Title
Signaling and self-organization in human pluripotent stem cell-based early embryonic models

Author
Vasic, Ivana

Publication Date
2022

Peer reviewed|Thesis/dissertation

Signaling and self-organization in human pluripotent stem cell based early embryonic models

by

Ivana Vasic

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
AND
UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

DocuSigned by:
Todd McDevitt
38EF3B1D3234412...

Todd McDevitt

Chair

DocuSigned by:
Valerie Weaver
[illegible]A7...

Valerie Weaver

DocuSigned by:
DAVID SCHAFFER
[illegible]410...

DAVID SCHAFFER

DocuSigned by:
Susan Fisher
3D7FD5EAA89E411...

Susan Fisher

Committee Members

Dedication

This work is dedicated to my parents, Dragana and Bane, and to my sister, Jelena, for their unconditional love and support. To my mother for treating my battles as her own, and for giving me the courage to take on the hardest challenges knowing that I would always have her by my side. To my father for showing me how to approach life with humor and curiosity, and for his exceptional ability to listen and understand. To my sister for teaching me life's most important lessons through example: be loyal to your friends, prioritize having fun, and always make time for your family. This body of work would not have been possible without them and the opportunities they provided for me.

This work must also be dedicated to my fiancée, Jon, who is a constant source of laughter and happiness, and who never tires of loving and re-encouraging me during times of doubt and uncertainty. With him I feel that anything is achievable.

Acknowledgments

I am very grateful to my family, friends, colleagues, and mentors who supported me, and who were there to endure the lows and celebrate the highs with me as I grew into a more resilient scientist and person.

To Todd, my advisor for the first 4 years of my PhD, who taught me to be creative and to take bold risks when formulating research questions. Aside from being an incredibly supportive mentor, Todd is also an incredibly kind person who always prioritized his mentees, and was able to cultivate a lab culture where lab members gave their time generously in help and support of their peers. I'd also like to thank the former McDevitt lab for my wonderful scientific upbringing, in particular: David, Oriane, Ariel, Ana, Vaishaali, Serah. A very special thanks goes to Ashley Libby for her continued mentorship throughout the years, the energy she devoted to answering my questions and showing me the ropes, and her determination in helping me fight off my impostor syndrome – I couldn't have done it without you, Ashley. To Jas – my partner in crime since day one in our PhD program. And last but certainly not least, to Will, Nick, and Emily – thank you for making lab a fun place, for all the dinners and nights out, and for listening to my science talks over and over and over again – you're the real MVPs.

To Shinya, my current PhD advisor, who kindly accepted me into his lab after Todd left Gladstone, and devoted time to ensure that I was taken care of during a time of great uncertainty. From Shinya I learned confidence and how to deliver compelling presentations, which has served me well as I transition into my new role as a founder. Shinya's mentorship helped me grow into a highly rigorous scientist, and I aspire to be a mentor like him one day. Thank you to the Yamanaka lab: Vanessa, Justin, Gay, Haruko, Akira, Sam, Jonathan, Eugenea. And a special thank you to Kiichiro – my unofficial PhD mentor, who has endless curiosity and energy for science.

To new friends that provided me with a home away from home in the Bay Area, and who made the last five years filled with happy memories. To Chris – for all the walks, lunches, and philosophical discussions that kept me going in the hardest moments. To Jake and Julian – who never judge, always make me laugh, and have seen it all. To Annie – the most empathic person I know, and an exceptional listener. To Katie – who is fiercely loyal, and who is always on my side whether I'm right or wrong. To Rami and Karen – for the stimulating debates, the dance parties, the trips, and the many wonderful memories throughout the years. To Cody – for all the private guitar sessions and serenades, and always being down for anything. To old friends who already know how much they mean to me, but especially Melissa, Gabi, William, and Lily for giving me love and support from afar during the last several years.

To my fiancé, Jon, who I met in my first year of the Bioengineering program. The last five years were the hardest of my life, but also the most fun and the most rewarding with you by my side. Thank you for being with me through it all, and choosing to grow alongside me as my partner in life. Finally, I'd like to thank my family: my parents, Dragana and Bane, and my sister, Jelena. Each in their own way has raised me and influenced who I am. Their opinion means more to me than anyone else's, and I wouldn't be where I am without the safety and comfort they provided me, even from afar.

Contributions

Some of the material in this dissertation has been, or soon will be, published in the references below. The co-authors listed in these publications assisted, directed, or supervised the research that forms the basis for this dissertation.

Chapter 1:

Vasic, I. & McDevitt, T. C. Engineering co-emergence in organoid models. Prog. Biomed. Eng. 3, 1–10 (2021).

IV and TCM both contributed in researching, organizing, and writing/editing the original review.

Chapter 2:

Libby, A. R.G.*, Vasic, I.*, Joy, D. A., Krakora, M. Z., Mendoza-Camacho, F. N., Conklin, B. R., McDevitt, T. C., "Silencing of E-cadherin in induced human pluripotent stem cells promotes extraembryonic fates accompanying multilineage differentiation" bioRxiv, pp. 1–19, 2020.

* denotes equal contribution

ARGL and TCM were responsible for the conceptualization of the project and designing experiments. Differentiations were conducted by ARGL and IV. Single cell RNA sequencing preparation and analysis were performed by ARGL, DAJ, and MZK. Image analysis script was written and implemented by DAJ and ARGL. Light sheet microscopy and analysis was conducted by ARGL, DAJ, and IV. Cell lines were generated by ARGL, DAJ, MZK, and NMC. ARGL and IV prepared the figures with input from all co-authors. All co-authors contributed to writing the original manuscript.

Chapter 3:

Vasic, I., Libby, A. R.G., Maslan, A., Bulger, E. A., Zalazar, D., Krakora Compagno, M. Z., Streets, A., Tomoda, K., Yamanaka, S., McDevitt, T. C., "Loss of tight junctions disrupts gastrulation patterning and increases differentiation towards the germ cell lineage in human pluripotent stem cells" Under Review at Developmental Cell, 2022.

IV, ARGL, and TCM were responsible for conceptualization. IV performed the majority of the experiments, with assistance from MZKC, DZ, and EAB. AM performed RNA sequencing and whole genome bisulfite sequencing data processing and analysis. IV wrote the manuscript with contributions from the other authors. AS, KT, SY, and TCM oversaw the project.

Signaling and self-organization in human pluripotent stem cell-based early embryonic models

Ivana Vasic

Abstract

Stem cells within the early embryo have the incredible capacity to self-organize and form higher order structures that provide the basis for large scale tissue morphogenesis. Understanding how embryonic signals are shaped, perceived, and transduced by cells early in embryonic development is critical for progress in the field of tissue engineering, so that we are able to recapitulate these cues *in vitro* and grow cells and tissue constructs outside of the body for high-throughput and animal-free drug/toxicological testing and cell/tissue replacement and therapy. However, efforts to study signaling and stem cell specification early in human development have largely been hindered by significant morphological differences in embryonic development of model organisms and ethical limitations regarding human embryo research. Thus, there is a need for tractable *in vitro* models which closely resemble native embryonic tissue, in both structure and behavior. Selecting an appropriate *in vitro* model that is tailored to hypothesis testing is critical for deriving conclusions about *in vivo* development. In this dissertation, I present a review on engineering co-emergence in organoid models, along with two studies which use novel 2D and 3D pluripotent stem cell models to study cell decision and symmetry breaking events during two different stages of early human embryonic development.

In Chapter 1, I discuss the current state of organoid engineering, and define stringent criteria for organoid models. I discuss how primitive patterns and asymmetries integrate with

endogenous cellular response systems to create autonomous and highly complex patterns that serve as a blueprint for higher order structure formation. In Chapter 2, I introduce cell structure as a key parameter which defines cell migration, organization, and specification. Specifically, I discuss how silencing of the cell-cell junction protein, CDH1, in a subpopulation of pluripotent stem cells within an embryoid body mimics zona pellucida hatching prior to gastrulation and prompts robust segregation and organization of CDH1- cells, along with their spontaneous differentiation to extraembryonic lineages. In Chapter 3, I focus on another cell-cell junction protein, ZO1, whose role is to both partition the plasma membrane into apical and basolateral functional domains, and assemble transmembrane tight junction proteins which prevent paracellular diffusion of macromolecules between the apical- and basolateral-facing lumens. Specifically, I demonstrate that cellular polarity, maintained by ZO1 in pluripotent stem cells, is a key regulator of morphogen signaling and pattern formation in a 2D peri-gastrulation model. In Chapter 4, I discuss the implications of these studies for signaling, specification, and migration of stem cells during embryonic development.

Table of Contents

List of Figures

List of Abbreviations

BMP4 – *bone morphogenic protein-4*

ECM – *extracellular matrix*

CRISPR - *clustered regularly interspaced short palindromic repeats*

hESC – *human embryonic stem cell*

hiPSC – *human induced pluripotent stem cell*

hPSC – *human pluripotent stem cell*

KD – *knockdown*

KO - *knockout*

PCR – *polymerase chain reaction*

PGC – *primordial germ cell*

PGCLC – *primordial germ cell like cell (derived in vitro)*

qPCR – *quantitative polymerase chain reaction*

WT – *wild type*

ZKD – *ZO1 knockdown*

ZO1 – *zonula occludens-1*

Chapter 1. Introduction: Engineering co-emergence in organoid models

Abstract

Pluripotent stem cell-derived organoids provide *in vitro* models of development and disease that can be used for a wide range of biomedical applications, including high-throughput screens or regenerative medicine. The ability of stem cells to self-renew and self-organize in three dimensions is the basis for creating highly structured multicellular organoid models. However, progress in clinical translation of organoid technologies has been stymied by the stochastic nature of stem cell differentiation within organoids, which leads to inconsistent cell type maturity, tissue function, reproducibility, and control over macroscale structure and phenotype(s). Advances in our understanding of developmental biology and the mechanisms which regulate symmetry breaking and pattern formation in the embryo have led to new approaches for engineering cooperative emergence (co-emergence) in organoid models to address these challenges.

Introduction

Our understanding of human development and tissue morphogenesis has been hindered by significant morphological differences between humans and model organisms, and ethical limitations regarding human embryo research. Organoids, which are three-dimensional microscale replicas of *in vivo* tissues, provide a powerful *in vitro* platform for studying development and disease. These models have proven to be powerful tools for high throughput drug discovery and competent surrogates for toxicological testing[1,2]. Organoid formation relies on the intrinsic ability of stem/progenitor cells to both propagate and self-organize into structures which mimic those

formed *in vivo*. Seminal studies describing intestinal crypt[3] and mammary gland[4] organoids formed from a single adult progenitor cell or from a homogeneous population of aggregated progenitor cells, respectively, demonstrated that even with minimal external input, progenitor cells contain sufficient intrinsic programs to guide their own organization and differentiation. Subsequent generation of organoid models from isolated adult human tissues have recapitulated aspects of the lung[5], stomach[6], liver[7], pancreas[8], and many other organ systems[9,10,11,12], providing fundamental insights into the developmental trajectory of these tissues and the translational potential of organoid technologies.

Since these original examples, advances in pluripotent stem cell (PSC) research have galvanized a new generation of organoid models. While previous models derived from isolated adult stem/progenitor cells were dominated by endodermal tissue types due to their ease of isolation and expansion *in vitro*, the use of PSCs has enabled expansion of organoids derived from all three germ lineages[13]. PSCs also have the potential to form complex multi-lineage organoids, are amenable to genome editing, and in the case of induced PSCs (iPSCs), can be reprogrammed directly from patients, enabling corrective editing for precision medicine and studies on how genetic diseases impact tissue. However, structure emergence in PSC-derived organoids is often not reproducible, with variability across multiple scales: between different organoids of the same batch, different batches, and in different cell lines[14,15,16,17]. While tissue organization may vary between organoids in the same batch, composition is typically uniform; however, studies report large variability between batches where organoids derived via identical protocols or bioreactor setups produce vastly different organoid compositions and phenotypes[16,17]. Even under the most stringent and reproducible conditions, where organoid culture is performed by the same researcher using the same media and materials, batch variability is frequently observed. Such variability is

most likely derived from stochastic symmetry breaking events that occur during both 2D and 3D differentiations, and result in either gain or loss of rare populations or differences in ratios of key cell types that drive self-organization and structure formation[18,19]. PSC-derived organoids also lack the functional maturity of their *in vivo* counterparts, often attaining behavioral and transcriptomic signatures resembling fetal tissues, but failing to progress further[20,21].

Irreproducibility and lack of maturity hinders clinical translation of PSC-derived organoids. In this review, we first define criteria for organoid models in reaching translation, and discuss the implicit criterium of cooperative emergence (co-emergence) of multiple cell types in replicating organogenesis and achieving maturity. We will discuss how engineering co-emergence may offer control over patterning, structure emergence, and reproducibility; and review progress that has been made toward engineering co-emergence in other model systems. We will focus specifically on how symmetry breaking events early in the differentiation process ultimately determine macroscale organoid phenotype, and introduce starting cell composition as a novel parameter in controlling the downstream events of symmetry breaking, co-emergence, and stem cell differentiation.

Criteria for Organoid Formation

Organoid engineering is performed by modulating macroscale parameters while adhering to a basic set of organoid specifications. These criteria include a three-dimensional structure which recapitulates native tissue and allows for autologous interaction between parenchymal cells and the extracellular matrix (ECM), a stem/progenitor cell source that both self-propagates and

differentiates in the correct ratios, self-organization of differentiated cell types, and most importantly, functionality which mimics aspects of healthy and/or diseased tissues.

Monolayer versus Three-Dimensional Culture

Studies in both mesenchymal and epithelial tissues underscore the importance of cell-cell and cell-matrix interactions in a 3D setting for tissue phenotype, maturity, and function. Differentiation of epithelial tissues is in many ways complementary to 2D culture conditions, as these protocols often create columnar sheets that mimic mesoscale organ properties. For example, monolayer differentiation of epithelial hepatocytes and cholangiocytes have allowed for a close interrogation of the biliary duct stem cell niche[22]; however, these culture parameters prevent formation of macroscale epithelial structures that are critical to tissue function. Only cholangiocytes cultured as 3D aggregates form cysts and secrete fluids, a critical function which is required to model the biliary disorders in Alagille syndrome and cystic fibrosis[23].

Cellular function and self-organization rely on the composition of endogenously-generated and tissue-specific ECM, which is continuously remodeled throughout development and can become pathogenic in disease states. For example, cardiomyocyte proliferation, sarcomere length and alignment, and adherence to ECM are significantly different between cardiomyocytes cultured on fetal versus aged decellularized cardiac matrices as well as typical matrix coatings, such as Matrigel, implicating ECM composition as a major regulator of tissue organization and function[24,25,26,27]. Consequently, fibrotic scarring that occurs post cardiac injury influences heart function via pathogenic changes to ECM composition[28]. These examples demonstrate how interactions between cells and tissue-specific ECM impact cell function and highlight the insufficiency of culturing cells using “one-size-fits-all” exogenous matrix protein coatings.

Transitioning from 2D to 3D culture ameliorates this problem by allowing differentiated cells to produce and regulate organization of their own matrix. In the context of cardiomyocytes, this generally occurs via co-emergence of tissue-specific fibroblasts and other supporting cell types[29,30].

Differentiation and Self-Organization

Tissue development and homeostasis require progenitor cells to both proliferate, in order to replenish the progenitor pool, and differentiate, resulting in the diversity of cell types found in a tissue. Controlling differentiation parameters to achieve the ideal proportion of self-renewal and differentiation events, as well as the intended ratio of terminally differentiated cell types has proven to be a major challenge in organoid engineering[31,32]. In many cases, the ratio of cell types also dictates the extent of self-organization, either through differential expression of physical adhesion proteins or intercellular signaling[33]. A simplified *in vitro* model of the mammary epithelium allows interrogation of the interplay between cell type ratio and self-organization. The mammary epithelium is comprised of two main subtypes: myoepithelial cells, which surround the cells of the lumen, and secretory luminal epithelial cells. Combining myoepithelial cells and luminal epithelial cells in a 1:1 ratio on 2D micropatterns causes clear partitioning and ring formation, with myoepithelial cells surrounding the luminal epithelial cells[34]. By contrast, when the cell types are combined in 1:2 or 1:3 ratios, cells are arbitrarily organized and do not undergo physical cell sorting. This definitive 2D pattern highlights the possibility that a threshold number of cells and ratio of cell types may be required for self-organization to occur in 3D organoid models. In 3D, ratios of epithelial versus mesenchymal cell types in intestinal aggregates have been shown to drive macroscale organoid properties. Specifically, aggregates which contain an inner mesenchymal mass surrounded by an epithelium (compared to pure epithelial or

mesenchymal aggregates) following aggregation form intestinal organoids instead of spheroids[33]. However, the complexity and multitude of required cell types (many of which are not characterized) impedes designing co-culture conditions that satisfy ratio requirements in other tissues. Furthermore, while organization in the micropatterned mammary gland seems to be primarily dictated by differential expression of the adhesion proteins N-cadherin and P-cadherin[34], other tissue types rely on a combination of juxtracrine and paracrine signals to induce organization, complicating efforts to engineer structure formation by co-culture. Thus, a fundamental understanding of stem cell fate decision programs is necessary to induce co-emergence and patterning of intermediate and developmentally-relevant cell types that instruct macroscale organization.

Function

The greatest challenge for clinical application of organoid models is the necessity that the model recapitulates functional properties of both healthy and diseased tissues. The definition of functionality is still a subject of debate among researchers. For example, cardiomyocytes derived from human PSCs in monolayer culture display calcium handling properties and excitation-contraction coupling, although the dynamics of these processes are immature and significantly slower than *in vivo*[35]. Nevertheless, human PSC-derived monolayer cardiomyocytes are used to screen drugs for drug-induced arrhythmic beating and have shown to be proficient at predicting clinical arrhythmias and QT prolongation[36], barring the need for more sophisticated functional properties. The level of functionality attained by an organoid model to some degree determines its use; however, a thorough understanding of existing tissue and disease models is required to make this assessment. For example, *in vitro* models of axial elongation derived from human PSC aggregates have been used to confirm known teratogenic compounds with a high success rate by

looking at which compounds inhibited elongation in the aggregates[37]. However, lack of endodermal cells within these models prevents characterization of compounds which may affect gastrulation or endodermal-related lineages specifically. For example, Bisphenol A (BPA), a common synthetic compound used in plastic manufacture has been shown in model organisms to affect migration and specification of primordial germ cells (PGCs) resulting in reproductive complications[38]. Using an organoid model of axial elongation to test the toxic effects of BPA in humans in this instance would be insufficient, as PGCs are an endoderm-related cell type.

The most reliable tests for functional competency are done by transplantation into an animal model. An exceptional example of a functional organoid model is the liver bud organoid developed by Takebe et al., which consist of hepatic endoderm precursors, human umbilical vein endothelial cells (HUVECs), and mesenchymal stem cells. Prior to transplantation, the endothelial cells self-organize into vascular structures that become integrated with mouse host vessels following transplantation to achieve perfusion, which in turn drives maturity of the hepatic precursors[39]. *In vitro*, the liver bud organoids have metabolic and secretory capabilities that parallel those of cultured human adult hepatocytes, and transplantation of the liver bud organoids can rescue drug-induced liver failure in mice. This type of functional evaluation is the ideal end-goal for assessing the success of organoid engineering.

Defining Cooperative Emergence

An implicit criterium of organoid formation is co-emergence: stem/progenitor cells within the organoid must generate multiple cell types in order for self-organization to occur. Co-emergence can refer to the emergence and interaction between two or more cell types from the

same progenitor cell, between two cell types from more distant progenitors, or even between cells from different germ lineages which interact and support each other's development (Figure 1)[40,41]. The principle of co-emergence is based on feedback loops of paracrine signals secreted by proximal cell and tissue types that develop synergistically[41]. Organoid models derived from primary adult stem/progenitor cells usually rely on co-emergence among closely-related ancestral progenitors. For example, in intestinal organoids derived from isolated adult Lgr5+ stem cells, Paneth cells, which are derived from Lgr5+ cells, are required for maintenance of the stem cell pool through their secretion of Wnt3 and TGFα[3,42]. Extending culture time allows for differentiation of other supporting cell types such as tufts cells, enterocytes, and goblet cells, which are needed to form crypt-villus structures. Tissue models which require co-emergence of cell types from more distant progenitors (or different germ lineages) would not mature over time if they lack the appropriate starting cell type composition, potentially explaining why some aggregates form non-organized spheroids instead of organoids under differentiation conditions[33].

Co-emergence in organoid models takes advantage of stem cells' intrinsic ability to give rise to multiple self-organizing cell types. By contrast, in co-culture methods, two or more cell types of interest are cultured together without necessarily being derived from a common progenitor. Co-culture places less emphasis on self-regulation and tissue organization, while focusing on understanding heterotypic cell interactions. As a result, co-culture studies have provided fundamental insights into how the development of two (or more) cell types may be interlinked. For example, in cultured pre-somitic mesoderm explants, removing ectodermal cells inhibits somitogenesis, demonstrating that interaction between the two cell types is required for structure formation[43]. In mesendodermal tissues, the presence of visceral endoderm cells enhances induction of cardiomyocytes in co-culture[44], and cardiomyocytes in turn promote maturation of

differentiated endoderm hepatocytes *in vitro*[45]. However, although co-culture studies have been used to demonstrate the complexity of organ formation and highlight the critical role of supporting cell type interactions in organogenesis, they are often insufficient to recapitulate tissue self-organization and function. For example, modeling the neuromuscular junction (NMJ) by co-culturing motor neurons and skeletal muscle cells in monolayer culture, microfluidic devices, or 3D culture has revealed the importance of dynamic cell functions in improving NMJ maturity, and the incorporation of optogenetic tools has strengthened modeling of neuromuscular diseases, such as myasthenia gravis and various muscular dystrophies[46,47,48]. However, co-cultures result in largely disorganized tissues which poorly recapitulate mesoscale tissue architecture. Alternatively, neuromuscular organoids developed from PSC-derived neuromesodermal progenitors undergo co-emergence of motor neurons, Schwann cells, and muscle cells, and display axial self-organization and central pattern generator-like circuits[49]. These neuromuscular organoids are contractile and electrophysiologically competent without the use of exogenous stimulation, demonstrating the importance of co-emergence in tissue organization and function of this model.

Like neuromuscular organoids, cardiac-gut organoids, which have been developed recently by our lab, offer insight into the distinction and benefit of co-emergence over co-culture. Cardiac-gut organoids are generated by culturing early mesoderm progenitors in a media that is permissive to both mesoderm and endoderm differentiation, resulting in distinct cardiac and intestinal structures within the same organoid[41]. When compared to purely cardiac microtissues, cardiomyocytes in cardiac-gut organoids exhibit higher structural and functional maturation, as demonstrated by the spatial compartmentalization of atrial and ventricular cells, enrichment of atrial/nodal cells, myocardial compaction, sarcomere alignment, and advanced calcium handling. In addition, the complexity of endoderm structures in cardiac-gut organoids also surpasses those

achieved in purely endodermal organoids by forming an enteric smooth muscle cell layer lining the gut epithelium that exhibits spontaneous peristaltic contractility. This advanced degree of cell and tissue maturity has previously only been achieved post-transplantation *in vivo* following embedding of intestinal organoids[3,50], demonstrating the self-sufficiency of endogenous programs within the cardiac-gut organoids to guide organization.

While co-culture of purified cell types may result in more reproducible microtissues, inducing co-emergence is necessary for recapitulating cellular organization and the signaling milieu of developing tissues that leads to improved structure and maturity. Furthermore, the vast number of different cell types that result from PSC differentiation and that are present in well-established PSC-derived models such as kidney[20], optic cup[51], and cerebral[52] organoids complicate the capacity to select and finely-tune cell types and ratios via co-culture. Therefore, engineering methods to control emergence and subsequent patterning of key cell types are needed to improve not only structure and function of PSC models, but also their robustness and reproducibility.

Engineering Parameters for Organoid Formation

To construct organized multicellular organoid models in 3D, parameters such as culture method (embedding, suspension, air-liquid interface); timing, concentration, and type of exogeneous differentiation signals; and starting cell composition must all be determined and optimized[53,40]. Each parameter has important consequences in directing organoid phenotype. For example, Matrigel embedding of ectodermal neurospheres results in cerebral organoid formation, whereas suspension culture forms optic cup organoids[18]. These parameters can be classified into extrinsic and intrinsic regulators: changing culture method and exogenous differentiation signals

in the media are extrinsic or "top-down" engineering approaches designed to globally instruct endogenous cellular programs. By contrast, intrinsic approaches, such as modulation of starting cell composition, provide "bottom-up" control by manipulating feedback elements of endogenous regulatory networks. Historically, tissue engineers have more commonly employed protocols that specify extrinsic parameters, and indeed, tools for modulation of intrinsic parameters are in short supply. Even ostensibly simple modifications such as controlling the types of stem/progenitor cells in the initial population in both 2D and 3D have proven quite technically challenging.

One reason starting cell composition has been perceived as an intractable engineering parameter is that differentiations result in a wide variety of cell types whose numbers and proportions change from batch-to-batch. Therefore, upon aggregation, batch variability among differentiations leads to variability in starting cell composition and subsequent variability in organoid phenotype and function. To improve outcomes, directed differentiation protocols have been developed to discourage or inhibit undesired cell types while enriching for specific cell types through addition of small molecule inhibitors and agonists. For example, in neural organoids, addition of LDN-193189 (a selective inhibitor of type I BMP receptors) prevents PSCs from differentiating to mesoderm, and subsequent addition of other small molecules, such as Purmorphamine (a Sonic Hedgehog agonist), specify regional cell fates, such as a ventral identity[54,55]. While these treatments are used to guide differentiation, they are imperfect in that they do not achieve complete selectivity for a cell type of interest and still generate some degree of heterogeneity. Flow cytometry sorting to further purify cell types can be performed at the cost of disrupting tissue architecture. Extensive purification and selection for specific cell types using directed differentiation or flow cytometry may remove supporting cell types, some of which are

uncharacterized and perform unknown functions, but are nevertheless needed for complex tissue development and function.

