2019

The Role and Control of WNT Signalling in an HESC Model of Human Primitive Streak

Iain Martyn

THE ROLE AND CONTROL OF WNT SIGNALLING IN AN HESC MODEL OF HUMAN PRIMITIVE STREAK

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Iain Martyn

June 2019

THE ROLE AND CONTROL OF WNT SIGNALLING IN AN HESC MODEL OF HUMAN PRIMITIVE STREAK

Iain Martyn, Ph.D.

The Rockefeller University 2019

In amniotes, gastrulation is marked by the creation of the primitive streak (PS) and is largely controlled by WNT, BMP, and ACTIVIN/NODAL signalling. Despite detailed characterization in model organisms, the human PS and the role these pathways play in its formation and patterning remains a mystery.

In this work I focused on understanding the role and control of the WNT pathway in human PS development. Due to the ethical limitations of working with human embryos, I used an *in vitro* human embryonic stem cell (hESC) micropatterned "gastruloid" system. I first showed that in the human PS there is a conserved BMP → WNT→ NODAL signalling initiation hierarchy, and that WNT is necessary and sufficient for PS formation. Next, I found that structured subpopulations of endoderm and mesoderm emerge and self-organize depending on different BMP, WNT, and ACTIVIN/NODAL levels, and that by comparison to the mouse embryo I could arrange these subpopulations along an anterior-posterior axis. With the development of a new cell tracking technique, I was also able to identify and characterize robust cell migrations from the PS region of each gastruloid that depended on which fates the cells would ultimately adopt. Putting these pieces together, I was able to derive a rudimentary first fate map of the human PS, as well as a rough picture of the BMP, WNT, and ACTIVIN/NODAL signalling gradients that determine it.

One interesting and unforeseen result from this fate map was the hint of a human "organizer" cell fate that emerged under joint WNT and ACTIVIN/NODAL stimulation. To characterize and functionally prove this organizer's existence, I devised an *ex ovo* cross-species transplantation strategy grafting treated gastruloids into chick embryos. The assay demonstrated that the human cells induce and contribute autonomously to a secondary axis while inducing neural fate in the host, thus fulfilling the most stringent criteria for an organizer. This work adds an important milestone to the research program begun in 1924 with the first famous organizer experiment of Hilde Mangold and Hans Spemann, and the methods I developed have opened a door to new functional explorations and tests of early human development.

Having learned more about the role of WNT in determining cell fates in the gastruloid model, I next endeavoured to understand how the spatial extent and duration of the WNT signal itself was controlled. With the use of various CRISPR/Cas9 knockout lines, I discovered that DKK1 and E-CADHERIN were the two dominant factors, with E-CADHERIN transducing boundary forces to focus WNT signalling to colony border at early times, and DKK1 controlling the late WNT pattern via cell non-autonomous negative feedback. With the help of time-lapse imaging of a fluorescent reporter line and mathematical modelling, I showed that these two factors mediate a wave of WNT signalling that spreads across the tissue to be patterned, and that this wave is a generic property of a bistable system and thus likely generalizable to other instances in development.

While limited by the use of hESCs, taken together my findings provide a first glimpse into the role and control of WNT signalling early on in our own, human development.

This thesis is dedicated to my family, especially to my Aunt Linda

ACKNOWLEDGMENTS

I have had the good fortune of working under and learning from not one great mentor but two. From the beginning Eric and Ali have not only given me all the support, advice, and attention to my development as a scientist I could dream of, but also always the freedom and space to explore on my own. From Ali I especially take away lessons in lab leadership and team-building, the belief that beautiful science should also be beautifully argued and presented, and the excitement and inclination to focus on the big questions and the boundary pushing experiments. From Eric I especially take away the value of meticulousness at every level of an experiment, the art of getting to the point, and the importance of listening to others and identifying exactly what it is that you (or they) do not know and are trying to answer.

I would also like to thank my dissertation committee members past and present for their support and guidance over the years. To Marc Tessier-Lavigne I am especially thankful for first introducing me to Rockefeller via his commencement speech at my undergraduate graduation ceremony. To Jim Hudspeth I am especially thankful for his attention to detail and his encyclopedic knowledge in fields far removed that helped to inspire new experiments. To Shai Shaham I am especially grateful stepping in to replace Marc and for the lessons in shrewd, calm creativity in tackling a research question, especially in the first-year class I took under him.

Other Rockefeller faculty I am indebted to for many useful discussions on many different topics that helped my thinking and development as a scientist include Marcelo Magnasco, Stan Liebler, and Jessie Ausubel. Outside of Rockefeller I was fortunate to have Kat Hadjantonakis and her lab

as a close neighbour and collaborator at Sloan Kettering, and to learn directly from the best embryologists in the world at the infamous embryology course at Woods Hole.

Aside from mentors at the faculty level, I wish to also thank the many post-doc mentors I have learned everything else I know about stem cell and developmental biology from. Thank you to Gist Croft, Albert Ruzo, Tomo Haremaki, Aryeh Warmflash, Benoit Sorre, and Zeeshan Ozair. I wish to especially thank my co-partner in "chuman" crime, Tatiane Kanno, who lightened every otherwise gruelling hour sacrificing egg after egg for human-chick chimeras with her limitless Brazilian optimism and music. To other past and present members of the Brivanlou lab, including Lauren Gerber, Anna Yoney, Mijo Simunovic, Danielle Little, Shu Li, Tiago Rito, Christian Markopoulos, Sam Khodursky, Tien Phan, Adam Marks, Caroline Lara, Jeremy Auerbach, Chris You, Jakob Metzger, Szilvia Galgoczi, Fred Etoc, Michael Heke, and Rohan Soman, thank you for your friendship, your advice and your criticisms, and for making the lab enjoyable and memorable each and every day.

On the administrative side I was constantly amazed at the daily magic performed by Peter Ingrassia, Jean-Marx Santel, Michael Heke, Shawn Boggs, Adam Souza, and Melanie Lee at making the necessary reagents or purchase orders appear. If only every lab could be so lucky. To Melanie Lee thank you also especially for your advice when I was starting out, and for making the Physics Center such a welcoming place.

Everywhere you turn at Rockefeller there is helpful person who is generous and happy to help you with your experiments, but in the case of my studies I am particularly thankful for the wonderful staff at the Bio-Imaging Resource Center, the Precision Fabrication Facility, and the Deans Office. Thank you to Alison North, Pablo Ariel, Kaye Thomas, and Christina Pyrgaki your professionalism and patience. Thank you to Kunal Shah and Jim Petrillo for your friendship and enthusiasm for my projects. Thank you to Cris Rosario, Marta Delgado, Kristen Cullen, Stephanie Fernandez, Sid Strickland and Emily Harms for being the stable and perennially supportive focal point of the graduate program. Thank you also to Emily for taking the time to meet with me on my first introductory trip to New York and to introduce me to Rockefeller.

To my friends at Rockefeller and beyond, thank you for a life for me outside of science.

Finally, and most importantly, thank you to my family, who not only supported me all these years but were the foundation that got me here in the first place. To my aunt and uncle Linda and Brian, thank you for teaching me the value of hard work and perseverance and for raising me to think both practically and radically. To my brothers, thank you for putting up with me. To my father, who would have thought high school science fairs would take me this far? To my mother, who heard about this fantastic place called Rockefeller and just knew, long before I did, that it was the right place for me: you were absolutely right.

STATEMENT OF CONTRIBUTIONS

Parts of Chapter 2-4 are adapted and modified from two published manuscripts that were co-written with Ali Brivanlou and Eric Siggia. Parts of Chapter 3 are adapted and modified from a book chapter that was co-written with Tatiane Kanno and Ali Brivanlou and edited by Insoo Hyun. The RUES2-GLR cell line described and used in Chapter 3 and 4 was created and validated by Albert Ruzo.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AE	Anterior endoderm
AP	Anterior-posterior
APS	Anterior primitive streak
β-CAT	β-catenin
BMP	Bone morphogenic protein
BRA	Brachyury
BSD	Blasticidin
C.C.	Carnegie Collection
CDX2	Caudal-type Homeobox 2
CER1	Cerberus 1
CHIR	CHIR-99021
CMFDA	5-chloromethylfluorescein diacetate green cell tracker dye
COLLAGEN IV	Type 4 collagen typically found primarily in the basal lamina
DAPI	Fluorescent stain that binds strongly to DNA
DE	Definitive endoderm
DKK(1/3)	Dickkopf WNT signalling pathway inhibitor
DNA	Deoxyribonucleic acid
DOX	Doxycycline
E-CAD	Epithelial cadherin
EMT	Epithelial to mesenchymal transition
EOMES	Eomesodermin

EpiLCs	Mouse epiblast-like cells
ExM	Extra-embryonic mesoderm
FGF	Fibroblast growth factor
FKPM	Fragments per kilobase million
FOX(A2/F1)	Forkhead Box
FZD	Frizzled class receptor
GBX2	Gastrulation brain homeobox 2
GLR	Germ layer reporter
H2B	Histone 2B
HAND1	Heart and neural crest derivatives expressed 1
HD	High density
hESCs	Human embryonic stem cells
HHD	High density
HNA	Human nuclear antigen
HOXB1	Homeobox B1
IF	Immunofluorescence
ISSCR	International Society for Stem Cell Research
ISL1	Islet-1
IVF	In vitro fertilization
IWP2	Small molecule that inhibits Wnt processing and secretion
kDa	kilodalton
KiKGR	Kikume Green-Red photoconvertible fluorescent protein

LD	Low density
LEF1	Lymphoid enhancer binding factor 1
LEFTY(1/2)	Left-Right determination factor
LM	Lateral mesoderm
LRP(5/6)	Low-density lipoprotein receptor-related protein
MEF-CM	Mouse embryonic fibroblast conditioned media
MSGN1	Mesogenin 1
N-CAD	Neural cadherin
OCT4	Octamer-binding Protein 4
OTX2	Orthodenticle homeobox 2
PBS	Phosphate-buffered saline
PITX2	Paired like homeodomain 2
PPS	Posterior primitive streak
PrEct	Primitive ectoderm
PS	Primitive streak
PSM	Presomitic mesoderm
PURO	Puromycin
qPCR	Quantitative polymerase chain reaction
RFP657	Far-red monomeric fluorescent protein with emission peak at 657 nm
RNA	Ribonucleic acid
ROCK	Rho Kinase
RUES2	Rockefeller university embryonic stem cell line 2

SB	SB431542
SFRP(1/2/5)	Secreted frizzled related protein
SMAD2	SMAD Family Member 2
SNAIL	Snail family transcriptional repressor 1
SOX(2/3/17)	SRY (Sex Determining Region Y)
TBX6	T-Box 6
TGFβ	Transforming growth factor beta
V5	Epitope tag derived from V protein of simian paramyxovirus virus 5
WNT	Wingless-related integration site
WT	Wild type
ZO1	Zonula occludens 1

CHAPTER 1: INTRODUCTION

The development of an embryo from a single fertilized cell can seem like the birth of a miniature universe: an exponential growth and series of divisions, a progressive breaking of symmetries, and a finally a long denouement and refinement of intermediate structures and elements, every step of which is self-contained and self-organized. Arguably one of the most dramatic points of this process is gastrulation. This is the key moment when the initially identical and interchangeable pluripotent cells of the early embryo rapidly break spatial symmetry to establish the main anterior-posterior body axis and transform themselves into a multi-layered structure composed of the three germ layers: ectoderm, mesoderm, and endoderm.

In amniotes, the first sign of gastrulation is the emergence of the primitive streak (PS), a transient structure that begins on the posterior edge of the epiblast and grows towards the center. As the streak grows, cells migrate through it and give rise to the different endodermal and mesodermal lineages of the future body plan, with the specification of these different cell types dependent on the position and time at which they transit through the streak. Although individuals and species vary widely in final size and form, only a few key evolutionarily conserved cell signalling pathways orchestrate fate acquisition and the emergence of discrete patterns in the PS. Among these, the Wnt pathway plays a pivotal role. Although much is known about the role of the Wnt pathway in the specification of the PS in model organisms, virtually nothing is known about its involvement in human PS specification. Uncovering what, if anything, the Wnt pathway does to initiate and pattern our own PS, the critical stage that every human that has ever lived has passed

through, is the central subject of this thesis. Before getting to that investigation, however, it is necessary to provide some relevant background. In this chapter I will first briefly introduce the Wnt pathway itself. Next, I will describe the main functions of Wnt signalling during gastrulation in mouse and chick embryos. Finally, I will outline what is known about human gastrulation and the formation of the PS *in vivo*, and show how recent experiments with human embryonic stem cells (hESCs) offer a promising alternative to study human development in a manner that avoids many of the difficult ethical limitations that come with working with real human embryos.

The canonical Wnt signalling pathway

Wnts are secreted ligands that activate a highly conserved signalling pathway that is reused time and time again throughout development and adult tissue homeostasis. In this latter role, the first Wnt ligand was identified by Roel Nusse and Harold Varmus who used oncogencic retroviruses to identify *int1* (integration site 1) as a site in the mouse genome whose disruption led to tumourogenesis[1]. In parallel, a screen of early development Drosophila mutants by Eric Wieschaus and Christiane Nüsslein-Volhard yielded a segment polarity loss-of-function mutant for *wingless* (*wg*)[2], a gene previously identified as responsible for disruptions to wing development when present as a hypomorph[3], hence its name. Continued research on *int1* led to the conclusion that *int1* and *wg* were one and the same[4], thus the portmanteau "Wnt" that brought these two identities together.

Since the initial discovery of the first Wnt gene, Wnt ligands have been discovered in every metazoan species to such an extent that they are a distinguishing metazoan feature, with all

species possessing at least one gene[5,6] and protists, fungi, and choanoflagelletes (the closest living common ancestor to all animals) possessing none[7,8]. In human and mice there are 19 genes encoding Wnt ligands. These secreted proteins are highly conserved, with each gene encoding for a secreted lipid-modified glycoprotein that is approximately 40 kDa in size and that can be defined by a nearly invariant positioning of 22 cysteine residues[9]. The most common lipid modification likely shared between all Wnts is a palmytolation[10], brought about by a special palmitoyl transferase which is called PORCUPINE (PORCN) in humans[11–13]. Because this modification renders the Wnt protein hydrophobic and thus "sticky" to cell membranes, Wnt proteins are generally thought of and visualized as acting at short range[14,15].

When a cell receives a Wnt signal there are several possible responses. The best studied is the so-called "canonical" Wnt response, but there are several other "alternative" Wnt responses that play important roles in development as well[16]. The key operative response during amniote PS formation and patterning, however, is the canonical Wnt pathway[17], and in all that follows, Wnt signalling is taken as implying "canonical Wnt signalling". In this modality WNT ligands first bind to FRIZZLED (FZD)[18] and low density lipoprotein receptors (LRP5 and LRP6)[19] to form a trimeric complex. The intracellular domain of this complex then recruits the scaffold protein AXIN away from a destruction complex that functions to continuously target free β-CATENIN (β-CAT) for degradation[20]. With the removal of AXIN from the destruction complex, free β-CAT ceases to be marked for degradation and is able to accumulate in the nucleus where it can act as a co-transcriptional activator with TCF/LEF proteins (amongst others)[21]. From a systems level view,

Wnt signalling thus acts as an inhibition of an inhibition, with β-CAT being the key molecular mediator (Figure 1.1).

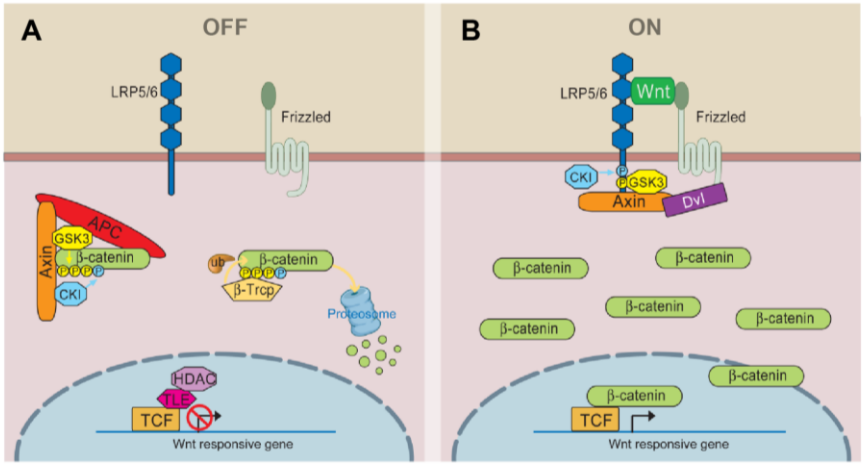


Figure 1.1 | Overview of Wnt/β-CAT signalling

Diagram showing core components of the canonical Wnt signalling pathway in (A) OFF state and (B) ON state. Adapted from MacDonald et al.,[22].

It must be noted, however, especially in anticipation of Chapter 4, in which it will be discussed further, that β-CAT is more than just the mediator of Wnt signalling. It is also plays an essential role in the structural organization and function of cadherins by linking cadherins (especially E-CADHERIN) through α-catenin to the actin cytoskeleton[23–25]. In fact, partly due to this dual nature β-CAT was discovered independently twice, once in its signalling function[26] and once in its structural function[27].

Wnt signalling pathway in the primitive streak of model organisms

One of the most conserved functions of Wnt signalling during development is to establish the main body axis of the developing embryo. Across the vast majority of animal species studied Wnt signalling is generally active in the posterior and inhibited in the anterior and Wnt pathway perturbation before or during axis formation results in dramatic axial consequences, such as axis duplication, truncation, or posteriorization[17]. Immediately after initial anterior-posterior (AP) axis formation, Wnt signalling also generally functions in posterior growth[28] and mesoderm specification[29].

The mouse and the chick are the two model organisms closest to human about which the most is known, and during their embryogenesis Wnt signalling is necessary to establish and pattern the PS. In mouse, expression of Wnt3 in the posterior side of the epiblast and posterior visceral endoderm is initiated by a Bmp4 signal from the extraembryonic ectoderm at ~E6.5[30–32]. This Wnt3 expression domain marks the extent of the PS and leads to high levels of β-CAT and Tcf-responsive promoter expression[33,34]. It is opposed by the Wnt inhibitor Dkk1 in the anterior visceral endoderm which prevents signalling in the anterior portion of the epiblast[35,36]. Wnt3 and β-cat knockout mice fail to form the primitive streak[33,37,38], whereas knockout of the Wnt inhibitor Dkk1 results in an expansion in Wnt signalling in the primitive streak and anterior truncation[35,39]. The Dkk1 knockout also results in a higher ratio of mesoderm to endoderm cells in the PS, consistent with the role of Wnt in mesoderm specification[39]. Ectopic Wnt signalling activation in the PS has also been shown to cause anterior truncation and expansion or bifurcations of the primitive streak, resulting in partial axial structure duplications[40–42].

In the chick Wnt is also necessary to form the primitive streak. During elongation and patterning Wnt8C expression domain marks the extent of the PS, and at stages XII-XIII, just prior to PS formation, Wnt8C is expressed in the marginal zone of the embryo in a gradient decreasing from posterior to anterior[43]. It is this early expression that gives competence to the marginal zone to form a streak when also provided with a Vg1 signal. Thus a normal streak will form from the posterior marginal zone, where Vg1 is naturally localized, and ectopic streaks will form when Vg1 is over-expressed in the marginal zone, or when both Wnt and Vg1 are co-overexpressed even in an area without normal Wnt expression, such as the area pellucida[43,44]. This Wnt activity is opposed in the anterior of the chick epiblast by the Wnt inhibitors Crescent, Cerberus, and Dkk1 that are secreted by the hypoblast[43,45,46], the chick analogue of the mouse visceral endoderm[47].

It is also important to emphasize what is not known about Wnt signalling in the PS of model organisms. Despite extensive knowledge of the intracellular aspects of the pathway and the outcome of genetic perturbations and classical embryological manipulations, many quantitative aspects of how Wnt creates and mediates the large-scale pattern in the PS remain unknown. For instance, in the embryo the interactions of the Wnt pathway with the other key signalling pathways involved in PS patterning, such as the BMP and Nodal pathways, are complex, with multiple morphogens and secreted inhibitors from overlapping regions acting at the same time. There are also radical differences in architecture, timing, and juxtaposition of extraembryonic and embryonic tissues between different species that makes direct comparison difficult[31,32]. Because of these factors there is a great utility for *ex vivo* assays that allow the precise control of

geometry, cell density, signalling strength, and timing so that one can take a reductionist approach to the complex interactions that confound experiments in the embryo. As I outline in the next section, although we may know the least about Wnt signalling in the human PS compared to other studied species, the human PS also offers the greatest potential for taking this kind of reductionist approach.

