 3 tested the likelihood of creating three desired patterns (Multi-Island, Bullseye, or Janus) within the limited parameter space that was

available in the *in vitro* system. Although the model was able to successfully predict an experimental design that yielded the Multi-Island and the Bullseye patterns, it was unable to find parameters that would give rise to the Janus pattern. Additionally, the computational model predicts the likelihood of achieving the target pattern, which in the case of the Bullseye was only 12% of the time. Therefore, it would be beneficial to examine how the model could be expanded and made more robust, so that the predicted patterns occurred more often and with less limitations and a wider diversity of patterns can be achieved.

Thus, analogous to the expansion of genetic targets *in vitro*, the *in silico* model can be expanded to incorporate additional parameters corresponding to genetic perturbations. The Cellular Potts Model has the advantage of being modular so that as new cellular behaviors are observed *in vitro*, they can be quantified and added as another “cell line” parameter in the computational model. This would allow for the expansion of the model to include multiple levels of control beyond just the two previously interrogated knockdowns (CDH1 and ROCK1). Including other adhesion molecules such as EPCAM or molecules involved in cellular repulsion such as ephrins (Taylor et al. 2017; Holmberg, Clarke, and Frisé 2000) may facilitate additional phenotypes and boundary formations that would provide the experimental parameters to create more robust and tunable patterns. For example, the Janus pattern may be achieved by a repulsive signal, the likelihood of obtaining the Bullseye pattern may be increased, or the number of islands in the Multi-Island pattern could be more robustly titrated. Overall, this allows for mechanistic interrogation of the necessary behaviors required for robust generation of developmental patterning and allows for expansion of our understanding of physical morphogenic regulation.

Furthermore, with more parameters to explore, the computational model could be pushed to interrogate the creation of more complex patterning events with multiple cellular populations or expanded to a 3D Cellular Potts Model (Wu 1982), simulating contexts similar to those present in organoid systems. Specifically, the computational model could interrogate the key events needed

to generate a primitive streak, or a closing neural tube, specifically focusing on what multi-cellular patterning events are required and how might they be replicated given an expanded set of parameters in the form of gene knockdown. Overall, the expansion of the computational model would enable many follow up studies interrogating key aspects of symmetry breaking, how they are robustly controlled, and potentially predict experimental conditions to mimic complex morphological events.

6.3 Interrogating Hippo Signaling in Pluripotent Stem Cell Transitions

In Chapter 4 of this dissertation, a model for early embryonic lineage emergence is discussed where the appearance of an extraembryonic-like population is seen with CDH1 knockdown and hydrogel encapsulation of 50 cell human iPSC aggregates. The development of the extraembryonic trophectoderm lineage is dependent on the Hippo signaling pathway and downstream TEAD activity to activate CDX2 and GATA3 driving trophoblast determination (Yagi et al. 2007; Nishioka et al. 2009; Ralston et al. 2010; Strumpf et al. 2005, 2). Hippo signaling is a mechanically sensitive signaling pathway distinguished by translocation of YAP to the nucleus where YAP then binds with TAZ and directs TEAD to appropriate downstream targets (Halder and Johnson 2011). Interestingly, the early embryo is encapsulated, similar to the experiments in Chapter 4, in a layer of ECM called the zona pellucida (Bleil and Wassarman 1980). The results in Chapter 4 suggest that encapsulation itself may provide a micro-environment conducive to extraembryonic lineage emergence, where encapsulation triggers Hippo signaling, either through mechanical restriction or sequestering of excreted signaling molecules. Therefore, an immediate follow up experiment for the model system described in Chapter 4 would be to determine whether there is increased nuclear YAP in encapsulated aggregates. Nuclear YAP would indicate that the Hippo signaling pathway may be responsible for the emergence of a trophoblast like population in the encapsulated aggregates. Further experiment could then use small molecule inhibitors to

determine whether canonical or non-canonical Hippo pathway is being activated by encapsulation.