Many organoid protocols begin with a seemingly homogeneous population of pluripotent stem cells, leaving limited opportunity for modulation of starting cell composition. For example, neuroectodermal organoids are typically derived from embryoid bodies[56,57,58,59], or spherical aggregates of pluripotent stem cells; whereas the majority of protocols for mesodermal or endodermal organoid formation begin with a 2D differentiation followed by aggregation. However, heterogeneity in cellular response to exogenous differentiation cues demonstrates that starting cell composition varies even in a seemingly homogeneous 2D monolayer or 3D aggregate of pluripotent stem cells. For example, a well-established cardiac differentiation protocol in which a 2D culture of hiPSCs is exposed to successive small molecule treatments[60] results in a wide range of cellular transcription profiles, evident through single cell sequencing[61]. In 3D, single cell sequencing of PSC-derived optic cup organoids also reveals a large degree of cell type heterogeneity[51]. In both cases, uniform exogenous signals give rise to non-uniform cellular responses, signifying that modulation of only extrinsic parameters is insufficient to achieve tight control over cell type emergence.

Whether performed in 2D or 3D, heterogeneity is ubiquitous in every type of germ lineage differentiation, revealing a more fundamental insight: pluripotent stem cells in the starting population have varying sensitivities to exogenous signals. These differences in sensitivity are important as they impact symmetry breaking events later in the differentiation. Nodes of cells that are more receptive to signals differentiate first then dictate the behavior of the rest of the tissue through mechanical or paracrine signaling, potentially explaining why discrete regions in 2D cardiac differentiations begin beating before confluent sheets of cells commence synchronous

beating, or why Rax expression, demarcating a budding optic cup structure, forms in discrete regions of neural organoids[62,63]. Stochastic symmetry breaking events impede reproducibility because they result in stochastic morphogen signaling patterns. Therefore, understanding what causes differences in sensitivity to signaling cues and how uncontrolled mosaic starting populations can lead to heterogeneity in differentiations is key to improving reproducibility and directing mesoscale tissue architecture to resemble macroscale organs.

Advances in Engineering Co-Emergence

Why are some cells differently sensitized to receiving signals than others, and how does this lead to mesoscale symmetry breaking? Noise in intracellular biochemical signaling and transduction, structural differences that lead to heightened receptor accessibility, or variations in epigenetic profiles ("naïve" versus "primed" pluripotency states or artifacts of the reprogramming process) are a few of the many possible explanations for how symmetry breaking occurs in a subset of cells. Morphological events that happen *in vivo*, however, are extremely robust, suggesting that the process of symmetry breaking can be tightly controlled to produce stable patterns that direct tissue morphogenesis and lead to viable organoid replicas. Perhaps the most convincing example of a robust and reproducible *in vitro* differentiation is the gastrulation-like model developed by Warmflash et al.[64], which has since been reproduced by multiple labs using different induced and embryonic pluripotent stem cell lines[65,66,67]. In this model, circular micropatterned colonies of PSCs exposed to BMP4 for 48 hours self-organize into radial patterns which express germ layer markers. The organization of these circular colonies is robust with a SOX2+ ectodermal region in the center, surrounded by rings of Brachyury (mesoderm) and SOX17-expressing cells

(endoderm), followed by a final trophectoderm-like CDX2+ outer ring. Phosphorylation of the downstream BMP4 effector SMAD1/5 is confined to the colony border, implying that sensitization to signal depends on spatial position in the colony and that the observed asymmetry between border and central cells is crucial to radial cell specification patterns. Because the pSMAD1/5 pre-pattern is the same in all micropatterned colonies, the cell specification pattern is robust. Applying this idea to 3D tissues and PSC-derived organoids suggests that robustness in patterning and differentiation could potentially be achieved by more precise regulation and replicability of symmetry breaking events.

Decades of developmental biology research have distilled fundamental theories about how stable differentiation and patterning emerge and propagate from symmetry breaking events. The principles of bistable switch, hysteresis, lateral inhibition, reaction diffusion, and many others have been used to computationally model patterning events observed *in vitro*[68] (summarized in Figure 2). At their core, these principles all require the interplay of positive and negative feedback elements to achieve stable pattern formation and subsequent differentiation. Advances in synthetic biology methods over the last twenty years have enabled scientists to create synthetic genetic circuits that test these ideas experimentally by endowing subpopulations of cells with different functions in the starting population. Most of these circuits have been integrated into bacteria for ease of culture and manipulation. The first of its kind was an *E. coli* network developed by Basu et al. to test principles of spatial patterning[69]. In this system, localized "sender" cells were programmed to form a gradient by secreting acyl-homoserine lactone (AHL). Receiver cells then responded to user-specified concentrations of AHL by expressing different fluorescent proteins, thereby creating ring-shaped patterns surrounding the sender cells, modelling a possible mechanism for cell fate specification. More recently, Matsuda et al. built a lateral inhibition circuit

in mammalian cells using Delta-Notch signaling, in which binding of Delta to endogenous Notch receptors in the neighboring cell resulted in translation of a Delta repressor. The circuit resulted in bifurcation of the cell population to cells only expressing Delta or only expressing Notch[70] in a salt-and-pepper pattern. The first example of a synthetic circuit to direct tissue morphogenesis in 3D was developed by Toda et al., based on SynNotch technology previously pioneered by Morsut and Roybal et al.[71,72]. Organization was achieved by modulating the starting cell stoichiometry of two different cell types engineered to recapitulate a classic feedback network. Binding of sender cells with receiver cells triggered ectopic transcription of the adhesion protein E-cadherin. The disparities between the degree of adhesion between the two cell types drive sorting events in which the receiver cells that express more E-cadherin cluster in the center and eventually become surrounded by the sender cells.

The idea behind a bottom-up approach to engineering organoid models is to transition from using synthetic circuits to understand principles of developmental patterning to using them to build organized multicellular tissue constructs. The majority of published synthetic circuits make use of a mosaic starting population (sender and receiver cells) combined with orthogonal feedback elements which do not interfere with endogenous cellular programs. Intrinsic organoid engineering approaches would also require a mosaic initial population but differ from the synthetic circuits in one key way: the feedback circuitry should leverage pre-existing interconnected endogenous signaling pathways. This could be accomplished through genome editing in order to direct a subpopulation of PSCs in the initial population to express specific lineage markers (transcription factor editing)[73], modulate patterning events *in vitro* (secreted factor editing)[65,67], or to achieve a mosaic starting population capable of symmetry breaking and developmental pre-patterning (structural protein editing)[74].

While closed-loop circuits capable of generating user-defined patterns have not yet been implemented in pluripotent stem cells or in organoid models, studies which synthetically modify expression of transcription factors in response to exogenous differentiation signals demonstrate one possibility of how to construct such a circuit. For example, double knockout of the mesendoderm transcription factors, Brachyury (T) and Eomesodermin (Eomes), in pluripotent stem cells leads the cells to differentiate to neuroectodermal lineages, despite the presence of exogenous cues for mesendodermal induction (Wnt, Nodal, and TGFβ)[73]. These types of studies provide a model for how to control initial cell type ratios. For example, T/EOMES double knockout cells could be combined with non-edited PSCs in a neuromesodermal progenitor differentiation and subsequently used to skew the proportion of cells that go to a neuronal versus mesodermal fate in neuromuscular organoids. Similar control could be exacted on cells already committed to a germ lineage; for example, knockout of the transcription factor Rax in a subpopulation of cells in ectodermal neurospheres could function to control the number of possible optic cup buds, which emerge from Rax-positive regions.

Using PSCs as mosaic elements in regulatory networks is desirable due to the relative ease with which genome editing can be performed on them compared with primary cells. Ease of genome editing varies between different primary cell lines, and is dependent on how well the cells are grown and passaged *in vitro*. Few groups have created organoids from direct genome editing in human primary cells, turning instead to primary cells from genetically modified mouse lines. However, a few notable examples include correction of the cystic fibrosis transmembrane conductor receptor (CFTR) locus in intestinal organoid models derived from primary cells of cystic fibrosis patients[75], and knock-in of fluorescent proteins to construct *KRT19* and *TUBB* reporter lines in human liver organoids from primary cells[76]. Epithelial cell types, especially those derived

from endodermal tissues, are amenable to *in vitro* culture and therefore represent the majority of cases of genome editing in primary cells. By contrast, examples of mesenchymal cell types that have undergone editing are notably lacking. With the exception of a study using genome-wide loss of function CRISPR screens on primary human T cells[77], cells from non-endodermal germ lineages are highly underrepresented, making PSCs an attractive option for engineering co-emergence in organoids.

Conclusions

PSC-derived organoid models have greatly expanded the scope of current organoid engineering possibilities by enabling the development of models from every type of germ lineage. PSC technology has furthered our understanding of the role of co-emergence in organogenesis, and how tissue structure and maturity depends on the interaction between different cell types during development. Engineering co-emergence, however, remains a challenging endeavor when confronted with the rigorous criteria for organoid models, which limit the list of modifiable engineering parameters. The high variability observed between independent biological replicates of organoid differentiation underscores the insufficiency of solely modifying exogenous parameters to enhance reproducibility. Instead, focus should be shifted on engineering systems of cells with controlled symmetry breaking and cell type emergence in the starting population for greater consistency and control. While these technologies are still nascent, progress in synthetic biological circuits which modulate expression of structural proteins, secreted morphogens, and transcription factors has demonstrated that controlled cell specification of multicellular patterning and organization is possible *in vitro*. Expansion of these methods to build complex functional 3D

tissues will require further translation to human PSCs and integration into endogenous gene regulatory networks.

Acknowledgements

We would like to thank Dr. Ana Silva for generating the illustrations, along with Dr. Kathryn Claiborn from the Gladstone Institutes Scientific Editing Department, and Emily Bulger, Oriane Matthys and Nicholas Elder for their help in providing ideas and revisions. Funding support includes: Emergent Behaviors of Integrated Cellular Systems (CBET 0939511) and the California Institute of Regenerative Medicine (LA1_C14-08015).

Figures

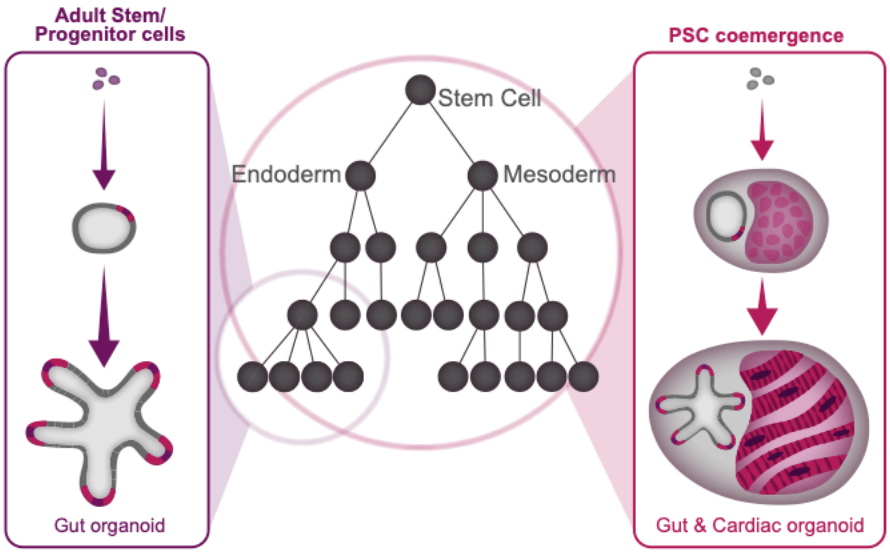

Figure 1.1. Differences between single and multi-lineage organoids.
Tissue-derived organoids, such as for the intestine, can be derived from a single clonal progenitor cell, demonstrating their inherent ability to organize and differentiate with few external cues. Pluripotent stem cells can produce complex multi-lineage organoids, containing cardiac and gut tissues, in which co-emergent tissues cooperatively secrete paracrine signals that enhance the development and maturity of their counterparts.

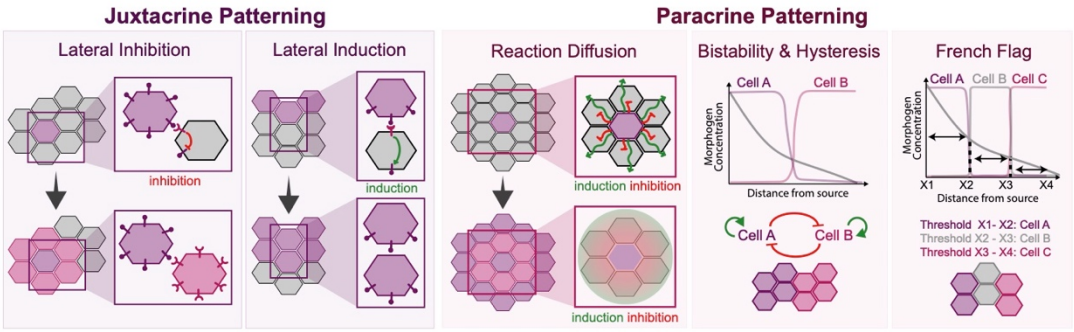

Figure 1.2. Common multicellular patterning mechanisms.
Multicellular patterning mechanisms rely on local physical and juxtracrine signals and/or longer-range paracrine signals. All patterning mechanisms require positive, negative, or both types of feedback circuits and are preceded by symmetry breaking events which distinguish at least one cell type as having properties that differ from the rest. In lateral inhibition and induction, emergence of a cell type that expresses more of a specific surface marker can cause population bifurcation by initiating a feedback response in surrounding cells. In reaction diffusion, emergence of a secretory cell type can cause formation of a spatial morphogen gradient. Two common models for cellular interpretation and reaction to morphogen gradients are bistability/hysteresis and the "French Flag" model. In a bistable system, cells use positive feedback loops to escalate their response to a signal and commit to a lineage. In the French Flag Model, cells use threshold-based sensing to respond to a signaling milieu, adopting cell fates based on where they fall in a range of morphogen concentrations.

References

1. Van De Wetering, M. *et al.* Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933–945 (2015).
2. Takahashi, T. Organoids for Drug Discovery and Personalized Medicine. *Annu. Rev. Pharmacol. Toxicol.* **59**, 447–462 (2019).
3. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
4. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993–997 (2006).
5. Sachs, N. *et al.* Long-term expanding human airway organoids for disease modeling. *EMBO J.* **38**, 1–20 (2019).
6. Bartfeld, S. *et al.* In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology* **148**, 126-136.e6 (2015).
7. Huch, M. *et al.* Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell* **160**, 299–312 (2015).
8. Loomans, C. J. M. *et al.* Expansion of Adult Human Pancreatic Tissue Yields Organoids Harboring Progenitor Cells with Endocrine Differentiation Potential. *Stem Cell Reports* **10**, 712–724 (2018).
9. Jung, P. *et al.* Isolation and in vitro expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–1227 (2011).
10. Karthaus, W. R. *et al.* Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell* **159**, 163–175 (2014).
11. Kessler, M. *et al.* The Notch and Wnt pathways regulate stemness and differentiation in

human fallopian tube organoids. *Nat. Commun.* **6**, (2015).

12. Turco, M. Y. *et al.* Trophoblast organoids as a model for maternal–fetal interactions during human placentation. *Nature* **564**, 263–281 (2018).
13. Dutta, D., Heo, I. & Clevers, H. Disease Modeling in Stem Cell-Derived 3D Organoid Systems. *Trends Mol. Med.* **xx**, 1–18 (2017).
14. Dahl-jensen, S. & Grapin-botton, A. The physics of organoids : a biophysical approach to understanding organogenesis. *Development* **144**, 946–951 (2017).
15. Si-tayeb, K. *et al.* Highly Efficient Generation of Human Hepatocyte–like Cells from Induced Pluripotent Stem Cells. *Hepatology* **51**, 297–305 (2011).
16. Quadrato, G. *et al.* Cell diversity and network dynamics in photosensitive human brain organoids. *Nat. Publ. Gr.* (2017). doi:10.1038/nature22047
17. Phipson, B. *et al.* Evaluation of variability in human kidney organoids. *Nat Methods* **16**, 79–87 (2019).
18. McCauley, H. A. & Wells, J. M. Pluripotent stem cell-derived organoids: Using principles of developmental biology to grow human tissues in a dish. *Dev.* **144**, 958–962 (2017).
19. Takebe, T. & Wells, J. M. Organoids by design. *Science (80-.).* **364**, 956–959 (2019).
20. Combes, A. N., Zappia, L., Er, P. X., Oshlack, A. & Little, M. H. Single-cell analysis reveals congruence between kidney organoids and human fetal kidney. *Genome Med.* **11**, 1–15 (2019).
21. Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. **112**, 1–6 (2015).
22. Dianat, N. *et al.* Generation of functional cholangiocyte-like cells from human pluripotent stem cells and HepaRG cells. *Hepatology* **60**, 700–714 (2014).

23. Ogawa, M. *et al.* Directed differentiation of cholangiocytes from human pluripotent stem cells. *Nat. Biotechnol.* **33**, 853–861 (2015).

24. Williams, C., Quinn, K. P., Georgakoudi, I. & Black, L. D. Young developmental age cardiac extracellular matrix promotes the expansion of neonatal cardiomyocytes in vitro. *Acta Biomater.* **10**, 194–204 (2014).

25. Kauss, M. A. *et al.* Cardiac Cell-Derived Matrices Impart Age-Specific Functional Properties to Human Cardiomyocytes. *bioRxiv Prepr.* (2020). doi:https://doi.org/10.1101/2020.07.31.231480

26. Ozcebe, S. G., Bahcecioglu, G., Yue, X. S. & Zorlutuna, P. Effect of Cellular and ECM Aging on Human iPSC-derived Cardiomyocyte Performance, Maturity and Senescence. *bioRxiv* (2020).

27. Zhang, Y. *Multi-scale Extracellular Matrix Mechanics and Mechanobiology*.

28. Frangogiannis, N. G. The extracellular matrix in myocardial injury , repair , and remodeling. **127**, (2017).

29. Ieda, M. *et al.* Cardiac Fibroblasts Regulate Myocardial Proliferation through β1 Integrin Signaling. *Dev. Cell* **16**, 233–244 (2009).

30. Hookway, T. A. *et al.* Bi-directional Impacts of Heterotypic Interactions in Engineered 3D Human Cardiac Microtissues Revealed by Single-Cell RNA-Sequencing and Functional Analysis. *bioRxiv* (2020). doi:10.1101/2020.07.06.190504

31. Kelava, I. & Lancaster, M. A. Stem Cell Models of Human Brain Development. *Stem Cell* **18**, 736–748 (2016).

32. Lewis, A., Keshara, R., Kim, Y. H. & Grapin-Botton, A. Self-organization of organoids from endoderm-derived cells. *J. Mol. Med.* (2020). doi:10.1007/s00109-020-02010-w

33. Arora, N. *et al.* A process engineering approach to increase organoid yield. *Development* **5**, 1128–1136 (2017).

34. Chanson, L. *et al.* Self-organization is a dynamic and lineage-intrinsic property of mammary epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3264–3269 (2011).

35. Karbassi, E. *et al.* Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. *Nat. Rev. Cardiol.* **17**, 341–359 (2020).

36. Guo, L. *et al.* Refining the human iPSC-cardiomyocyte arrhythmic risk assessment model. *Toxicol. Sci.* **136**, 581–594 (2013).

37. Marikawa, Y., Chen, H. R., Menor, M., Deng, Y. & Alarcon, V. B. Exposure-based assessment of chemical teratogenicity using morphogenetic aggregates of human embryonic stem cells. *Reprod. Toxicol.* **91**, 74–91 (2020).

38. Safura, S. S. *et al.* Evaluating the Effects of Bisphenols F and S with Respect to Bisphenol A on Primordial Germ Cell Migration in Zebrafish (Danio rerio) Embryos Using Immunofluorescence Microscopy How has open access to Fisher Digital Publications benefited you ? 69–77 (2019).

39. Takebe, T. *et al.* Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* **499**, 481–484 (2013).

40. Matthys, O. B., Silva, A. C. & Mcdevitt, T. C. Engineering human organoid development ex vivo — challenges and opportunities. *Curr. Opin. Biomed. Eng.* **13**, 160–167 (2020).

41. Silva, A. C. *et al.* Developmental co-emergence of cardiac and gut tissues modeled by human iPSC-derived organoids. *bioRxiv* 2020.04.30.071472 (2020). doi:10.1101/2020.04.30.071472

42. Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature*

469, 415–418 (2011).

43. Rifes, P. *et al.* Redefining the role of ectoderm in somitogenesis: A player in the formation of the fibronectin matrix of presomitic mesoderm. *Development* **134**, 3155–3165 (2007).

44. Nijmeijer, R. M., Leeuwis, J. W., DeLisio, A., Mummery, C. L. & Chuva de Sousa Lopes, S. M. Visceral endoderm induces specification of cardiomyocytes in mice. *Stem Cell Res.* **3**, 170–178 (2009).

45. Fair, J. H. *et al.* Induction of hepatic differentiation in embryonic stem cells by co-culture with embryonic cardiac mesoderm. *Surgery* **134**, 189–196 (2003).

46. Bakooshli, M. A. *et al.* A 3d culture model of innervated human skeletal muscle enables studies of the adult neuromuscular junction. *Elife* **8**, 1–29 (2019).

47. Steinbeck, J. A. *et al.* Functional Connectivity under Optogenetic Control Allows Modeling of Human Neuromuscular Disease. *Cell Stem Cell* **18**, 134–143 (2016).

48. Machado, C. B. *et al.* In Vitro Modeling of Nerve–Muscle Connectivity in a Compartmentalized Tissue Culture Device. *Adv. Biosyst.* **3**, 1–14 (2019).

49. Faustino Martins, J. M. *et al.* Self-Organizing 3D Human Trunk Neuromuscular Organoids. *Cell Stem Cell* **26**, 172-186.e6 (2020).

50. Múnera, J. O. *et al.* Differentiation of Human Pluripotent Stem Cells into Colonic Organoids via Transient Activation of BMP Signaling. *Cell Stem Cell* **21**, 51-64.e6 (2017).

51. Sridhar, A. *et al.* Single-Cell Transcriptomic Comparison of Human Fetal Retina, hPSC-Derived Retinal Organoids, and Long-Term Retinal Cultures. *Cell Rep.* **30**, 1644-1659.e4 (2020).

52. Kanton, S. *et al. Organoid single-cell genomic atlas uncovers human-specific features of brain development*. *Nature* **574**, (2019).

53. Rossi, G., Manfrin, A. & Lutolf, M. P. Progress and potential in organoid research. *Nat. Rev. Genet.* **19**, 671–687 (2018).

54. Libby, A. R. G. *et al.* Axial Elongation of Caudalized Human Pluripotent Stem Cell Organoids Mimics Neural Tube Development. *bioRxiv* (2020).

55. Butts, J. C. *et al.* Differentiation of V2a interneurons from human pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4969–4974 (2017).

56. Sakaguchi, H. *et al.* Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue. *Nat. Commun.* **6**, (2015).

57. Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K. & Sasai, Y. Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep.* **10**, 537–550 (2015).

58. Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).

59. Kadoshima, T. *et al.* Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20284–20289 (2013).

60. Lian, X. *et al.* Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proc. Natl. Acad. Sci. U. S. A.* **109**, (2012).

61. Friedman, C. E. *et al.* Single-Cell Transcriptomic Analysis of Cardiac Differentiation from Human PSCs Reveals HOPX-Dependent Cardiomyocyte Maturation. *Cell Stem Cell* **23**, 586-598.e8 (2018).

62. Mummery, C. L. *et al.* Differentiation of Human ES and iPS Cells to Cardiomyocytes: A

Methods Overview. *Circ. Res.* **111**, 344–358 (2012).

63. Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–58 (2011).

64. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* **11**, 847–854 (2014).

65. Etoc, F. *et al.* A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Dev. Cell* **39**, 302–315 (2016).

66. Tewary, M. *et al.* A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning. *Development* dev.149658 (2017). doi:10.1242/dev.149658

67. Martyn, I., Brivanlou, A. H. & Siggia, E. D. A wave of WNT signalling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. *Development* dev.172791 (2019). doi:10.1242/dev.172791

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

69. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005).

70. Matsuda, M., Koga, M., Woltjen, K., Nishida, E. & Ebisuya, M. Synthetic lateral inhibition governs cell-type bifurcation with robust ratios. *Nat. Commun.* **6**, 1–12 (2015).

71. Morsut, L. *et al.* Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164**, 780–791 (2016).

72. Toda, S., Blauch, L. R., Tang, S. K. Y., Morsut, L. & Lim, W. A. Programming self-

organizing multicellular structures with synthetic cell-cell signaling. *Science (80-.).* **361**, 156–162 (2018).

73. Tosic, J. *et al.* Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat. Cell Biol.* **21**, 1518–1531 (2019).

74. Libby, A. R. G. *et al.* Spatiotemporal mosaic patterning of pluripotent stem cells using CRISPR interference. *bioRxiv* 1–23 (2018). doi:https://doi.org/10.1101/252189

75. Schwank, G. *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658 (2013).

76. Artegiani, B. *et al.* Fast and efficient generation of knock-in human organoids using homology-independent CRISPR–Cas9 precision genome editing. *Nat. Cell Biol.* **22**, 321–331 (2020).