Early human development and stem cell models

Human gastrulation and primitive streak development occur during what may be termed a research "blackout period". This is because these events occur both after the stage when it is ethically permissible to conduct experiments with donor human fertilized eggs[48–50] (the appearance of the streak at day 14-15 post-fertilization *in vivo* in fact defines this ethical red line[51,52]) and before the embryo is old enough or large enough to be available in sufficient quantities as donor fetal tissue. Thus, unlike in model systems, the information derived from genetic studies, time-lapse imaging, fate maps, or gene maps are almost nonexistent for the human embryo. Instead, the vast majority of knowledge of human development during this period is based purely on morphology and what can be inferred from serial light and electron microscopy sectioning of the few samples that exist in the various collections of early human embryos from around the world[53]. This is a valuable resource and amongst other things has led to the staging of human embryos[54,55] and identification of when the human PS starts, but it is not a substitute for modern molecular techniques and for instance offers no information on the possible role for Wnt in establishing and patterning the streak.

A path forward to understanding human development during this blackout period is with hESCs. These cells are derived from the day 5-6 pre-implantation blastocyst and demonstrate the hallmark stem cell characteristics of renewal and pluripotency[56]. They exist in a primed state on the cusp of gastrulation[57] and can maintain pluripotency indefinitely in culture as well as differentiate into embryoid bodies[58], form teratomas[59], and contribute to chimeras[60]. Over the past decade the use of hESCs to study human developmental processes has been most powerfully demonstrated with self-organizing "organoids"[61,62] that exhibit strong similarities to *in vivo* organs such as gut, lung, kidney, and even brain[63–67]. This allows a highly quantitative approach to understand self-organization of cells, tissues, and organs. More recent work from our labs and others have shown that hESCs can also be used to generate self-organizing "gastruloids" that model a gastrulating embryo[68].

More specifically, in our lab we have shown that hESCs confined to micropatterned colonies of 1 millimeter diameter can be used as *in vitro* assay to model the human epiblast[69,70]. These micropatterns self-organize in response to 48 hours of BMP4 stimulation and recapitulate the patterning of germ layers observed during mammalian gastrulation, with concentric rings corresponding to ectoderm, mesoderm, endoderm, and extraembryonic tissue arranged from the center to edge. Critically, in addition to providing a path to visualize and measure aspects of human gastrulation, these gastruloid models also permit single-cell quantification and control over geometry, density, signalling strength, and genetics. Exploiting these features, follow-up work in my lab was able to deduce how the BMP pathway contributes to this patterning, discovering that BMP receptors in cells in the epiblast are localized to baso-lateral surfaces to

create a geometrically defined “pre-pattern” that is reinforced by the secreted BMP inhibitor NOGGIN[71].

Based on these earlier promising demonstrations, I believe that the gastruloid model system thus offers not only an ethically acceptable way to investigate early human development, but a powerful assay to ask and answer general quantitative questions about signalling pathways and patterning that are exceedingly difficult or impossible in an *in vitro* developing embryo. In the chapters that follow, I will use this assay to investigate Wnt signalling in the formation and patterning of the human PS, and will be guided by two general questions: what is the role of Wnt in the human PS, and what controls Wnt in the human PS.

CHAPTER 2: THE ROLE OF WNT SIGNALLING IN THE HUMAN PRIMITIVE STREAK

As described in the previous chapter, our group discovered that it was possible to create *in vitro* "gastruloid" models of early human development by stimulating monolayer disc-shaped micropatterns of hESCs with BMP4 supplied in the media[69]. The inspiration to use BMP4 came from studies of the mouse embryo, which have shown that BMP4 from the extra-embryonic ectoderm is at or near the top of a PS initiation hierarchy where BMP signalling activates the WNT pathway which in turn activates the ACTIVIN/NODAL pathway[30] (Figure 2.1A). Since it has also been shown in mouse and other amniotes that WNT is especially critical for initiating and patterning the PS (as detailed in the previous chapter), an immediate follow-up question was what is the role of WNT in establishing the gastruloid?

In this chapter I focus on uncovering the role the WNT pathway plays in initiating and patterning the human PS. We discover that there is a conserved BMP4 to WNT to ACTIVIN/NODAL signalling hierarchy that initiates PS formation in human, and that WNT is necessary and sufficient for this formation. We also discover that WNT induces an epithelial to mesenchymal transition (EMT), as one would expect to find in a gastrulating embryo. Looking beyond initiation of the PS to the role WNT plays in the context of the BMP and ACTIVIN/NODAL pathways in fate specification and patterning within the PS, we find that different subpopulations of endoderm or mesoderm emerge robustly from each set of gastruloids depending on BMP, WNT, and NODAL levels, and that by comparison to the mouse embryo we find that we can arrange these subpopulations along an anterior-posterior axis. We also find that there are robust cell migrations from the PS region of each gastruloid and that the character of these migrations depends on what fates the

differentiating cells commit to, with fast single-cell migrations in the case of endoderm for instance, or slower group migration in the case of mesodermal populations. Combining this data with our anterior-posterior fate characterizations, we are able to determine the role WNT plays in patterning the human PS and sketch out a rudimentary fate map of the *in vivo* human PS.

A conserved BMP4→WNT→ACTIVIN/NODAL signalling initiation hierarchy

Given the signalling hierarchy in the mouse and the gastruloid result with BMP4 stimulation, the first task in focusing on the role of the WNT signalling pathway in human PS formation was to determine its relationship to the BMP and ACTIVIN/NODAL branches of the TGFβ signalling pathways. More specifically, we first set out to test whether the interaction network of these pathways from mouse held true in human as well. Using RNA-Seq, we found that of all the 19 *WNT* ligands present in the human genome, only one, *WNT3*, is significantly and immediately induced upon BMP4 presentation (Figure 2.1B). qPCR analysis shows that activation of WNT signalling directly induces *NODAL* expression (Figure 2.1C). Further qPCR analysis showed that NODAL induction was reduced when the NODAL inhibitor SB431542 (SB) was present, and taken together with the observation that ACTIVIN induces *NODAL* expression, suggests the presence of a NODAL feedback loop, as also noted in the mouse[72]. Additionally, no direct *BMP4* induction by either WNT or NODAL signalling was observed (Figure 2.2B). Thus the transcriptional hierarchy of BMP→WNT→NODAL is evolutionarily conserved in hESCs.

To test if the hierarchy of signalling activity was also conserved, we challenged BMP4 self-organizing activity with two small inhibitors: SB and the WNT inhibitor IWP2. Either BMP4+ IWP2

or BMP4+ SB led to a loss of mesoderm (BRA) and endoderm (SOX17; Figure 2.1D and Figure 2.2C). Thus both WNT and ACTIVIN/NODAL signalling are necessary for mesendodermal induction and patterning downstream of BMP4.

Figure 2.1 | Primitive streak signalling in hESCs follows BMP→Wnt→Nodal hierarchy

(A) Model of the proposed hierarchy in hESCs that initiates the PS in hESCs, along with indication at which step the inhibitors SB and IWP2 act. Like in mouse, BMP acts on WNT, and then WNT acts on NODAL. There is also positive feedback between WNT and NODAL.

(B) RNA-seq expression of all known WNT ligands in pluripotency and after 4 h of BMP4 in hESCs on 500 µm diameter micropatterns. The results show that overall WNT transcription is low in pluripotency and that WNT3 is the only strong and direct WNT ligand target of BMP4 stimulation. Data is from previously published data set[71] (GEO accession number GSE77057). FKPM, Fragments per kilobase million. (C) qPCR analysis showing expression of WNT3 and NODAL in small colonies of hESCs after 4 h stimulation with each condition shown on x-axis. Data are mean ± s.d. of *n* = 3 biologically independent replicates. (D) Pie sections are of representative 1000 µm diameter micropatterned colonies stimulated with BMP4, BMP4+IWP2, or BMP4+SB and fixed and stained for germ layer molecular markers after 48 h. All micropattern experiments were performed on at least *n*=3 separate occasions with similar results, and unless mentioned otherwise, all other micropatterns are 1000 µm in diameter. Staining is quantified in Figure 2.2C.

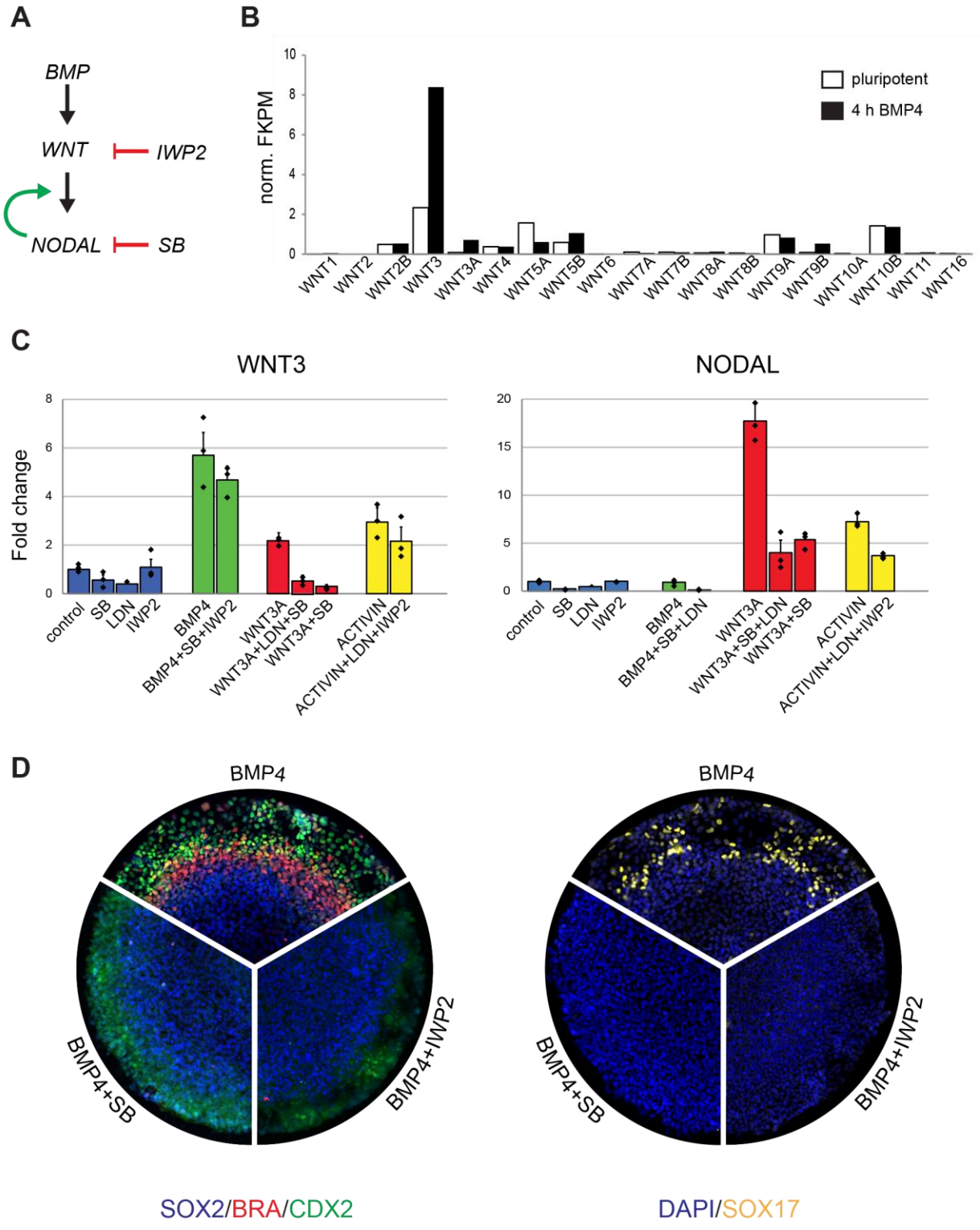
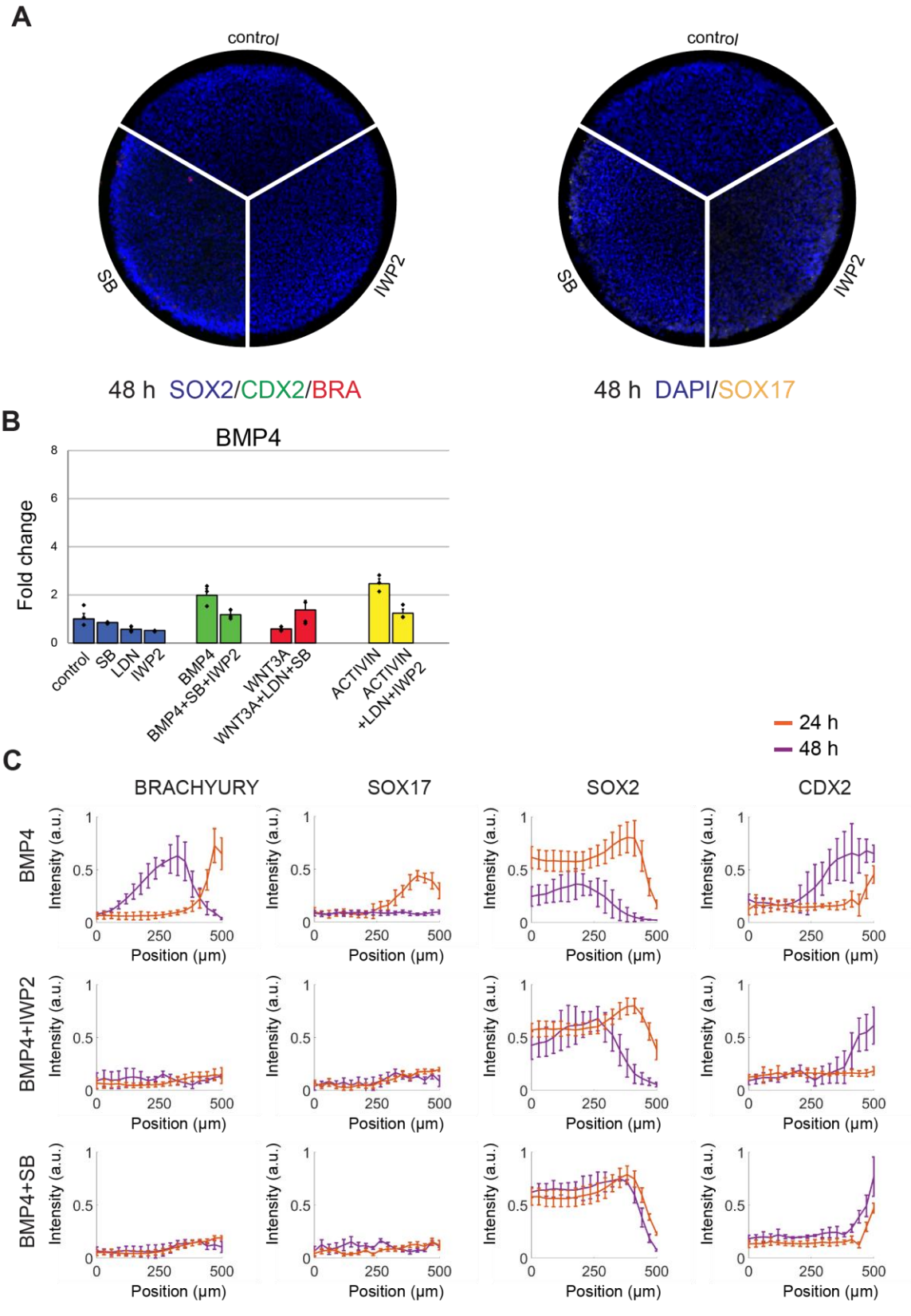

A
BMP
WNT
IWP2
NODAL
SB
B
norm. FKPM
pluripotent
4 h BMP4
WNT1
WNT2
WNT2B
WNT3
WNT3A
WNT4
WNT5A
WNT5B
WNT6
WNT7A
WNT7B
WNT8A
WNT8B
WNT9A
WNT9B
WNT10A
WNT10B
WNT11
WNT16
C
WNT3
NODAL
Fold change
control
SB
LDN
IWP2
BMP4
BMP4+SB+IWP2
WNT3A
WNT3A+LDN+SB
WNT3A+SB
ACTIVIN
ACTIVIN+LDN+IWP2
BMP4+SB+LDN
WNT3A+SB+LDN
D
BMP4
BMP4+SB
BMP4+IWP2
SOX2/BRA/CDX2
DAPI/SOX17

Figure 2.2 | Controls for investigating hESC PS initiation hierarchy

(A) Micropatterned colonies stimulated with IWP2, SB, or blank media, fixed and stained for germ layer molecular markers after 48 h. This experiment was repeated at least three times independently with similar results. (B) qPCR for BMP4 of unpatterned small colonies stimulated for 4 h conditions arraigned on the x-axis. As consistent with model hierarchy, there is no significant induction of BMP4 by ACTIVIN, WNT3A, or itself. Data are mean ± s.d. of three biologically independent replicates. (C) Quantification of Figure 2.1D. In this and in all other analysis unless stated otherwise, nuclei were segmented using DAPI and intensity of immunofluorescence signal for each marker was normalized to the DAPI intensity. Single cell expression data was binned radially and averaged. The final radial profile represents the mean ± s.d. of n = 25 colonies.


A
control
SB
IWP2
48 h SOX2/CDX2/BRA
control
SB
IWP2
48 h DAPI/SOX17
B
BMP4
Fold change
8
6
4
2
0
control
SB
LDN
IWP2
BMP4
BMP4+SB+IWP2
WNT3A
WNT3A+LDN+SB
ACTIVIN
ACTIVIN +LDN+IWP2
C
24 h
48 h
BRACHYURY
SOX17
SOX2
CDX2
BMP4
BMP4+IWP2
BMP4+SB
Intensity (a.u.)
Position (μm)
0
250
500
0.5
1

WNT signalling is necessary and sufficient for human PS formation

To ask if WNT or ACTIVIN/NODAL signalling alone was sufficient as well as necessary, hESC colonies were stimulated with either WNT3A or ACTIVIN. After 48 hours of treatment, WNT3A led to differentiation of the periphery into mesoderm (BRA) and endoderm (SOX17; Figure 2.3A and Figure 2.4A). The center cells maintained their pluripotent epiblast fate (SOX2 and NANOG) rather than differentiating into ectoderm (Figure 2.3A and B). After 48 hours of ACTIVIN treatment, however, no cells showed any sign of differentiation or self-organization, and all maintained the same morphology and expression of the pluripotency markers (Figure 2.3A and B).

As it is unlikely that ACTIVIN/NODAL has no effect during human gastrulation, we presented WNT3A in two combinations that represent the opposite extremes of an ACTIVIN/NODAL gradient: WNT3A+ACTIVIN and WNT3A+SB. In accordance with studies in model systems and human and mouse embryonic stem cells[73,74], we found that ACTIVIN/NODAL signalling acts as a modifier of mesoderm and endoderm patterning, with all the cells on the periphery converting to endoderm (SOX17+) with no mesoderm (BRA-) in WNT3A+ACTIVIN, and all cells converting to mesoderm (BRA+) with no endoderm (SOX17-) in WNT3A+SB (Figure 2.3A and Figure 2.4A). In addition, we found evidence of an epithelial-to-mesenchymal transition (EMT) with SNAIL expression and an E-CADHERIN (E-CAD) to N-CADHERIN (N-CAD) switch where the mesendodermal fates later establish themselves (Figure 2.3C). Comparison of WNT3A, WNT3A+ACTIVIN, and WNT3A+SB conditions also revealed that while WNT3A is sufficient to induce this EMT, different ACTIVIN/NODAL levels leads to differences in the timing of this

transition. Increased ACTIVIN/NODAL levels lead to an earlier transition and lower levels lead to a delayed transition (Figure 2.4B). Taken together, we found that WNT signalling is necessary and sufficient to induce PS, and that ACTIVIN/NODAL signalling acts as a modifier that controls timing of EMT and patterning of mesoderm versus endoderm.