Furthermore, the emergence of an extraembryonic-like cell lineage largely in human iPSCs that have lost CDH1 expression leaves additional questions to be addressed. The regulation of CDH1 in trophoblast development is often described as a redistribution of cadherin junctions in the polar trophoblast cells (Vestweber et al. 1987), helping to create a barrier surrounding the cyst of the developing blastocyst. Moreover, this process involves the non-canonical Wnt signaling pathway to regulate the changes in cell polarity (Ducibella et al. 1975; Hirate et al. 2013; Vestweber et al. 1987). It is possible that the knockdown of CDH1 activating the Wnt signaling pathway, leading to the emergence of extraembryonic lineages (Stephenson, Yamanaka, and Rossant 2010). To interrogate whether CDH1 loss promotes trophoblast emergence by enabling increased Wnt signaling, first nuclear localization of β -catenin and phosphorylated ATF2 could be examined, marking canonical and noncanonical Wnt signaling, respectively. Then a combination of follow up experiments using knockdown to target other members of both the canonical (TCF4, β -catenin) and non-canonical (AP-1, ATF2) Wnt signaling pathway would then illuminate how the pathways might regulate trophoblast emergence. Overall, the proposed future experiments would highlight the role of Hippo and Wnt signaling in the emergence of the trophoblast lineage and how it relates to the micro-environment provided by encapsulation, suggesting novel methods by which the embryo regulates population emergence.

6.4 Examining Neuromesodermal Progenitor Fate Regulation

In Chapter 5, an organoid model for axial extension and neuromesodermal progenitors (NMPs) is described to mimic many aspects of early neural tube formation and neuronal population emergence. However, a remarkable aspect of NMPs is their ability to contribute to tissues across germ lineages to form of the posterior neural tube and neighboring somites (Henrique et al. 2015; Martin 2016; Wilson, Olivera-Martinez, and Storey 2009). The formation of

the neural tube and surrounding somites are coordinated events in embryonic development where disruption to either process often results in malformations in the parallel tissue (Harris and Juriloff 2007; Vermot et al. 2005). It is proposed that the axial stem cell pool of NMPs coordinates the balanced formation of the posterior neural tube and neighboring somites to generate a functional spine. However, how this pool of progenitors mechanistically balances production of two paired tissues is poorly understood, because of a lack of techniques to either track NMP dynamics and cellular organization *in vivo*, or to define intrinsic genetic regulators of the NMP lineage specification, such as through a large-scale *in vivo* screen. Therefore, the system developed in Chapter 5 offers a wealth of potential follow up experiments to interrogate questions about NMP regulation that were previously limited due to the difficulties of working *in vivo*. For example, the system lends itself to quickly generated genetic knockdown of desired targets that may affect the regulation of NMP differentiation such as but not limited to, HOX genes, Wnt and FGF pathway regulators, adhesion molecules, and NOTCH signaling pathway regulators. Overall, this would enable mechanistic interrogation of regulatory genes and pathways specific to neural tube formation, expanding our insight on what regulates robust formation of the spine.

Furthermore, organoid culture facilitates larger scale screening technologies both at the genetic and small molecule level. Follow up experiments could focus on small molecule inhibitor screens to interrogate the necessary pathways involved in NMP establishment, maintenance, and differentiation. Additionally, the extension phenotype could be used as a read out to screen for key genetic determinants of morphogenic extension. Overall the axial extension organoid system enables many avenues of experimentation interrogating morphogenesis and lineage specification of the spinal cord and surrounding tissues that were previously limited due to the technical constraints of working with *in vivo* systems.

6.5 Conclusions

The results of this dissertation significantly contribute to the fields of developmental biology and bioengineering by establishing platforms to interrogate multicellular organization in the presence of symmetry breaking events such as pattern formation and multi-lineage specification. The development of computational models that inform systematic pattern recognition to identify and predict symmetry breaking events enables the interrogation of developmental biology concepts in a high throughput, quantitative manner. The establishment of multiple organoid systems in which symmetry breaking events can be mechanistically interrogated, either in the form aggregate protrusion morphologies or axial extension as a result of NMP population emergence, will help to further illuminate concepts in developmental biology. Finally, the described studies use these platforms to highlight mechanisms by which heterogeneous cell-cell adhesion contributes to lineage fate decisions and demonstrate preliminary results on the role of microenvironment in early lineage fates.