77. Shifrut, E. *et al.* Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* **175**, 1958-1971.e15 (2018).

Chapter 2. Silencing of E-cadherin in induced human pluripotent stem cells promotes extraembryonic fates accompanying multilineage differentiation

Abstract

In embryonic development, symmetry breaking events and the mechanical milieus in which they occur coordinate the specification of separate cell lineages. Here, we use 3D aggregates of human pluripotent stem cells (hPSCs) encapsulated in alginate microbeads to model the early blastocyst prior to zona pellucida hatching. We demonstrate that 3D confinement combined with modulation of cell-cell adhesions is sufficient to drive differentiation and collective migration reminiscent of the pre-implantation embryo. Knockdown of the cell adhesion protein CDH1 in encapsulated hPSC aggregates resulted in protrusion morphologies and emergence of extra-embryonic lineages, whereas unencapsulated CDH1(-) aggregates displayed organized radial delamination and mesendoderm specification bias. Transcriptomic similarities between single-cell RNA-sequencing data of early human embryos and encapsulated CDH1(-) aggregates establishes this *in vitro* system as a competent surrogate for studying early embryonic fate decisions and highlights the relationship between cell-cell adhesions and the mechanical microenvironment in directing cell fate and behavior.

Highlights

- Generation of embryonic scale 3D morphogenesis using hydrogel encapsulation
- Manipulating adhesion triggers emergence of specific morphologies and cell fates
- Acquisition of germ layer cell fates mimics early human embryonic diversity

Introduction

In embryonic development, symmetry breaking events characterize the repeated emergence of different cell fates, which self-assemble into complex embryonic tissues via coordinated changes in cell adhesion. To this end, key three-dimensional states in which specific adhesions are relevant have been examined in work with model organisms and limited studies of human embryos. During compaction, the embryo forms a cystic cavity adjoined by the inner cell mass: a nonpolar group of cells that highly express the cell-cell adhesion molecule E-cadherin (CDH1) (1–5). The embryo then undergoes zona pellucida hatching, which involves epithelial sheet protrusion and subsequent expulsion of the early epiblast from its casing. Finally, during gastrulation, CDH1 is sharply downregulated and replaced by N-cadherin (CDH2), resulting in delamination and invasion of primitive streak into the space between the epiblast and developing yolk sac . Dysregulation of symmetry breaking during these key transitional periods often results in congenital disease or embryonic lethality, highlighting the importance of attaining a more mechanistic understanding of morphogenesis in the earliest stages of development (6,7). Despite extensive research into how intercellular adhesion molecules regulate the movement of cell populations, the connection between changes in intercellular adhesions and lineage fate decisions has remained poorly characterized. Furthermore, studies of human embryos *ex vivo* have been limited, and the dynamics of human development have been difficult to interrogate *in vivo* due to the physical restrictions, optical opacity, and complex signaling milieu inherent to the developing embryo. Therefore, to study how human specific symmetry breaking events direct morphogenesis, it is essential to establish *in vitro* human systems that promote the coincident development of analogous heterogeneous cell 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 (8–11). However, these models manifest inherently stochastic and un-reproducible differentiation (12,13), limiting our understanding of the mechanisms that control and coordinate human morphogenesis. Therefore, new approaches to reliably control the emergence and organization of multiple germ cell types would greatly advance tissue modeling and organ developmental studies.

Recent studies have demonstrated that mechanical signals are important regulators of cell fate, and have implicated CDH1 not only as a regulator of cell-cell adhesions and sorting, but as a common transducer of external mechanical forces which contribute to mesendoderm versus ectoderm fate decisions (14,15). In this study, we interrogate the interplay between mechanics and cell-cell adhesions in directing symmetry breaking events in 3D. Specifically, we employ hydrogel encapsulation to mimic the physically confined environment of the zona pellucida and induce heterogenous changes in adhesion through mosaic knockdown of CDH1 using CRISPRi in small cell aggregates. We show that hiPSC aggregates undergo changes in 3D structure as well as population emergence reminiscent of the three germ lineages of pre-implantation embryos. Additionally, the combination of CDH1 knockdown and encapsulation leads to specific emergence of extraembryonic populations in aggregates after 6 days of culture, highlighting potential regulation of trophoblast development triggered by the physical microenvironment. Recent studies demonstrating that mechanical signals are important regulators of cell fate have implicated CDH1 not only as a regulator of cell-cell adhesions and sorting, but as a common transducer of external mechanical forces that contribute to mesendoderm vs. ectoderm fate decisions (14–16). In this study, we interrogate the interplay between mechanics and cell-cell adhesions in directing symmetry breaking events in 3D pluripotent human cell aggregates. Specifically, we employ

hydrogel encapsulation to mimic the physically confined environment of the zona pellucida and induce heterogenous changes in adhesion through mosaic knockdown of CDH1 using CRISPRi in hiPSC aggregates. We show that the aggregates undergo changes in 3D structure as well as population emergence of the three germ lineages similar to pre-implantation embryos. Additionally, the specific combination of CDH1 knockdown and encapsulation leads to emergence of extraembryonic populations in aggregates after 6 days of culture, highlighting potential regulation of trophoblast development by integrated microenvironmental factors.

Results

Loss of CDH1 Promotes Protrusion Morphology when Encapsulated

To mimic the size-scale and environment of the pre-implantation blastocyst, 50-cell human hiPSC aggregates were encapsulated in 1.5% alginate beads mixed with laminin and cultured for 6 days. Alginate encapsulation served as a proxy for zona pellucida encapsulation by providing both a physical barrier (unfunctionalized alginate) and a signaling competent ECM (laminin) (Figure 1A). To model CDH1 loss during gastrulation, CDH1 knockdown was induced in either 0%, 25% or 100% of the aggregate population using a previously established CRISPRi system, where knockdown is triggered by doxycycline inducible production of dead Cas9-KRAB (14,17). Post aggregation, both encapsulated and unencapsulated CDH1(+), CDH1(-), and mosaic aggregates displayed polarized behavior, creating a single layer of cells surrounding a cystic cavity in the center (Figure 1B). Over time, a bilayer of cells in CDH1(+) aggregates assembled at the surface of the growing aggregate and the central cyst remained intact, indicating that the polarity of human iPSCs in monolayer culture is initially maintained in suspension culture.

At day 3 of CDH1 knockdown, all aggregates consisted of bilayers surrounding a cystic cavity. In mixed aggregates, CDH1(+) and CDH1(-) cells physically sorted away from one another. The CDH1(-) cells clustered within the outer layer of the cystic aggregates, recapitulating the same spatial organization of cells engineered to express either high or low levels of CDH1 (18). Taken together, our results suggest that modulation of cell-cell adhesions alone is sufficient to induce radial organization of cell subtypes in hPSC aggregates (Figure 1C). Interestingly, while the CDH1(+) aggregates and the mixed aggregates retained a spherical shape, the CDH1(-) aggregates formed oblong protrusions extending away from the main body of the aggregate that reflected collective migration characteristics (Figure 1B,D). These extensions displayed similar morphologies to those present during zona pellucida hatching (1,19,20), in that extensions remained epithelial and no single cells were observed to migrate away from the aggregate into the alginate hydrogel material.

Encapsulation Affects the Morphology of CDH1(-) Aggregates

To determine whether the observed extensions were a result of CDH1 knockdown or a response to encapsulation itself, human iPSC aggregates (50 cells each) were cultured in round bottom wells for 6 days without added extracellular matrix or alginate encapsulation (Figure 2A,B). Similar to encapsulated aggregates, cystic cavities formed in unencapsulated aggregates by day 1 persisted throughout culture. In mixed aggregates after 3 days, CDH1(+) cells segregated from CDH1(-) cells, mirroring the behavior seen in encapsulated mixed aggregates (Figure 2B, arrows mark segregated CDH1(-) cells). By day 5, CDH1(-) aggregates underwent a sheet delamination event characterized by an outer layer of mesenchymal-like cells peeling away from an inner epithelial core, displaying a cystic cavity (Figure 2C,D). Moreover, in the CDH1(-)

condition, the entire aggregate lost OCT4 expression (Figure S1), the outer layer displayed EOMES expression, and the inner layer displayed SOX2 expression (Figure 2D), indicating a loss of pluripotency and subsequent divergence of cell populations into mesendoderm and ectoderm lineages, respectively. This lineage emergence is in contrast to the CDH1(+) aggregates which remained largely pluripotent (OCT4+SOX2+) or ectodermal (SOX2+) (FigureS1). These results suggest that silencing of CDH1 accelerates the loss of OCT4 in human iPSCs while still allowing for emergence of the mesendoderm and ectoderm lineages.

Emergence of Extraembryonic-like Cell Fates in Encapsulated Aggregates

Due to the marked differences in the morphology and cell type emergence between encapsulated/unencapsulated CDH1(+)/CDH1(-) aggregates, we used single-cell RNA sequencing to interrogate the diversity of cellular populations generated within the differentiating aggregates. The transcriptomes of encapsulated and unencapsulated aggregates at days 1, 3, and 6 post-aggregation were examined between CDH1(+), CDH1(-), and mixed aggregates (Figure 3A,B). The resulting clustered data set of single cell transcriptomes represented 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 3C), whereas after 3 and 5 days aggregates transitioned through multiple lineage states, eventually clustering into populations expressing markers of mesendoderm, ectoderm, endoderm, and mesoderm (Figure 3C). Interestingly, clusters 8 and 10 displayed markers of the extraembryonic trophectoderm (CDX2, GATA3, CDH3, HAND1) (Figure 3C). Gene Ontology analysis of clusters 8 and 10 revealed gene networks associated with placental development and reproductive system development (Figure 3D). Furthermore, these clusters were largely populated by CDH1(-) cells that were encapsulated

(Figure 3E), and contained highly expressed genes in pathways associated with trophectoderm development, such as Hippo signaling and non-canonical Wnt signaling (2,21–23) (Figure 3F).

To computationally investigate whether this extraembryonic-like population (clusters 8 and 10) diverged from the other three germ lineages early on in the spontaneous aggregate differentiation, we examined cellular trajectories indicating transitions between states. Cellular trajectory reconstruction revealed that the differentiating population of encapsulated and non-encapsulated aggregates as a whole bifurcated after day 1, with one branch transitioning through primitive streak and then mesoderm lineage states, while a separate trajectory proceeded toward an extra-embryonic lineage state (Figure S1A). The mesendoderm, endoderm, and ectoderm germ lineages clustered separately from the lineage trajectories of pluripotency to mesoderm or to extraembryonic lineages (Figure S1A), indicating that the transition away from pluripotency may occur on a faster time scale that was not captured by sampling at days 1,3, and 5, and therefore preventing computational reconstruction of the complete lineage trajectory from pluripotency. This rapid time scale of differentiation is consistent with previous reports of fast switch-like exit from pluripotency (24), meaning that in order to capture lineage trajectories of cells as they lose pluripotency, more frequent sampling must occur. Overall, the early bifurcation between the germ lineages specific to the epiblast (mesoderm, endoderm, ectoderm) and the extraembryonic-like lineage in pseudo-time resembles the divergent population emergence at compaction of the *in vivo* embryo (Figure S1B), suggesting that this 3D human iPSC-based system can be used to model early developmental lineage diversification.

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

To interrogate whether the emergence of lineages in our encapsulated and unencapsulated cell aggregates resembled that of the early human embryo (Figure 4A), our *in vitro* data set was directly contrasted with three previously published single cell transcriptome data sets from preimplantation embryos (25–27). *In vitro* aggregate cell transcriptomes were co-clustered with single-cell RNA sequencing of preimplantation embryo stages from oocyte through late blastocyst from Yan *et al*. (Figure 4B). Embryonic stages from the oocyte to 4-cells clustered far from our *in vitro* aggregates (Figure 4B, red arrows), with only 33% clustering overlap between 4-cell embryos and 3D hiPSC aggregates, reflecting the lack of totipotent cell-like transcriptional profiles *in vitro*. Cells from the morula and late blastocyst co-clustered with the embryonic-like fractions of *in vitro* aggregates (Figure 4B, blue arrow heads) including in regions of primitive streak and germ-lineage restricted cells, demonstrating that *in vitro* aggregates reflect similar lineage transitions to those which occur *in vivo.* Interestingly, cells from 8-cell and morula stages had similar transcriptomes to *in vitro* extraembryonic-like cells from the encapsulated aggregates, marked by high CDX2 and HAND1 expression, possibly reflecting the *in vivo* transition from totipotency towards lineage-restricted embryonic and extra-embryonic cell types. Overall, co-clustering with the Yan *et al*. dataset indicated that *in vitro* aggregate cell transcriptomes recapitulate aspects of pre-implantation embryos from approximately the 8-cell stage through late blastocyst, with the majority of cells reflecting a late-blastocyst-like phenotype.

To confirm the approximate staging of *in vitro* aggregates with the late blastocyst, the transcriptional profiles were subsequently aligned with two independent single cell sequencing data sets of late blastocyst embryos (Figure 4C,D). In both Blakeley *et al*. and Petropoulous *et al*.,

cells from the trophectoderm clustered close to the extraembryonic-like fraction, confirming that *in vitro* aggregates contain a trophectoderm-like population of cells. There was 100% alignment overlap between cells from the primitive endoderm and epiblast in the Blakeley *et al*. data set and the embryonic-like fraction of *in vitro* aggregates, with epiblast cells aligning better to the pluripotent and transitioning fractions, while primitive endoderm cells aligned more with *in vitro* cells in the differentiating compartment (Figure 4C). Additionally, 18.8% of the trophectoderm cell transcriptomes clustered with the extra-embryonic like compartment of cells from the *in vitro* aggregates (Figure 4C, blue arrow heads). Alignment with primitive endoderm and epiblast cells was poorer in the Petropoulous *et al*. data set, possibly reflecting differences in the developmental stage of the *in vitro* and *in vivo* cells, where *in vitro* iPSCs presumably reflect a population of cells in a primed pluripotency state whereas the *in vivo* population of cells in the epiblast would represent a more naïve state of pluripotency. Despite this reduced overlap between populations, 36.1% of primitive endoderm cells and 28.9% of epiblast cells still aligned with the differentiating compartment of the *in vitro* aggregates, again suggesting that *in vitro* aggregates recapitulate the developmental trajectories of cells in the developing embryo (Figure 4D). Overall, the transcriptomic similarities between our *in vitro* results and the three human embryo data sets suggest that hydrogel encapsulation coupled with CDH1 knockdown enhances activity of gene regulatory networks that mimic developmental processes in pre-implantation embryos.

Discussion

The emergence and coincident organization of multiple lineages in the developing embryo is essential for the formation of the overall body plan and the development of functional tissues.

However, the process by which cell-cell adhesions, ECM/mechanical signals, and morphogen cues coordinate to direct symmetry breaking and subsequent tissue formation *in vivo* is not well understood. In this study, we demonstrate that culture of individual encapsulated human iPSC aggregates within hydrogel environments yields aspects of early human development and lineage emergence, even in the absence of adding any exogenous morphogens. Additionally, attenuation of CDH1 within encapsulated aggregates enabled the emergence of an extraembryonic-like population of cells reminiscent of trophoblast cells. These results demonstrate that the physical structure of iPSC culture is sufficient to initiate symmetry breaking, and that adhesion changes in the developing embryo may directly control lineage decisions, versus simply emerging as a result of fate commitment. Furthermore, this study suggests that the physical microenvironment plays a valuable role in the coordination of lineage fate decisions, highlighting a possible regulatory mechanism of both adhesions and mechanical cues employed by the early embryo to control population emergence.

As the embryo develops, it undergoes a series of cell polarity changes, where cells are segregated into an apical and basal domain, that dictate developmental transitions. For example, compaction relies on the polarization of the trophoblasts to allow for cavity formation (2,23), while gastrulation relies on the loss of polarity as invaginating cells begin their migration across the embryo 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 (2,23), whereas gastrulation relies on the loss of polarity as invaginating cells begin their migration across the embryo (21,28). This previous work in combination with our studies suggest that the loss of adhesions in aggregates due to CDH1 knockdown triggers a signaling network that regulates polarity-dependent lineage transitions.

The mechanical microenvironment has been shown to regulate lineage fate emergence in pluripotent stems cells *in vitro,* suggesting that a similar mechanism may regulate embryonic cell fate (15,29,30). The emergence of a trophoblast-like population, more so in encapsulated CDH1 knockdown aggregates, suggests that the physical microenvironment provided by encapsulation in tandem with changing surface adhesion properties provides necessary cues for extraembryonic fate specification. Trophoblast development is primarily regulated by Hippo signaling, a well-established mechanically responsive signaling pathway (22,31). 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 a similar mechanism *in vitro* may regulate embryonic cell fate (15,29,30). The emergence of a trophoblast-like population, particularly in encapsulated CDH1 knockdown aggregates, suggests that the physical microenvironment provided by encapsulation provides necessary cues for extraembryonic fate specification. Interestingly, trophoblast development is primarily regulated by Hippo signaling, a mechanically responsive signaling pathway (22,31). 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.

Overall, this study demonstrates how physical environmental parameters and intercellular adhesive properties of iPSCs can cooperatively impact developmental processes that regulate 3D morphogenesis and lineage emergence in the early embryo. Our results demonstrate that 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 that robustly regulate development. Ultimately, this work provides insights that are relevant to stem cell biology and human embryonic development, facilitating routes to improve

the robustness of *in vitro* differentiation methods as well as potentially illuminating strategies relevant for therapeutics, such as *in vitro* fertilization.

Methods

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 Allen Institute WTC11-LaminB cell line (AICS-0013 cl.210). All cell lines were karyotyped by Cell Line Genetics and by qPCR 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 mTeSR™-1 medium (STEMCELL Technologies) (32). Cells were passaged by dissociation with Accutase (STEM CELL Technologies) and re-seeded in mTeSR™-1 medium supplemented with the small molecule Rho-associated coiled-coil kinase (ROCK) inhibitor Y-276932 (10 μM; Selleckchem)(33) at a seeding density of 12,000 cell per cm^2.

The CRISPRi CDH1 cell line was generated via insertion of a previously published Tet-ON system inserted into the AAVS1 locus via TALENS (Addgene plasmid # 73498; Mandegar et al., 2016) into the AICS-0013 cl.210 parent line. The previously published guide RNA sequence used to target CDH1 (GCAGTTCCGACGCCACTGAG) was cloned into the gRNA expression vector (addgene plasmid # 73501) using a BsmBI restriction enzyme cloning strategy described in Mandegar et al. The guide RNA vector was then electroporated into the parent line containing the

CRISPRi system using the Human Stem Cell Nucleofector Kit 1 solution with the Amaxa nucleofector 2b device (Lonza). Nucleofected cells were then seeded into a 6-well plate in mTeSR™-1 supplemented with Y-27632 (10 μM) and underwent blasticidin (ThermoFisher Scientific; 10 μg/ml) selection for 6 days. Surviving cells were then colony picked, expanded, and knockdown tested via qPCR after 5 days CRISPRi induction via addition of doxycycline (2μM) to the culture media (Figure S2).

Encapsulated Mixed Aggregate Generation

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 (34–36). Dissociated iPSC cultures were resuspended in mTeSR™-1 supplemented with Y-27632(10μM), 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 MVG alginate (Pronova) mixed with Laminin from Engelbreth-Holm-Swarm murine sarcoma (6μg/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 400μm 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 mTeSR™-1 medium (STEMCELL Technologies). Encapsulated aggregates were allowed to recover for 24 hours in mTeSR™-1 medium (STEMCELL

Technologies) in rotary suspension, then fed daily with mTeSR™-1 medium supplemented with doxycycline (DOX)(2μM) to induce CDH1 knockdown.

Un-encapsulated Mixed Aggregate Generation

Unencapsulated 50 cell aggregates were created by dissociation of human iPSCs with Accutase (STEM CELL Technologies) and reseeding in mTeSR™-1 medium (STEMCELL Technologies) supplemented with Y-276932 (10μM; Selleckchem) into 96-well non-adherent round bottom plates, with a total of ~50 cells were seeded per well. After 18 hours of aggregate formation, Y-276932 was removed from the media and DOX(2μM) was supplemented into the mTeSR™-1 medium to induce CDH1 knockdown. Aggregates were then fed daily with mTeSR™-1 medium supplemented with DOX (2μM).

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 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 95ºC 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.

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 diameter imaging capillaries and subsequently imaged on the Zeiss Z.1 Light sheet Microscope equipped with a PCO.edge SCMOS camera.

Single Cell RNA Sequencing Sample and Library Preparation

Multiple organoid samples were combined and processed together using the MULTI-Seq technology (37). Organoids were singularized using Accutase (STEMCELL Technologies) and washed with cold PBS. Cells were resuspended in PBS with lipid-modified Anchor and Barcode oligonucleotides (kindly provided by Dr. 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 (37). 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 (37).

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 (38). First cells with fewer than 200 detected genes, fewer than 1,000 total detected transcripts, or greater than 10% mitochondrial gene expression were removed. 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.

Data Acquisition, Processing, and Merging

Human single cell RNA sequencing data sets from previously published papers (25–27) 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 (39) for GEO66507 and GSE36552 and using TopHat2 (40) for E-MTAB-3929 prior to generating a counts matrix using the FeatureCounts software (41). 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 (38). Batch correction was performed using the Harmony algorithm (42).

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. (43) 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 (44).

Acknowledgements

We would like to thank the Gladstone Light Microscopy and Histology Core, the Gladstone Stem Cell Core and The Gladstone Flowcytometry Core for their support and experimental expertise. Additionally, we would like to specifically thank Dr. Vaishaali Natarajan for her expertise and assistance with encapsulation and Dr. Zev Gartner and his lab for reagents and expertise in the MULTI-seq assay.

This work was supported by the National Science Foundation Center: Emergent Behaviors of Integrated Cellular Systems (CBET 0939511) and the California Institute of Regenerative Medicine (LA1_C14-08015). ARGL was supported by the National Heart, Lung, and Blood Institute (1F31HL140907-01).

Figures

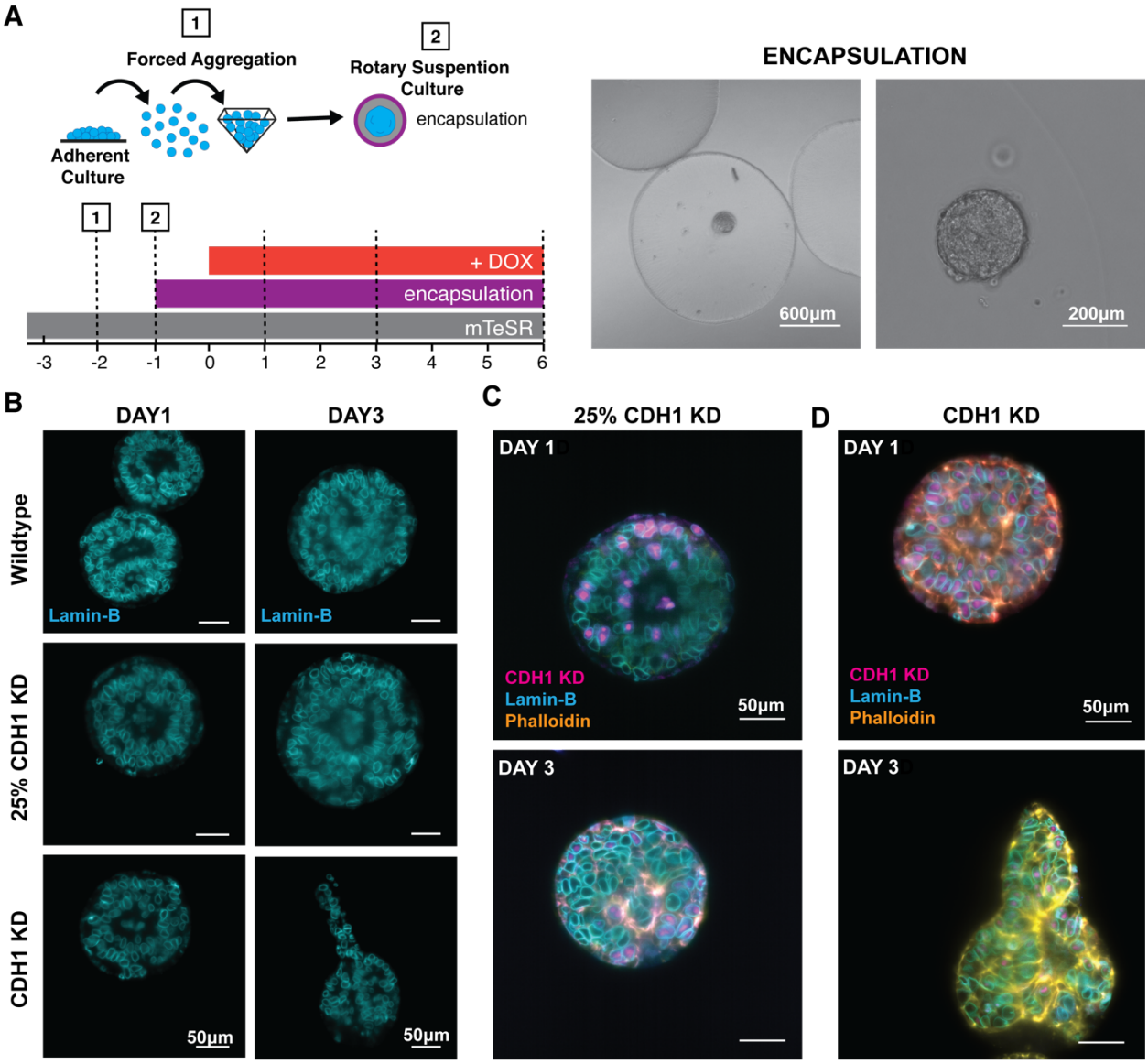

Figure 2.1: Loss of CDH1 promotes protrusion morphology.
(A) A schematic of the experimental set up of encapsulation (left) and example images of encapsulated 50 cell human iPSC aggregates in alginate gel 20 minutes post encapsulation. **(B)** Optical sections of aggregates demonstrating evolution of morphologies on day 1 and day 3. **(C)** Optical section of mixed aggregates with 25% CDH1 knockdown. **(D)** Optical sections of aggregates lacking CDH1.