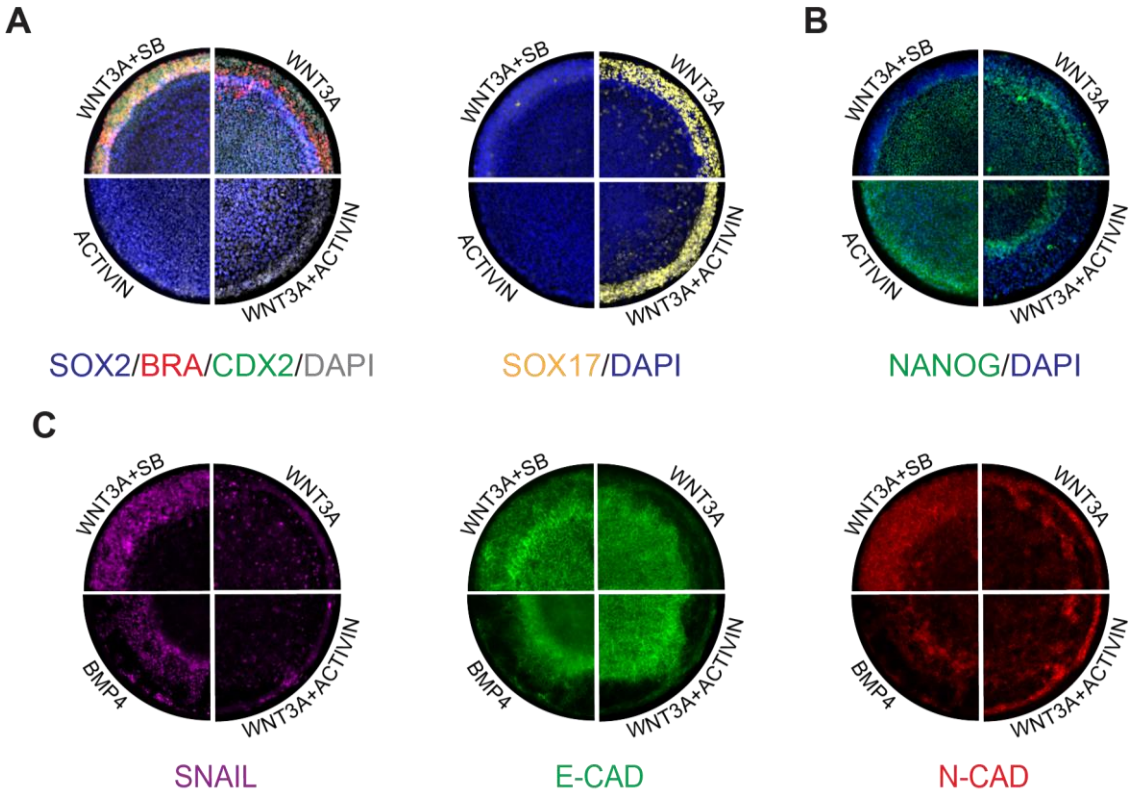
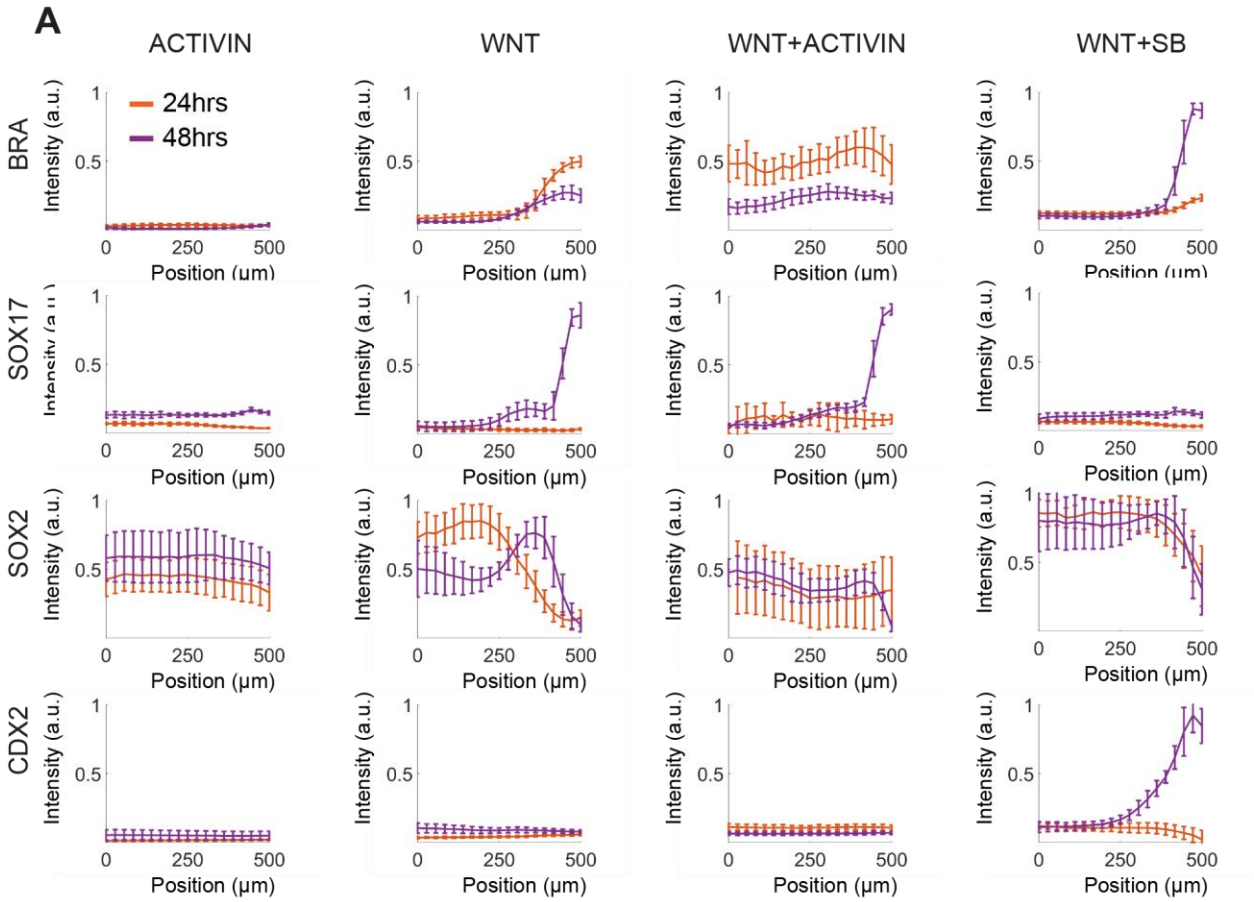


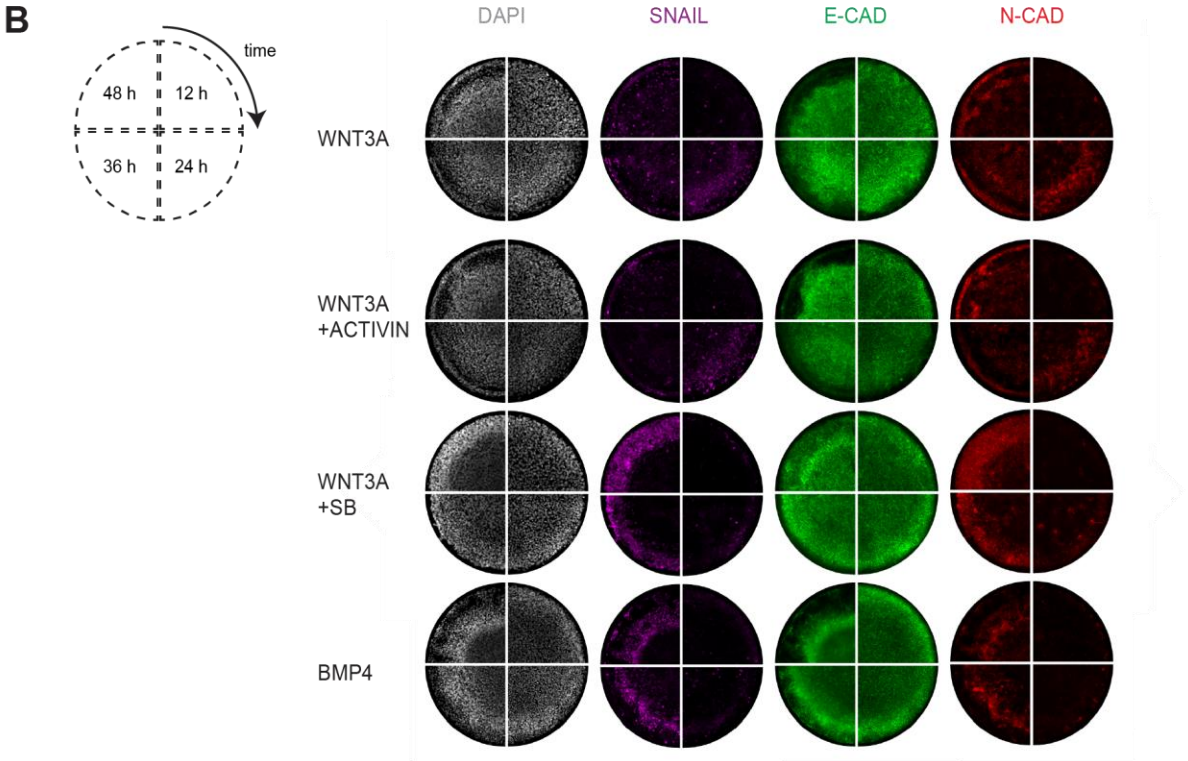
Figure 2.3 | WNT is necessary and sufficient to induce PS markers and morphology

(A) Micropatterned colonies stimulated with WNT3A, WNT3A+ACTIVIN, WNT3A+SB, or ACTIVIN and fixed and stained for germ layer molecular markers after 48 h. Staining is quantified in Figure 2.4A. (B) Micropatterned colonies stimulated with WNT3A, WNT3A+ACTIVIN, WNT3A+SB, or ACTIVIN and fixed and stained for pluripotency marker NANOG after 48 h. (C) Micropatterned colonies stimulated with BMP4, WNT3A, WNT3A+SB, or WNT3A+ACTIVIN and fixed and stained for EMT markers SNAIL, E-CAD, and N-CAD after 48 h. Note that WNT3A and WNT3A+ACTIVIN stimulated colonies show less SNAIL at 48 h because they started their EMT earlier, at 24 h (see Figure 2.4B for 12, 24, and 36 h timepoints).

Figure 2.4 | PS germ layer quantification and EMT timing

(A) Quantification of Figure 2.3A. The radial profile represents the mean ± s.d. of *n* = 25 colonies. (B) Micropatterned colonies stimulated with BMP4, WNT3A, WNT3A+SB, or WNT3A+ACTIVIN and fixed and stained for primitive streak molecular markers SNAIL, E-CAD, and N-CAD after 12, 24, 36, or 48 h. Note that WNT3A and WNT3A+ACTIVIN stimulated colonies turn on EMT markers faster than BMP4 or WNT3A+SB stimulated colonies, and have mostly downregulated SNAIL by 48 h. This experiment was repeated at least three times independently with similar results.


A
ACTIVIN
WNT
WNT+ACTIVIN
WNT+SB
BRA
SOX17
SOX2
CDX2
24hrs
48hrs
Intensity (a.u.)
Position (μm)
0
250
500
0.5
1


B
time
48 h
12 h
36 h
24 h
DAPI
SNAIL
E-CAD
N-CAD
WNT3A
WNT3A
+ACTIVIN
WNT3A
+SB
BMP4

Anterior-posterior fate specification in human gastruloids

Building on our previous characterization of gastruloid cell fates[69] and incentivized by recent work successfully mapping cell types in mouse gastruloids to mouse embryos[75], we hypothesized that primitive streak cell subtypes were present in our gastruloids and that they could be compared to the anterior-posterior axis of the mouse embryo. In the mouse gastruloid studies specific combinations of cell type specific transcription factors were used to identify discrete fates and compare their pattern with the mouse embryo. Here we follow the same strategy and analyze our BMP4, WNT3A, WNT3A+SB, or WNT3A+ACTIVIN induced human gastruloids for anterior-posterior identity and compare them with the mouse gene map and fate map at E7.5 to shed some light on spatial structure of the human PS (Figure 2.5A-B).

Strikingly, we found self-organized and largely homogenous anterior-posterior subpopulations that arose distinctly in one set of stimulation conditions and not the others. For instance, only BMP4 induced expression of HAND1, CDX2, and GATA3, and these markers were all present in the same set of cells on the periphery of the gastruloid (Figure 2.5C). In the mouse HAND1 is first expressed at E7.5 in the trophectoderm and extra-embryonic mesoderm, including the amnion, chorion, allantois and visceral yolk sac[76]. GATA3 is expressed in the mouse and human pre-implantation trophoblast[48,77,78] and in the mouse E7.5 extra-embryonic ectoderm and allantois[78,79]. CDX2 is also expressed in the mouse and human pre-implantation trophoblast[48] and in mouse is restricted to the extra-embryonic ectoderm, mesoderm, and posterior endoderm until E8.5[75,80,81]. FOXF1, which beginning at E7.5 in the mouse turns on and marks the lateral plate mesoderm and yolk sac and allantois[82,83], is also most highly expressed in cells in this region.

Based on these comparisons, we take this region of the BMP4 induced gastruloid to most closely resemble the mouse E7.5 extra-embryonic mesoderm. This is further supported by the fact that there is a BMP source from the extra-embryonic ectoderm immediately adjacent to it *in vivo*. The fact that we do not see significant BRA expression in these cells as was found in mouse[75] may be due to species-specific timing differences (for example we have shown previously that there is a wave of BRA expression earlier in this region at 12-36 h[69]). Radially interior to this extra-embryonic mesoderm population are three other readily identifiable subpopulations that are unique to the BMP4 gastruloid. First, in the region adjacent to the extra-embryonic mesoderm there is a population of BRA+/GATA6+/ISL1+ cells. In the mouse at E7.5 GATA6 marks the parietal and definitive endoderm plus the lateral mesoderm[75,84], while ISL1 first appears at E8.5 and also marks cells in the lateral mesoderm[85,86]. Thus we identify this subpopulation as lateral mesoderm. Second, staining for SOX17 (a marker of definitive endoderm first apparent in mouse at E7-7.5[75]), NANOG (maker of definitive endoderm and epiblast), and OTX2 (marker of anterior epiblast and anterior PS in mouse from E7[75]) detects a population of the SOX17+/NANOG-/OTX2- cells. Based on these markers we identify this population as posterior endoderm. Finally, staining for SOX2 (marker of primitive ectoderm), and OCT4 (marker of epiblast) reveals a SOX2+/NANOG-/OCT4- subpopulation indicative of primitive ectoderm. Together, these four subpopulations in the BMP4 stimulated gastruloid all approximately match the E7.5 proximal posterior primitive streak in mouse.

With WNT3A+SB stimulation we found selective expression of TBX6 in the region that co-expresses CDX2 and BRA (Figure 2.5C). TBX6 did not appear in the other stimulation conditions,

and using qPCR we also found that MSGN1 was selectively induced with WNT3A+SB only (Figure 2.6A). In the mouse, both TBX6 and MSGN1 are first expressed in the primitive streak in the same region as BRA at E7.5, only to become restricted to the paraxial mesoderm by E8.5[87–89]. The fact that we do not detect significant TBX6 or MSGN1 levels at earlier times in any of the other gastruloids where we also see BRA (data not shown) may reflect a species specific difference between human and mouse. Additionally, although we use CDX2 in our panel of markers for the BMP4 induced gastruloids, CDX2 has also been shown to be critical for paraxial mesoderm development in the mouse and is detectable there from E8.5 onwards[80,90]. The union of these three molecular markers is thus highly suggestive of paraxial mesoderm, and a corresponding time of ~E7.5-8.5 in the mouse.

In the case of WNT3A and WNT3A+ACTIVIN, stimulation led to selective co-expression of the transcription factors FOXA2 and OTX2 in the SOX17+ region at the edge (Figure 2.5C). In the mouse FOXA2 begins to be expressed in the anterior primitive streak at E7, and becomes restricted to the anterior definitive endoderm and axial mesoderm by E7.75[75,91]. Thus the FOXA2+/OTX2+/SOX17+/BRA- provides the signature of anterior endoderm. Additionally, at 24 hours with WNT3A+ACTIVIN, but not WNT3A alone, we detected the organizer marker GSC. WNT3A+ACTIVIN also leads to the highest expression of key secreted inhibitors known to be produced by the organizer and its derivatives[92], such as CHORDIN, DKK1, CER1, LEFTY1, and LEFTY2, as well as to the highest expression of NODAL, which at later stages in mouse is also specific to the organizer (Figure 2.6B). As the organizer also has several well defined *in vivo* functional properties it offers a more stringent test of correspondence than just molecular

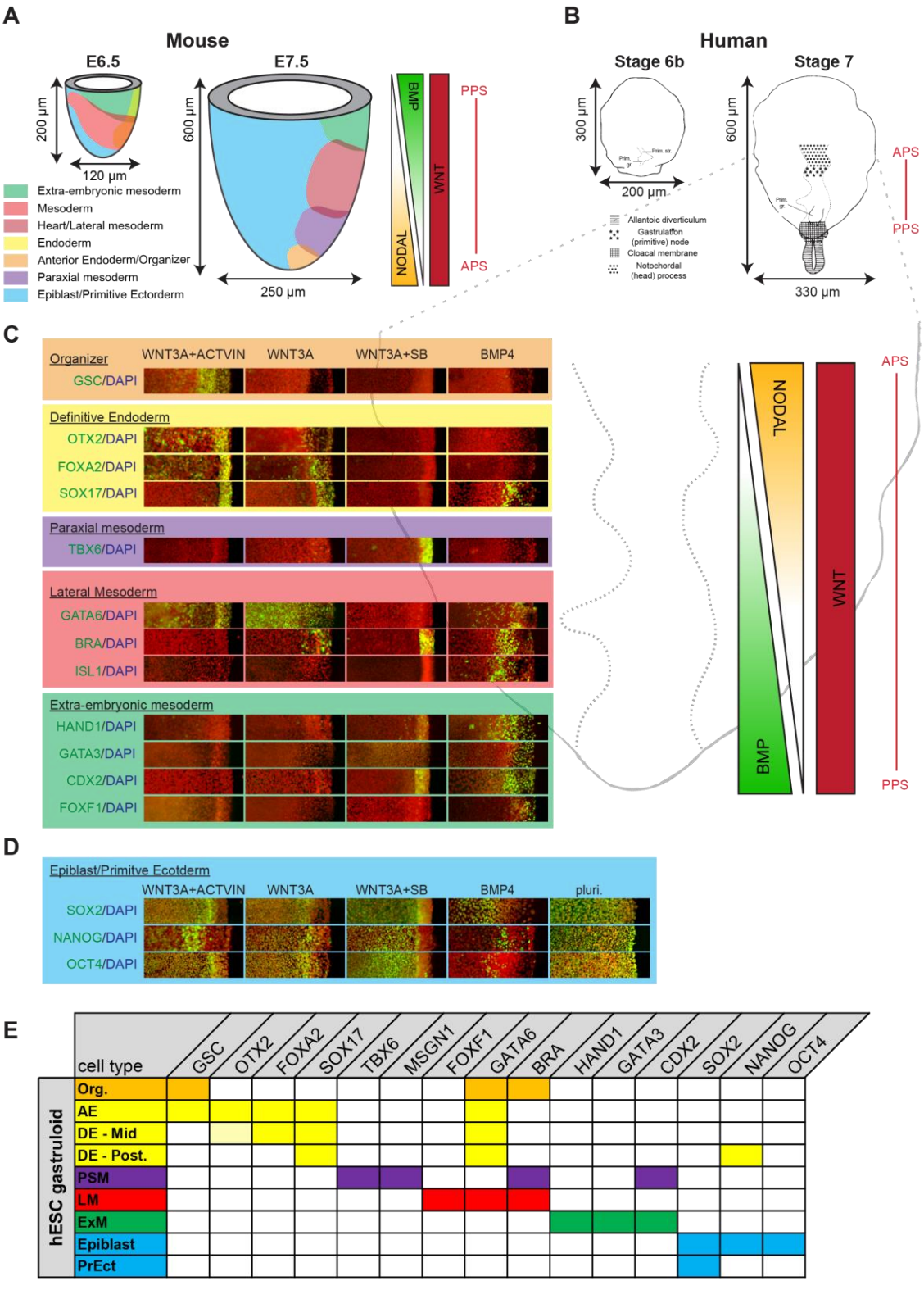
markers, and as such I will return to it for this test in Chapter 3. Keeping the focus on molecular markers, finally the centers of the WNT3A, WNT3A+Activin, and WNT3A+SB stimulated gastruloids differ from the center region of the BMP4 stimulated gastruloids in that they still express NANOG and OCT4, albeit at a lower level than in pluripotency (Figure 2.5C). We thus categorize these regions as epiblast and not as primitive ectoderm.

A summary of all of the readily identifiable subpopulations is provided in Figure 2.5E. Overall, we find good agreement between the mouse embryo and the gastruloid subpopulations.

Figure 2.5 | Mapping gastruloid fates to the human primitive streak

(A) Reproduction of the mouse primitive streak fate map from Tam and Behringer 1997. On the right are the inferred signalling gradients of BMP, WNT3A, and Nodal[31,32].

(B) Dorsal graphical representation of the human primitive streak fate map from the Carnegie Collection[54] (Carnegie Fig. 6-5. and Carnegie Fig. 7-4). (C) Mapping of gastruloids stimulated with either BMP4, WNT3A, WNT3A+SB, or WNT3A+ACTIVIN to the Carnegie Collection (C.C.) stage 7 human primitive streak. Gastruloids were fixed after 48 h stimulation and stained for the indicated sets of markers. Since each staining is radially symmetric, only a section from r=0 to r=R (500 μm) is shown. (D) Table summarising the expression of each marker for each classified sub-type. The OTX2/DE-Mid box with lighter yellow indicates that expression of OTX2 is less than that observed in other cells in other conditions. Org. = organizer, AE = anterior endoderm, DE - Mid = mid-streak definitive endoderm, DE - Post. = posterior definitive endoderm, PSM = presomitic mesoderm, LM = lateral mesoderm, ExM = extra-embryonic mesoderm, PrEct = primitive ectoderm.