6.6 Bibliography

- Adams, David, Richard Baldock, Shoumo Bhattacharya, Andrew J. Copp, Mary Dickinson, Nicholas D. E. Greene, Mark Henkelman, et al. 2013. "Bloomsbury Report on Mouse Embryo Phenotyping: Recommendations from the IMPC Workshop on Embryonic Lethal Screening." *Disease Models & Mechanisms* 6 (3): 571–79. <https://doi.org/10.1242/dmm.011833>.
- Ahn, J., and A. Fire. 1994. "A Screen for Genetic Loci Required for Body-Wall Muscle Development during Embryogenesis in *Caenorhabditis Elegans*." *Genetics* 137 (2): 483–98.
- Bassik, Michael C., Martin Kampmann, Robert Jan Lebbink, Shuyi Wang, Marco Y. Hein, Ina Poser, Jimena Weibezahn, et al. 2013. "A Systematic Mammalian Genetic Interaction Map Reveals Pathways Underlying Ricin Susceptibility." *Cell* 152 (4): 909–22. <https://doi.org/10.1016/j.cell.2013.01.030>.
- Bleil, Jeffrey D., and Paul M. Wassarman. 1980. "Structure and Function of the Zona Pellucida: Identification and Characterization of the Proteins of the Mouse Oocyte's Zona Pellucida." *Developmental Biology* 76 (1): 185–202. [https://doi.org/10.1016/0012-1606\(80\)90371-1](https://doi.org/10.1016/0012-1606(80)90371-1).
- Boettcher, Michael, and Michael T. McManus. 2015. "Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR." *Molecular Cell* 58 (4): 575–85. <https://doi.org/10.1016/j.molcel.2015.04.028>.
- Bredenoord, Annelien L., Hans Clevers, and Juergen A. Knoblich. 2017. "Human Tissues in a Dish: The Research and Ethical Implications of Organoid Technology." *Science* 355 (6322): eaaf9414. <https://doi.org/10.1126/science.aaf9414>.
- Clamons, Samuel E., and Richard M. Murray. 2017. "Modeling Dynamic Transcriptional Circuits with CRISPRi." *BioRxiv*, November, 225318. <https://doi.org/10.1101/225318>.

- Ducibella, Thomas, David F. Albertini, Everett Anderson, and John D. Biggers. 1975. "The Preimplantation Mammalian Embryo: Characterization of Intercellular Junctions and Their Appearance during Development." *Developmental Biology* 45 (2): 231–50. [https://doi.org/10.1016/0012-1606\(75\)90063-9](https://doi.org/10.1016/0012-1606(75)90063-9).
- Halder, Georg, and Randy L. Johnson. 2011. "Hippo Signaling: Growth Control and Beyond." *Development* 138 (1): 9–22. <https://doi.org/10.1242/dev.045500>.
- Harris, Muriel J., and Diana M. Juriloff. 2007. "Mouse Mutants with Neural Tube Closure Defects and Their Role in Understanding Human Neural Tube Defects." *Birth Defects Research Part A: Clinical and Molecular Teratology* 79 (3): 187–210. <https://doi.org/10.1002/bdra.20333>.
- Henrique, Domingos, Elsa Abranches, Laure Verrier, and Kate G. Storey. 2015. "Neuromesodermal Progenitors and the Making of the Spinal Cord." *Development* 142 (17): 2864–75. <https://doi.org/10.1242/dev.119768>.
- Hirate, Yoshikazu, Shino Hirahara, Ken-ichi Inoue, Atsushi Suzuki, Vernadeth B. Alarcon, Kazunori Akimoto, Takaaki Hirai, et al. 2013. "Polarity-Dependent Distribution of Angiotensin Localizes Hippo Signaling in Preimplantation Embryos." *Current Biology* 23 (13): 1181–94. <https://doi.org/10.1016/j.cub.2013.05.014>.
- Holmberg, Johan, Diana L. Clarke, and Jonas Frisén. 2000. "Regulation of Repulsion versus Adhesion by Different Splice Forms of an Eph Receptor." *Nature* 408 (6809): 203–6. <https://doi.org/10.1038/35041577>.
- Kampmann, Martin, Michael C. Bassik, and Jonathan S. Weissman. 2014. "Functional Genomics Platform for Pooled Screening and Generation of Mammalian Genetic Interaction Maps." *Nature Protocols* 9 (8): 1825–47. <https://doi.org/10.1038/nprot.2014.103>.
- Lancaster, Madeline A., and Juergen A. Knoblich. 2014. "Organogenesis in a Dish: Modeling Development and Disease Using Organoid Technologies." *Science* 345 (6194): 1247125. <https://doi.org/10.1126/science.1247125>.