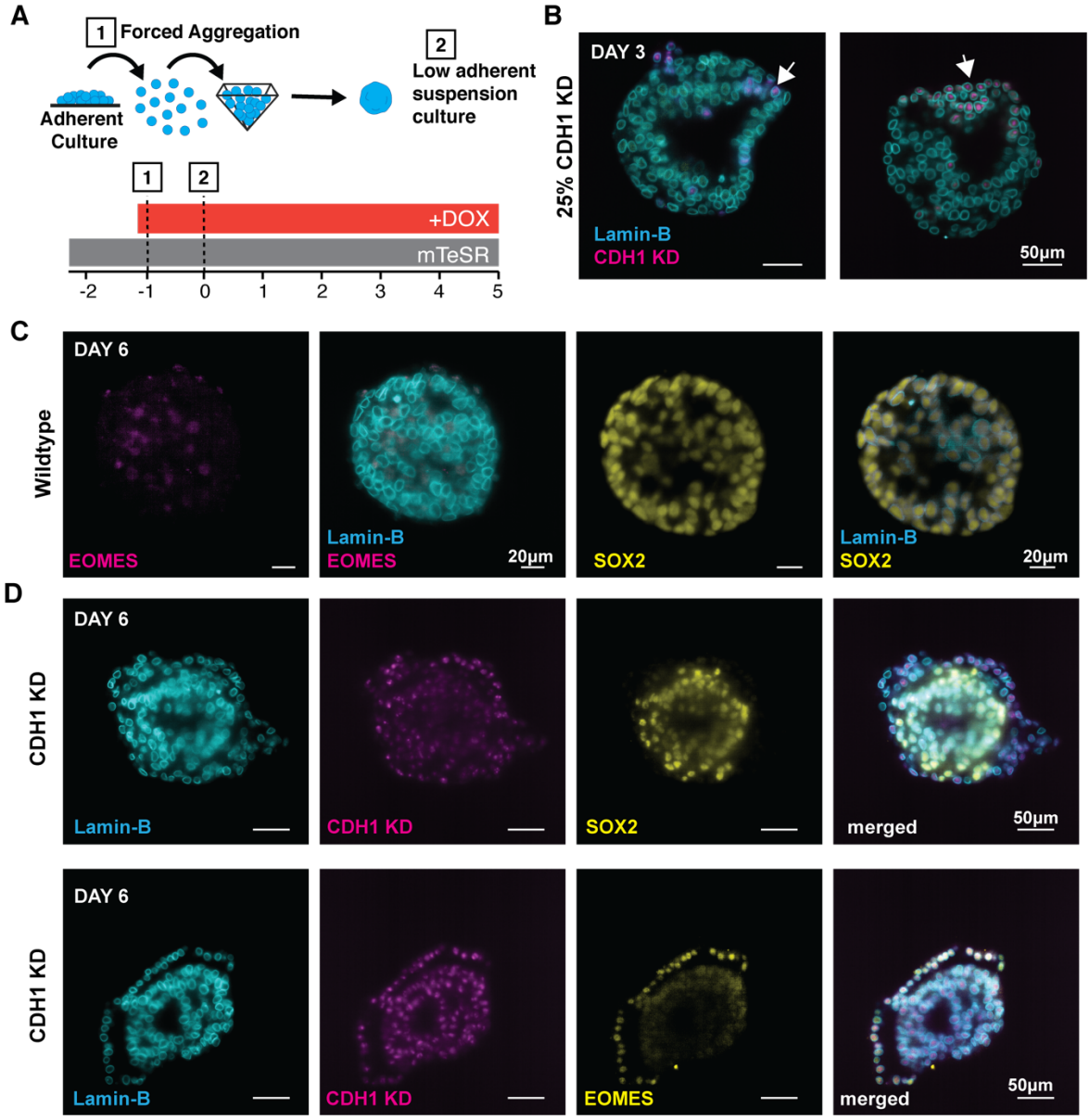

Figure 2.2: Encapsulation affects the morphology of CDH1(-) aggregates.
(A) A schematic of the experimental set up. **(B)** Optical sections of mixed aggregates with 25% CDH1 knockdown demonstrating maintenance of cysts and segregation of CDH1(-) cells by day 3. **(C)** Optical sections of wildtype unencapsulated aggregates stained for EMOES and SOX2. **(D)** Optical sections of aggregates lacking CDH1 demonstrating changes in morphology and lineage fate spatial segregation.

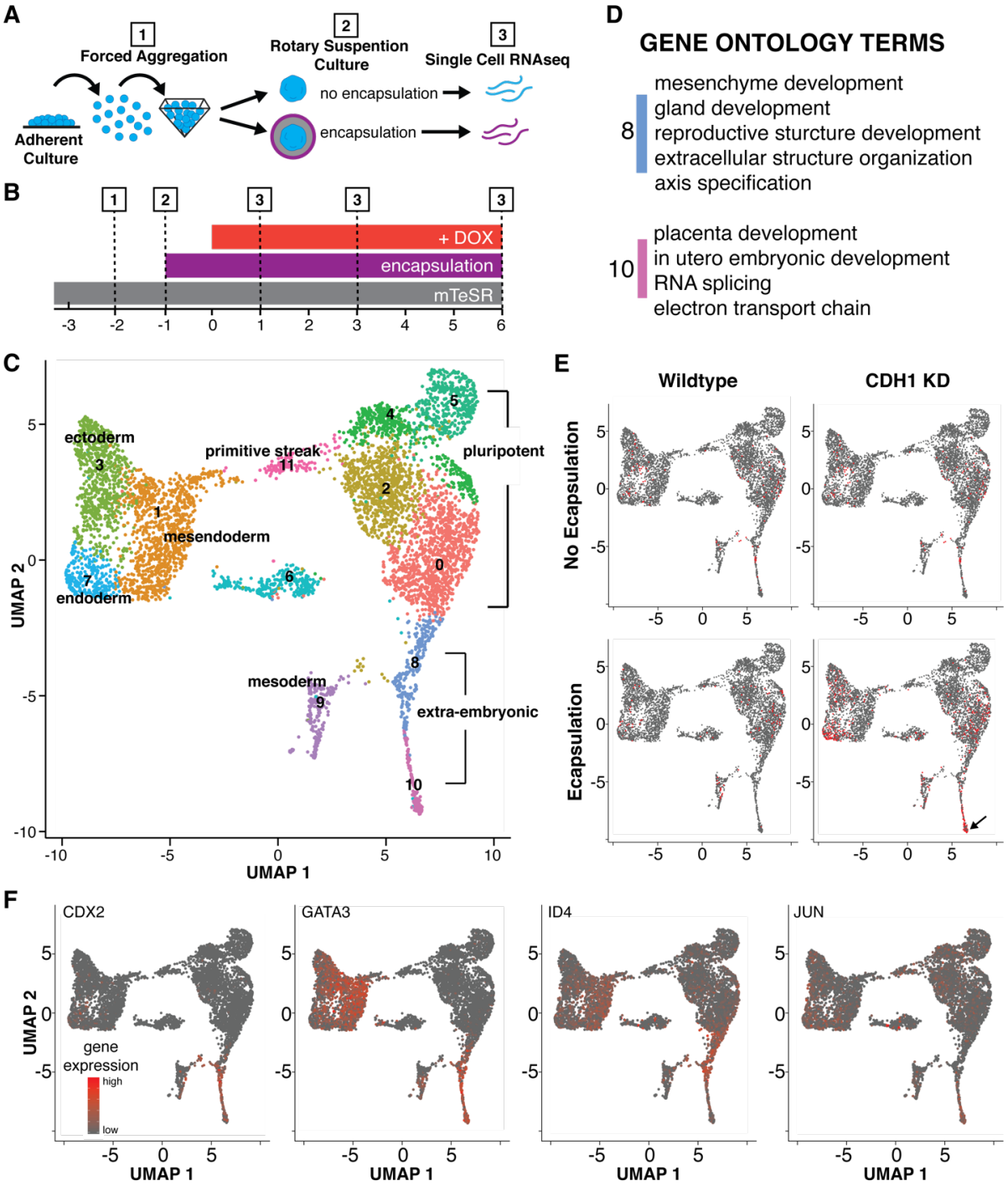

Figure 2.3: Single cell sequencing analysis of encapsulated and unencapsulated aggregates. **(A,B)** A schematic of the experimental set up. **(C)** UMAP demonstrating 12 clusters of cell populations at day 1,3, and 6. **(D)** Gene ontology terms for extraembryonic-like clusters 8 and 10 ($p < 0.05$). **(E)** Distribution of encapsulated and CDH1 knockdown cells within the UMAP clusters of all samples. **(F)** Distribution of genes associated with trophectoderm.

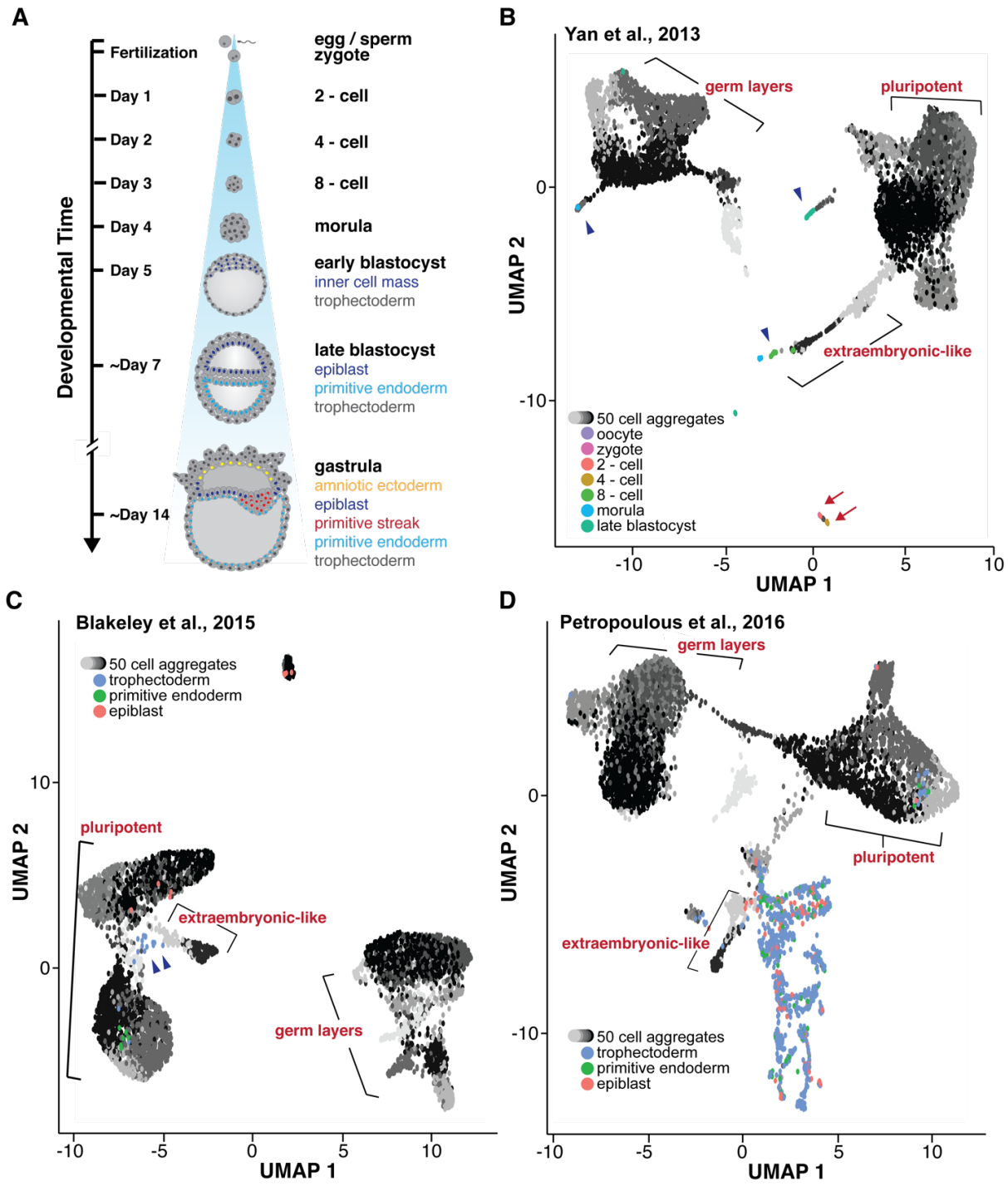

Figure 2.4: Comparison of *in vitro* aggregate transcriptome with preimplantation human embryos.

(A) Schematic of early human embryonic development. **(B-D)** 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 Petropoulous *et al.*, respectively) where the extraembryonic-like population is labeled with brackets.

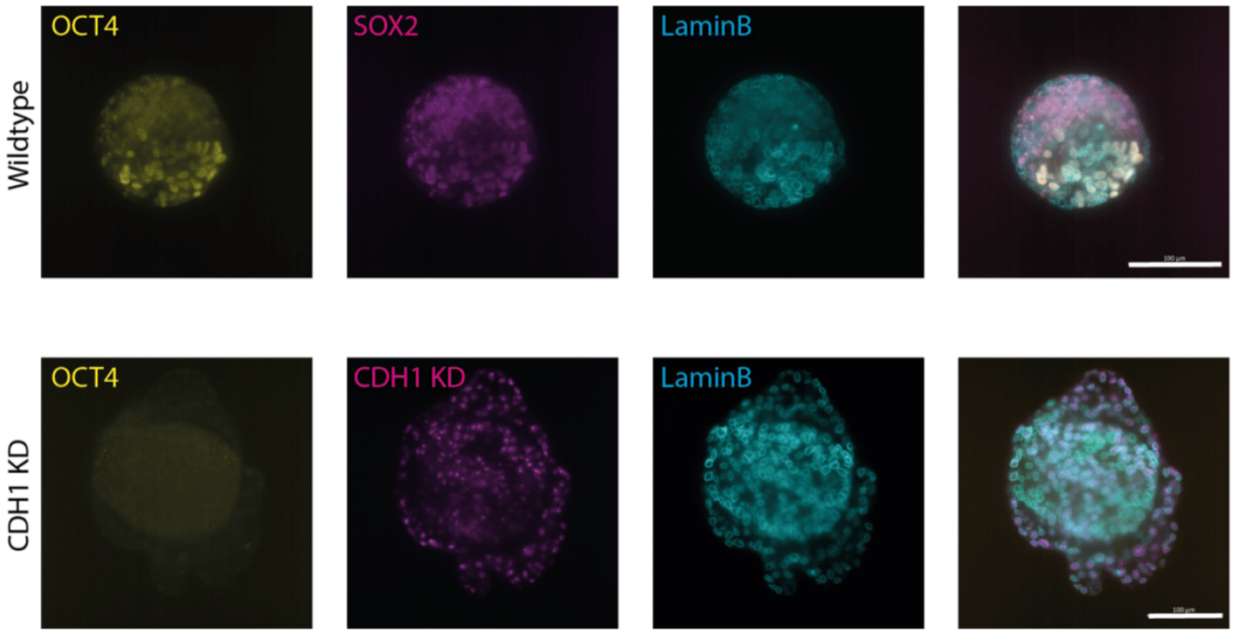

Figure S2.1: OCT4 staining in encapsulated aggregates.
(A) Immunofluorescence images of OCT4 staining in encapsulated aggregates at day 5 (scale bars = 100um).

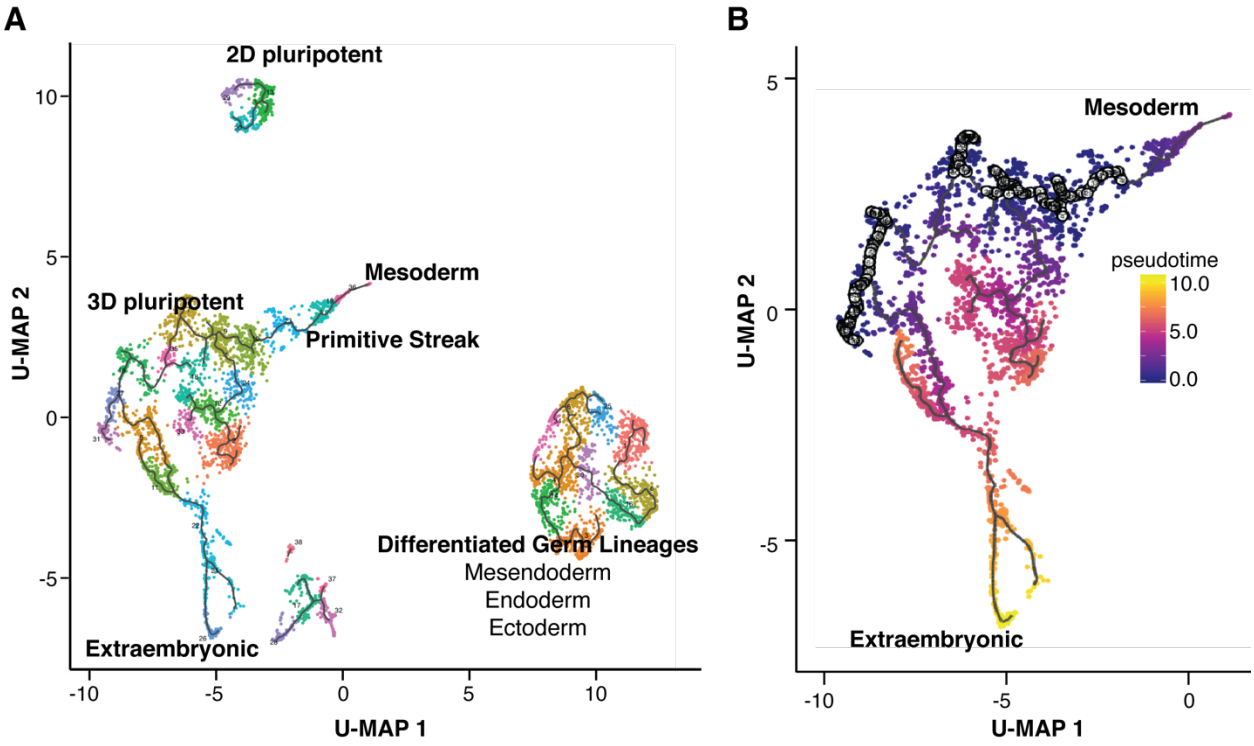

Figure S2.2: Lineage trajectories of 50 cell aggregates.
(A) Lineage trajectory reconstruction using Monacle demonstrating Mesoderm and Extraembryonic cell fates on separate lineage branches. **(B)** Pseudotime reconstruction of the mesoderm and extraembryonic lineages.

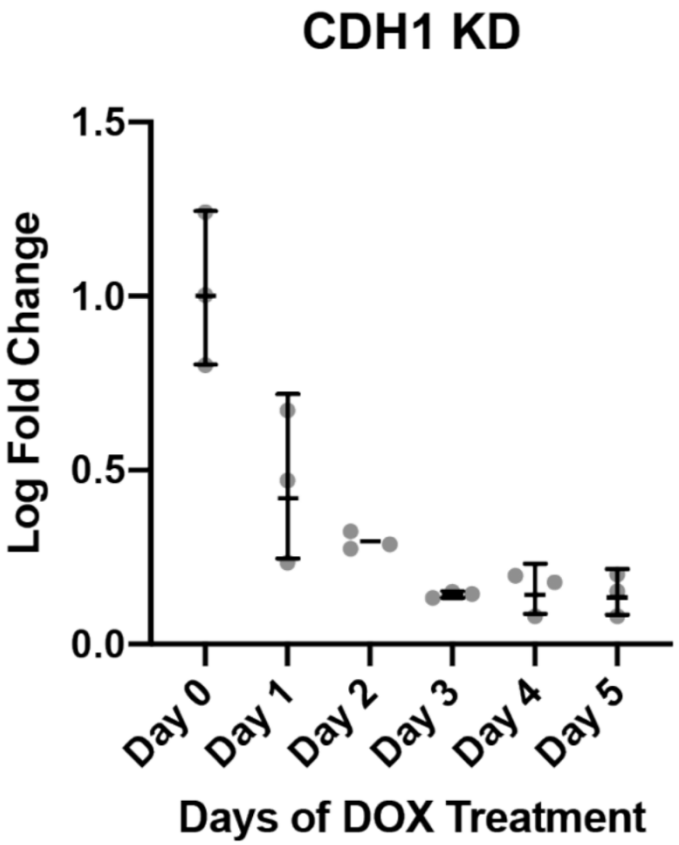

Figure S2.3: Knockdown testing of CRISPRi CDH1 line.
(A) qPCR plot of loss of CDH1 expression over a 5 day time course of DOX addition to cell culture media (n=3).

References

1. Hildebrand S, Hultin S, Subramani A, Petropoulos S, Zhang Y, Cao X, et al. The E-cadherin/AmotL2 complex organizes actin filaments required for epithelial hexagonal packing and blastocyst hatching. Sci Rep. 2017 Aug 25;7(1):9540.
2. Hirate Y, Hirahara S, Inoue K, Suzuki A, Alarcon VB, Akimoto K, et al. Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos. Curr Biol. 2013 Jul 8;23(13):1181–94.
3. Campbell S, Swann HR, Seif MW, Kimber SJ, Aplin JD. Integrins and adhesion mlecules: Cell adhesion molecules on the oocyte and preimplantation human embryo. Hum Reprod. 1995 Jun 1;10(6):1571–8.
4. Johnson MH, Ziomek CA. Cell interactions influence the fate of mouse blastomeres undergoing the transition from the 16- to the 32-cell stage. Dev Biol. 1983 Jan 1;95(1):211–8.
5. Fleming TP. A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. Dev Biol. 1987 Feb 1;119(2):520–31.
6. Kaplan KM, Spivak JM, Bendo JA. Embryology of the spine and associated congenital abnormalities. Spine J. 2005 Sep 1;5(5):564–76.
7. Srivastava D. Congenital Heart Defects. Circ Res. 2000 May 12;86(9):917–8.
8. Nakano T, Ando S, Takata N, Kawada M, Muguruma K, Sekiguchi K, et al. Self-Formation of Optic Cups and Storable Stratified Neural Retina from Human ESCs. Cell Stem Cell. 2012 Jun 14;10(6):771–85.

9. Eiraku M, Watanabe K, Matsuo-Takasaki M, Kawada M, Yonemura S, Matsumura M, et al. Self-Organized Formation of Polarized Cortical Tissues from ESCs and Its Active Manipulation by Extrinsic Signals. Cell Stem Cell. 2008 Nov 6;3(5):519–32.

10. Spence JR, Mayhew CN, Rankin SA, Kuhar MF, Vallance JE, Tolle K, et al. Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*. Nature. 2010 Dec 12;470(7332):105.

11. Dedhia PH, Bertaux-Skeirik N, Zavros Y, Spence JR. Organoid Models of Human Gastrointestinal Development and Disease. Gastroenterology. 2016 May 1;150(5):1098–112.

12. Kurosawa H. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. J Biosci Bioeng. 2007 May 1;103(5):389–98.

13. Carpenedo RL, Sargent CY, McDevitt TC. Rotary Suspension Culture Enhances the Efficiency, Yield, and Homogeneity of Embryoid Body Differentiation. STEM CELLS. 2007;25(9):2224–34.

14. Libby AR, Joy DA, So P-L, Mandegar MA, Muncie JM, Mendoza-Camacho FN, et al. Spatiotemporal mosaic self-patterning of pluripotent stem cells using CRISPR interference. eLife. 2018 Oct 9;7:e36045.

15. Przybyla L, Lakins JN, Weaver VM. Tissue Mechanics Orchestrate Wnt-Dependent Human Embryonic Stem Cell Differentiation. Cell Stem Cell. 2016 Oct 6;19(4):462–75.

16. Sagy N, Slovin S, Allalouf M, Pour M, Savyon G, Boxman J, et al. Prediction and control of symmetry breaking in embryoid bodies by environment and signal integration. Development [Internet]. 2019 Oct 15 [cited 2020 Oct 22];146(20). Available from: https://dev.biologists.org/content/146/20/dev181917

17. Mandegar MA, Huebsch N, Frolov EB, Shin E, Truong A, Olvera MP, et al. CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. Cell Stem Cell. 2016 Apr 7;18(4):541–53.

18. Toda S, Blauch LR, Tang SKY, Morsut L, Lim WA. Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science. 2018 May 31;eaat0271.

19. Ducibella T, Albertini DF, Anderson E, Biggers JD. The preimplantation mammalian embryo: Characterization of intercellular junctions and their appearance during development. Dev Biol. 1975 Aug 1;45(2):231–50.

20. Perona RM, Wassarman PM. Mouse blastocysts hatch in vitro by using a trypsin-like proteinase associated with cells of mural trophectoderm. Dev Biol. 1986 Mar 1;114(1):42–52.

21. Tada M, Concha ML, Heisenberg C-P. Non-canonical Wnt signalling and regulation of gastrulation movements. Semin Cell Dev Biol. 2002 Jun;13(3):251–60.

22. Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, et al. The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. Dev Cell. 2009 Mar 17;16(3):398–410.

23. Gueth-Hallonet C, Maro B. Cell polarity and cell diversification during early mouse embryogenesis. Trends Genet. 1992 Aug 1;8(8):274–9.

24. Gunne-Braden A, Sullivan A, Gharibi B, Sheriff RSM, Maity A, Wang Y-F, et al. GATA3 Mediates a Fast, Irreversible Commitment to BMP4-Driven Differentiation in Human Embryonic Stem Cells. Cell Stem Cell. 2020 May 7;26(5):693-706.e9.

25. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, et al. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell. 2016 May 5;165(4):1012–26.

26. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol. 2013 Sep;20(9):1131–9.

27. Blakeley P, Fogarty NME, Valle I del, Wamaitha SE, Hu TX, Elder K, et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development. 2015 Sep 15;142(18):3151–65.

28. Montero J-A, Heisenberg C-P. Gastrulation dynamics: cells move into focus. Trends Cell Biol. 2004 Nov;14(11):620–7.

29. Krieg M, Arboleda-Estudillo Y, Puech P-H, Käfer J, Graner F, Müller DJ, et al. Tensile forces govern germ-layer organization in zebrafish. Nat Cell Biol. 2008 Apr;10(4):429–36.