A
Mouse
E6.5
E7.5
200 µm
120 µm
600 µm
250 µm
Extra-embryonic mesoderm
Mesoderm
Heart/Lateral mesoderm
Endoderm
Anterior Endoderm/Organizer
Paraxial mesoderm
Epiblast/Primitive Ectorderm
BMP
WNT
NODAL
PPS
APS
B
Human
Stage 6b
Stage 7
300 µm
200 µm
600 µm
330 µm
Allantoic diverticulum
Gastrulation (primitive) node
Cloacal membrane
Notochordal (head) process
C
Organizer
GSC/DAPI
WNT3A+ACTVIN
WNT3A
WNT3A+SB
BMP4
Definitive Endoderm
OTX2/DAPI
FOXA2/DAPI
SOX17/DAPI
Paraxial mesoderm
TBX6/DAPI
Lateral Mesoderm
GATA6/DAPI
BRA/DAPI
ISL1/DAPI
Extra-embryonic mesoderm
HAND1/DAPI
GATA3/DAPI
CDX2/DAPI
FOXF1/DAPI
D
Epiblast/Primitve Ecotderm
pluri.
SOX2/DAPI
NANOG/DAPI
OCT4/DAPI
E
cell type
GSC
OTX2
FOXA2
SOX17
TBX6
MSGN1
FOXF1
GATA6
BRA
HAND1
GATA3
CDX2
SOX2
NANOG
OCT4
hESC gastruloid
Org.
AE
DE - Mid
DE - Post.
PSM
LM
ExM
Epiblast
PrEct

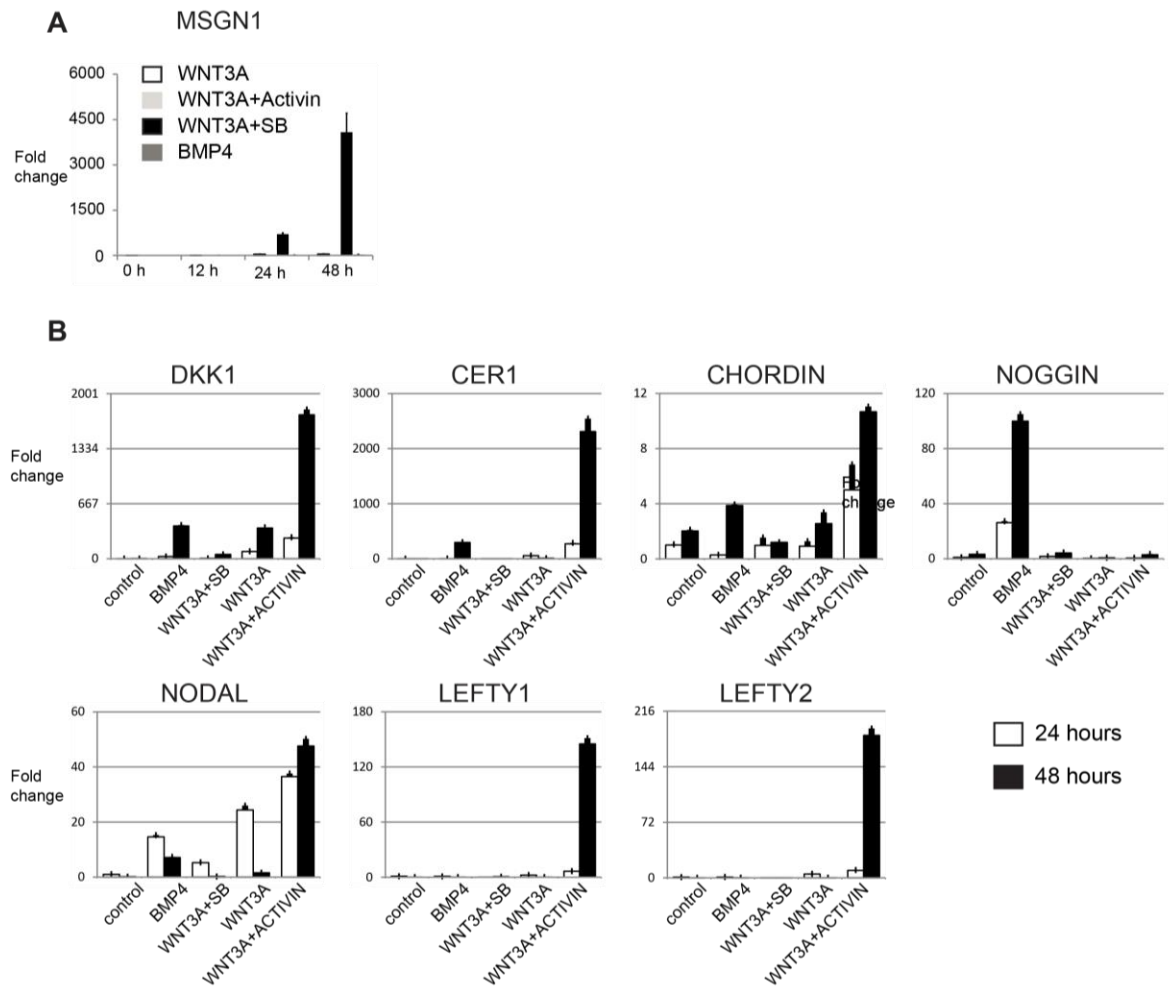


Figure 2.6 | Additional markers of early endoderm and mesoderm subtypes

(A) qPCR for paraxial mesoderm marker MSGN1 shows that it is most highly expressed in WNT3A+SB treated micropatterns at 48 h. Error bars represent the standard deviation of three technical replicates for each condition. (B) qPCRs of additional organizer markers, taken from RNA collected from 500 μm diameter micropatterns stimulated with either BMP4, WNT3A, WNT3A+SB, or WNT3A+ACTIVIN for 24 or 48 h. With the exception of NOGGIN, the characteristic organizer secreted inhibitors DKK1, CER1, CHORDIN, LEFTY1, and LEFTY2, are all most highly expressed in WNT3A+ACTIVIN conditions. The high NOGGIN induction by BMP4 in hESCs has been noted before[71], and may represent a human-mouse species difference. NODAL, which in mouse is restricted to the organizer later in gastrulation, is also most highly expressed in WNT3A+ACTIVIN conditions. Error bars represent the standard deviation of *n*=3 biologically independent replicates and the measure of the center represents the mean.

Cell migration

In addition to the emergence of distinct mesoderm and endoderm subtypes in different anterior-posterior positions along the primitive streak, vertebrate gastrulation is also characterized by highly orchestrated cell migrations through the streak and under the epiblast. Indeed, fate specification and migration occur concomitantly.

To track cells in our gastruloid system, we used the ePiggybac transposable element system[93] to derive clonal RUES2-KiKGR-RFP657-H2B cell lines that contain the photo-convertible protein KikGR and the far-red histone localized fluorescent protein RFP657-H2B. KikGR protein normally fluoresces in green but permanently converts to red upon UV excitation. KikGR also has a long life-time, enabling the detection of cells in which the protein has been switched to the red state even after two days. This tool allows photo-conversion of cells in specific regions of gastruloids and determination of their location after a window of time. More specifically, we used a digital micro-mirror to direct a 405 nm laser to illuminate one of three different annular regions: A_1, all cells <50 μm from the colony center; A_2, all cells in a ring >200 μm and <250 μm from the colony center; and A_3, all cells >400 μm from colony center (Figure 2.7B). Immediately after photo-conversion, micropatterns were stimulated with either control medium, BMP4, WNT3A, WNT3A+ACTIVIN or WNT3A+SB, and imaged to establish the starting point. The same colonies were imaged again at 24 h (Figure 2.8), and again at 52 h (Figure 2.7C and quantified in D).

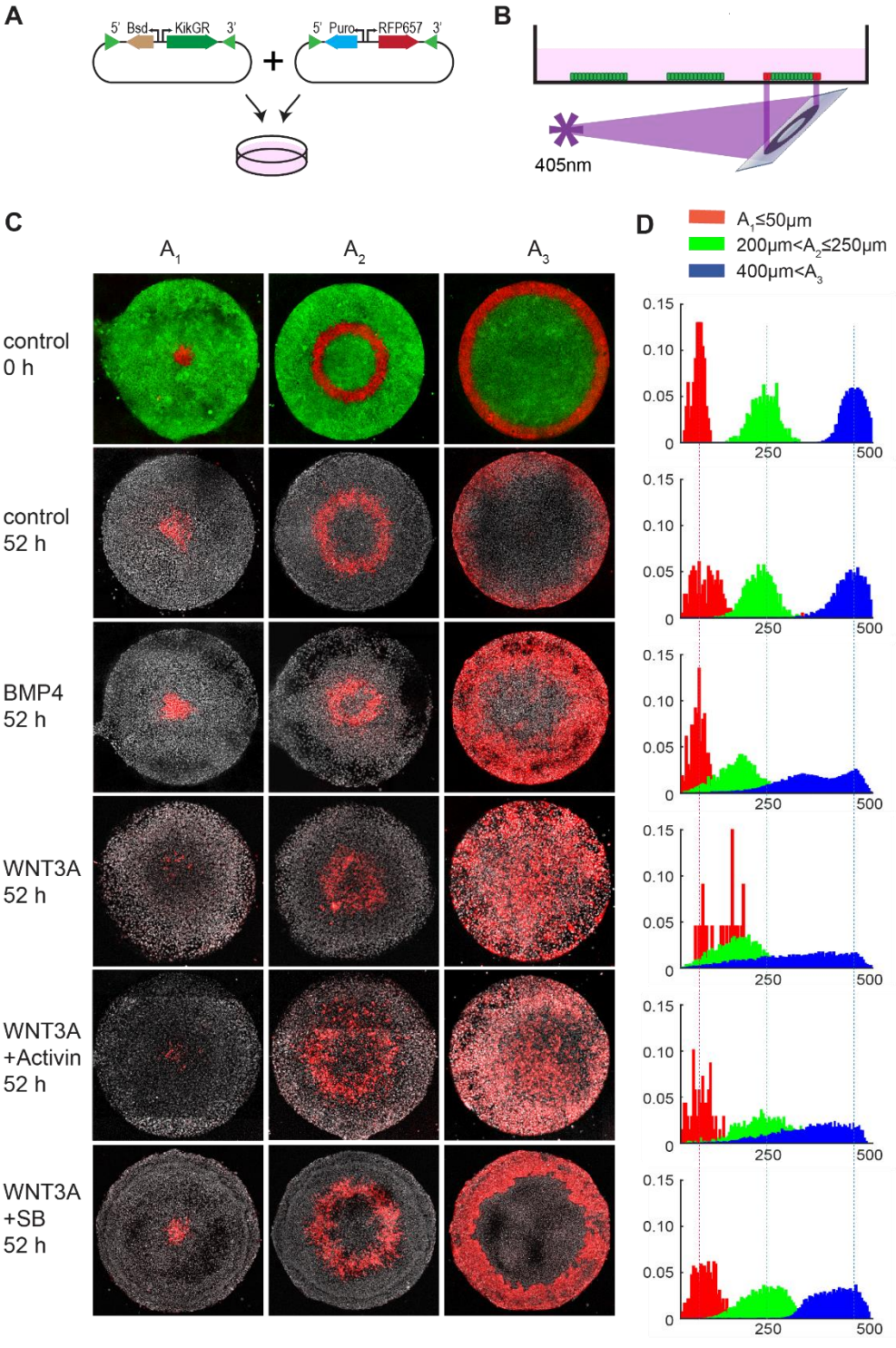
We found that in the unstimulated micropatterns the photoconverted cells retained their original position even after 52 h (Figure 2.7C, 1st and 2nd row). Stimulation with BMP4, WNT3A,

WNT3A+ACTIVIN, or WNT3A+SB, however, led to migration of cells localized at the edge (A_3) towards the center (Figure 2.7C, rows 3-6), and the onset of these migrations correlated well with the epithelial-to-mesenchymal transition (EMT) onset observed in each condition (Figure 2.4B). Under WNT3A and WNT+ACTIVIN stimulation migration started shortly after 24 hours of stimulation and cells migrated in a dispersed, individual manner, travelling long distances between the edge of the colony to the center (Figure 2.7C, 4^{th} and 5^{th} rows, and Figure 2.8). In contrast, a slower, shorter, and more compact migration was observed in the BMP4 and WNT3A+SB induced gastruloids (Figure 2.7C, 3^{rd} and 6^{th} rows). Quantification of the photoconverted cells in the BMP4 treatment revealed two distinct populations: one that remained on the outer edge, and another that migrated inwards (Figure 2.7C and D, 3^{rd} row). Finally, while no migration was observed in A_1 regardless of the stimulation, cells in the A_2 region shifted slightly inward by 52 h following WNT3A+ACTIVIN, WNT3A and BMP4 stimulation. However, these cells do not express EMT markers early on (Figure 2.4B), and it is hard to differentiate between active movement and passive movement as the result of being pushed in by the migration of cells from A_3. For instance, we speculate that as the A_2 region in BMP4 gastruloids is more compact than the WNT treated gastruloids this is more the result of pushing from the exterior cells rather than autonomous movement.

Figure 2.7 | Directed cell migration in the PS region

(A) Cloning strategy to make the RUES2-KiKGR-RFP657-H2B cell line.

(B) Using a digital micromirror, annular regions of micropatterned RUES2-KiKGR-RFP657-H2B colonies were selectively exposed to 405 nm light for 3 seconds and permanently switched from green to red fluorescence. (C) Three different annular regions were photoconverted: A_1 (<50 µm from colony center), A_2 (>200 µm and <250 µm from colony center), and A_3 (>400 µm from colony center). After photo-conversion cells were stimulated with either WNT3A, WNT3A+Activin, WNT3A+SB, BMP, or blank media and imaged at 0 h, 24 h (Figure 2.8), and 52 h. First row shows unconverted KikGR fluorescence (green) and converted KikGR fluorescence (red) at 0 h. All other rows show just converted KikGR fluorescence (red) and the far-red histone nuclear marker (grey) at 52 h. In all conditions significant movement of cells in the A_3 region is observed. (D) Quantification of C (see Methods).


A
5' Bsd KikGR 3'
5' Puro RFP657 3'
B
405nm
C
A1
A2
A3
control 0 h
control 52 h
BMP4 52 h
WNT3A 52 h
WNT3A +Activin 52 h
WNT3A +SB 52 h
D
A1≤50μm
200μm<A2≤250μm
400μm<A3
0.15
0.10
0.05
0
250
500

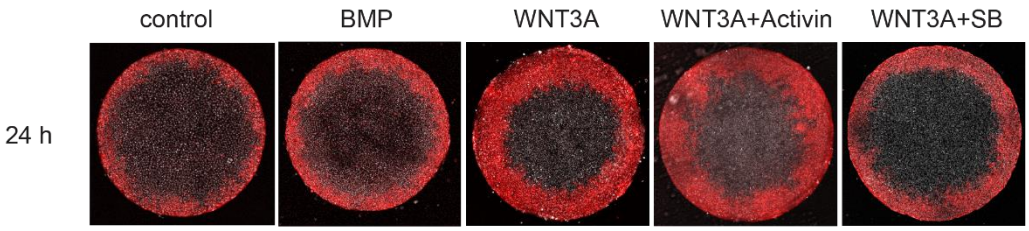


Figure 2.8 | Onset of directed cell migration

Micropatterns with photoconverted cells in region A_3 (>400 µm from colony center) and stimulated with either WNT3A, WNT3A+Activin, WNT3A+SB, BMP, or blank media were imaged at 24 h. As can be seen by comparison with the blank stimulated colony, cells in the outer region of the WNT3A and WNT3A+Activin micropatterns have started moving inwards at this time, ahead of the corresponding cells in the BMP4 and WNT3A+SB micropatterns.

To better understand how the cells migrate in each of the conditions we also examined the 3D structure of the gastruloids and what the corresponding fate markers of the migrating cells are. In the WNT3A and WNT3A+ACTIVIN gastruloids the migrating cells express SOX17 and so belong to the Anterior DE subpopulation identified previously. In the BMP4 gastruloids the migratory cells express BRA and so mostly belong to the LM subpopulation. In the WNT3A+SB gastruloids the migratory cells also express BRA and so are the PSM fated cells (Figure 2.9A and B). In all cases the migrating cells appear to push under the inner epiblast section en route towards the center of the gastruloid (Figure 2.9A). The nature of this attachment and the interaction of these cells with the migratory cells is also related to the COLLAGEN IV layer that we detect separating these layers in the WNT3A, WNT3A+ACTIVIN, and WNT3A+SB gastruloids (Figure 2.9C). In the mouse embryo the formation of a COLLAGEN IV basement membrane precedes gastrulation, but here it is unclear if the layer also exists prior to stimulation, or it is produced from one or both populations of cells as differentiation proceeds.

The fact that the observed cell migrations are robust, concurrent with EMT, and dependent on the fate the cells adopt, suggests that we are seeing movements attempting to fulfill the *in vivo* human gastrulation program. In support of this is the fact that in the mouse mesoderm first migrates as compact "wings"[94,95] behind a leading edge of more dispersed definitive endoderm fated cells[96,97], since this is consistent with the rates and behavior of mesoderm and endoderm migrating cells in our gastruloids. Compared with cell migration in the avian primitive streak[98–102] not much else is known about mammalian primitive streak cell migration or the mechanisms and chemical cues behind it[103]. We believe that our gastruloid model offers a glimpse of this difficult to study *in vivo* process, and moving forward may present a fruitful alternative approach to dissect the molecular mechanisms underlying cell-migration during a pivotal time of human gastrulation.

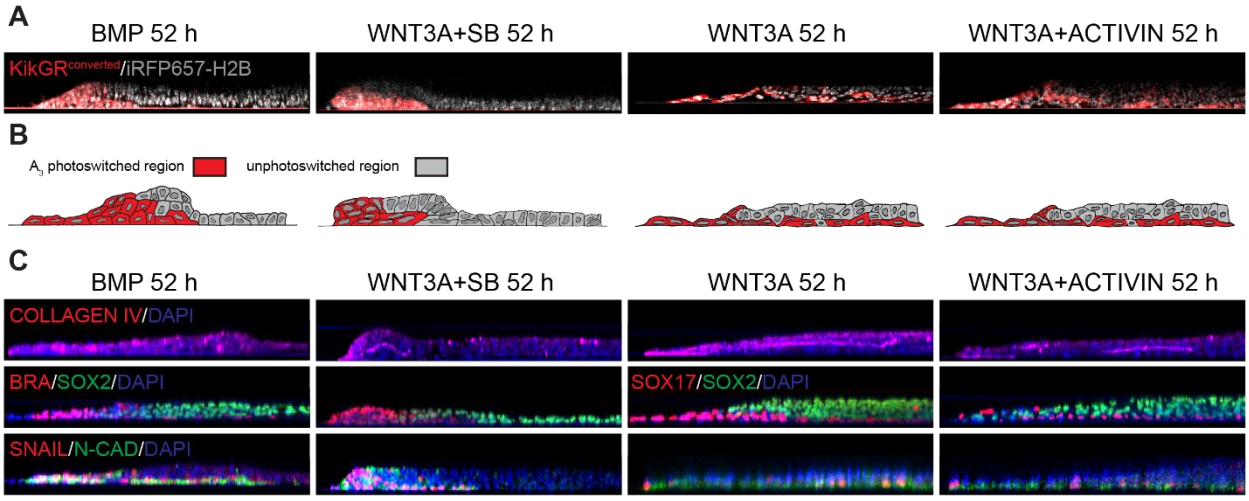


Figure 2.9| 3D gastruloid morphology

(A) Radial cross-sections of RUES2-KiKGR-RFP657-H2B micropatterns photoconverted in region A_3 and stimulated with WNT3A, WNT3A+Activin, WNT3A+SB, or BMP4 for 52 h. In all conditions the photoconverted (red) cells can be seen to be migrating under the inner epiblast or ectoderm-like region. (B) Hand-drawn depiction of 3D structure of gastruloids inferred from (A). (C) Radial cross-sections of RUES2 gastruloids stimulated with WNT3A, WNT3A+Activin, WNT3A+SB, or BMP4 and fixed and stained at 52 h for the indicated markers. As can be seen by comparison with (A) and (B), the migratory cells are also differentiated to either mesoderm or endoderm and express PS markers. In all the conditions except BMP4, one can also see that a basement layer of COLLAGEN IV separates the migrating cells from the undifferentiated epiblast-like cells overtop, as would be expected in vivo.

Mapping cell migrations and fates to the human primitive streak

Putting together our gene maps and anterior-posterior signatures, our cell migration patterns, and 3D cross-sections, we are able to suggest a detailed graphical representation at what gastrulation may look like in human PS at various anterior-posterior positions (Figure 2.10). We propose that the edges of the Epibalst/PrEct region of each gastruloid correspond to the median of the PS, while the centers of each gastruloid are positioned laterally relative to this median. In this schema the direction of migration of differentiating cells (indicated by arrows) is from the medial line of the streak out laterally, underneath the COLLAGEN IV and epiblast or primitive ectoderm layers. The uncovered region of differentiated cells on the edge of our gastruloids we believe would be covered in the embryo since in that anterior-posterior polarized streak geometry we would expect the Epibalst/PrEct region to grow (as in mouse) or flow (as in chick) to fill in that space.