- Larson, Matthew H., Luke A. Gilbert, Xiaowo Wang, Wendell A. Lim, Jonathan S. Weissman, and Lei S. Qi. 2013. "CRISPR Interference (CRISPRi) for Sequence-Specific Control of Gene Expression." *Nature Protocols* 8 (11): 2180–96. <https://doi.org/10.1038/nprot.2013.132>.
- Martin, Benjamin L. 2016. "Factors That Coordinate Mesoderm Specification from Neuromesodermal Progenitors with Segmentation during Vertebrate Axial Extension." *Seminars in Cell & Developmental Biology*, Bone development and disease, 49 (January): 59–67. <https://doi.org/10.1016/j.semcdb.2015.11.014>.
- Nishioka, Noriyuki, Ken-ichi Inoue, Kenjiro Adachi, Hiroshi Kiyonari, Mitsunori Ota, Amy Ralston, Norikazu Yabuta, et al. 2009. "The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass." *Developmental Cell* 16 (3): 398–410. <https://doi.org/10.1016/j.devcel.2009.02.003>.
- Ralston, Amy, Brian J. Cox, Noriyuki Nishioka, Hiroshi Sasaki, Evelyn Chea, Peter Rugg-Gunn, Guoji Guo, Paul Robson, Jonathan S. Draper, and Janet Rossant. 2010. "Gata3 Regulates Trophoblast Development Downstream of Tead4 and in Parallel to Cdx2." *Development* 137 (3): 395–403. <https://doi.org/10.1242/dev.038828>.
- Stephenson, Robert Odell, Yojiro Yamanaka, and Janet Rossant. 2010. "Disorganized Epithelial Polarity and Excess Trophectoderm Cell Fate in Preimplantation Embryos Lacking E-Cadherin." *Development* 137 (20): 3383–91. <https://doi.org/10.1242/dev.050195>.
- Strumpf, Dan, Chai-An Mao, Yojiro Yamanaka, Amy Ralston, Kallayane Chawengsaksophak, Felix Beck, and Janet Rossant. 2005. "Cdx2 Is Required for Correct Cell Fate Specification and Differentiation of Trophectoderm in the Mouse Blastocyst." *Development* 132 (9): 2093–2102. <https://doi.org/10.1242/dev.01801>.
- Taylor, Harriet B., Anaïs Khuong, Zhonglin Wu, Qiling Xu, Rosalind Morley, Lauren Gregory, Alexei Poliakov, William R. Taylor, and David G. Wilkinson. 2017. "Cell Segregation and Border Sharpening by Eph Receptor–Ephrin-Mediated Heterotypic Repulsion." *Journal of the Royal Society Interface* 14 (132). <https://doi.org/10.1098/rsif.2017.0338>.

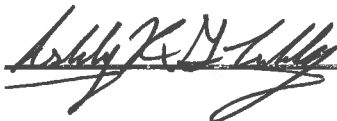
- Vermot, Julien, Jabier Gallego Llamas, Valérie Fraulob, Karen Niederreither, Pierre Chambon, and Pascal Dollé. 2005. "Retinoic Acid Controls the Bilateral Symmetry of Somite Formation in the Mouse Embryo." *Science* 308 (5721): 563–66. <https://doi.org/10.1126/science.1108363>.
- Vestweber, Dietmar, Achim Gossler, Klaus Boller, and Rolf Kemler. 1987. "Expression and Distribution of Cell Adhesion Molecule Uvomorulin in Mouse Preimplantation Embryos." *Developmental Biology* 124 (2): 451–56. [https://doi.org/10.1016/0012-1606\(87\)90498-2](https://doi.org/10.1016/0012-1606(87)90498-2).
- Wilson, Valerie, Isabel Olivera-Martinez, and Kate G. Storey. 2009. "Stem Cells, Signals and Vertebrate Body Axis Extension." *Development* 136 (10): 1591–1604. <https://doi.org/10.1242/dev.021246>.
- Wu, F. Y. 1982. "The Potts Model." *Reviews of Modern Physics* 54 (1): 235–68. <https://doi.org/10.1103/RevModPhys.54.235>.
- Yagi, Rieko, Matthew J. Kohn, Irina Karavanova, Kotaro J. Kaneko, Detlef Vullhorst, Melvin L. DePamphilis, and Andres Buonanno. 2007. "Transcription Factor TEAD4 Specifies the Trophectoderm Lineage at the Beginning of Mammalian Development." *Development* 134 (21): 3827–36. <https://doi.org/10.1242/dev.010223>.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature



Date

Dec 2nd 2019