30. Muncie JM, Ayad NME, Lakins JN, Weaver VM. Mechanics Regulate Human Embryonic Stem Cell Self-Organization to Specify Mesoderm [Internet]. Rochester, NY: Social Science Research Network; 2020 Mar [cited 2020 Oct 21]. Report No.: ID 3543833. Available from: https://papers.ssrn.com/abstract=3543833

31. Janmey PA, Wells RG, Assoian RK, McCulloch CA. From tissue mechanics to transcription factors. Differentiation. 2013 Oct 1;86(3):112–20.

32. Ludwig TE, Bergendahl V, Levenstein ME, Yu J, Probasco MD, Thomson JA. Feeder-independent culture of human embryonic stem cells. Nat Methods. 2006 Aug;3(8):637–46.

33. Park S, Kim D, Jung Y-G, Roh S. Thiazovivin, a Rho kinase inhibitor, improves stemness maintenance of embryo-derived stem-like cells under chemically defined culture conditions in cattle. Anim Reprod Sci. 2015 Oct;161:47–57.

34. Hookway TA, Butts JC, Lee E, Tang H, McDevitt TC. Aggregate formation and suspension culture of human pluripotent stem cells and differentiated progeny. Methods [Internet]. 2015 [cited 2016 Apr 21]; Available from: http://www.sciencedirect.com/science/article/pii/S1046202315301675

35. Ungrin MD, Joshi C, Nica A, Bauwens C, Zandstra PW. Reproducible, ultra high-throughput formation of multicellular organization from single cell suspension-derived human embryonic stem cell aggregates. PloS One. 2008;3(2):e1565.

36. Wilson JL, Najia MA, Saeed R, McDevitt TC. Alginate encapsulation parameters influence the differentiation of microencapsulated embryonic stem cell aggregates. Biotechnol Bioeng. 2014;111(3):618–31.

37. McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods. 2019 Jul;16(7):619–26.

38. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018 May;36(5):411–20.

39. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan 1;29(1):15–21.

40. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions | SpringerLink [Internet]. [cited 2020 Oct 21]. Available from: https://link.springer.com/article/10.1186/gb-2013-14-4-r36

41. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 Apr 1;30(7):923–30.

42. Korsunsky I, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, et al. Fast, sensitive, and accurate integration of single cell data with Harmony. bioRxiv. 2018 Nov 5;461954.

43. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS J Integr Biol. 2012 Mar 28;16(5):284–7.

44. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014 Apr;32(4):381–6.

Chapter 3. Loss of tight junctions disrupts gastrulation patterning and increases differentiation towards the germ cell lineage in human pluripotent stem cells

Abstract

Biological patterning events that occur early in development establish proper tissue morphogenesis. Identifying the mechanisms which guide these patterning events is necessary for understanding the molecular drivers of development and disease, and to build tissues *in vitro*. In this study, we used an *in vitro* model of gastrulation to study the role of tight junctions and apical/basolateral polarity in modulating bone morphogenic protein-4 (BMP4) signaling and gastrulation-associated patterning in colonies of human pluripotent stem cells (hPSCs). Disrupting tight junctions via knockdown (KD) of zonula occludens-1 (ZO1) allows BMP4 to robustly and ubiquitously activate pSMAD1 signaling over time and results in loss of the gastrulation patterning phenotype, establishing ZO1 as a key regulator of patterning in this system and clarifying previous mechanistic discrepancies. We find that ZO1 KD causes marked differentiation bias of pluripotent stem cells to primordial germ cell-like cells (PGCLCs). We recapitulate the effects of ZO1 KD in control hPSCs through prolonged basolateral BMP4 stimulation, and demonstrate that PGCLC bias can largely be attributed to loss of barrier function and increases in pSMAD1 activation in ZO1 KD cells. Our results demonstrate that contrary to prevailing theories, primed hPSCs have the capacity to efficiently differentiate to PGCLCs without pre-conditioning or resetting to a more naïve pluripotent ground state. Together, these findings give important insights into how signaling events are regulated and lead to spatial emergence of diverse cell types *in vitro* and *in vivo*.

Introduction

Early in embryonic development, gradients of signaling molecules guide the spatial organization and specification of stem cells to generate a blueprint for axis formation and subsequent tissue growth[1–3]. One of the earliest biological patterning events occurs in the epiblast, when pluripotent cells differentiate and segregate into the three somatic germ layers (endoderm, mesoderm, and ectoderm), in a process known as gastrulation. Understanding how signaling gradients first emerge, take shape, and are interpreted by cells during gastrulation is important for understanding the fundamental principles which guide patterning and morphogenesis of tissues.

Prior to gastrulation, the epiblast is an epithelial tissue reinforced by tight junction complexes. These complexes are comprised of an intracellular scaffold, which partitions the phospholipid bilayer into apical and basolateral domains, and a set of transmembrane proteins, which prevent paracellular diffusion of macromolecules between the apical-facing and basolateral-facing lumens[4–6]. Signaling protein/receptor families responsible for regulating the gastrulation program (including BMP4, WNT, NODAL, NOGGIN, etc.) have conserved basolateral or apical trafficking motifs[7–10]. This receptor orientation implicates epithelial structure and cell polarity as critical determinants of morphogen gradient shape and subsequent cell type patterning during gastrulation via asymmetric attenuation of cellular response to morphogen signaling[9,11]. In addition to spatial patterning imposed by epithelial structure, previous studies have proposed that morphogens can autonomously form patterns that lead to germ layer segregation during gastrulation, as described by Alan Turing's reaction diffusion (RD) and Lewis Wolpert's positional information (PI) models[12–18]. However, a lack of mechanistic insight hinders our understanding of the possible interplay between tight junction expression and RD/PI in the context of gastrulation patterning.

Seminal work shows that gastrulation-associated patterning events, including the emergence of ectoderm-like, mesendoderm-like, and trophectoderm-like cells in a radial pattern, can be recapitulated in culture by confining hPSCs to circular micropatterns and stimulating with BMP4[19,20]. BMP4, which acts through phosphorylation of the signal transducer SMAD1, is known to be an important initial cue in the gastrulation cascade[21,22]. Follow-up studies suggest that patterning in this system follows an RD/PI process, as BMP4 driven phosphorylation of SMAD1 and subsequent pathway activation causes cells in the colony to secrete both BMP4 and its inhibitor (NOGGIN) in a feedback loop[18]. Differences in the diffusivities between NOGGIN and BMP4 are thought to create a steady-state gradient of effective BMP4 concentrations across the colony, and cells are presumed to sense positional information and differentiate based both on this concentration gradient and its overlap with other members of the BMP4-induced feedback loop, including WNT and NODAL[17,18].

Contradictory reports confound our understanding of the extent to which tight junctions influence ligand-receptor accessibility, signaling pathway activation, and subsequent RD/PI driven germ layer patterning in this system. For example, Etoc et al. demonstrate that changes in apical/basolateral polarity occur on the edge of micropatterned colonies, presumably due to regional loss of tight junctions. They propose that these polarity changes lead to differential receptor accessibility and preferential pathway activation on the colony edge, and that this edge activation is necessary for radial patterning[17]. However, disruption of tight junction assembly throughout hPSC colonies via calcium chelation does not result in disruption of the radial patterning phenotype. Calcium chelation in addition to ROCK inhibition does disrupt patterning[17]; however, these macroscale and non-specific perturbations make it difficult to decipher whether and how tight junctions are involved in signaling and patterning. Alternatively, Tewary et al. show

that pathway activation is ubiquitous throughout micropatterned colonies at earlier timepoints. These results negate the role of differential receptor accessibility caused by tight junction expression and implicate secreted inhibitors as the primary regulators of patterning[18].

In this study, we use a modified *in vitro* hPSC gastrulation model which maintains epithelial structure over time and recapitulates the characteristics of the epiblast prior to gastrulation. We target tight junctions specifically by knocking down ZO1, a critical component of the tight junction complex[23,24], to study the effects of tight junction expression on signaling pathway activation, gastrulation-associated patterning, and cell specification. We demonstrate that ZO1 KD increases cellular receptiveness to BMP4 and causes ubiquitous and sustained signaling pathway activation, which appears to override endogenous NOGGIN inhibitor activity throughout the colony. Significant changes in multicellular patterning and proportions of different cell types in the ZO1 KD (ZKD) versus control hPSC (ZWT) colonies establish ZO1 as a key regulator of patterning in our gastrulation model. Additional characterization of cell type emergence demonstrates that ZKD colonies predominantly differentiate towards PGCLC fates. By comparing the ground state of ZWT and ZKD cells, we show that PGCLC fate bias is predominantly a result of heightened and sustained BMP4 signaling pathway activation that can be recapitulated in ZWT cells. These results provide key insights both into how tight junction complexes within the embryo may shape signaling and morphological patterning, and the importance of epithelial structure in influencing hPSC differentiation *in vitro*.

Results

BMP Pathway Activation Correlates with Regional Loss of ZO1

hPSCs confined to circular micropatterns and treated for 42-48 hours with BMP4 undergo radial patterning of gastrulation-associated makers CDX2 (trophectoderm-like), TBXT (mesendoderm-like), and SOX2 (ectoderm-like)[19,20]. Our lab and others have demonstrated that similarly-sized colonies whose growth is not confined by micropatterns undergo analogous radial patterning in response to BMP4 stimulation[25,26,27] (Figure 1A). In this modified protocol, ~50-100 hPSCs are aggregated overnight within pyramidal microwells, and the following day these 3D aggregates are re-plated sparsely and allowed to grow into distinct 2D colonies that reach 300-500μm in diameter. Compared with micropatterned colonies, unconfined colonies maintain a relatively uniform cell density and a robust epithelial morphology over time (Figure S1A, S1B, S1C). This is critical given that epithelial integrity is a direct function of cell density, and previous reports have linked changes in signaling and cell specification with regional changes in cell density[17,28,29,30].

Others have demonstrated that low cell densities prevent proper tight junction formation and presumably enhance permeability to signaling proteins[17]. Interestingly, we find the opposite is also true: in monolayer culture at high cell densities, the honeycomb-like intercellular protein expression pattern of ZO1, which is indicative of an intact epithelium, becomes disrupted and punctate (Figure S1D). Moreover, regions with punctate ZO1 expression, which increase in frequency as cell density increases, overlap with regions of BMP4-induced signaling pathway activation (phosphorylation of SMAD1). Therefore, both very low and very high cell densities can cause increases in epithelial permeability. In our experience, punctate ZO1 expression is also

present in micropatterned colonies; regions of high density lose ZO1 and overlap with pSMAD1 activation upon BMP4 stimulation (Figure S1A). Discrepancies in previously reported pSMAD1 pre-patterns may therefore be partially attributed to regional changes in cell density that perturb epithelial structure.

ZO1 expression inversely correlates with pSMAD1 activation even in the context of unconfined colonies with uniform density. For example, upon induction with BMP4, pSMAD1 activity is primarily limited to the cells on the edge at earlier timepoints (15 min – 1 hour). We and others noticed that ZO1 expression does not fully extend to the edge of the colony, and tapers off a distance of approximately one cell layer before reaching the edge[31]. Co-staining of ZO1 and pSMAD1 in unconfined colonies after 1 hour of BMP4 stimulation showed an anti-correlation between pSMAD1 positive and ZO1 positive regions (Figure 1B). We used CellProfiler[32] (see Methods) to visualize and quantify pSMAD1 and ZO1 fluorescence signal at different radial distances from the colony center and normalized these values to LMNB1 nuclear signal, demonstrating the inverse relationship between pSMAD1 and ZO1 (Figure 1C, 1D). Given that initial pSMAD1 pre-patterning has been implicated in regulating subsequent gastrulation-associated patterning in micropatterned colonies[17,18], we aimed to elucidate the effect of tight junctions on pluripotent morphogenic signaling and gastrulation patterning.

ZO1 Knockdown Leads to Ubiquitous and Sustained Pathway Activation

In vitro, hPSCs are cultured as epithelial sheets that have tight junctions and display apical/basolateral polarity[33], with most morphogen receptors, including BMP receptors BMPR1A, BMPR2, and ACVR2A, localized to the basolateral side[9,17]. These receptors are physically partitioned away from morphogens present in the soluble media on the apical side. As a result,

tight junction expression presumably attenuates cellular response to exogenous morphogen signals *in vitro* (Figure 2A).

In order to explore how tight junctions affect cellular receptiveness to signaling in unconfined colonies and, in turn, pattern formation, we sought to knockdown a tight junction component in hPSCs using a DOX inducible CRISPR interference (CRISPRi) system[34]. We targeted the zonula-occludens (ZO) family of scaffolding proteins, because they form dual purpose adhesion plaques that are critical both for maintaining apical/basolateral polarity and barrier function. Preliminary RNA sequencing data showed that ZO1 is much more highly expressed in cultured hPSCs than ZO2 or ZO3, so we decided to target ZO1 specifically (Figure S2A). We created both male (WTC) and female (WTB) hPSC ZO1 KD lines. The WTC line also contains a LMNB1-GFP fusion reporter for live nuclear visualization. Both hPSC ZO1 CRISPRi lines are karyotypically normal (Figure S2C), and ZO1 RNA and protein expression are significantly depleted after five days of DOX treatment, as shown by qPCR, RNA sequencing, immunofluorescence (IF), and western blot (Figures 2B, S4A, 2C, S2B). We performed most of the characterization in the WTC ZO1 CRISPRi line with and without DOX (referred to in the text as ZKD and ZWT, respectively); however, several supplemental figures show phenotype reproducibility in the WTB ZO1 CRISPRi line.

Prior to BMP4 exposure, we observed distinct morphology between ZWT and ZKD cells in standard culture. ZKD cells grew in denser colonies and exhibited changes in nuclear shape (Figure S2D, S2E). Where ZWT nuclei were stretched and flat, ZKD nuclei were taller and rounder, presumably as a result of severed connections between the cell-cell junctions and the actin cytoskeleton/nuclear lamina. When grown as unconfined colonies and exposed to BMP4, ZWT largely limited pSMAD1 expression to the colony edge at early timepoints (15 min – 1 hour). At

later timepoints (6 hours), pSMAD1 was detectable in cells located centrally within the colony; however, due to well-known inhibitor feedback loops[17,18], this pathway activation was shut off by 48 hours (Figure 2D, 2E). Strikingly, at early timepoints, the ZKD colonies displayed pSMAD1 throughout the colony. Furthermore, ZKD cells maintained pSMAD1 activation over time, despite significant increases in transcription of the secreted BMP inhibitor NOGGIN (Figure S3C), which is implicated in SMAD1 inactivation in ZWT cells over time[18]. In ZWT cells, NOGGIN is secreted apically and is trafficked transepithelially with assistance from glycoproteins on the apical surface[10]. The maintenance of pSMAD1 pathway activation despite increased NOGGIN expression in ZKD colonies suggests that ZO1 is not only important for preventing ligands such as BMP4 from accessing basolateral receptors, but may also be necessary to render the cells sensitive to some inhibitors. Presumably, this occurs by maintaining expression of the apical surface glycoproteins[31] that enable transepithelial trafficking of apically secreted inhibitors such as NOGGIN[10,35] or sequestration and concentration of other basolaterally secreted morphogen inhibitors within the colony interior. The idea that ZO1 helps maintain expression of apical glycoproteins is reinforced by the fact that ZKD cells also exhibit loss of apical Ezrin expression (Figure S3F), which has been shown to be important in tethering apical glycoproteins to the actin cytoskeleton[36].

Signaling Changes are a Result of Increased Permeability in ZKD Cells

In order to confirm basolateral sequestration of BMP receptors within an epithelium, hPSCs were grown on a transwell membrane, where apical and basolateral sides of the media are independently accessible. As expected, basolateral presentation of BMP4 is required for pSMAD1 activation in ZWT cultures (Figure S3A). Alternatively, both apical and basolateral stimulation

activates pSMAD1 in ZKD cells (Figure S3A). Given that BMP receptor gene expression does not considerably differ between ZWT and ZKD cells (Figure S3B), two main possibilities could explain this phenomenon. The first is that ZKD causes mixing of apical/basolateral domain elements through the plasma membrane and disrupts trafficking of receptors to their proper domains (loss of apical/basolateral polarity). The second is that ZKD causes increased permeability to signaling molecules (loss of barrier function). To test these possibilities, we first looked to characterize apical/basolateral polarity between ZWT and ZKD cells.

In polarized cells, the Golgi apparatus faces the apical (secretory domain)[37,38], therefore, we examined the positioning of the Golgi in ZWT and ZKD cells. Z-stacks of confocal microscopy images revealed that in both cell types, the Golgi sits on top of the nucleus facing the apical side of the cell, suggesting that polarity of the ZKD cells is still intact (Figure S3D, S3E). However, staining for the apical marker Ezrin revealed significant eradication of the apical domain in ZKD cells, characterized by punctate Ezrin localization. This localization is consistent with previous reports that Ezrin expression is diminished on the edge of regular hPSC colonies[31]. IF images showed that swaths of ZKD cells lost apical Ezrin. Even in regions where Ezrin was present, Ezrin signal overlapped significantly with BMPR1A (a basolateral BMP receptor), suggesting potential changes in localization of apical/basolateral elements (Figure S3F, S3G). Our results suggest that polarity-associated changes do not occur in cytoplasmic elements within the cell, but may result in polarity changes of elements bound to the plasma membrane.

Next, we performed a FITC-based diffusion assay to look for differences in permeability of ZWT and ZKD hPSC monolayers. To do this, we grew each cell type on a transwell membrane and added a 40kDa dextran conjugated with FITC to the apical compartment (Figure 2F). 40kDa-FITC was selected due to its similarity in size and hydraulic radius to BMP4. Fluorescence

measurements of the basolateral compartment over time allowed us to quantify permeability of the ZKD compared to ZWT cells. We found that significant increases in FITC diffusion through ZKD cell layers could be observed as early as 30 minutes following treatment (Figure 2G). Similarly, transepithelial resistance (TEER) measurements performed on ZWT and ZKD monolayers confirmed that ZKD cells were not able to form a continuous epithelium that resists passage of ions through the paracellular space (Figure 2H). Therefore, while some changes in apical/basolateral polarity may occur, our results suggest that definitive changes in molecular permeability permit heightened signaling pathway activation seen in ZKD cells.

ZKD Causes Changes in Cell Fate Proportions in Unconfined Gastrulation Models

The initial pSMAD1 edge pre-pattern is assumed to dictate both the shape of a BMP-NOGGIN RD gradient and the subsequent spatiotemporal pSMAD1 and gastrulation-like pattern[17,18]. pSMAD2 activation, which occurs through NODAL signaling, is also known to be important for the patterning phenotype, as pharmacological inhibition of NODAL disrupts the emergence of CDX2+ and TBXT+ cells in micropatterned hPSCs stimulated with BMP4[15]. However, Tewary et al. demonstrate that stimulation of micropatterned hPSC colonies solely with NODAL does not lead to gastrulation-like patterning[18], implicating the BMP-NOGGIN gradient and subsequent spatiotemporal pSMAD1 activation as the primary driver of patterning in accordance with PI principles[15].

In ZWT colonies, cells on the edge that remain pSMAD1 positive throughout BMP4 stimulation eventually acquire CDX2+ trophectoderm-like fates. Etoc et al. model the edge region as being simultaneously pSMAD1+ and pSMAD2+, a combination predicted to yield CDX2 fates. Therefore, if the current RD/PI paradigm established by Tewary et al. and Etoc et al. is correct,

ZKD colonies, which mimic the "edge" phenotype and maintain ubiquitous and sustained pSMAD1 activation throughout the entire colony for 48 hours, would ubiquitously differentiate to the CDX2 lineage following BMP4 treatment (Figure 3A). Accordingly, our results show that ZKD colonies treated with BMP4 have increased CDX2 expression across the colony interior. In addition, these colonies display a stark decrease in central SOX2 expression, and disruption of the TBXT ring pattern (Figure 3B, 3C). These results establish ZO1, and therefore tight junction stability, as a key component of BMP4-induced cell fate and spatial patterning.

RNA Sequencing of BMP4-Treated ZKD Colonies Reveals PGCLC Bias

Unexpectedly, we observed that like CDX2, TBXT expression was substantially increased throughout the center of the colony (Figures 3B). Many progenitor cell types express TBXT; therefore, to better identify this population and quantify changes in ZKD-induced lineage bias, we performed RNA sequencing on pluripotent and BMP4-treated (48 hours) ZWT and ZKD cells. RNA sequencing confirmed IF staining results: CDX2 and TBXT transcripts were highly expressed, whereas SOX2 were lowly expressed in ZKD cells treated with BMP4, compared with ZWT. Analysis of a panel of well-known gastrulation-associated lineage markers in ZWT and ZKD cells revealed that ZKD cells have the tendency to express mesendoderm, PGC, and extraembryonic markers at the expense of ectodermal-like lineages (Fig 3D). Gene ontology (GO) analysis performed on clusters 2 and 3 of the top 150 differentially expressed genes between ZWT and ZKD showed upregulation of endoderm and sex cell-related pathways in ZKD colonies (Figure 3E). Similarly, when considering top most highly expressed genes between ZWT and ZKD, analysis revealed significant increases in NANOS3, SOX17, and WNT3: genes, which when expressed together, are associated with the human PGC specification program[39] (Figure 3F).

Subsequent IF staining for PGC markers BLIMP1, TFAP2C, and SOX17 at 48 hours showed increases in expression of these markers in ZKD colonies compared with the ZWT controls (Figure 4A, 4B). PGCLC differentiation of ZKD cells can also be observed in cells grown under standard monolayer conditions (not aggregated or grown into circular colonies) with 48 hours of BMP4 stimulation (3E-3F). By 72 hours, clear triple positive expression of BLIMP1/TFAP2C/SOX17 was seen in nearly half (~47%) of ZKD cells (Figure 4G-H) in monolayer culture, a phenotype that was also confirmed in the female (WTB) iPSC line (Fig S4A-D). Together, these results suggest that BMP4 stimulation of hPSCs lacking tight junctions dramatically augments cell receptiveness to signals needed for PGCLC emergence.

Decoupling Signaling and Structural Changes in ZKD PGCLCs

Upon the discovery of a nascent PGCLC population within our ZKD colonies, we next sought to decouple the effects of structural changes due to tight junction instability and ubiquitous pSMAD1 activation that might enable the robust emergence of a PGCLC population. We first looked to existing literature to understand the required components enabling PGCLC differentiation *in vitro*. Two seminal papers describe different protocols for generating human PGCLCs[39,40]. In the first protocol (Sasaki et al.), hPSCs are pre-induced into an incipient mesoderm-like (iMeLC) state that renders the cells poised for PGCLC specification. In the second protocol (Irie et al.), hPSCs are first reset from a primed to a naïve pluripotency state, as primed hPSCs are thought to have lost the developmental potential to generate PGCLCs. Indeed, without iMeLC or naïve pluripotency pre-induction, both protocols fail to efficiently generate PGCLCs (1-2% efficiency). However, in our differentiation, ZKD cells do not undergo any form of pre-induction yet are able to produce a robust PGCLC population. Two possibilities potentially explain

this PGCLC specification bias: 1) ZKD is causing a change in pluripotent ground state (to a naïve- or iMeLC- like state), or 2) signaling changes caused by ZKD recapitulate *in vivo* PGC specification, and are sufficient to drive PGCLC differentiation *in vitro*.

We first looked to characterize pluripotency in ZWT and ZKD cells in the absence of BMP4. RNA sequencing results indicated that aside from ZO1 and ZNF10 (which is part of the CRISPRi machinery), few genes were both significantly and substantially differentially expressed between ZWT and ZKD cells (Figure S4F) in the pluripotent state, and no significant changes were detected in canonical pluripotency markers (Figure 4C). Whole genome bisulfite sequencing demonstrates that while several gene regions were differentially methylated (Figure 4D, S4E), there were no global changes in methylation between ZWT and ZKD cells, which would be expected if a resetting process to a more naïve pluripotency state occurred. GO analysis also did not reveal any significant regulatory or functional relationship between differentially methylated genes. Together, these data suggest that the transcriptome and methylome were not greatly affected by ZKD and thus, a potential change in the ground state would not explain the ZKD predisposition to adopt PGCLC fates.

We next tested the hypothesis that ZKD cells are predisposed to PGCLC fates because, unlike ZWT cells which are known to experience NOGGIN-related BMP4-pathway inhibition at later timepoints, ZKD cells experience sustained BMP4-pathway activation[18]. To decouple changes in signaling from potential structural changes that result from ZO1 knockdown, we looked to recapitulate the pSMAD1 signaling dynamics in hPSCs without ZO1 knockdown. To do this, we grew ZWT cells on a transwell membrane. Bi-directional stimulation of hPSCs with BMP4 resulted in ubiquitous and sustained activation of pSMAD1 over the course of 48 hours, much like when ZKD cells are stimulated in standard culture (Figure 5A, 5B). RNA sequencing of stimulated

ZWT and ZKD cells grown on transwells showed remarkable similarities in marker expression between the two samples, demonstrating that most of the observed changes in cell fate are a direct result of increased signal pathway activation. The total number of differentially expressed genes between ZWT and ZKD samples was significantly higher in standard culture (3150) than in transwell (53) culture, highlighting the magnitude of the expression changes dependent solely on changes in pSMAD1 signaling. Of these 53 genes, unbiased clustering and GO analysis demonstrated that ZKD cells retain a bias towards mesendodermal lineages (Figure 5C).