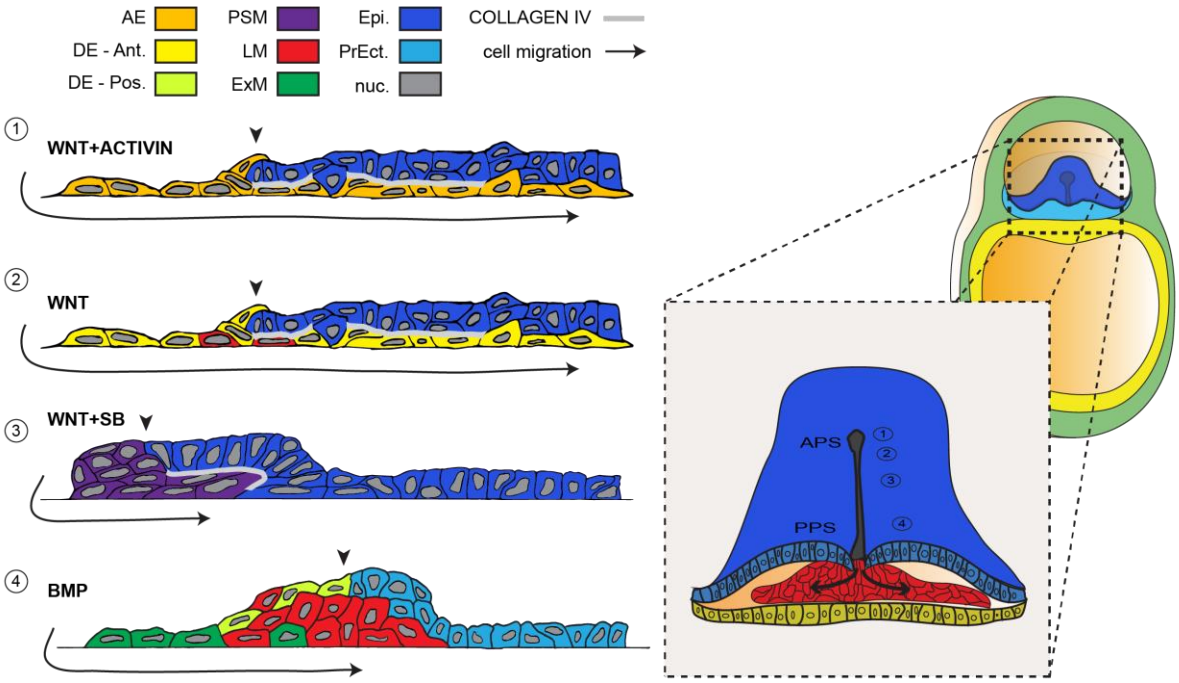


Figure 2.10| Mapping gastruloid cell migrations and fates to the human PS

(A) Cartoons summarizing the fates and 3D structure of each type of gastruloid at 52 h and mapping to the human embryo. As indicated by the arrowheads in the diagram, we believe the edge of the Epiblast/PrEct region in each gastruloid corresponds to the medial part of the *in vivo* primitive streak, and that our migrations (indicated by arrows) therefore occur medially to laterally.

Discussion

In this chapter I have shown that the necessary role of WNT signalling in initiating the PS is evolutionarily conserved in humans. Furthermore, I have shown that the human signalling initiation hierarchy between BMP, WNT, and ACTIVIN/NODAL pathways closely resembles the mouse rather than the chick, and WNT is necessary and sufficient to induce a streak of mesoderm and endoderm with EMT characteristics. Moving on to investigating the role of WNT in patterning the human PS, I was able to find and largely separate the highly structured subpopulations of PS fates that emerged in the gastruloids according to their corresponding position along an anterior-posterior axis and their BMP, WNT, and ACTIVIN/NODAL requirements. I was also able to discern and characterize robust cell migrations from the PS region of each gastruloid that depended on which fates the cells would ultimately adopt. Putting these pieces together, I was able to derive a rudimentary first fate map of the human PS. There is no doubt that this map lacks details and features that could be observed in the developing *in vivo* human embryo. We anticipate that missing cell types, such as germ cells or intermediate mesoderm, for example, might be revealed in the future with the use of single cell RNA-seq of gastruloids and sets of markers informed by new efforts to acquire single cell RNA-seq data from gastrulating primate embryos[104]. There is also the limitation that, unlike the *in vivo* case, our anterior-posterior streak is a composite of separate differently stimulated gastruloids. That said, given what we have learned about the required stimulation conditions for each fate subpopulation, it may be possible with advances in micropatterning techniques or localized ligand sources to recreate the entire anterior-posterior streak in a single micropattern. With further investigation it should also be possible, given the robust nature of the observed cell

migrations, to determine more detail about the mechanisms governing these migrations; the identity of chemoattractants or repellents and their sources for example. This would be a superior model and allow much better understanding of the relative timing of EMT, fate specification, and migrations in the *in vivo* human PS. Regardless of the limitations of our current studies, however, we believe our results represent a significant first step forwards to observing, mapping, and understanding the role of Wnt signalling in this crucial stage of human development.

CHAPTER 3: DISCOVERY OF A FUNCTIONAL HUMAN ORGANIZER

Much of the significance of the fate mapping results in the previous section depends on the correspondence of the hESC gastruloid model to the *in vivo* human primitive streak. As embryonic studies of the human primitive streak are ethically prohibited, the proof of this correspondence so far has relied on indirect comparisons of the gastruloid to closely related model organisms, looking for instance for similarities in molecular markers, 3D structure, and collective cell behaviour. It may be argued, however, that there is yet a stronger test for correspondence that could be carried out: the classical embryological test of "functionality", i.e. whether one cell or tissue type can substitute for another *in vivo* and carry out all of its normal functions for proper development.

The Organizer

Of the many cell types in the gastruloid model to possibly assess functionality for, the organizer cell type is the most obvious target. This is because in many ways the organizer sets the paradigm for functionality. It was discovered in 1924 through the pioneering experiments of Spemann and Mangold with the demonstration that a small group of cells located on the dorsal side of the early amphibian embryo have the ability to induce and "organize" a complete secondary axis when transplanted to the ventral side of another embryo[105]. Later discovery of embryonic tissue with similar organizer activity in fish, birds, and rodents[106–109] demonstrated that this early embryonic activity was evolutionarily conserved, and could work even in cross-species transplants. In order to functionally qualify as an organizer, cells must fulfill two stringent criteria:

(i) they must contribute autonomously to axial and paraxial mesoderm, including head process and notochord; and (ii) they must induce neural fate non-autonomously in their neighbors.

Besides having such a well-defined test, another reason for choosing to test the functionality of the gastruloid organizer cells is because the human organizer remains undefined. In the almost 100 years since the famous Spemann-Mangold experiment, technical and ethical difficulties have prevented researchers from demonstrating the presence of an organizer in human embryos. Thus a successful functional test of the organizer cells in our gastruloid would not only be a stronger validation of our previous results, but would also be a major developmental milestone in human embryology as well.

Chick chimeras

An hESC to human embryo test of functionality is impossible due to ethical limitations, but interspecies chimeric assays have been used in the past to validate hESC results from cell culture, such as for example previous work from my lab featuring the first human-mouse chimera[60]. Attempts have been made to graft GSC expressing hESC embryoid bodies into early amphibian embryos[110], but the method employed was highly problematic and susceptible to misinterpretation[111] and the grafts failed the test of functionality. Complicating matters further, the embryoid bodies used also only expressed a low percentage of GSC positive cells, unlike the WNT+ACTIVIN treated gastruloids. In contrast to amphibian, the mouse model may seem like the more obvious choice for attempting an interspecies graft to test the hESC organizer, given its much closer evolutionary distance to human, but early pre-PS mouse embryos at the organizer

stage are difficult to isolate and even mouse intraspecies organizer grafting operations present considerable challenges to the most experienced mouse embryologists[108]. In contrast to both amphibians and mammals, chick embryos are excellent models for interspecies transplantation studies and provide a technically less challenging platform[112–114]. The great advantages of the chick model for studying developmental biology are its affordability, versatility, and ease of access[115,116]. Armed with only an incubator and a nearby farmyard or poultry facility one can examine almost any stage of vertebrate development, from pre-gastrulation to neurulation to hatching, simply by cracking open an egg at the appropriate time and examining its contents. The chick has precise classifications of embryonic stages[117], and is amenable to classic experimental embryological manipulations, such as tissue grafting, ablation, tissue recombination, and genetic perturbations. Due to its transparency the chick embryo can also be imaged live *in ovo* or *ex ovo* using simple culture systems. Chick embryos are excellent models for interspecies transplantation studies, providing a technically less challenging platform than the mammalian counterpart.

Indeed, the chick experimental model has been used for decades as a xenograft host for chimera experiments[118–122], beginning with the classical work of Nicole le Douarin and colleagues who grafted quail neural crest cells at the neural plate boundary of the chick embryo, generating chick-quail chimeras[112–114]. Successful transplantations of mammalian cells and tissues into the chick embryos has also been reported and is facilitated by the fact that they can grow in similar temperature condition. Mouse-chick chimeras have been used to understand and dissect the mechanism of specific cell types and tissue such as neural crest, motor neurons, and somite

specification[123–125]. Mammalian cells thus respond to local signalling and contribute to the chick anatomy. Finally, human embryonic and adult stem cells have been previously transplanted in chick embryos and shown to be adaptable to local signals[119,126,127].

Inspired by the potential of the chick system as a host where I could functionally test the gastruloid organizer micropatterns, I devised an *ex ovo* cross-species transplantation strategy based on previous mammalian organizer studies[128,129], grafting fluorescent reporter hESC micropatterns treated for 24 or 48 hours with WNT3A+ACTIVIN into the marginal zone of Early Chick (EC) culture embryos[130] (stage HH 2 to 3+). I used 500 µm diameter rather than 1000 µm diameter micropatterns as these gave a purer population of GSC+ cells, and I grafted at 24 hours post-treatment as well as at 48 hours post-treatment as 24 hours is when GSC first becomes apparent and is also co-expressed with BRA (Figure 3.1). For the reporter line, I used the CRISPR-Cas9 generated RUES2-GLR (Germ Layer Reporter) cell line created in our lab and reported previously[131]. Figure 3.2 shows the preparation of the chick embryo for grafting, and Figure 3.3 shows the transplantation of the hESC WNT+ACTIVIN micropattern into the chick embryo.

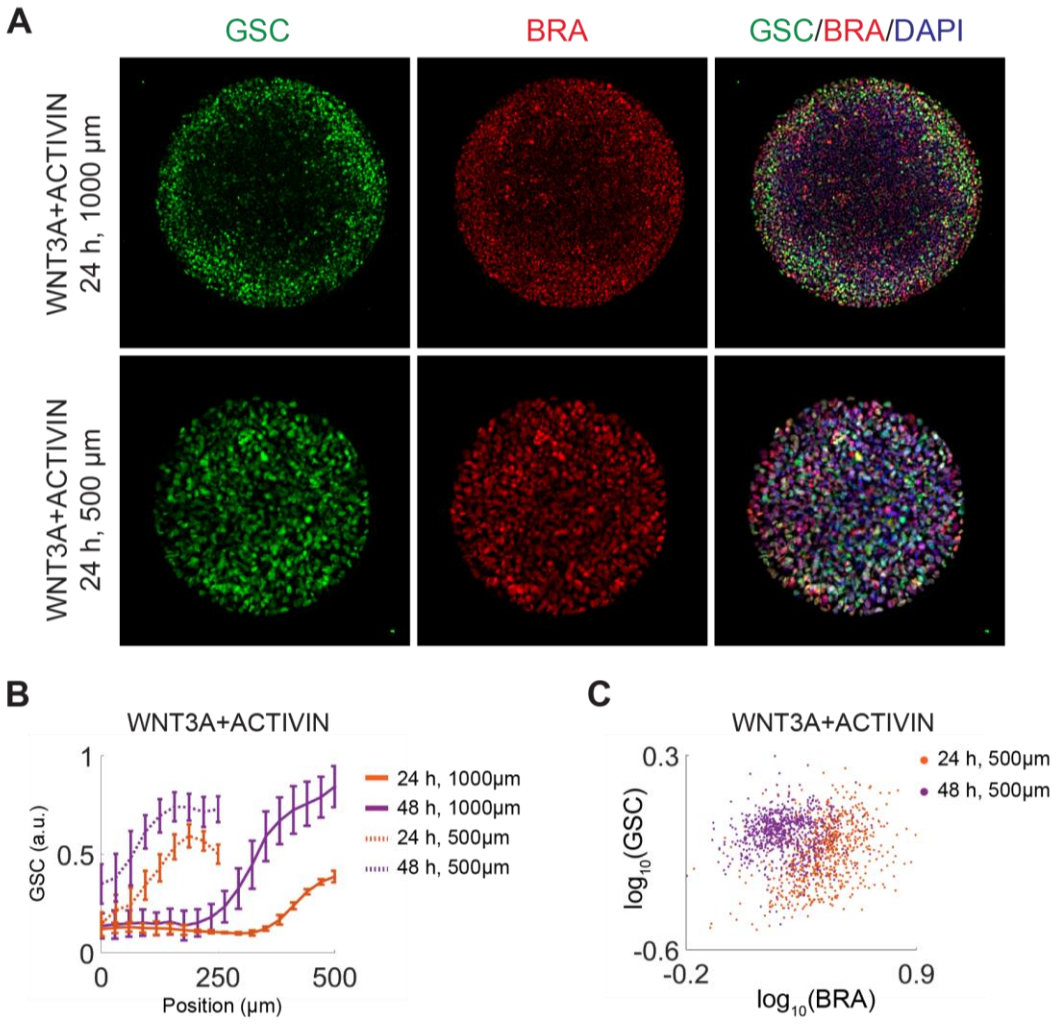


Figure 3.1 | Further hESC organizer characterization

(A) 1000 µm and 500 µm diameter micropatterned colonies stimulated with WNT3A+ACTIVIN and fixed and stained for GSC and BRA at 24 hours. Note that as observed by Warmflash et al.,[69] for BMP induction, shrinking the colony size results in removal of center micropattern fate region, thus resulting here in a higher proportion of GSC expressing cells. This experiment was repeated at least *n*=3 times independently with similar results. (B) Quantification of (A): the radial profile represents the average of *n*=25 colonies and errors bars represent the standard deviation. (C) Scatterplot of single-cell expression of GSC vs BRA. Note that at 24 hours most cells co-express BRA and GSC, but that by 48 hours GSC expression is increased and BRA expression is decreased. Because of this we grafted micropatterns at 24 hours as well as at 48 hours post-stimulation, reasoning that earlier co-expression of BRA and GSC would result in greater graft contribution to axial mesoderm structures.

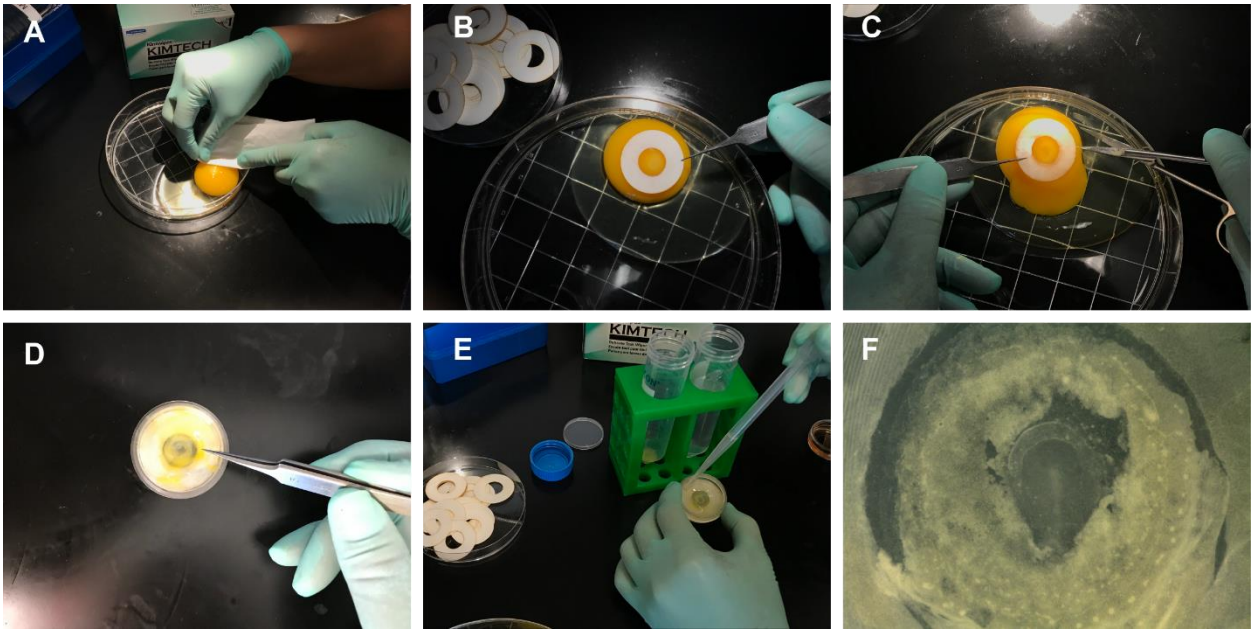


Figure 3.2 | Preparation of host embryo

(A) Cleaning albumin from area around the embryo. (B-D) Attaching filter paper frame, cutting vitelline membrane around the frame, and removing frame from yolk and placing in 35 mm dish with agar mount. (E) Washing the EC culture host. (F) Ventral view of HH 3 stage EC culture embryo ready for donor graft.

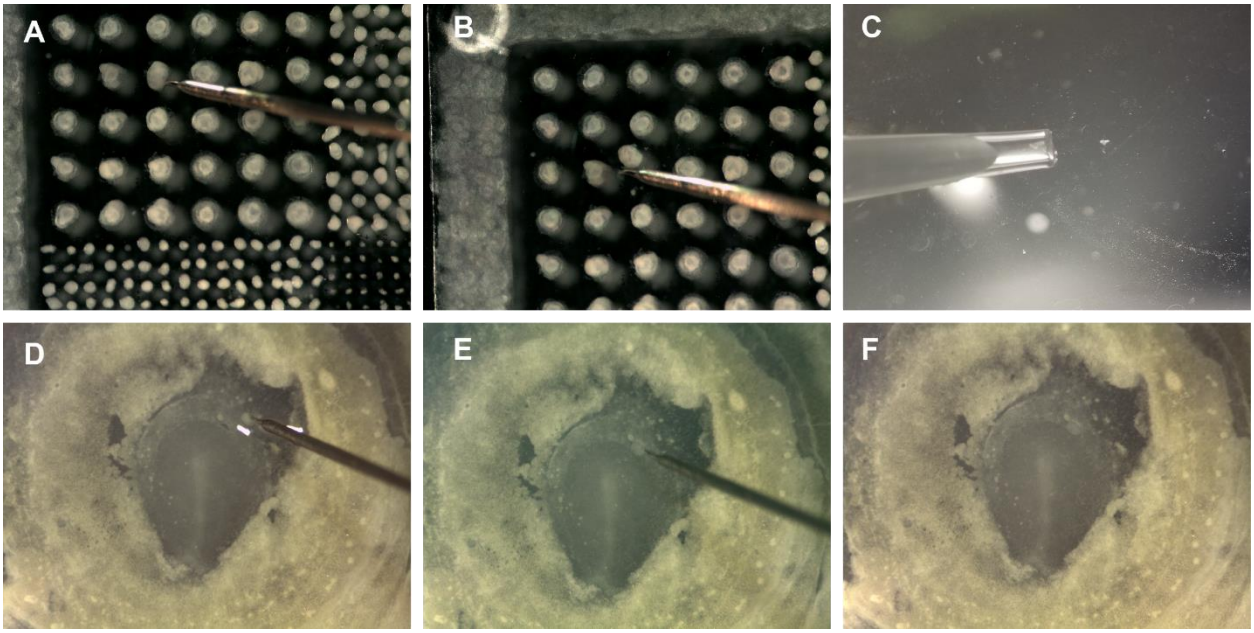


Figure 3.3 | Grafting manipulations

(A) Detaching a suitable micropattern from the coverslip. (B) Detached free floating micropattern. (C) Positioning the washed detached micropattern at the bottom of a 10 µl pipette tip. (D-F) Locating the transferred micropattern and sliding it into the pocket created in the marginal zone.