Interestingly, neither ZWT nor ZKD cells grown on transwell membranes and treated for 48hours with BMP4 (50ng/mL) efficiently differentiated to PGCLC fates, as seen for ZKD cells on standard plates. We hypothesized that this could be due to excessive signal pathway activation from bi-directional stimulation on the transwell membrane, or mechanical differences between the two plates, which has previously been shown to affect differentiation[41]. By testing a range of BMP4 concentrations, we discovered robust and ubiquitous PGCLC differentiation of ZWT cells on the transwell membranes (Figure 5D, 5E). We found that bi-directional stimulation with 10ng/ml BMP4 was optimal for PGCLC specification. Taken together, these results indicate that changes in cell identity in the absence of ZO1, and specifically the emergence of a PGCLC population, are largely due to increased susceptibility to BMP4 signaling.

Discussion

In this study, we definitively link epithelial structure, maintained by tight junction proteins, with signaling pathway activation and multicellular patterning. We use a gastrulation-like patterning platform which maintains epithelial structure over time to clarify the role of tight

junctions and apical/basolateral polarity in shaping the pSMAD1 pre-pattern and in directing germ layer patterning. By stably knocking down ZO1 in our system, we disrupt the function of tight junctions and recapitulate "edge" phenotypes in every cell throughout the colony (i.e. loss of apical/basolateral polarity, loss of epithelial barrier function). We demonstrate that perturbing tight junction structure causes ubiquitous pSMAD1 activation and loss of the pSMAD1 edge asymmetry postulated by Etoc et al. to drive the formation of the RD gradient. ZO1 loss also renders cells insensitive to pSMAD1 pathway inactivation via NOGGIN, and as a result, disrupts the patterning of germ layer markers. Aside from repression of NOGGIN, few other studies to date have reported perturbations that lead to such a significant loss of patterning as we observe with ZO1 KD, distinguishing ZO1 as a critical regulator of signaling and patterning.

These findings have interesting implications for the role of epithelial structure in the maintenance and controlled emergence of cell types during embryonic development. In the embryo, the significance of the epiblast's epithelial structure has remained a mystery; however, epithelialization is a well-known hallmark of early development in amniotes, as the epiblast starts as a non-polar aggregate of cells, passes through a semi-polar rosette stage, and eventually forms a laminar epithelial structure thought to be a requisite for subsequent gastrulation[42]. *In vitro* capture of pluripotent stem cells from each of these stages reveals that these cells are transcriptionally and epigenetically distinct, and are maintained by different signaling milieus[43–46]. While we show that changes in epithelial structure do not influence the pluripotent state, our findings suggest that different classes of signaling proteins might, by design, be used in signaling of non-polar versus epithelial tissues.

For example, we observe that both NOGGIN and WNT expression increase in ZKD colonies, however, the activity of NOGGIN appears to be negligible, as there is no change in the pSMAD1

activation profile over time. However, high induction of TBXT-positive cells implies that WNT activity remains high in the colonies. One explanation for this discrepancy could be that compared to WNT, NOGGIN is a long-range inhibitor[47,48] that is secreted apically and relies on trans-epithelial transport to interact with BMP4 ligands which are secreted basolaterally[10]. Previous findings show that despite basolateral BMP4 stimulation, apical NOGGIN is capable of shutting off pSMAD1 activation. Loss of the apical domain in ZKD cells may interfere with shuttling functions, which might explain inactivity of NOGGIN in the colonies. By contrast, previous evidence suggests that WNT is a short-range, juxtacrine signaling protein[49]. Increases in density that are correlated with loss of epithelial structure are therefore suitable for prolific WNT signaling, and may account for increases in TBXT fates throughout ZKD colonies. In other words, epithelial structure may be necessary for signaling by some types of proteins and may interfere with signaling by others, depending on their size, chemical structure, and mechanism of action.

Additionally, regional loss of tight junctions due to increases in density, embryo geometry, or cell-activated breakdown could be an important mechanism for increasing sensitivity to apical signaling cues in specific locations in the embryo. For example, previous studies in the mouse embryo show that there is a break in tight junction expression between the extraembryonic ectoderm (ExE) and the epiblast[9]. The ExE is responsible for secreting BMP4 signals necessary for both PGC differentiation and gastrulation initiation[22]. In mouse and cynomolgus monkeys, PGCs arise in the proximal epiblast directly adjacent to the ExE[50], bordering the region that is reported to have lower expression of tight junctions and higher pSMAD1 activity[9]. Given our finding that loss of ZO1 primes hPSCs for PGCLC fates, it is possible that this specific and unique position in the embryo permits future PGCs to be exposed to higher and more sustained levels of BMP compared with the rest of the embryo proper, thereby promoting their specification.

Understanding how ZO1 affects cells' receptiveness to signaling molecules enables us to "unlock" *in vitro* stem cell differentiation protocols for previously intractable cell types, as we did in generating PGCLCs. Our results show that hPSCs in regular culture are largely unresponsive to BMP4 presented apically in the media, an observation that could apply to other types of morphogens as well. RNA sequencing data reveal that hPSCs stimulated for 48 hours with BMP4 are transcriptionally similar to unstimulated cells. Therefore, while many differentiation protocols aim to recapitulate exposure to developmentally relevant morphogens, hPSCs in culture are likely not receiving these signals. For example, genetic studies in mouse null mutants lacking genes for BMP4, SMAD1, and SMAD5 fail to develop PGCs, demonstrating that their specification is dependent on the acute activation of the canonical BMP4-SMAD1/5 pathway[50,51]. Previous protocols to generate PGCLCs have relied on excessively high concentrations of BMP4 that far exceed physiological levels and require a dissociation step, which is perhaps inadvertently used to overcome signaling barriers present in standard culture of hPSCs with an epithelial phenotype. However, we and others have shown that tight junctions re-assemble very quickly after aggregation in hPSCs, after which they would confer the same signaling barriers as standard culture and may result in inconsistencies in differentiation. Using temporary pharmacological inhibition of ZO1 or culturing cells on transwells prior to differentiation can enable precise, controlled, and multiplexed signaling pathway activation, as well as standardization and optimization of many hPSC differentiation protocols.

Patterning events that drive tissue morphogenesis are coordinated by a series of fundamental and interconnected pathways, which span many modes of cellular communication (e.g. paracrine, juxtacrine, extracellular matrix signaling, mechanical push and pull) and are conserved in the development of different tissue types. Identifying key parameters that drive morphogenesis is

important for gaining a mechanistic understanding of how tissues develop and therefore also for engineering tissues *in vitro*. Our findings show that tissue structure is an important factor to consider in understanding how morphogen gradients are shaped *in vivo* and in developing robust and reproducible differentiation protocols *in vitro.*

Acknowledgements

This study was supported by the on-site laboratory initiative launched by Kyoto University. Graphical illustrations and figure formatting were done by Tami Tolpa, Gladstone Institutes.

Methods

Cell Culture

The hPSC lines used in this study are the male WTC-LMNB1-mEGFP line from the Allen Institute (Cell Line ID: AICS-0013 cl.210, passage 32) obtained from Coriel, and the female WTB CRISPRi-Gen1B line obtained from Dr. Bruce Conklin's lab at the Gladstone Institutes (Gladstone Stem Cell Core, passage 40). For routine culture, hPSCs were grown feeder-free on growth factor reduced Matrigel (BD Biosciences) and fed daily with mTESR1 medium (Stem Cell Technologies). Cells were passaged every 3-4 days with Accutase (Stem Cell Technologies) and seeded at a density of 100 cells/mm^2. ROCK inhibitor Y-276932 (10μM; Selleckchem) was added to the media to promote cell survival after passaging. All generated cell lines were karyotyped prior to expansion and confirmed normal both by Cell Line Genetics and by using the hPSC Genetic Analysis Kit (Stem Cell Technologies Cat. # 07550). The cells were also regularly tested for mycoplasma using a MycoAlert Mycoplasma Detection Kit (Lonza).

Generation of CRISPRi Lines

ZO1 KD was achieved using a doxycycline (DOX) inducible CRISPR interference (CRISPRi) system, which is comprised of two components: a dCas9-KRAB repressor driven by a Tet-on-3G promoter knocked in into the AAVS1 safe harbor locus and expressed only under DOX treatment, and a constitutively expressed guide RNA (gRNA) which targets the transcriptional start site of a gene (Figure S2A), described previously by Mandegar et al.[34]. Briefly, to generate the parental line containing CRISPRi machinery, approximately 2 million WTC or WTB derived cells were nucleofected with the knock-in vector (1μg) along with TALENS targeting the AAVS1 locus (0.5μg for each pair) and cultured in mTESR1 and ROCK inhibitor (10μM). Knock-in selection was performed with Genticin (Life Technologies) over the course of 10 days, and a clonal population was generated through colony picking under the EVOS picking microscope (Life Technologies) in sterile conditions. To achieve ZO1 KD, we designed gRNAs which bind within 150bp of the transcription start site of ZO1 and cloned them into the gRNA-CKB vector at the BsmB1 restriction site, following the protocol described in Mandegar et al. Vectors containing each gRNA sequence were individually nucleofected into the WTC-LMNB1-mEGFP line (containing the CRISPRi-KRAB construct) using the Human Stem Cell Nucleofector Kit 1 solution with the Amaxa nucleofector 2b device (Lonza). Nucleofected cells were subsequently seeded at a density of 8,000 cells/cm^2 and recovered in mTESR1 supplemented with ROCK inhibitor Y-276932 (10μM) for two days. Guide selection was performed with blasticidin (10μg/mL, ThermoFisher Scientific) for seven days, and clonal populations were generated through colony picking. Knockdown efficiency was evaluated through exposure to Doxycycline (2μM) for five days, after which mRNA was isolated, and relative levels of ZO1 were assessed

through qPCR. Levels of ZO1 were normalized to copy numbers from the same line without CRISPRi induction, and the most effective guide was selected (sequence: CCGGTTCCCGGGAAGTTACG). After validation, this guide was subsequently introduced into the WTB CRISPRi-Gen1B line, which was selected and validated using the same methods.

BMP4 Differentiation in Unconfined Colonies

To generate unconfined colonies of a defined size, we first force aggregated hPSCs into 400x400mm PDMS microwell inserts (24-well plate sized, ~975 microwells/insert) using previously published protocols[25,52,53]. Briefly, PSCs were dissociated, resuspended in mTESR1 supplemented with ROCK inhibitor (10μM), seeded into the microwell inserts at a concentration of ~50-100cells/well, centrifuged at 200 rcf for 3 minutes, and left overnight to condense into aggregates. Next, the aggregates (~50-100 cells in size) were resuspended in mTESR1 supplemented with ROCK inhibitor (10μM) and transferred to Matrigel-coated 96 well plates at a concentration of approximately ~10 aggregates/well, where they were allowed to attach and form 2D colonies. After 24 hours, we removed ROCK inhibitor and the colonies were fed with mTESR1. mTESR1 supplemented with BMP4 (200μl/well, 50ng/ml, R&D Systems) was added another 24 hours later to start the differentiation. Unconfined colonies of a defined size were also generated using an alternative protocol. Briefly, dissociated hPSCs were seeded at 2 cells/mm^2, and fed with mTESR1 supplemented with ROCK inhibitor for 4 days, after which they were fed for 2 days with regular mTESR1 or until they reached an appropriate size (approximately 300-500μm in diameter), after which they were treated with BMP4 as described above.

Transwell Culture of hPSCs and FITC Diffusion Assay

Corning Costar Transwell plates with a 6.5 mm diameter and 3 μm pore size (Cat. # 07-200-147) were used. Transwell membranes were coated overnight with Matrigel. Prior to seeding, the Matrigel was removed and the membrane was rinsed 3X with PBS+/+ and then put into mTESR1 supplemented with ROCK inhibitor Y-27632 (10μM). Cells were then immediately seeded onto the transwell membranes at a density of 1,500cells/mm^2 (49,500cells/well). 24 hours later, ROCK inhibitor was removed, and the cells were fed with fresh mTESR1. 24 hours after ROCK inhibitor removal, the membranes were imaged on an EVOS fluorescence microscope at 10X to visualize whether the GFP labelled cellular nuclei reached confluence and were completely covering the membrane. We have previously determined that this protocol generates intact epithelia at this timepoint.

To visualize pSMAD1 activity in BMP4 stimulated transwells over time, BMP4 (50ng/ml) was added to either the apical (top) or basolateral (bottom) compartments of the transwell. The transwells were fixed at the appropriate time points by transferring the insert to a new 24well plate, rinsing with PBS, and fixing with 4% PFA.

To perform the FITC diffusion assay, we added FITC conjugated to 40-kDa dextran (Sigma-Aldrich) to the apical compartment and collected 10μl of media from basolateral compartment at various timepoints, which was mixed with 90μl of PBS and placed into a 96-well dark-sided plate. Fluorescence measurements were taken using a plate reader.

Immunofluorescent Staining

hPSCs were rinsed with PBS 1X, fixed in 4% paraformaldehyde (VWR) for 15 minutes, and subsequently washed 3X with PBS. The fixed cells were permeabilized and blocked in 0.3% Triton X-100 (Sigma Aldrich) and 5% normal donkey serum for an hour, and then incubated with primary antibodies overnight (also in 0.3% Triton, 5% normal donkey serum). The following day, samples were washed 3X with PBS and incubated with secondary antibodies in 0.3% Triton and 1% BSA at room temperature for 2 hours. Secondaries used were conjugated with Alexa 647, Alexa 405, and Alexa 555 (Life Technologies), used at a dilution of 1:400.

RNA Sequencing and Data Analysis

ZWT and ZKD cells were seeded sparsely onto standard culture 6-well plates in mTESR1 supplemented with ROCK inhibitor (10μM) and allowed to grow into colonies 300-500μm in size. At this time, cell lysate for the pluripotent condition was taken by putting 1.5mL RLT buffer/well for 3 minutes, and freezing this lysate at -80C for subsequent RNA extraction. Simultaneously, BMP4 (50ng/ml) was added to the remaining wells (BMP4+ condition). After 48 hours of BMP4 treatment, cell lysates were taken as described above. RNA extraction was performed using Qiagen's RNeasy kit, and samples were subsequently sent to Novogene for library preparation and sequencing (Illumina, PE150, 20M paired reads).

An RNA-seq data analysis pipeline was created using Snakemake (v3.13.3) with python (v3.6.10). Adapters were trimmed using trimmomatic (v0.36) in paired end mode with the following parameters: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Quality control was performed using FastQC (v0.11.9) and MultiQC (v1.9). Transcripts were quantified using Salmon (v0.14.2) with parameters --

validateMappings -l A with the GRCh38 reference transcriptome (downloaded from http://ftp.ensembl.org/pub/release-104/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz). Differential expression analysis was performed using the voom function in limma. Differential expression was called based on logFC significantly greater than 1 and adjusted p-value < 0.01. All raw sequencing data are available upon request.

Whole Genome Bisulfite Sequencing and Data Analysis

ZWT and ZKD cells were seeded and cultured as described in the RNA sequencing section. Only pluripotent samples were sent for sequencing. To do this, cells were dissociated using Accutase and resuspended in 200ul PBS + proteinase K, and frozen at -20C for subsequent DNA extraction. DNA extraction was performed using Qiagen's DNA extraction kit. Samples were subsequently sent to CD Genomics for whole genome bisulfite sequencing (Illumina, PE150, 250M paired reads). 1 μg of genomic DNA was fragmented by sonication to a mean size of approximately 200-400 bp. Fragmented DNA was end-repaired, 5'-phosphorylated, 3'-dA-tailed and then ligated to methylated adapters. The methylated adapter-ligated DNAs were purified using 0.8× Agencourt AMPure XP magnetic beads and subjected to bisulfite conversion by ZYMO EZ DNA Methylation-Gold Kit (zymo). The converted DNAs were then amplified using 25 μl KAPA HiFi HotStart Uracil+ ReadyMix (2X) and 8-bp index primers with a final concentration of 1 μM each. The constructed WGBS libraries were then analyzed by Agilent 2100 Bioanalyzer and quantified by a Qubit fluorometer with Quant-iT dsDNA HS Assay Kit (Invitrogen), and finally sequenced on Illumina Hiseq X ten sequencer.

A bisulfite sequencing data analysis pipeline was created using Snakemake (v3.13.3) with python (v3.6.10). Quality control was performed using FastQC (v0.11.9) and MultiQC (v1.9). Bisulfite analysis was performed using Bismark (v0.22.3) with bowtie2 (v2.3.4.1). First, bismark_genome_preparation --bowtie2 was run to bisulfite convert and index the genome. The reference genome was GRCh37, downloaded from: http://ftp.ensembl.org/pub/grch37/current/fasta/homo_sapiens/dna/Homo_sapiens.GRCh37.dna.primary_assembly.fa.gz. Then mapping was performed with bismark --bowtie2 --bam in paired end mode. Reads were deduplicated with deduplicate_bismark --bam. Finally, methylation calls were extracted with bismark_methylation_extractor. All raw sequencing data are available upon request.

Monolayer PGCLC Induction with BMP4

ZWT and ZKD cells were seeded in mTESR1 supplemented with ROCK inhibitor Y-27632 (10μM) into 96 well plates at a density ~100cells/mm^2. The following day, they were fed with 100μl of mTESR1. On day 2, they were induced with BMP4 (50ng/mL) in mTESR1, media with BMP4 was refreshed daily. At 48 and 72 hours after induction with BMP4, the cells were fixed prior to staining for PGCLC lineage markers.

Transwell PGCLC Induction with BMP4

Transwells were prepared as previously described. Cells were seeded onto the transwell membrane at a density of 500-1,000cells/mm^2 (16,600-33,200cells/well). Twenty-four hours later, ROCK inhibitor was removed, and the cells were fed with fresh mTESR1. Twenty-four hours after

ROCK inhibitor removal, BMP4 (1-50ng/mL) was added to both the apical (top) and basolateral (bottom) compartments. 48hours after BMP4 induction, the transwells were fixed prior to staining for PGCLC lineage markers. Prior to imaging, the transwell membrane was removed and mounted onto a glass coverslip. 10ng/mL BMP4 on transwells with a cell density of 750-1,000 cells/mm^2 was optimal for PGCLC induction.

Quantification and Statistical Analysis

Data in the figure panels are presented as mean and ± standard deviation. Comparisons between two groups were evaluated using two-tailed, unpaired Student's t-tests using GraphPad Prism 9 software. Significant differences are noted at $p < 0.05$, and range of statistical significance is shown by an asterisk within the figure panels: ns = not significant, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. For colony images (Figure 1D, 2D, 3C, 4A, 5A), immunofluorescence quantification includes results from three experiments, and at least three colonies were imaged from each experiment and timepoint (unless otherwise noted). For epithelial images in Figures 4E, 4G, S4A, S4C, immunofluorescent quantification includes results from three experiments, and one region was imaged for each timepoint. For epithelial images in Figure 5D, immunofluorescent quantification includes results from three experiments, and at least three randomized regions were imaged for each timepoint. For RNA sequencing, samples were collected from three experiments per experimental group (12 samples total: 3 ZWT pluripotent, 3 ZWT BMP4-treated, 3 ZKD pluripotent, 3 ZKD BMP4-treated). For whole genome bisulfite sequencing, samples were collected from three experiments per experimental group (6 samples total: 3 ZWT pluripotent, 3 ZKD pluripotent).

Code Availability

Quantification of marker expression at various radial distances from colony center was done using both CellProfiler and custom ImageJ scripts. For the CellProfiler pipeline (used in Figure 1D), to create a colony object with which to perform subsequent localization measurements, the nuclear LMNB1 channel was first thresholded, dilated, and holes were removed. LMNB1, pSMAD1, and ZO1 object intensity distributions were then measured across the colony object in radial slices each with a width of approximately 8μm. pSMAD1 and ZO1 fluorescence intensities were then normalized to LMNB1 signal. For the ImageJ pipeline (used in Figure 3C and 4B), the nuclear LMNB1 channel was first thresholded, dilated, and holes were removed. Stepwise erosion was performed and generated radial slices approximately 20um in diameter. Measurements of fluorescence intensity were taken from each slice for LMNB1, pSMAD1, and ZO1. pSMAD1 and ZO1 fluorescence intensities were then normalized to LMNB1 signal. Code for ImageJ scripts can be found at: https://github.com/ivanavasic/stepwise-erosion. Quantification of nuclei with positive marker expression (Figures 2D, 4F, 4H, S4B, S4D, 5E) was done using CellProfiler, using a default pipeline for percent positive marker expression[32]. Code for RNA-seq data analysis can be found at: https://github.com/amaslan/rna-seq-pipeline-ivana. Code for bisulfite sequencing data analysis can be found at: https://github.com/amaslan/bs-seq-pipeline-ivana.

Data Availability

RNA Sequencing data have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE213911. Whole genome bisulfite sequencing data have been deposited in the Sequence Read Archive (SRA) under accession number PRJNA880895.

Figures

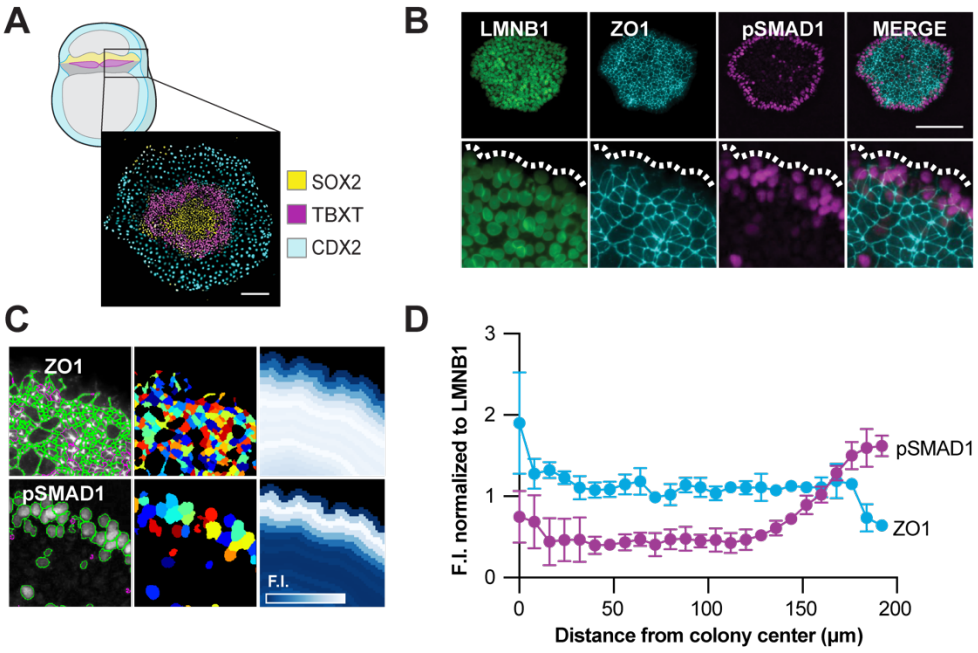

Figure 3.1: Unconfined hPSC colonies undergo radial gastrulation-like patterning and lose ZO1 on the colony edge.
1A: Unconfined circular colonies of hPSCs undergo radial patterning of gastrulation-associated markers after 48 hours of BMP4 stimulation. Scalebar depicts 200μm. 1B: IF image of the colony edge, showing regional loss of ZO1 which overlaps with pSMAD1 activity. Scalebar depicts 200μm. 1C: CellProfiler used to visualize and quantify expression of various proteins within the colony. 1D. Quantification of loss of ZO1/gain of pSMAD1 on the colony edge. N = 11 (4,4,3) colonies. Error bars depict standard deviation.

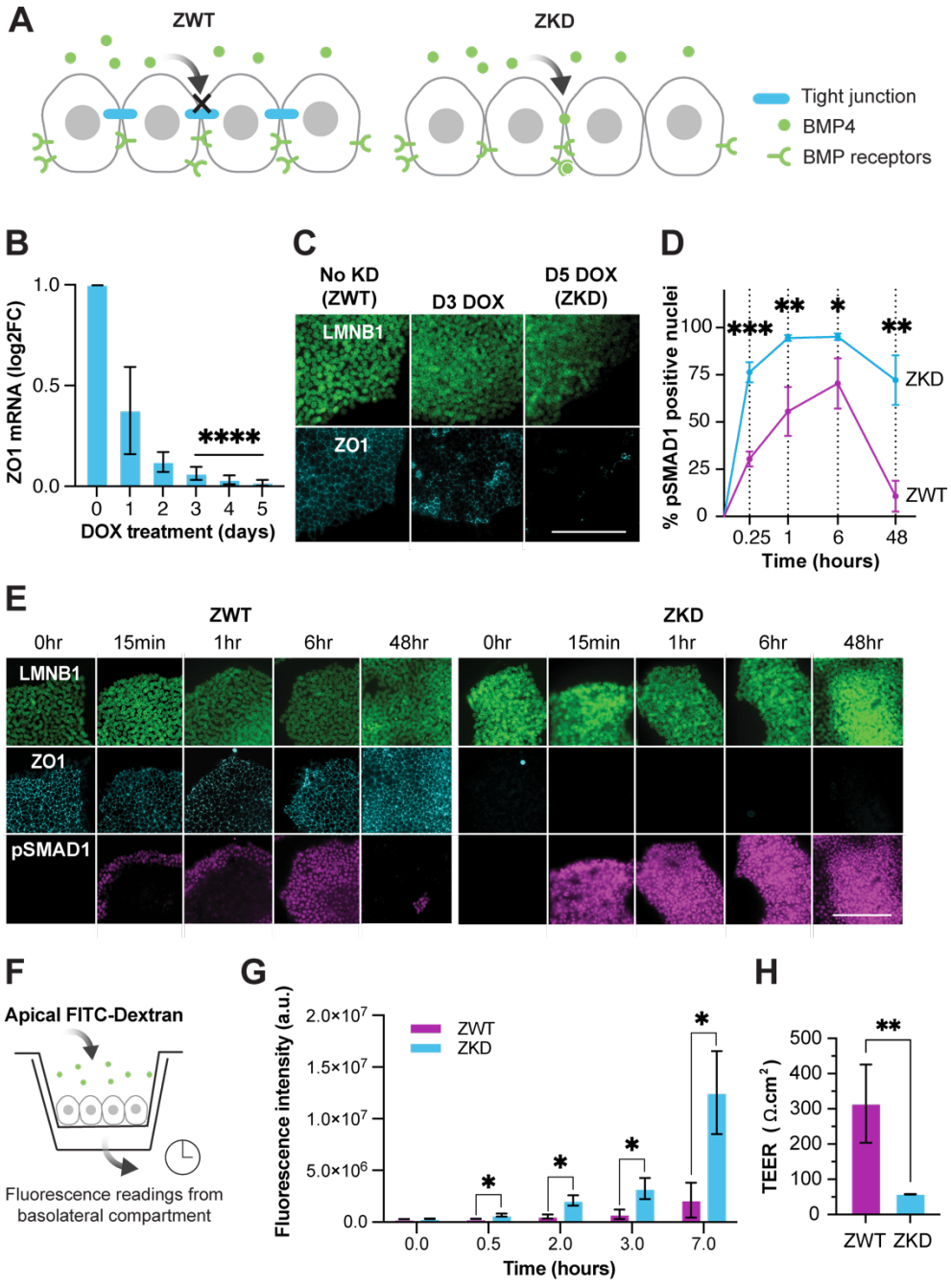
A
ZWT
ZKD
Tight junction
BMP4
BMP receptors
B
ZO1 mRNA (log2FC)
1.0
0.5
0.0
****
0 1 2 3 4 5
DOX treatment (days)
C
No KD (ZWT)
D3 DOX
D5 DOX (ZKD)
LMNB1
ZO1
D
% pSMAD1 positive nuclei
100
75
50
25
0
***
**
*
**
ZKD
ZWT
0.25 1 6 48
Time (hours)
E
ZWT
ZKD
0hr 15min 1hr 6hr 48hr 0hr 15min 1hr 6hr 48hr
LMNB1
ZO1
pSMAD1
F
Apical FITC-Dextran
Fluorescence readings from basolateral compartment
G
Fluorescence intensity (a.u.)
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
ZWT
ZKD
*
*
*
*
0.0 0.5 2.0 3.0 7.0
Time (hours)
H
TEER (Ω.cm^2)
500
400
300
200
100
0
**
ZWT ZKD

Figure 3.2: ZKD causes ubiquitous and sustained phosphorylation of SMAD1 throughout the colony.
2A: Schematic depicting z-direction view of a cellular monolayer, with apical presentation of BMP4. BMP receptors are sequestered in the basolateral domain by tight junctions, rendering them inaccessible to BMP4 ligands. 2B: qPCR showing expression of ZO1 over time with DOX treatment and subsequent induction of dCas9 expression in the WTC ZKD line, N = 5. 2C: IF images showing loss of ZO1 in the ZKD cell line over the course of 5 days of DOX treatment. 2D: Quantification of pSMAD1+ cells over time using CellProfiler nuclear segmentation and co-localization. Plot indicates significant increases in pSMAD1 activation in ZKD compared with ZWT cells at all timepoints. N = 119, [15min: ZKD (7,7,4) ZWT (6,5,4)], [1 hour: ZKD (5,6,6) ZWT (9,6,3)], [6 hour: ZKD (6,6,4), ZWT (8,6,4)], [48 hour: ZKD (3,3,3), ZWT (2,3,3)], error bars depict standard deviation. 2E: IF images showing sustained and ubiquitous phosphorylation of SMAD1 in ZKD cells over the course of 48 hours. Scalebar depicts 200μm. 2F-2G: Schematic of FITC-dextran diffusion assay. ZWT and ZKD cells are cultured in a transwell plate, 40kDa FITC is applied to the apical side, and fluorescence measurements are taken from the basolateral compartment over time. Plot indicates significant increases in diffusion of FITC-dextran in ZKD compared with ZWT monolayers at all timepoints. N = 9 transwells [ZKD (3), ZWT (6)], error bars depict standard deviation. 2H: TEER measurements in ZWT and ZKD monolayers, indicating increase in ion permeability in ZKD compared with ZWT monolayers. N = 8 transwells [ZKD (3), ZWT (5)].

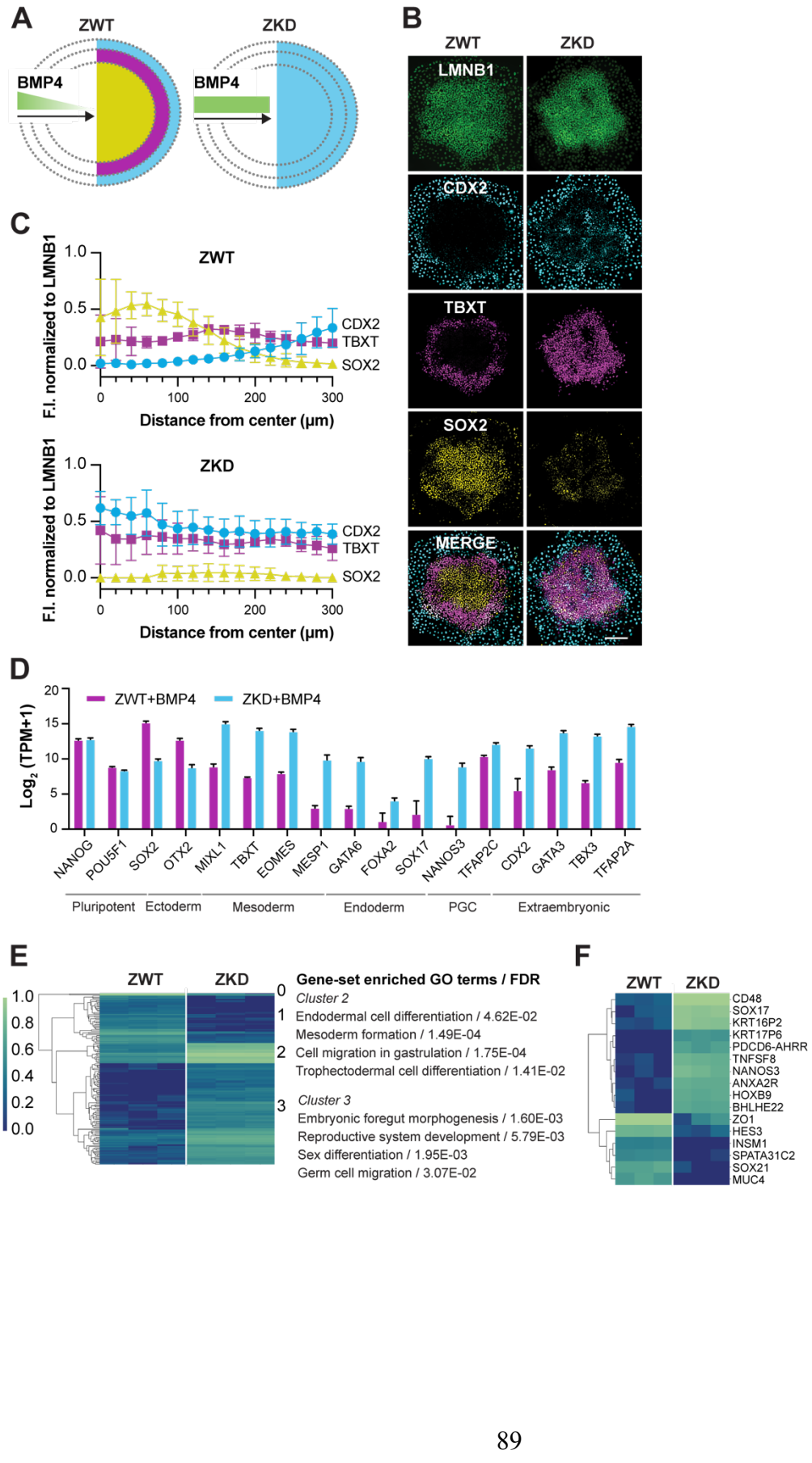
A
ZWT
ZKD
BMP4
BMP4
B
ZWT
ZKD
LMNB1
CDX2
TBXT
SOX2
MERGE
C
ZWT
F.I. normalized to LMNB1
1.0
0.5
0.0
CDX2
TBXT
SOX2
0
100
200
300
Distance from center (μm)
ZKD
F.I. normalized to LMNB1
1.0
0.5
0.0
CDX2
TBXT
SOX2
0
100
200
300
Distance from center (μm)
D
ZWT+BMP4
ZKD+BMP4
Log2 (TPM+1)
20
15
10
5
0
NANOG
POU5F1
SOX2
OTX2
MIXL1
TBXT
EOMES
MESP1
GATA6
FOXA2
SOX17
NANOS3
TFAP2C
CDX2
GATA3
TBX3
TFAP2A
Pluripotent
Ectoderm
Mesoderm
Endoderm
PGC
Extraembryonic
E
ZWT
ZKD
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
Gene-set enriched GO terms / FDR
Cluster 2
Endodermal cell differentiation / 4.62E-02
Mesoderm formation / 1.49E-04
Cell migration in gastrulation / 1.75E-04
Trophectodermal cell differentiation / 1.41E-02
Cluster 3
Embryonic foregut morphogenesis / 1.60E-03
Reproductive system development / 5.79E-03
Sex differentiation / 1.95E-03
Germ cell migration / 3.07E-02
F
ZWT
ZKD
CD48
SOX17
KRT16P2
KRT17P6
PDCD6-AHRR
TNFSF8
NANOS3
ANXA2R
HOXB9
BHLHE22
ZO1
HES3
INSM1
SPATA31C2
SOX21
MUC4

Figure 3.3: ZKD causes changes in patterning and proportion of somatic germ lineages.
3A: Reaction diffusion and positional information paradigm and prediction in ZWT (left) and ZKD (right) unconfined colonies. 3B: IF images of LMNB1, CDX2, TBXT, SOX2 in ZWT and ZKD colonies after 48 hours of stimulation with BMP4. Scalebar depicts 200μm. 3C: Quantification of fluorescence intensity of CDX2, TBXT, and SOX2 (from 3B) at various radial positions from the colony edge indicating loss of patterning phenotype in ZKD colonies. N = 17 colonies [ZKD (2,3,3), ZWT (3,3,3)]. 3D: RNA sequencing data showing expression of canonical gastrulation markers. TPM: transcripts per million, N = 6 [ZKD (3), ZWT (3)]. 3E: Unbiased clustering and GO analysis of top 150 differentially expressed genes between ZWT and ZKD cells. 3F: Unbiased clustering of top 16 differentially expressed genes between ZWT and ZKD cells highlighting increases in PGCLC related genes. Colorbar scale in 3E, 3F represents standardized scale of $Log_2(TPM+1)$ across each individual gene row.

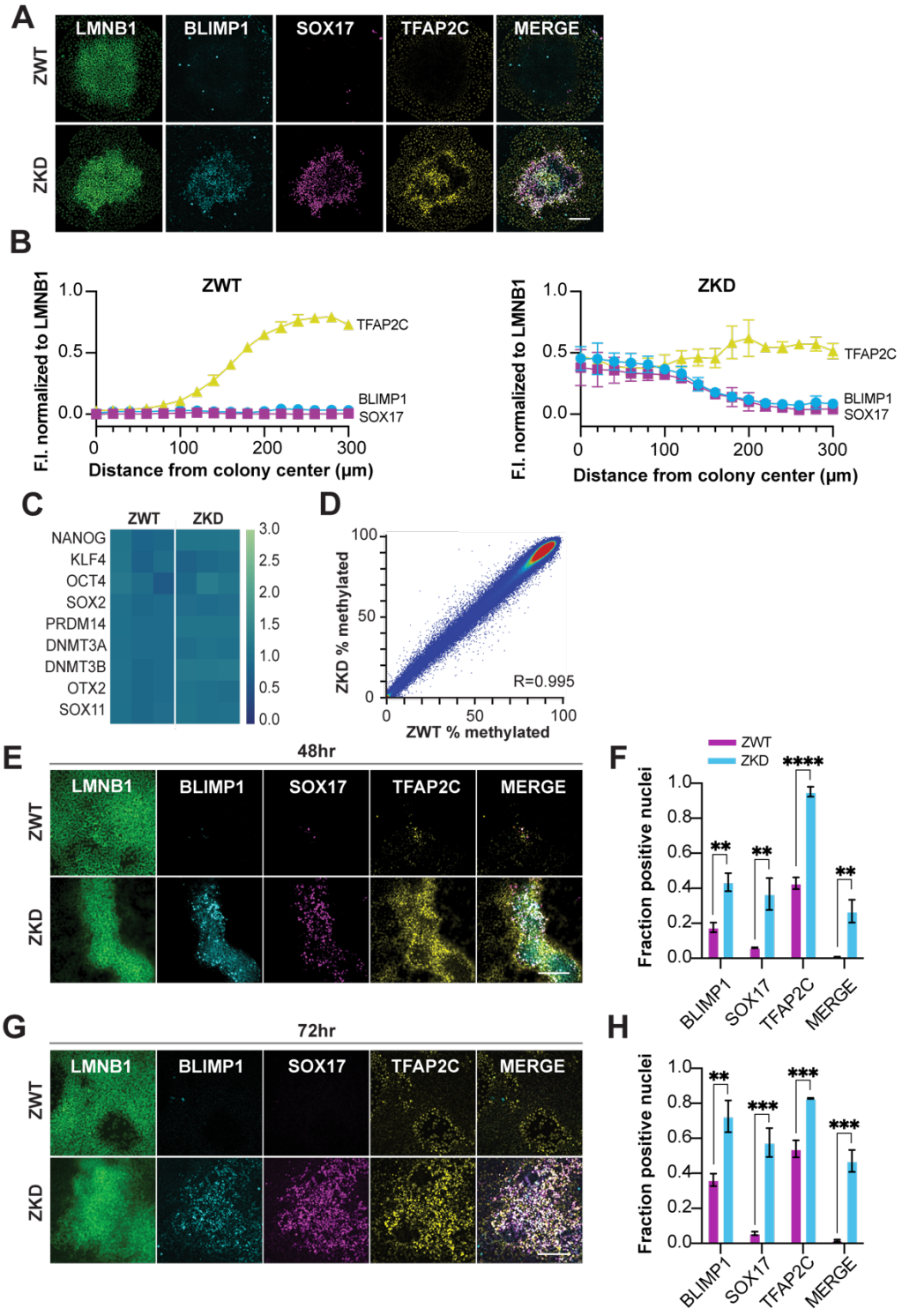
A
LMNB1
BLIMP1
SOX17
TFAP2C
MERGE
ZWT
ZKD
B
ZWT
ZKD
F.I. normalized to LMNB1
TFAP2C
BLIMP1
SOX17
Distance from colony center (μm)
C
ZWT
ZKD
NANOG
KLF4
OCT4
SOX2
PRDM14
DNMT3A
DNMT3B
OTX2
SOX11
D
ZKD % methylated
ZWT % methylated
R=0.995
E
48hr
F
ZWT
ZKD
Fraction positive nuclei
BLIMP1
SOX17
TFAP2C
MERGE
G
72hr
H

Figure 3.4: ZKD cells have a bias for PGCLC differentiation, without experiencing changes in pluripotency.
4A: IF images of LMNB1, BLIMP1, SOX17, TFAP2C, in ZWT and ZKD unconfined colonies. Scalebar depicts 200μm. 4B: Quantification of fluorescence intensity of BLIMP1, SOX17, TFAP2C at various radial positions from the colony edge N=18 [ZKD (3,3,3), ZWT (3,3,3)], error bars depict standard deviation. 4C: Expression of canonical pluripotency markers in ZWT and ZKD cells prior to BMP4 stimulation. Colorbar scale represents Log_2(TPM+1). 4D: Probe methylation levels between ZWT and ZKD cells gathered from whole genome bisulfite sequencing data (N=3). 4E,4G: IF images of LMNB1, BLIMP1, SOX17, TFAP2C in ZWT and ZKD cells in standard monolayer culture after 48 hours and 72 hours of stimulation with BMP4. Scalebar depicts 200μm. 4F,4H: Quantification of PGC marker expression in ZWT and ZKD cells. N=12 [48 hour: ZKD (3), ZWT (3)], [72 hour: ZKD (3), ZWT (3)], error bars indicate standard deviation.

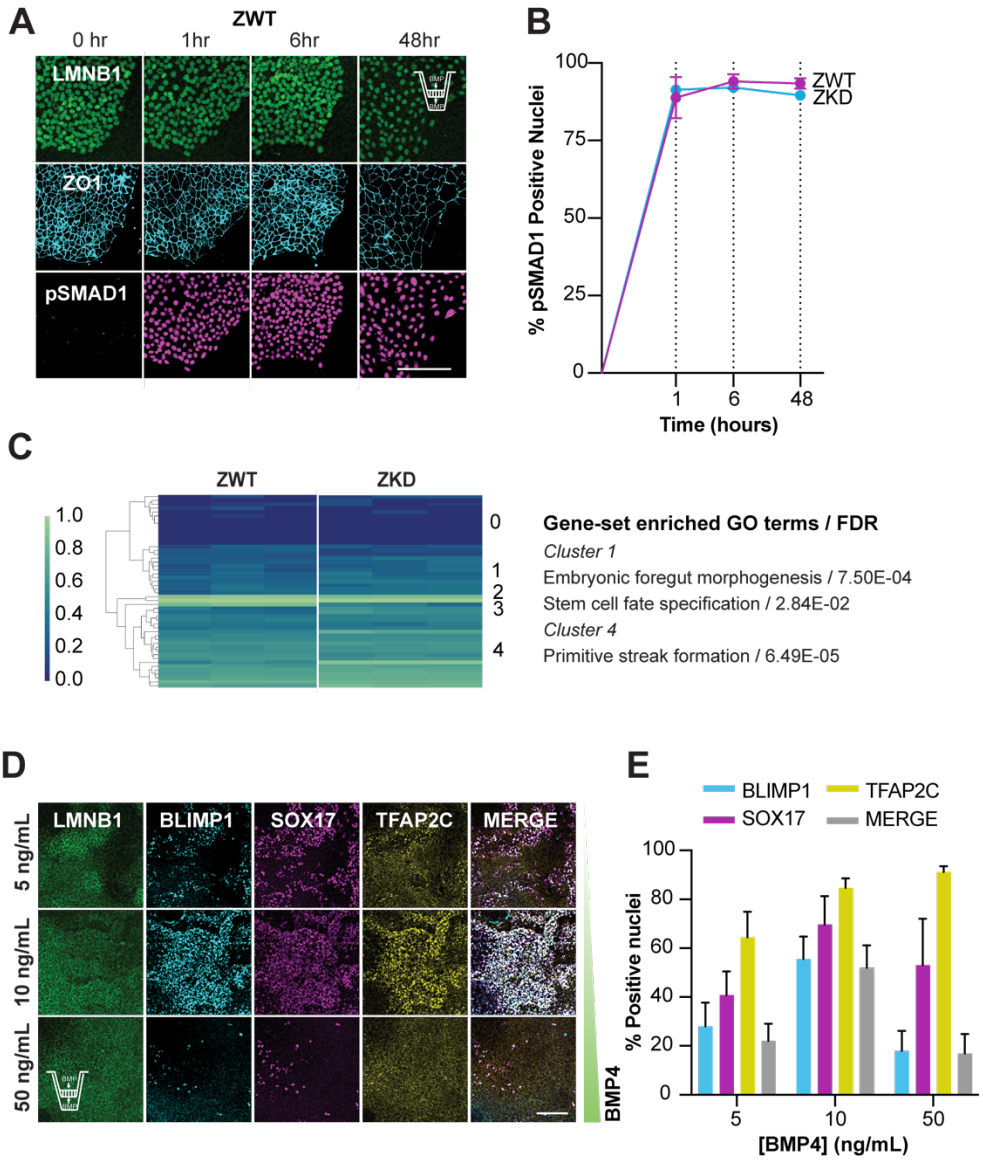

Figure 3.5: ZO1 KD-related PGCLC bias is a product of increased signaling.
5A: IF images of pSMAD1 after bi-directional BMP4 stimulation of ZWT unconfined colonies grown on transwell membranes from timepoints between 0-48 hours. 5B: Quantification of pSMAD1 positive cells in ZWT and ZKD unconfined colonies grown on transwells and stimulated with BMP4 for 0-48 hours. 5C: Unbiased clustering and GO analysis of all differentially expressed genes between ZWT and ZKD cells. Colorbar scale represents standardized scale of $Log_2(TPM+1)$ across each individual gene row. 5D: IF images of LMNB1, BLIMP1, SOX17, TFAP2C in ZWT and ZKD cells grown as a monolayer at a seeding density of 1,000cells/mm^2 on transwell membranes after 48 hours of bi-directional (apical and basolateral) stimulation with BMP4 (concentrations between 5-50ng/ml). Scalebar depicts 200µm. 5E: Quantification of % positive nuclei for each marker (BLIMP1, SOX17, TRAP2C, and triple positive MERGE). N = 29, [5ng/mL: (3,3,4)], [10ng/mL: (4,3,3)], [50ng/mL: (3,3,3)].

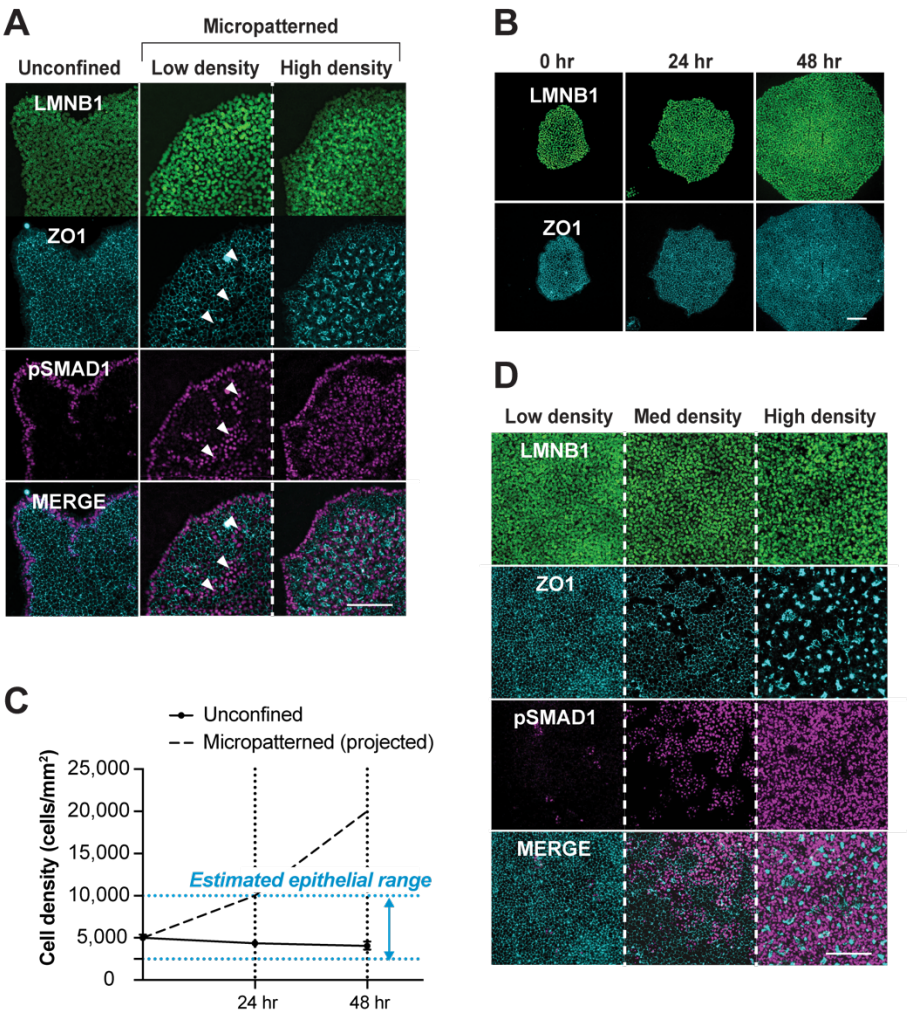

Figure S3.1: Micropatterned colonies experience disruptions in ZO1 expression as a result of regional increases in density .

S1A: Comparison between ZO1 and pSMAD1 expression in unconfined colonies vs low (150 cells/mm^2) and high (1000 cells/mm^2) seeding density micropatterned colonies, cultured 3 days prior to 1 hour BMP4 stimulation. Scalebar depicts 200µm. S1B: IF images of unconfined colonies over time showing maintenance of honeycomb ZO1 expression. Scalebar depicts 200µm. S1C: Density measurements in unconfined colonies, and projected density curve for micropatterned colonies (assuming density of 5,000 cells/mm2 upon induction with BMP4). Epithelial range, based on structure of cell-cell junction pattern, is estimated in the range between 3,000 – 10,000 cells/mm^2. Values estimated based on various seeding densities (partial data shown in S1D) that produce intact epithelium. S1D: ZO1 and pSMAD1 expression as a function of cell density in monolayer culture. Low (1000 cells/mm^2), med (2500 cells/mm^2), and high (5000 cells/mm^2), seeding density cultured 2 days. Epithelial structure (honeycomb cell-cell junction pattern) is lost and pSMAD1 activation is increased as cell density increases. Scalebar depicts 200µm.