Combined WNT and NODAL signalling generate a functional human organizer

We found that RUES2-GLR grafts survived, mingled with host cells, and induced and contributed to a secondary axis that became obvious between 24-48 hours (Figure 3.4A-K). Both the live cell reporter and a human-specific nuclear antigen revealed that the human cells directly contributed to the ectopic axis autonomously and continued differentiating in their new location, contributing both BRA and SOX17 cells (Figure 3.4G, K-L). This mirrors previous observations in mouse-to-mouse organizer grafting experiments[108]. Confocal cross-sectioning of these secondary axes often revealed self-organizing features directly resembling those found in the early chick and mouse embryo, for example correct layering of germ layers, and central elongated notochord-like structures composed partially or wholly of graft derived cells (Figure 3.4M-Q and Figure 3.5). Analysis of molecular markers also established that the human cells induced neural tissue in the chick non-autonomously: SOX2 and SOX3 were ectopically induced in chick cells that surrounded the human cells (Figure 3.4D-F, K-L, R). Additional *in situs* and antibody staining for HOXB1, GBX2, and OTX2 established that the neural tissue was predominantly posterior in nature (Figure 3.4S-U). Since in the mouse the early-gastrula-organizer and late-streak node also does not induce anterior neural structures when grafted to another mouse embryo, this result suggests that our human organizer is closer to these organizer stages than to the mouse mid-gastrula-organizer[92,109]. As controls, RUES2-GLR grafts treated instead with WNT3A, WNT3A+SB, BMP4, or blank media showed less overall survival and never induced chick neural markers (Table 3.1 and Figure 3.6). Taken together with the morphological, cellular, and molecular evidence described in the previous chapter, this functional test in an embryonic environment provides the most stringent evidence for the induction of a human organizer.

Figure 3.4 | Human organizer induces secondary axis in chick embryo

(A-F) Secondary axis induced by 24 h stimulated RUES2-GLR colony into HH stage 3 chick; (A-C) SOX17-tdTomato (red) live marker at 0, 24, and 38 h post-graft; (E-F) SOX2 (green) is prominent in the tip of the secondary axis 48 h post-graft, and does not overlap with the hESCs (D and F, red, Human Nuclear Antigen). Scale bar is 200 μm. (G-J) Another example of a 24 h stimulated hESC micropattern inducing a secondary axis in a chick host, 27 h post-graft; (G) live image of SOX17-tdTomato hESC cells (red); (H and J) fixed stains for Human Nuclear Antigen (HNA, red) and SOX2 (I and J, green). Scale bars are 500 μm (G) and 200 μm (H-J). (K-Q) Example of secondary axis induction from a 24 h stimulated hESC micropattern with more complete self-organizing structures, 27 h post-graft; (K) live image of SOX17-tdTomato hESC cells (red); (L) confocal slice of secondary axis for DAPI (grey), HNA (red), and BRA (green); (M-Q) confocal cross-section of indicated region in (m), with the same channels plus SOX17 (blue). Note in the merged image (Q) how the secondary axis is layered, with epiblast chick cells on top of a layer of human BRA cells which in turn are on top of a layer of human SOX17 cells, exactly how the epiblast, mesoderm, and endoderm layers would arrange themselves in a gastrulating mouse or chick embryo. Scale bars are 500 μm (K), 100 μm (L), and 50 μm (M-Q). (R) In situ for chicken *SOX3* shows expression in the host chick throughout the neural tube and head, as well as in the induced secondary axis. (S) *OTX2* is expressed in the host forebrain but is absent in the graft induced tissue (indicated by arrow). (T) *HOXB1* is expressed in the host and the graft induced secondary axis. (U) *GBX2* is expressed in the host and the graft induced secondary axis.

(V-W) Zoom of region indicated in (U): (V) shows secondary axis and tdTomato-hESCs (red) after fixation; (W) shows *GBX2* expression after in situ. The arrow shows the location of the graft hESCs before and after. All experiments were performed at least *n*=3 times with similar results, for exact numbers and measure of reproducibility please see Table 3.1.

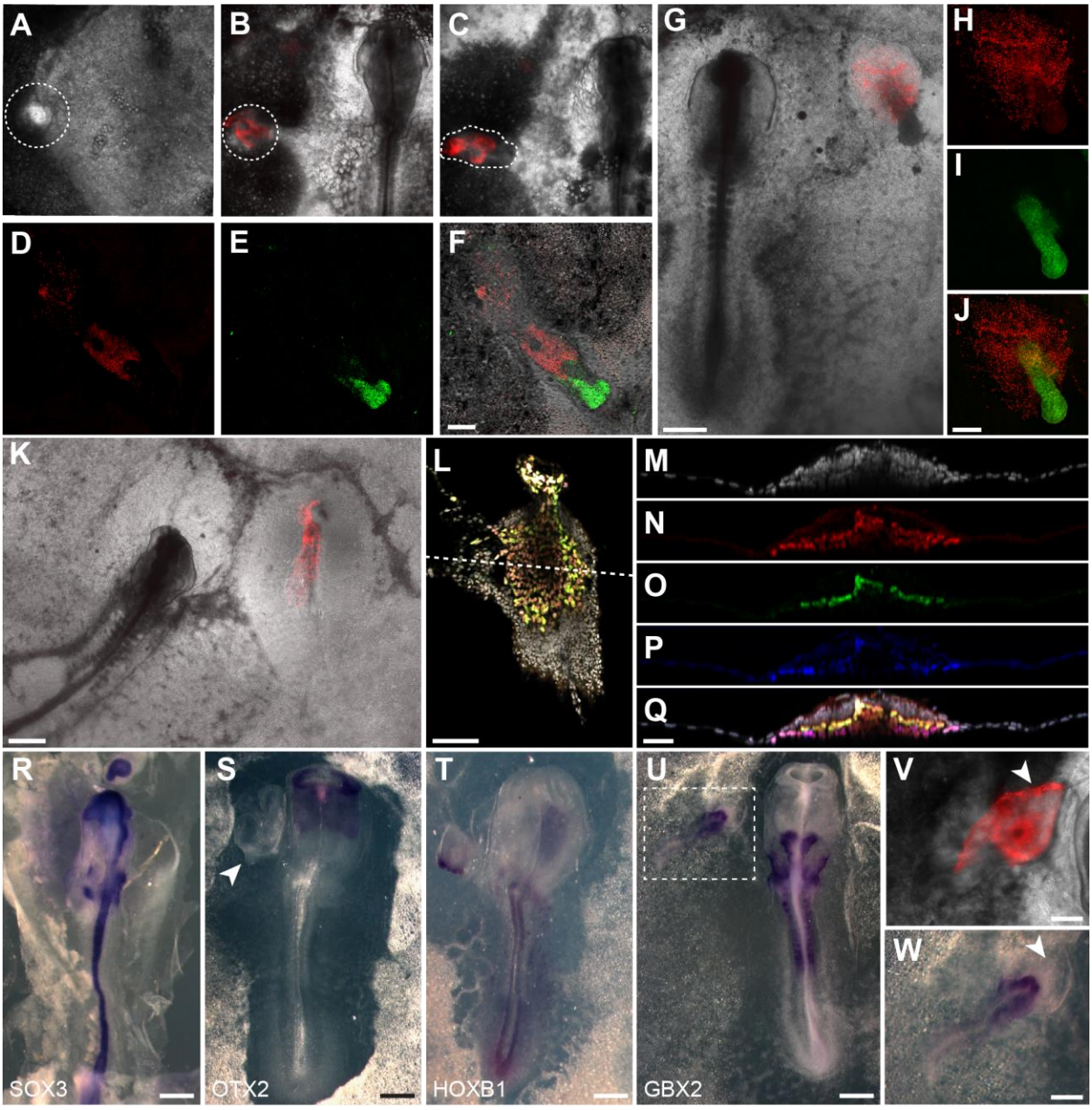

A
B
C
G
H
I
J
D
E
F
K
L
M
N
O
P
Q
R
S
T
U
V
W
SOX3
OTX2
HOXB1
GBX2

Table 3.1 | Induction of chick neural tissue by hESC micropatterns

	Survival	SOX2	*SOX3*	*GBX2*	*HOXB1*	*OTX2*
control (blank)	4/15	0/15	-	-	-	-
WNT3A+ACTIVIN 24 h	273/300****	10/19****	-	5/6	9/10	0/14
WNT3A+ACTIVIN 48 h	37/40****	6/15****	6/8	-	-	-
WNT3A 48 h	8/15	0/15	-	-	-	-
WNT3A+SB 48 h	7/15	0/15	-	-	-	-
BMP4 48 h	14/15***	0/15	-	-	-	-

Survival column represents whether a treated and grafted hESC micropattern was detected in the host chick 24 h post-graft, as measured live with the RUES2-GLR cell line. Many of these grafts were used to optimize and test various versions of the antibodies and probes listed in the remaining columns. Statistical analysis (χ^2, 2x2contingency test, compared to the control condition): *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Figure 3.5 | Further characterisation of the induced secondary axis

(A) Examples of classifying the notochord-like feature (NLF) based on morphology. For z=+19 µm, one can discern the NLF as a tighter and brighter rod of cells running north-south that is also distinct and somewhat separated from the surrounding chick epiblast. For z=+46 µm, one sees that paired elongated cells stick out ahead of the other cells in a continuation of the originally identified NLF. Other cells belonging to the NLF between z=+46 µm and z=+19 µm are obscured at these slices or out of focus, but can be easily identified slice-by-slice at the other z positions. Scale bar is 100 µm. (B) Top to bottom: yellow shows co-Sox17:tdTomato (blue) with human (red) cells; cross-section shows that chick and human cells arrange themselves into germ layers properly, and that they flank the central notochord-like feature indicated by the arrow (cyan); a proportion of human mesoderm cells contribute to part of the notochord-like structure, while the cyan-coloured cells without HNA (red) shows that the remainder of the NLF is composed of host cells. (C) Examples of donor hESC graft contributing to the induced notochord-like feature, imaged live 27 h (left) and 23 h (right) post-graft. Scale bars are 200 µm (left) and 100 µm (right). Similar notochord-like features were observed in at least *n*=10 independent biological replicates.

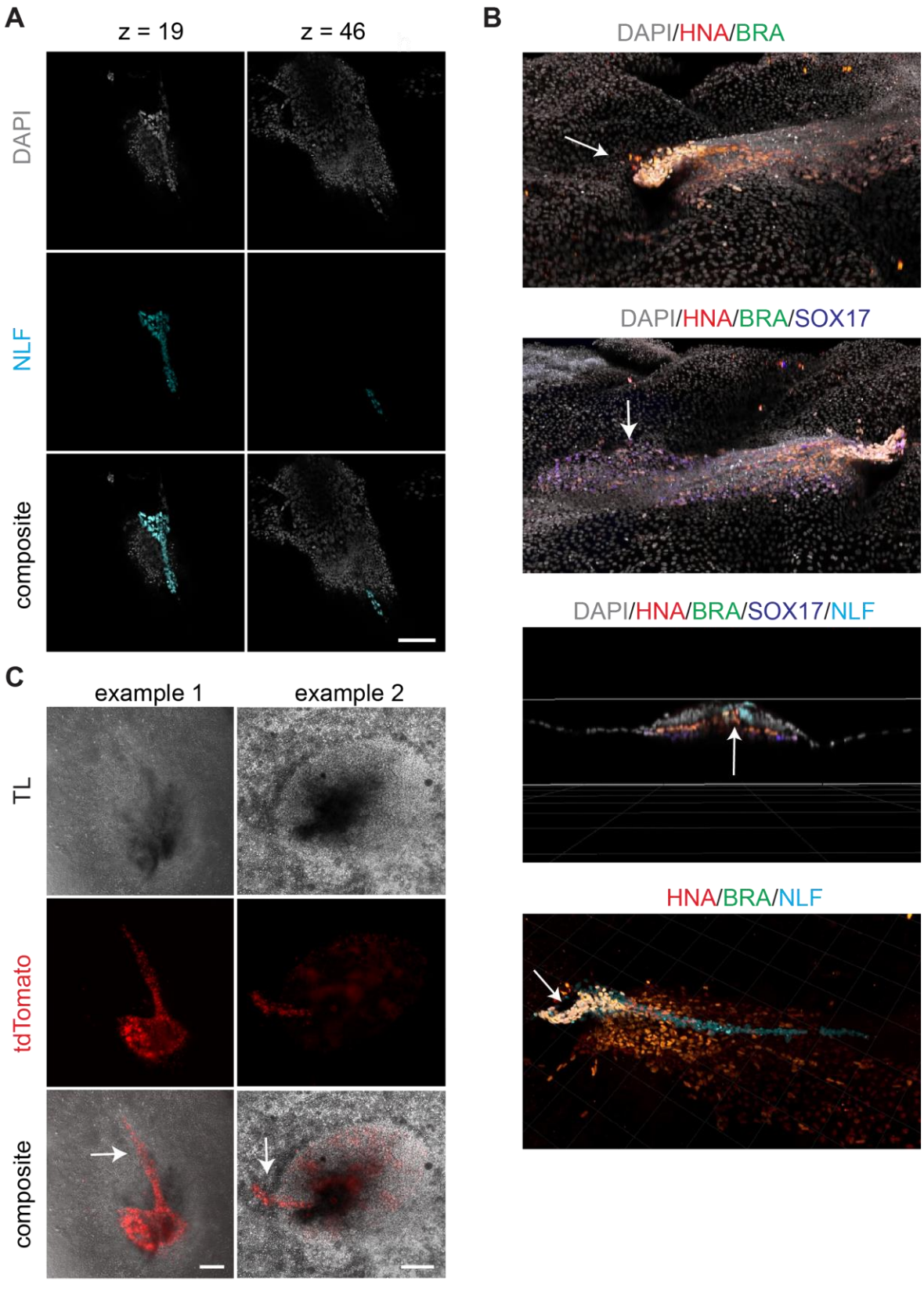
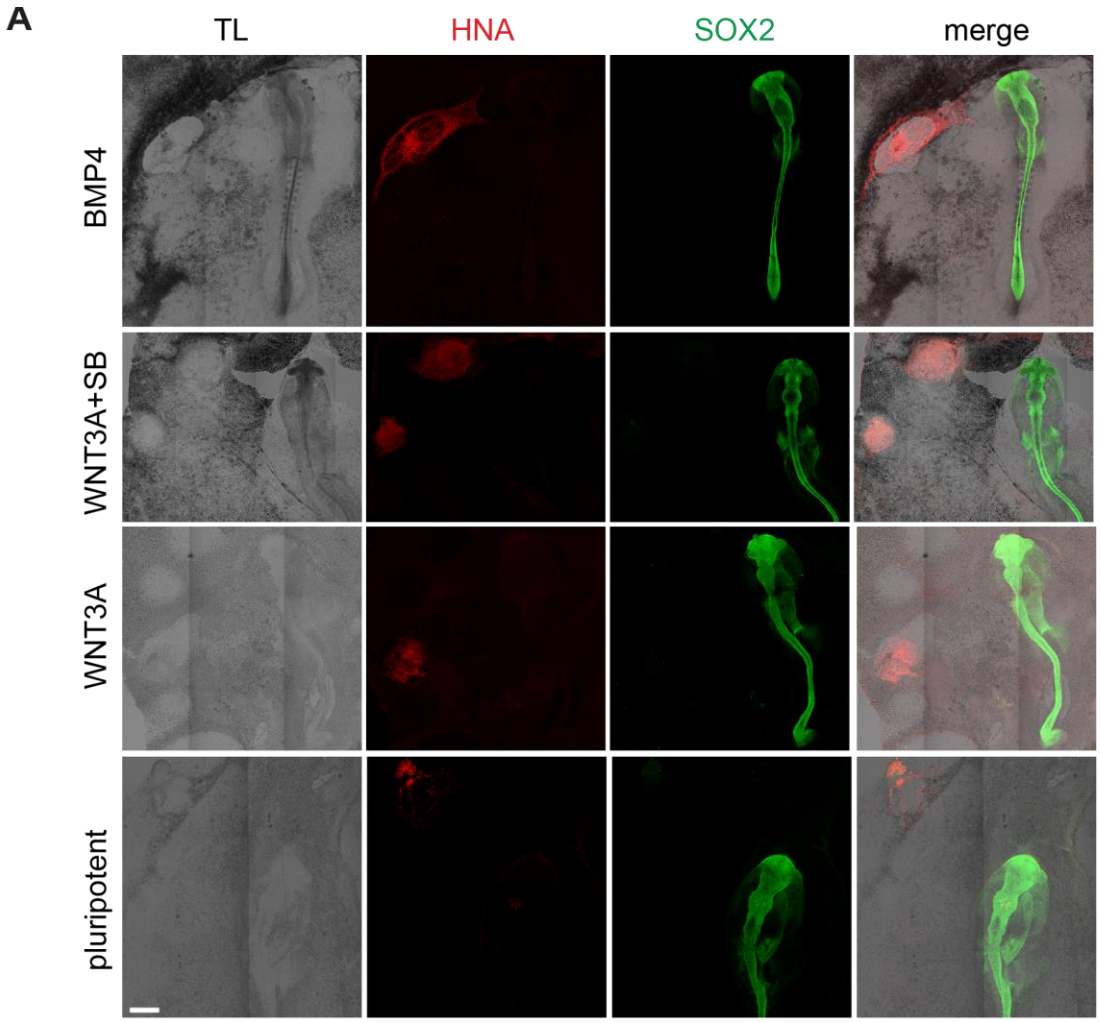
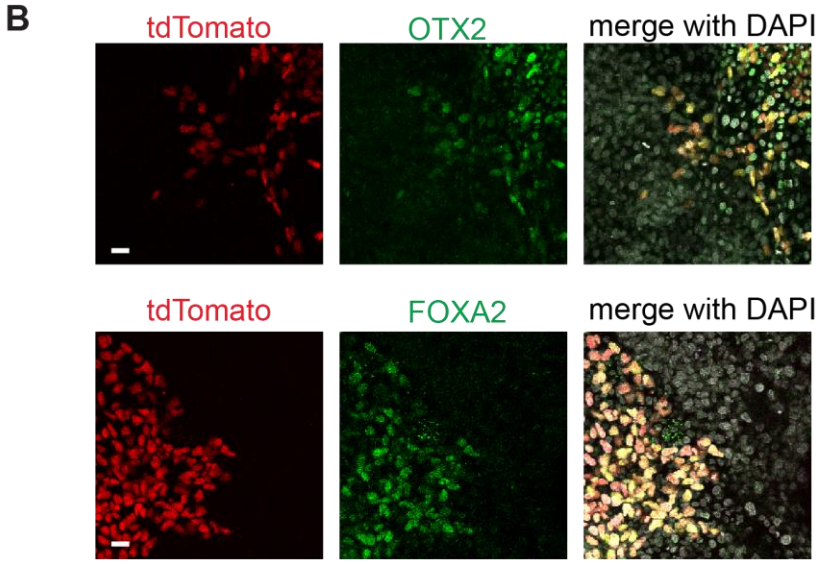

A
z = 19
z = 46
DAPI
NLF
composite
B
DAPI/HNA/BRA
DAPI/HNA/BRA/SOX17
DAPI/HNA/BRA/SOX17/NLF
HNA/BRA/NLF
C
example 1
example 2
TL
tdTomato
composite

Figure 3.6 | Control chick grafts

(A) Representative grafts for control conditions. With the exception of the BMP4 control condition, grafted hESC colonies were static, with the colonies either growing or dying in place. With BMP4, often the colonies were elongated, possibly due to hESC migration. In all control conditions, however, there was never induction of SOX2 in the host cells. Note that in the case of the WNT3A+SB graft shown, two colonies were grafted into two different locations. Scale bar is 500 µm. Experiments were repeated at least *n*=3 times independently with similar results. (B) Confocal cross-sections showing co-expression of SOX17 (tdTomato) and FOXA2 or OTX2 in human cells that contribute to the secondary axes induced by a 24 h WNT3A+ACTIVIN stimulated hESC micropattern. Scale bar is 20 µm. Experiments were repeated at least *n*=3 times independently with similar results.


A
TL
HNA
SOX2
merge
BMP4
WNT3A+SB
WNT3A
pluripotent


B
tdTomato
OTX2
merge with DAPI
tdTomato
FOXA2
merge with DAPI

Discussion

Our ability to generate a human organizer closes the loop initiated by classical experimental embryologists working on amphibian systems, nearly 100 years ago, and demonstrates that the concept of the "organizer" is evolutionarily conserved from frogs to humans. The fact that the chick and human cells are able to signal to one another and self-organize correctly across 350 million years of divergent evolution is also a striking demonstration of the conservation and importance of the signalling pathways involved at this critical phase of development. More generally, our chick experiments also define a novel *in vivo* platform to validate results obtained in an *in vitro* gastruloid platform, and may be generally applicable to test and explore other aspects of early human development.

CHAPTER 4:THE CONTROL OF WNT SIGNALLING IN THE HUMAN PRIMITIVE STREAK

In the previous chapters I focused on understanding the role of WNT signalling in the human PS, its function in forming and patterning the germ layers that emerge from it. I have so far largely ignored the question of what instead controls WNT signalling in the human PS, what determines its extent and magnitude in time and space. The question is important not just because nothing is known about these patterning dynamics in the human embryo, but because throughout development long range signalling by morphogens and their inhibitors define embryonic patterning yet quantitative data and models are rare. Our gastruloid system, with its self-organized patterned response to a uniformly presented WNT ligand, offers an ideal environment to explore how WNT signalling leads to spatial organization, and specifically how the human PS forms and is spatially confined.