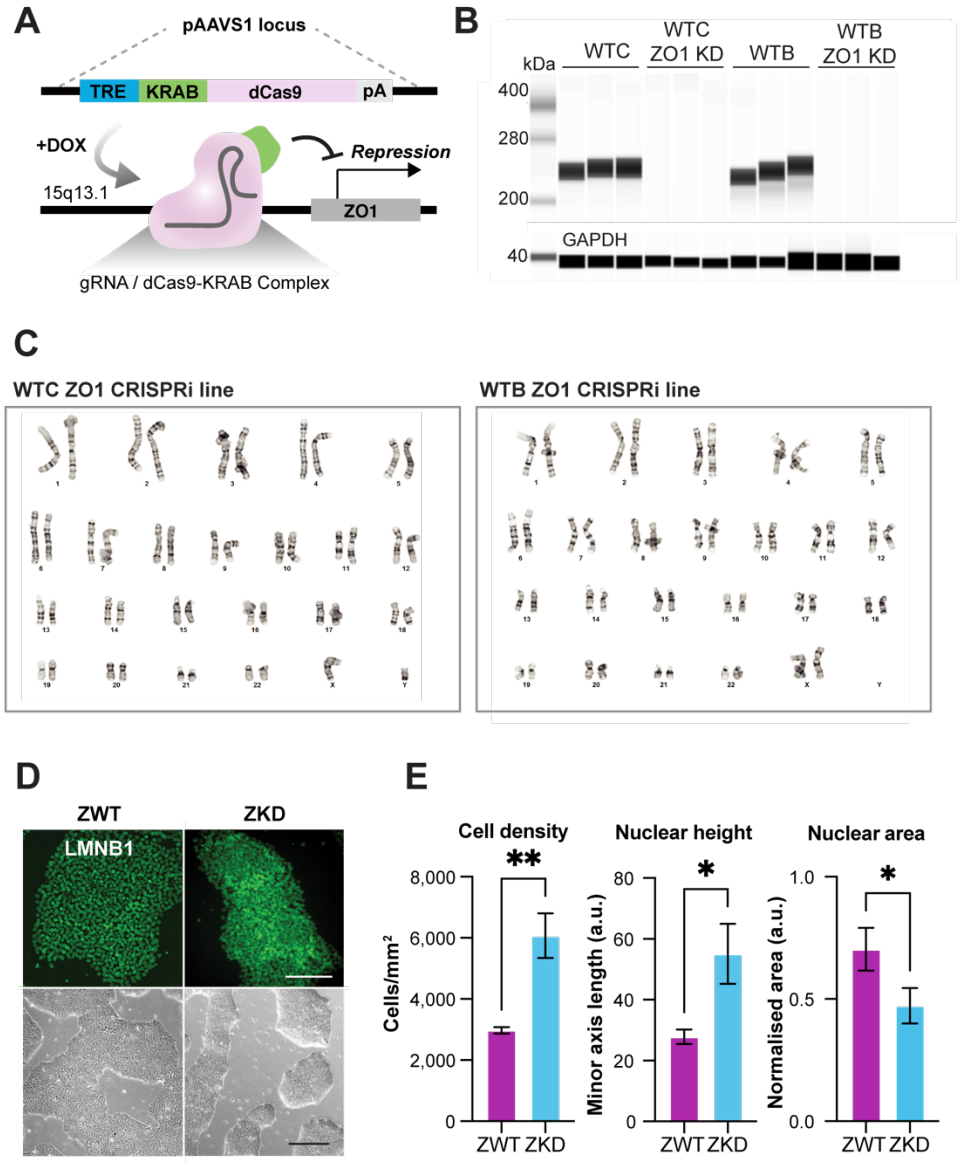

Figure S3.2: CRISPRi knockdown of ZO1 cells leads to changes in colony and cell morphology

S2A: CRISPRi construct used for ZO1 KD. TRE: Tet responsive element, KRAB: Krüppel associated box, pA: polyA signal. S2B: Western blot showing protein loss in the ZKD cell lines. S2C: Karyotype of the ZKD cell line in both the WTC and WTB lines. S2D: IF (top) and brightfield images (bottom) showing morphological differences between ZWT and ZKD cells. Scalebar depicts 200µm and 500µm respectively. S2E: Quantification of changes in cell density, nuclear height, and nuclear area between ZWT and ZKD cells. N = 4 [ZKD (3), ZWT (2)].

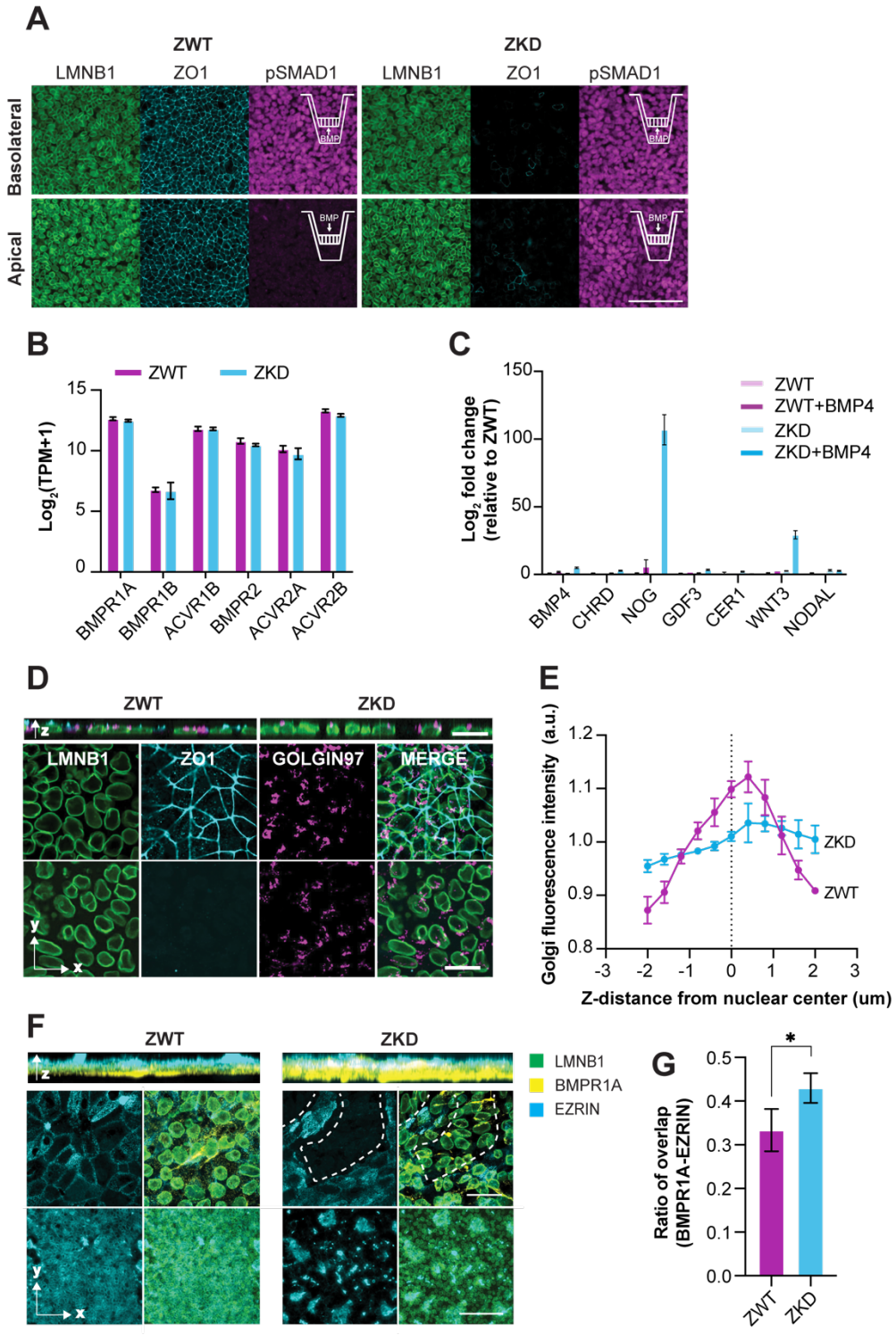
A
ZWT
ZKD
LMNB1
ZO1
pSMAD1
Basolateral
Apical
BMP
B
ZWT
ZKD
Log2(TPM+1)
BMPR1A
BMPR1B
ACVR1B
BMPR2
ACVR2A
ACVR2B
C
Log2 fold change (relative to ZWT)
ZWT
ZWT+BMP4
ZKD
ZKD+BMP4
BMP4
CHRD
NOG
GDF3
CER1
WNT3
NODAL
D
ZWT
ZKD
LMNB1
ZO1
GOLGIN97
MERGE
E
Golgi fluorescence intensity (a.u.)
Z-distance from nuclear center (um)
ZKD
ZWT
F
ZWT
ZKD
LMNB1
BMPR1A
EZRIN
G
Ratio of overlap (BMPR1A-EZRIN)
*
ZWT
ZKD

Figure S3.3: Loss of ZO1 induces changes to directional receptivity to morphogen signaling, and results in changes to polarity of membrane-bound proteins.
S3A: IF images of pSMAD1 after basolateral (top row) and apical (bottom row) BMP4 stimulation for 1 hour in ZWT and ZKD cells grown on transwell membranes. Scalebar depicts 100μm. S3B: BMP receptor expression from RNA sequencing data. S3C: qPCR of secreted morphogens at 12 hours of BMP4 stimulation, N = 12 [ZKD (3), ZKD+BMP4 (3), ZWT (3), ZWT+BMP4 (3)]. S3D: IF images labeling Golgi in ZWT and ZKD cells. Scalebar depicts 20μm. S3E: Quantification of Golgi position relative to the nucleus. S3F: IF image showing punctate Ezrin expression in ZKD cells, top scalebar: 20μm, bottom scalebar: 200 μm. S3G: Quantification of overlap between Ezrin and BMPR1A.

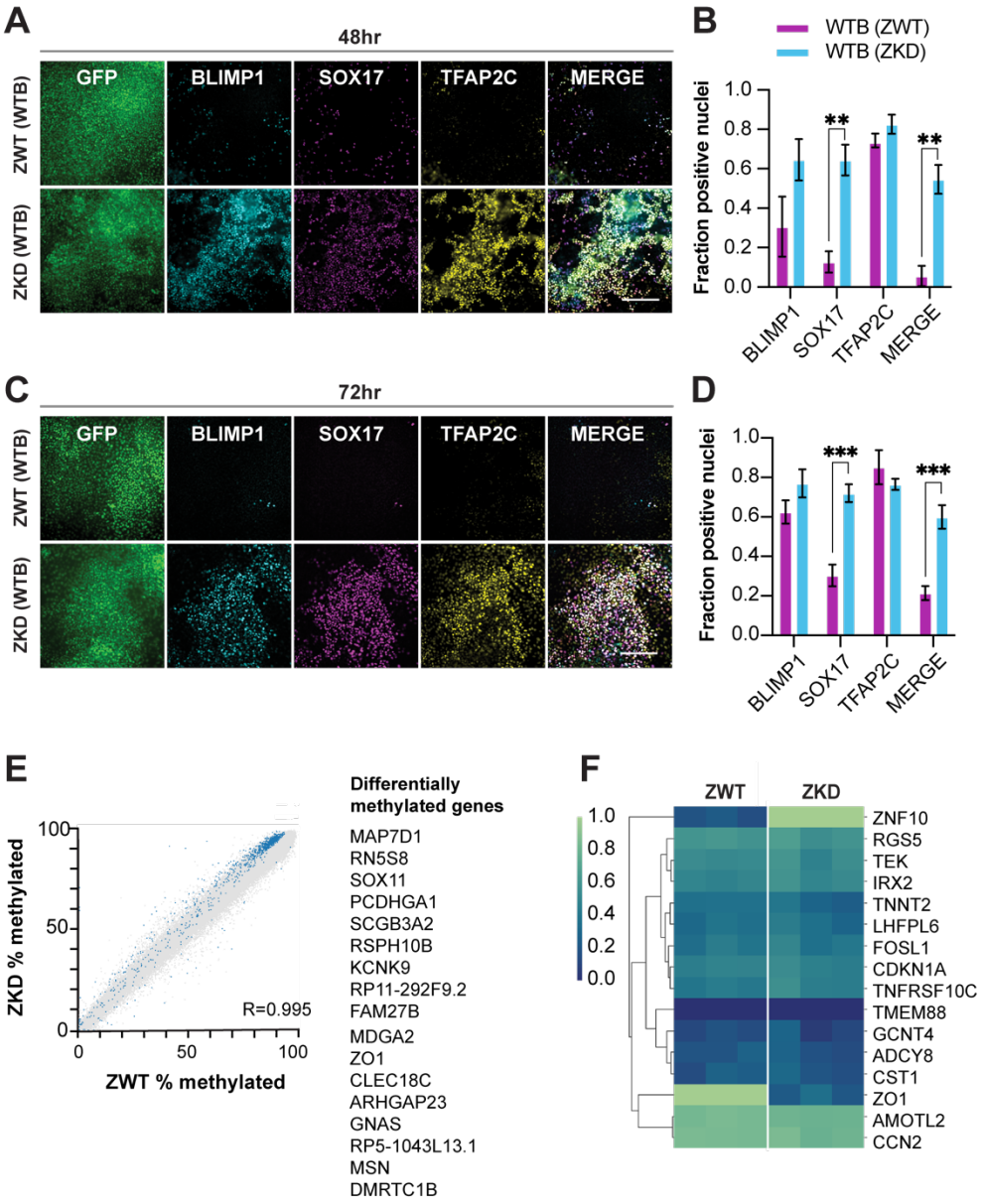

Figure S3.4: WTB-ZKD cells are predisposed to PGCLC differentiation.
S4A,S4C: IF images of GFP (nuclear label), BLIMP1, SOX17, TFAP2C in WTB ZWT and ZKD cells in standard monolayer culture after 48 hours and 72 hours of stimulation with BMP4. Scalebar depicts 200µm. S4B,S4D: Quantification of PGC marker expression in WTB ZWT and ZKD cells. N=12 [48 hour: ZKD (3), ZWT (3)], [72 hour: ZKD (3), ZWT (3)], error bars indicate standard deviation. S4D: Probe methylation levels between ZWT and ZKD cells gathered from whole genome bisulfite sequencing data, probes with significant differences in methylation are labeled in blue (p-values<0.05). S4F: Unbiased clustering of top 16 differentially expressed genes between ZWT and ZKD cells in the pluripotent condition. Colorbar scale represents standardized scale of $Log_2(TPM+1)$ across each individual gene row.

References

1. Tam, P. P. L. & Loebel, D. A. F. Gene function in mouse embryogenesis : get set for gastrulation. *Nat. Rev.* **8**, 368–381 (2007).
2. Rossant, J. & Tam, P. P. L. Blastocyst lineage formation , early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701–713 (2009).
3. Tam, P. P. L. & Behringer, R. R. Mouse gastrulation : the formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25 (1997).
4. Farquhar, M. G. & Palade, G. Junctional complexes in various epithelia. *J. Cell Biol.* **17**, 375–412 (1963).
5. Claude, P. & Goodenough, D. A. Fracture faces of zonulae occludentes from 'tight' and 'leaky' epithelia. *J. Cell Biol.* **58**, 390–400 (1973).
6. Zihni, C., Mills, C., Matter, K. & Balda, M. S. Tight junctions : from simple barriers to multifunctional molecular gates. *Nat. Rev.* **17**, 564–580 (2016).
7. Murphy, S. J. *et al.* Differential Trafficking of Transforming Growth Factor-Receptors and Ligand in Polarized Epithelial Cells. *MBoC* **15**, 2853–2862 (2004).
8. Yin, X. *et al.* Basolateral delivery of the type I transforming growth factor beta receptor is mediated by a dominant-acting cytoplasmic motif. *MBoC* **28**, 2701–2711 (2017).
9. Zhang, Z., Zwick, S., Loew, E., Grimley, J. S. & Ramanathan, S. Mouse embryo geometry drives formation of robust signaling gradients through receptor localization. *Nat. Commun.* **10**, (2019).
10. Phan-Everson, T. *et al.* Differential compartmentalization of BMP4/NOGGIN requires NOGGIN trans-epithelial transport. *Dev. Cell* **56**, 1930–1944 (2021).

11. Xiang, L. *et al.* A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, (2020).

12. Turing, A. M. The chemical basis of morphogenesis. *Philos. Trans. R. Soc.* 37–72 (1952). doi:https://doi.org/10.1098/rstb.1952.0012

13. Wolpert, L. Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* **25**, 1–47 (1969).

14. Green, J. B. A. & Sharpe, J. Positional information and reaction-diffusion: Two big ideas in developmental biology combine. *Development* **142**, 1203–1211 (2015).

15. Schier, A. F. Nodal morphogens. *Cold Spring Harb. Perspect. Biol. Biol.* **1**, (2009).

16. Müller, P. *et al.* Differential diffusivity of nodal and lefty underlies a reaction-diffusion patterning system. *Science (80-.).* **336**, 721–724 (2012).

17. Etoc, F. *et al.* A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Dev. Cell* **39**, 302–315 (2016).

18. Tewary, M. *et al.* A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning. *Development* dev.149658 (2017). doi:10.1242/dev.149658

19. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* **11**, 847–854 (2014).

20. Deglincerti, A. *et al.* Self-organization of human embryonic stem cells on micropatterns. *Nat. Protoc.* **11**, 2223–2232 (2016).

21. Ben-Haim, N. *et al.* The Nodal Precursor Acting via Activin Receptors Induces Mesoderm by Maintaining a Source of Its Convertases and BMP4. *Dev. Cell* **11**, 313–323 (2006).

22. Arnold, S. J. & Robertson, E. J. Making a commitment: Cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10**, 91–103 (2009).

23. Fanning, A. S. & Anderson, J. M. Zonula Occludens-1 and -2 Are Cytosolic Scaffolds That Regulate the Assembly of Cellular Junctions. *Ann. N. Y. Acad. Sci.* **1165**, 113–120 (2009).

24. Mcneil, E., Capaldo, C. T. & Macara, I. G. Zonula Occludens-1 Function in the Assembly of Tight Junctions in Madin-Darby Canine Kidney Epithelial Cells □. *MBoC* **17**, 1922–1932 (2006).

25. Libby, A. R. G. *et al.* Spatiotemporal mosaic patterning of pluripotent stem cells using CRISPR interference. *Elife* **7**, 1–23 (2018).

26. Joy, D. A., Libby, A. R. G. & McDevitt, T. C. Deep neural net tracking of human pluripotent stem cells reveals intrinsic behaviors directing morphogenesis. *Stem Cell Reports* **16**, 1317–1330 (2021).

27. Gunne-Braden, A. *et al.* GATA3 Mediates a Fast, Irreversible Commitment to BMP4-Driven Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell* **26**, 693–706 (2020).

28. Nallet-Staub, F. *et al.* Cell Density Sensing Alters TGF-β Signaling in a Cell-Type-Specific Manner, Independent from Hippo Pathway Activation. *Dev. Cell* **32**, 640–651 (2015).

29. Smith, Q. *et al.* Cytoskeletal tension regulates mesodermal spatial organization and subsequent vascular fate. *PNAS* **115**, 8167–8172 (2018).

30. Manfrin, A. *et al.* Engineered signaling centers for the spatially controlled patterning of human pluripotent stem cells. *Nat. Methods* **16**, 640–648 (2019).

31. Kim, Y. *et al.* Cell position within human pluripotent stem cell colonies determines apical specialization via an actin cytoskeleton-based mechanism. *Stem Cell Reports* **17**, 68–81 (2022).

32. Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 1–11 (2021).

33. Krtolica, A. *et al.* Disruption of Apical-Basal Polarity of Human Embryonic Stem Cells Enhances Hematoendothelial Differentiation. *Stem Cells* **25**, 2215–2223 (2007).

34. Mandegar, M. A. *et al.* CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell* **18**, 541–553 (2016).

35. Paine-Saunders, S., Viviano, B. L., Economides, A. N. & Saunders, S. Heparan sulfate proteoglycans retain Noggin at the cell surface. A potential mechanism for shaping bone morphogenetic protein gradients. *J. Biol. Chem.* **277**, 2089–2096 (2002).

36. Granés, F., Ureña, J. M., Rocamora, N. & Vilaró, S. Ezrin links syndecan-2 to the cytoskeleton. *J. Cell Sci.* **113**, 1267–1276 (2000).

37. Yadav, S., Puri, S. & Linstedt, A. D. A Primary Role for Golgi Positioning in Directed Secretion , Cell Polarity , and Wound Healing. *MBoC* **20**, 1728–1736 (2009).

38. Rodriguez-Boulan, E. & Macara, I. G. Organization and execution of the epithelial polarity programme. *Nat. Rev. Mol. Cell Biol.* **15**, 225–242 (2014).

39. Irie, N. *et al.* SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253–268 (2015).

40. Sasaki, K. *et al.* Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. *Cell Stem Cell* **17**, 178–194 (2015).

41. Muncie, J. M. *et al.* Mechanical Tension Promotes Formation of Gastrulation-like Nodes and Patterns Mesoderm Specification in Human Embryonic Stem Cells. *Dev. Cell* 1–16 (2020). doi:10.1016/j.devcel.2020.10.015

42. Sheng, G. Epiblast morphogenesis before gastrulation. *Dev. Biol.* **401**, 17–24 (2015).

43. Kinoshita, M. *et al.* Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* **28**, 1–19 (2021).

44. Guo, G. *et al.* Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437–446 (2016).

45. Takashima, Y. *et al.* Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).

46. Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).

47. Larripa, K. & Gallegos, A. A mathematical model of Noggin and BMP densities in adult neural stem cells. *Lett. Biomath.* **4**, 1–22 (2017).

48. Jones, C. M. & Smith, J. C. Establishment of a BMP-4 Morphogen Gradient by Long-Range Inhibition. *Dev. Biol.* **194**, 12–17 (1998).

49. Farin, H. F. *et al.* Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**, 340–354 (2016).

50. Lawson, K. A. *et al.* Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424–436 (1999).

51. Arnold, S. J., Maretto, S., Islam, A., Bikoff, E. K. & Robertson, E. J. Dose-dependent Smad1, Smad5 and Smad8 signaling in the early mouse embryo. *Dev. Biol.* **296**, 104–118 (2006).

52. Hookway, T. A., Butts, J. C., Lee, E., Tang, H. & McDevitt, T. C. Aggregate formation and suspension culture of human pluripotent stem cells and differentiated progeny. *Methods* **101**, 11–20 (2016).

53. Ungrin, M. D., Joshi, C., Nica, A., Bauwens, C. & Zandstra, P. W. Reproducible, ultra high-throughput formation of multicellular organization from single cell suspension-derived human embryonic stem cell aggregates. *PLoS One* **3**, (2008).

Chapter 4. Conclusion: Cell structure underlies morphogenesis within the embryo

In this dissertation, I showcase two examples which highlight the importance of cell structure in guiding tissue morphogenesis in models of early embryonic development. Cellular self-organization and patterning phenotypes in both studies rely on interactions between two subpopulations of cells with different structural properties. Within the embryo, these interactions stem organically from an initial symmetry breaking event, where a subpopulation of cells spontaneously emerges with different structural characteristics than the original population.

Cell-cell adhesion drives self-organization and specification

In Chapter 2, symmetry breaking events that eventually result in separation of the inner cell mass and the trophoblast are modelled via mixing of WT and CDH1- cells. Following this study, other groups subsequently demonstrated that hPSCs are capable of similar sophisticated self-organization behaviors with minimal external cues – but necessarily prompted by mixing with a subpopulation with different characteristics. For example, in generating synthetic mouse embryos, Tarazi et al. and Bao et al. demonstrate that mixing of pre-differentiated trophectoderm, primitive endoderm, and naïve pluripotent cells into embryoid bodies results in coordinated and complex morphogenic events that resemble gastrulation and formation of the neural tube, foregut, somites, and a beating heart structure[1,2]. In another study, Minn et al. demonstrate that a heterogeneous and well-mixed population of stem cells expressing germ layer markers T, SOX2, and CDX2 will spatially re-group after seeding[3]. Single cell RNA sequencing analysis shows that each germ lineage expresses different combinations of adhesion markers, and implicates these differences in driving sorting phenotypes. By contrast, our study provides compelling evidence that differences in adhesion can solely drive self-organization and even specification bias; whereas

before it was assumed that expression of transcription factor identity markers initiates cell fate changes and accompanying structural changes that are important for migration and organization.

Epithelial cell structure regulates receptivity to morphogen signaling

In Chapter 3, mixing of cell types is not required to mimic a symmetry breaking event. Instead, we see that regional loss of ZO1 expression on the border of monolayer hPSC colonies renders cells on the edge of colonies more receptive to signaling by BMP4 and results in pSMAD1 activation in these cells. By knocking down ZO1 in all hPSCs within a colony, we show that we can remove this initial asymmetry and render all cells sensitive to soluble BMP4 signals in the media. These experiments demonstrate that this initial symmetry breaking event is required for gradient formation within the colony and radial patterning of diverse germ layers. Our study, which definitively demonstrates that expression of ZO1 attenuates response to morphogenic signaling, has important implications for how gradients are shaped within the embryo and how the embryo might regulate organized cell type emergence via regional loss or weakening of tight junctions. For example, emergence of PGCs occurs in a unique and spatially conserved location in mouse and cynomolgus monkey embryos[4,5]: on the border between epiblast and BMP4-secreting extraembryonic tissues. Studies in mouse show that there is a break in tight junction expression between the epiblast and the extraembryonic ectoderm[6]. Our study, which demonstrates that PGC specification relies on sustained pSMAD1 activation, provides both a link between these two publications and an explanation for the conserved spatial location of PGC emergence.

References

1. Tarazi, S. *et al.* Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. *Cell* **185**, 3290-3306.e25 (2022).
2. Bao, M. *et al.* Stem cell-derived synthetic embryos self-assemble by exploiting cadherin codes and cortical tension. *Nat. Cell Biol.* **24**, 1341–1349 (2022).
3. Minn, K. T. *et al.* High-resolution transcriptional and morphogenetic profiling of cells from micropatterned human ESC gastruloid cultures. 1–34 (2020).
4. Irie, N., Tang, W. W. C. & Azim Surani, M. Germ cell specification and pluripotency in mammals: a perspective from early embryogenesis. *Reprod. Med. Biol.* **13**, 203–215 (2014).
5. Sasaki, K. *et al.* The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. *Dev. Cell* **39**, 169–185 (2016).
6. Zhang, Z., Zwick, S., Loew, E., Grimley, J. S. & Ramanathan, S. Mouse embryo geometry drives formation of robust signaling gradients through receptor localization. *Nat. Commun.* **10**, (2019).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Ivana Vasic
DDE73B08C3774E5...
Author Signature

12/6/2022
Date