In this chapter I address the molecular mechanisms underlying WNT-mediated self-organization of human PS. We show that two primary factors control patterning: E-CAD and DKK1. First, E-CAD establishes a pre-pattern by limiting the initial WNT response to the boundary. Secondly, and in parallel to the NOGGIN dynamics in the BMP case, the secreted inhibitor DKK1 is upregulated by a combination of WNT and NODAL signalling and is required to ultimately confine the PS to the colony boundary. Multiple single and double combinations of homozygous CRISPR/Cas9 knockouts of secreted inhibitors of the WNT and NODAL pathways confirmed that only DKK1 plays a major role in the spatial restriction of the PS. We found that CERBERUS1 (CER1) is also highly upregulated by a combination of WNT and NODAL signalling, but that in our cells it functions as a NODAL inhibitor rather than dual WNT/NODAL inhibitor. CER1 thus does not

influence the size of the PS, but instead serves to bias the mesoderm versus endoderm fate decision in this region. We found also that in our DKK1$^{-/-}$ cells E-CAD not only establishes a pre-pattern, but, via its mutual antagonism with WNT, generates a cooperative EMT wave that travels from the micropattern periphery to the center. We show via quantitative modelling that this EMT wave is a generic property of a bistable system with diffusion and a single quantitative model describes both the wave and our knockout data.

WNT response is edge and density dependent and apically-basally symmetric

In Chapter 2 we showed that uniform application of WNT3A ligand to hESC micropatterns is sufficient to self-organize a PS-like structure, with mesoderm and endoderm emerging from an EMT on the colony periphery after 48 hours and with ACTIVIN/NODAL level biasing the choice of endoderm versus mesoderm (Figure 2.3A). During this time the transcription factor SOX17 demarcates the endoderm and the transcription factor BRA demarcates mesoderm. Changes in the EMT markers SNAIL, E-CAD, and N-CAD can also be used to identify the PS, but since these markers are more transient and harder to measure than BRA or SOX17[131], we will use the union of SOX17 and BRA to define the spatial extent of the induced PS. We also showed previously that despite the uniform application of WNT, the interior of the colony remains pluripotent, expressing both NANOG and SOX2 (Figure 2.3B). This pattern, with pluripotent cells on the interior and mesoderm and endoderm cells on the periphery, represents the terminal spatial pattern we seek to understand.

In order to decipher the molecular mechanism underlying this spatial pattern, we first attempted to use the detection of nuclear β-CAT as an early readout for canonical WNT signalling[132–134]. However, commercially available antibodies did not have adequate resolution on our dense epithelia (Figure 4.1A), so instead we used LEF1, a co-factor of β-CAT that is a direct target of WNT signalling with the same response profile as AXIN2 (Figure 4.1B) and is localized to the nucleus[21,135]. We found that the LEF1 response profile depended on colony density, with nuclear expression throughout the colonies at low density and restriction to the periphery at high density (Figure 4.2A). Co-presentation of the SMAD2 pathway inhibitor SB, together with WNT3A did not change the outcome, demonstrating that the density dependence of the LEF1 pattern is specifically due to WNT, and not caused by secondary ACTIVIN/NODAL signalling.

Similar to previous work studying the effect of colony density on BMP signalling[71], low density represents a pre-epithelial state before tight junctions have completely established themselves, while high density represents an epithelial state with complete tight junctions (Figure 4.2B) where the SMAD1 response due to BMP stimulation would also be edge restricted (Figure 4.3C). Because density is an important variable, for all experiments in this chapter we consistently use two defined values of seeding density “low” or “high” (Figure 4.3B).

As one of the factors involved in BMP4 induced self-organization was due to polarized signal reception[71], we first examined the localization of the WNT receptors[136] FRIZZLED and LRP6 in our micropatterns. We find that while some WNT-receptors were detected on the apical side, they were predominantly and homogenously located basolaterally underneath the tight junctions (Figure 4.2C). We also find little distinction between edge and center. To functionally test for

signal reception, we cultured cells on transwell-filter culture dishes, where cells can be selectively stimulated from the apical or basal side. Cells were cultured at the same density as the high density micropatterns and stimulated them with WNT3A for 12 hours. While a stronger response was detected with basal rather than apical stimulation, the mean basal response on filters, however fell below the edge response on colonies (Figure 4.2D and E), suggesting that additional factors were involved in setting up the WNT response on the colony boundary.

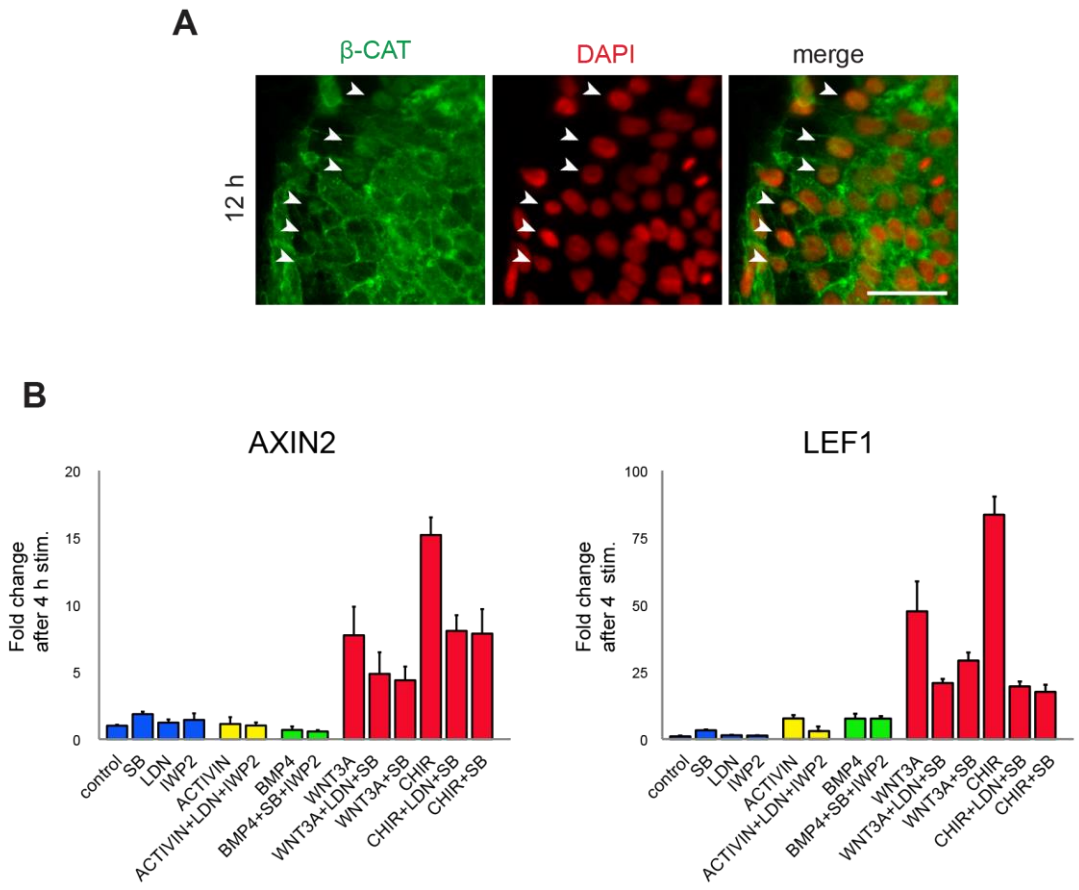
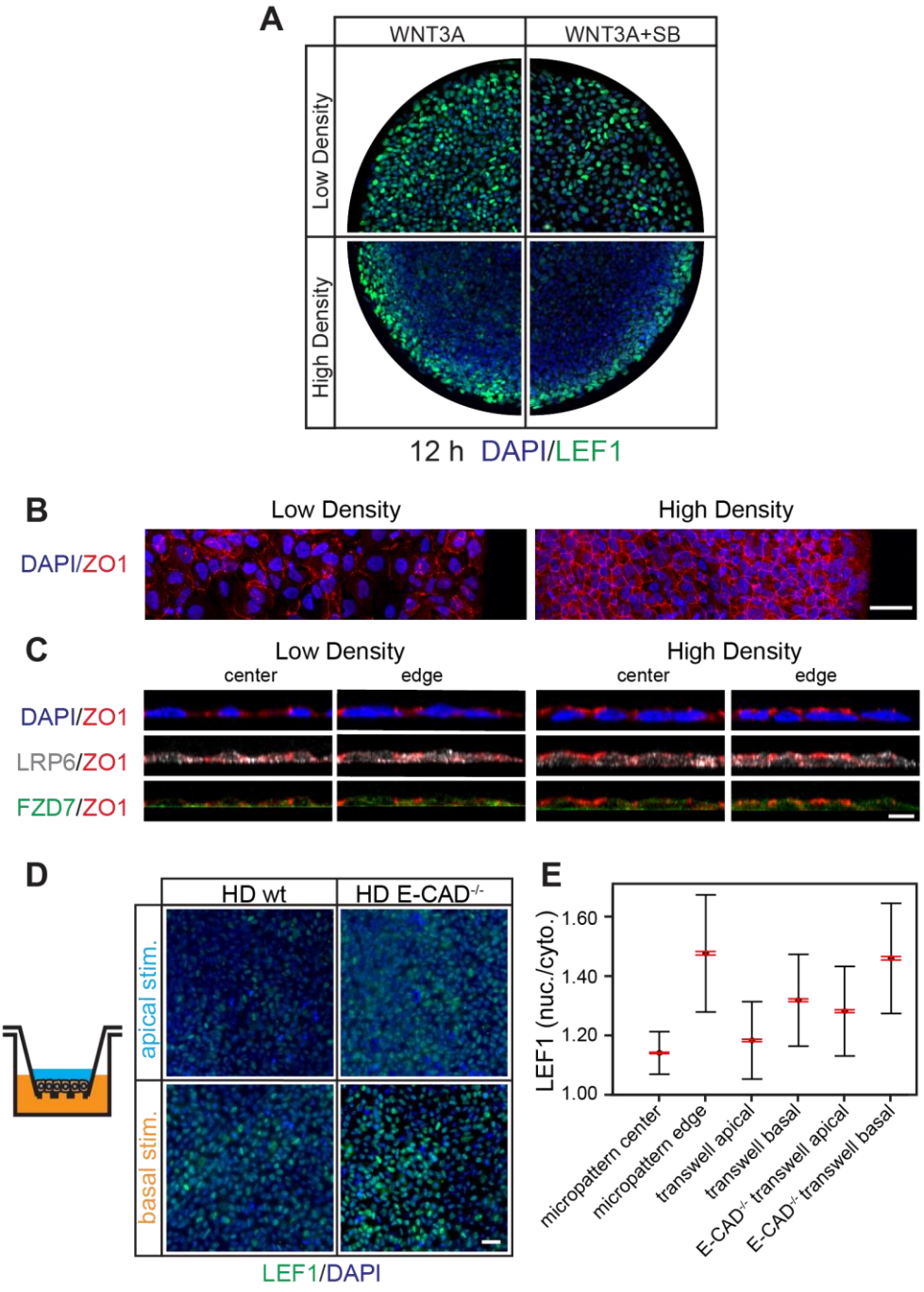


Figure 4.1 | Early WNT3A response markers

(A) Inset of an abnormally low density micropattern colony stimulated with WNT3A and fixed and stained at 12 h for active β-CAT. Even at this low density (seeded with $6x10^5$ cells, stimulated 2 hours after seeding) an increase in β-CAT in the nucleus is only visible in loosely connected cells on the colony edge (arrows). Nuclear β-CAT in cells away from the periphery or at higher density (including standard low density conditions) is not easily observed, largely due to signal from membrane-bound or cytoplasmic regions. Scale bar, 50 µm. (B) qPCR of AXIN2 and LEF1 as a function of inputs arrayed on x-axis. The results show that the LEF1 and AXIN2 response is dominated by WNT3A, though synergism with the NODAL pathway is significant as well, as can be seen by comparing WNT3A or CHIR with WNT3A+SB or CHIR+SB. Thus LEF1 and AXIN2 can be used as proxies for early WNT3A response. Also note that LEF1 gives a greater positive signal than AXIN2, reaching 84-fold induction compared with 15-fold induction for CHIR condition.

Figure 4.2 | WNT3A response is edge and density dependent

(A) Micropatterns stimulated with WNT3A or WNT3A+SB at high density (22 h after seeding, 3474 ± 430 cells/mm^2) or low density (8 h after seeding, 1810 ± 236 cells/mm^2) and fixed and stained for LEF1 at 12 h. (B) Maximum intensity projection of tight junction marker ZO1 and nuclear marker DAPI in low density and high density micropatterns immediately prior to stimulation. Note that the network of tight junctions is only fully formed in the high density micropatterns. Scale bar, 50 µm. (C) Cross-sections showing the apical-basal position of WNT receptors relative to DAPI and ZO1 (apical marker). At high density the cells are in an apically-basally polarized epithelial state, as judged by the relative position ZO1 and DAPI. While not polarized themselves, the majority of the WNT receptors in this state do lie underneath the tight junctions, and so presumably are not as accessible from the apical side as the basal side. This is in contrast to the low density state where cells are not epithelized and receptors are visible on both sides of ZO1. Additionally, there is no significant edge-vs-center expression of the receptors at either density. Scale bar, 10 µm. (D) and (E): basally stimulated high density hESCs in transwell filters show a higher WNT3A response than when apically stimulated. However, judging by the quantification of LEF1 nuclear to cytoplasmic ratio in (F), this difference is not enough to explain the edge-vs-center difference in micropatterns. It takes the knockout of E-CAD and basal stimulation to reach the micropattern edge level of WNT activation. The black error bars in (E) represent the standard deviation on 1,000 cells and the red error bars represent the standard error on the mean, illustrating the significant difference of the mean between each sample. The scale bar in (D) is 50 µm.


A
WNT3A
WNT3A+SB
Low Density
High Density
12 h DAPI/LEF1
B
Low Density
High Density
DAPI/ZO1
C
Low Density
High Density
center
edge
center
edge
DAPI/ZO1
LRP6/ZO1
FZD7/ZO1
D
HD wt
HD E-CAD-/-
apical stim.
basal stim.
LEF1/DAPI
E
LEF1 (nuc./cyto.)
1.60
1.40
1.20
1.00
micropattern center
micropattern edge
transwell apical
transwell basal
E-CAD-/- transwell apical
E-CAD-/- transwell basal

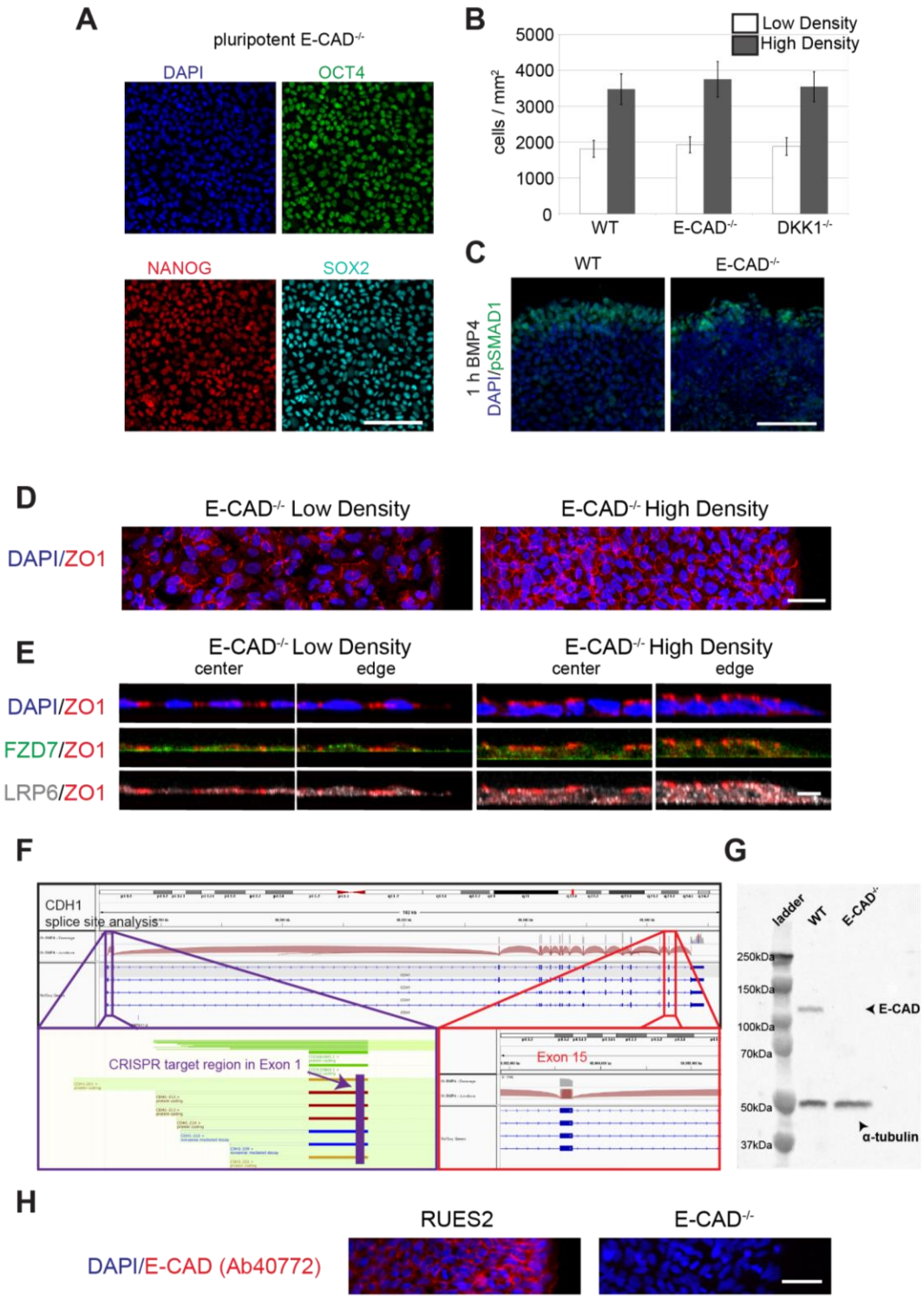
E-CAD knockdown sensitizes cells to WNT

Given the fact that E-CAD is classically considered to antagonize WNT signalling via its binding and sequestration of β-CAT[23,24,27,137], and the fact that cells on the periphery of our micropatterns have fewer neighbouring cells and so presumably fewer E-CAD junctions, we hypothesized that E-CAD may contribute to the early WNT pattern. To test this hypothesis, we made clonal CRISPR/Cas9 E-CAD knockout cell lines. We used a guide RNA targeting a region present in all isoforms of the protein and confirmed the result with western blot and immunofluorescence using two separate antibodies (Figure 4.3F-H). We found that E-CAD$^{-/-}$ hESCs could be passaged and seeded as per normal hESC culture, grew at the same rate as wild type cells, maintained pluripotency markers as single cells, unpatterned colonies, and micropatterned colonies, and still apical-basally polarized at high density to form intact epithelia (Figure 4.3A-D). E-CAD$^{-/-}$ WNT receptor localization was indistinguishable from wild type (Figure 4.3E), and the response to BMP4 continued to be edge only. Interestingly, N-CAD protein was up regulated in the knockout (Figure 4.4), and in other contexts can substitute for the loss of E-CAD[138].

Figure 4.3 | Epithelial integrity is preserved in E-CAD$^{-/-}$ cells

(A) E-CAD$^{-/-}$ cells maintain pluripotency markers even with continual passaging (i.e. >20 passages). Images are of unpatterned, standard hESC culture colonies. Scale bar, 100 µm. (B) E-CAD$^{-/-}$ cell seeding efficiency and growth is similar to wild type and DKK1$^{-/-}$ cell lines. (C) Test of epithelial integrity and BMP4 response. Edge of high density wild type and E-CAD$^{-/-}$ micropatterns stimulated with BMP4 and fixed and stained for pSmad1 after 1 h. As in the wild type, pSMAD1 expression is restricted to the periphery in E-CAD$^{-/-}$ colonies. Scale bar, 100 µm.

(D) Maximum intensity projection of ZO1 and DAPI in low density and high density E-CAD$^{-/-}$ micropatterns immediately prior to stimulation. The network of tight junctions is the same as in the wild type (Figure 4.2B). Scale bar, 50 µm. (E) Cross-sections showing the apical-basal position of WNT receptors relative to DAPI and ZO1. Result is the same as for wild type micropatterns (Figure 4.2C). Scale bar, 10 µm. (F) Top: sashimi plot in Integrative Genomics Viewer of E-CAD region from previously published RNA-seq data set[71] showing exon splicing pattern observed in RUES2 cells in pluripotency conditions. Bottom left: zoom of Exon 1 region in Ensembl viewer showing the four different E-CAD isoforms and the sgRNA CRISPR targeting site used. Bottom right: zoom of Exon 15 region that the Cell Signalling 3195 E-CAD antibody targets. As one can see from the plot, there is no isoform that skips this exon. (G) Western blot of pluripotent wild type and E-CAD$^{-/-}$ cells for E-CAD and α-tubulin. One can see that while E-CAD and α-tubulin are detected in the wild type cells, only α-tubulin is visible in the knockout cell line. (H) Staining of pluripotent wild type and E-CAD$^{-/-}$ cells with Abcam anti-E-CADHERIN ab40772 antibody that targets amino acids 600-700. One can see that while E-CAD is detected in the wild type cells, no E-CAD is visible in the knockout cell line. Scale bar, 50 µm.


A
pluripotent E-CAD-/-
DAPI
OCT4
NANOG
SOX2
B
Low Density
High Density
cells / mm2
5000
4000
3000
2000
1000
0
WT
E-CAD-/-
DKK1-/-
C
WT
E-CAD-/-
1 h BMP4
DAPI/pSMAD1
D
E-CAD-/- Low Density
E-CAD-/- High Density
DAPI/ZO1
E
E-CAD-/- Low Density
E-CAD-/- High Density
center
edge
center
edge
DAPI/ZO1
FZD7/ZO1
LRP6/ZO1
F
CDH1
splice site analysis
CRISPR target region in Exon 1
Exon 15
G
ladder
WT
E-CAD-/-
250kDa
150kDa
100kDa
70kDa
50kDa
37kDa
E-CAD
α-tubulin
H
RUES2
E-CAD-/-
DAPI/E-CAD (Ab40772)

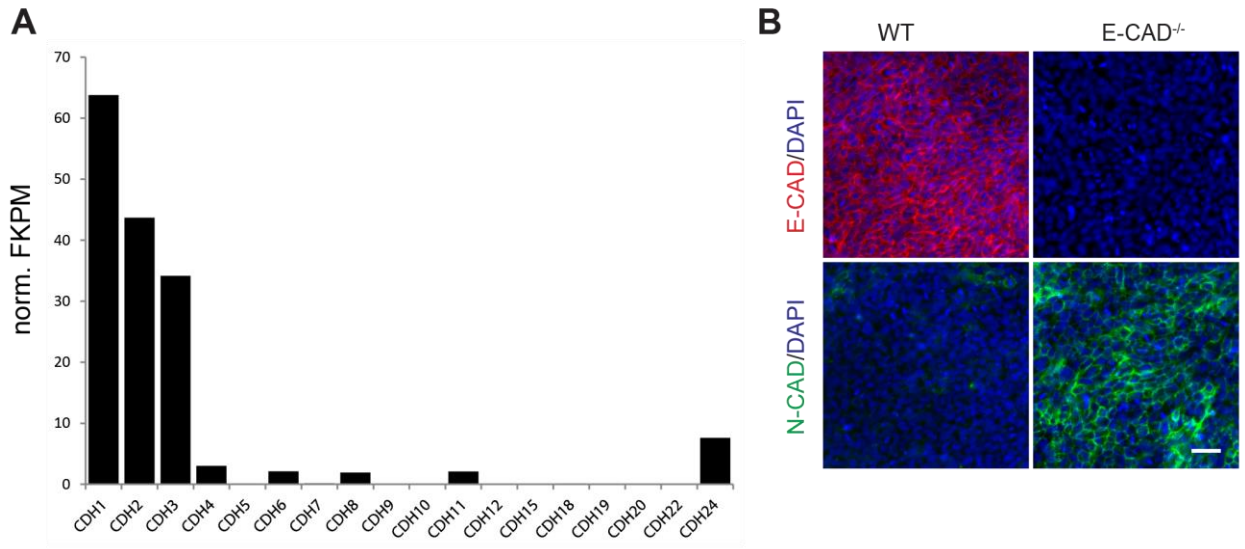


Figure 4.4 | Change in protein expression of N-CAD occurs in E-CAD$^{-/-}$ cell line

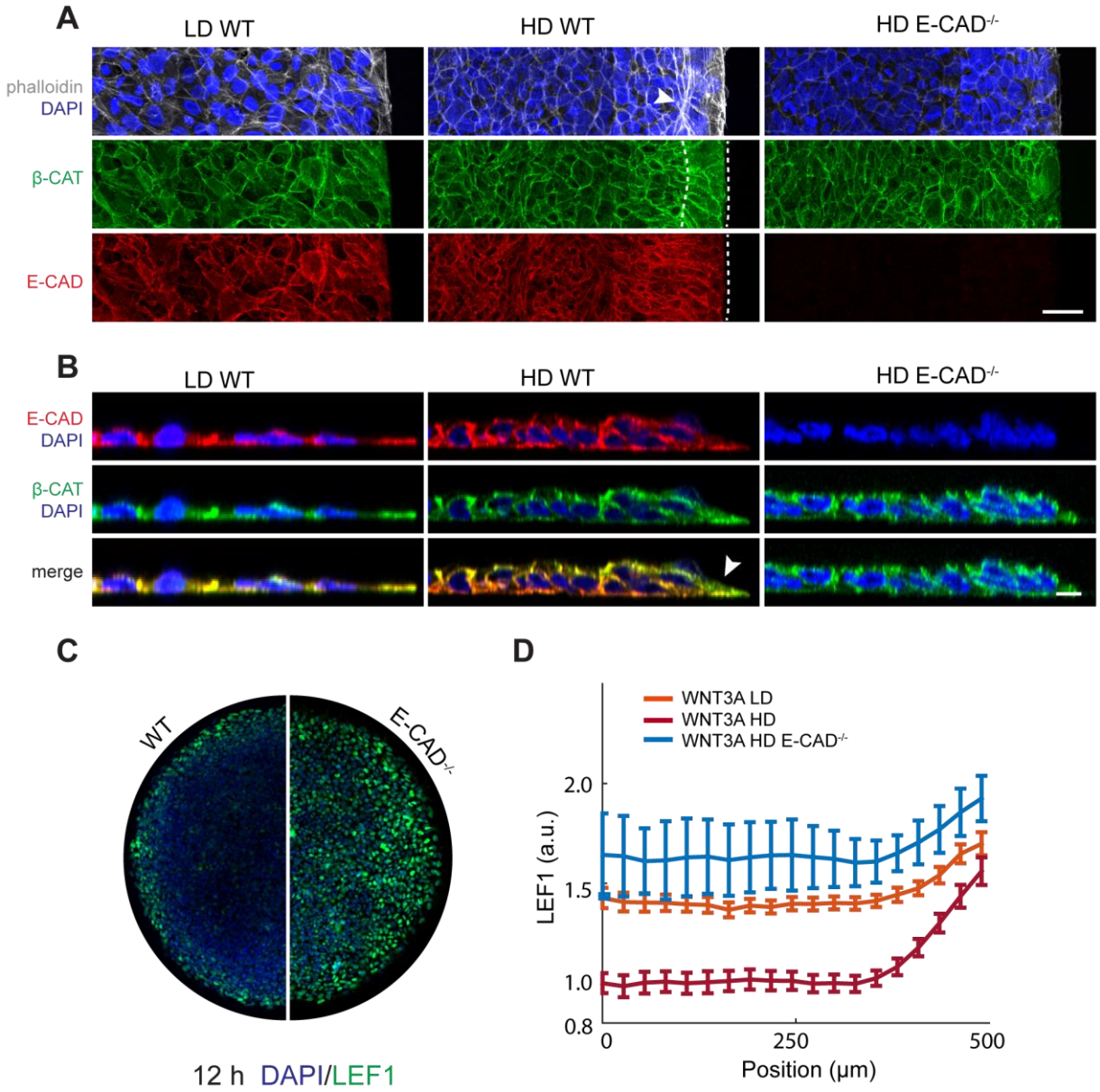
(A) RNA-seq profiling for all classic cadherins in pluripotent hESCs. Since the E-CAD (CDH1) knockout is lethal in mouse at the post-compaction stage, we were somewhat surprised at the lack of an E-CAD$^{-/-}$ pluripotent phenotype. However, both N-CADHERIN and P-CADHERIN (CDH2, CDH3) are substantially expressed at the mRNA level in hESCs in pluripotency. (B) Stain for N-CAD and E-CAD in wild type and E-CAD$^{-/-}$ in unpatterned pluripotent colonies. Antibody stain for E-CAD confirms that gene is knocked out in E-CAD$^{-/-}$ cells. More interestingly, while N-CAD is barely visible in wild-type cells, N-CAD is highly expressed and membrane localized in E-CAD$^{-/-}$ cells. Thus there is a reservoir of N-CAD message in hESC that is only expressed in the absence of E-CAD, which may be a consequence of the same pathway that up regulates N-CAD during EMT when the transcription of E-CAD is abrogated by SNAIL. Interestingly, other research has shown that the artificial replacement E-CAD protein by N-CAD protein in the mouse intestine after gastrulation showed that N-CAD could fulfill the structural role of E-CAD, but the replacement lead to an up-regulation of WNT signalling that is also consistent with our findings[138]. Scale bar, 50 µm.

The most striking phenotype of the E-CAD$^{-/-}$ cells, however, was that the early WNT pattern was abolished in E-CAD$^{-/-}$ micropatterns (Figure 4.5). Cells at the center of high density colonies now showed nuclear localization of LEF1. Quantifying our results over multiple micropatterns, we observed that the WNT response is modestly biased to the edge, and generally comparable to the level in low density colonies (Figure 4.5D). Our transwell filter assay still showed an apical-basal asymmetry, but now the basal response was elevated to match the edge response of the parental cells (Figure 4.2D and E). These results demonstrate that the early WNT pattern is primarily due to E-CAD activity, and a minor influence exerted by WNT receptor accessibility.

If all cells express E-CAD why are there spatial differences in the WNT response? We hypothesized that spatial differences in E-CAD localization or in the state of E-CAD junctions and their binding partners could account for spatial WNT signalling differences. E-CAD and β-CAT stains support this hypothesis as they show that in high density wild type micropatterns E-CAD is reduced in cells on the periphery and there is observable cytoplasmic β-CAT here as well (Figure 4.5A and B). Actin stress fibers have also been observed on hESC micropattern boundaries[139,140] and have been implicated in E-CAD dysregulation[141]. Phalloiden staining in our wild type micropatterns revealed that there were indeed actin stress fibers and they were restricted to the region that is WNT responsive. Furthermore, these stress fibers are absent from low density or E-CAD$^{-/-}$ micropatterns (Figure 4.5A), which suggests a connection between mechanics and WNT signalling.

Figure 4.5 | E-CAD knockdown eliminates early WNT3A pattern

(A) Maximum intensity projection of DAPI, actin (marked via phalloiden), β-CAT, and E-CAD in low density wild type, high density wild type, and high density E-CAD$^{-/-}$ micropatterns immediately prior to WNT3A stimulation. Note the thick actin stress fibers (arrow) and cytoplasmic β-CAT (between the two dashed lines) only apparent on the edge of the high density wild type micropatterns, in the same region that shows the highest LEF1 response to WNT3A. Note also that outer facing side of cells on the micropattern edge are lower in E-CAD compared with sides of the same cells that that join with neighbouring cells (bottom dashed line). Scale bar, 50 µm. (B) Cross-section of micropatterns from (A) showing the overlap of E-CAD and β-CAT. In low density wild type micropatterns there is no significant asymmetry in E-CAD or β-CAT localization (superposition is yellow), but at high density one can see unmatched free β-CAT (green) on cells on the periphery of the micropattern (arrow). Scale bar, 10 µm. (C) 12 h WNT3A response measure by LEF1 in high density wild type and E-CAD$^{-/-}$ micropatterns. Knockdown of E-CAD allows the WNT3A response to penetrate into the center of the micropattern. (D) Quantification (C) and comparison to the low density wild type micropatterns. Single cell expression data was binned radially and averaged. The final radial profile represents the average of *n*=25 colonies. Error bars here and on all following graphs represents the standard deviation among colonies.


A
LD WT
HD WT
HD E-CAD-/-
phalloidin
DAPI
β-CAT
E-CAD
B
LD WT
HD WT
HD E-CAD-/-
E-CAD
DAPI
β-CAT
DAPI
merge
C
WT
E-CAD-/-
12 h DAPI/LEF1
D
WNT3A LD
WNT3A HD
WNT3A HD E-CAD-/-
LEF1 (a.u.)
2.0
1.5
1.0
0.8
0
250
500
Position (μm)

Disruption of E-CAD/β-CAT binding or actin cytoskeleton also sensitizes cells to WNT

To further test whether the classical connection between E-CAD, β-CAT, and the actin cytoskeleton was responsible for our early WNT pattern we performed two additional experiments. In the first we inserted into our E-CAD$^{-/-}$ cells either constitutively expressed full length E-CAD or constitutively expressed E-CAD that lacked the β-CAT binding domain[142] into the AAVS1 locus using TALENS. Clonal lines were cultured in micropatterns and stimulated for 12 hours to examine the WNT response. We found that the constitutively expressed full length E-CAD rescued the edge restricted phenotype but that the E-CAD without the β-CAT binding domain did not (Figure 4.6A-B). This shows that E-CAD binding to β-CAT is essential for the early WNT pattern.

In the second experiment we disrupted the actin cytoskeleton across the entire micropattern with the small molecule inhibitors blebbistatin or cytochalasin B while stimulating with WNT3A. Blebbistatin has been shown to dislodge E-CAD from the membrane into the cytoplasm in hESCs in pluripotency[143], and cytochalasin B has an even more direct action on dissociating the actin cytoskeleton. We found that both reduced the edge restriction, with blebbistatin broadening the size of the LEF1 band and cytochalasin B permitting a WNT response even in the center of the colony (Figure 4.6C). Taken together, our results demonstrate that colony geometry acts via the cytoskeleton and E-CAD to bias WNT signalling to the colony boundary.

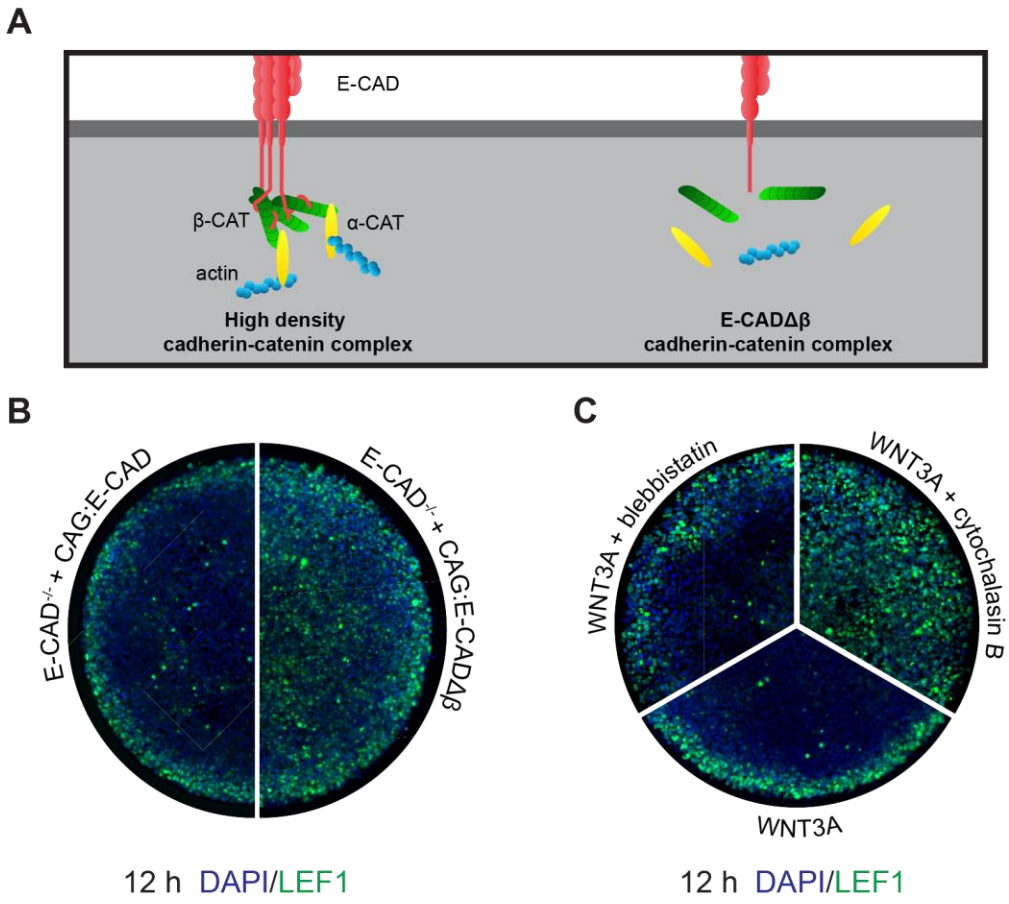


Figure 4.6 | Other perturbations to the E-CAD-β-CAT-Actin complex

(A) Cartoon showing the conventional model of how E-CAD, β-CAT, and the actin cytoskeleton connect to one another. (B) E-CAD$^{-/-}$ cells rescued with constitutively expressed full length E-CAD or E-CADΔβ (missing the C terminal β-CAT binding domain) were put in high density micropatterns and stimulated with WNT3A for 12 h and then stained for LEF1. The full length E-CAD micropatterns recovered the wild type phenotype whereas those with E-CADΔβ did not, showing that the β-CAT link to E-CAD is essential for the 12 h WNT pattern. (C) Wild type high density micropatterns stimulated with WNT3A, WNT3A+blebbistatin, or WNT3A+cytochalasin B, and stained for LEF1 after 12 h. Blebbistatin blocks myosin II controlled actin contraction and cross-linking, whereas cytochalasin B more directly interferes with the actin cytoskeleton by reducing actin polymerization. Corresponding to this difference of degree of perturbation, we see a minor increase in the width of the LEF1 region with blebbistatin, and a more dramatic elimination of the LEF1 edge restriction with cytochalasin B.

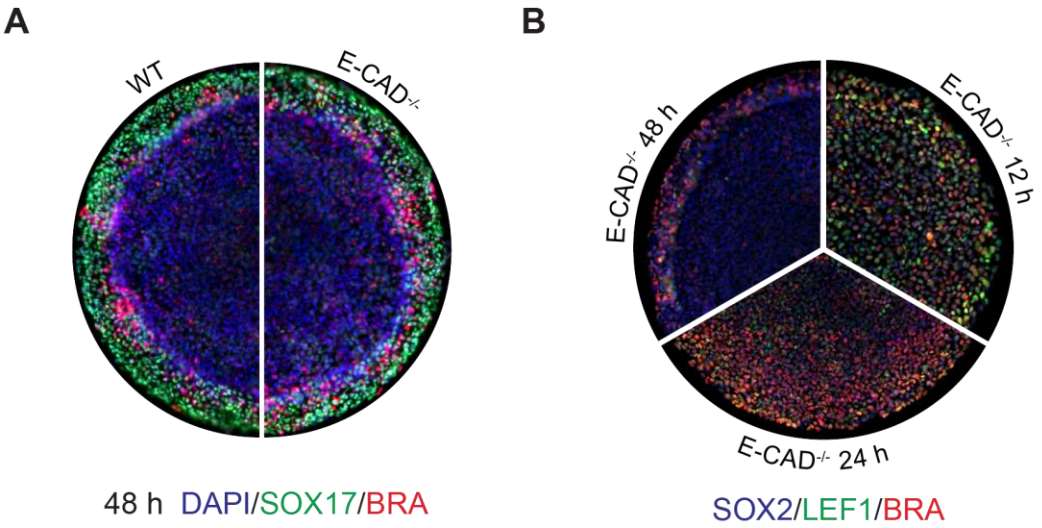
Self-organization of the PS to the edge still occurs in E-CAD$^{-/-}$ cells

Having understood the reasons for the edge asymmetry in the initial response to WNT, we were surprised to see that the location of the PS was virtually the same in wild type and E-CAD$^{-/-}$ colonies at 48 hours (Figure 4.7A). Examination of LEF1, SOX2 and BRA expression at intermediate times showed that a homogenous early expression of these markers gradually becomes restricted to the edge (Figure 4.7B). Since cells in these colonies continue to grow and divide throughout this timecourse we checked whether increasing density could be responsible for this effect. However, E-CAD$^{-/-}$ colonies at a higher starting density (matching that of the Figure 4.7A colonies after 36 hours stimulation) also showed a WNT response in the center (Figure 4.7C). Thus we rule out cell proliferation and increasing density as a major contributing factor for the progressive edge restriction of the WNT response, and instead note that these dynamics are suggestive of a WNT induced secreted inhibitor of WNT that is highest in the center and progressively restricts WNT activity to the boundary.

We had previously shown that BMP4 directly induced the expression of its own inhibitor NOGGIN, which in turn was necessary and sufficient to restrict BMP signalling to the colony edge after 48 hours of stimulation[71]. To address if a similar mechanism of WNT3A inducing its own inhibitor was involved in this case as well, we activated the pathway with CHIR-99021, a small molecule compound that acts cell autonomously and will skip receptors and secreted inhibitors. After 48 hours stimulation the compound edge restriction was abolished (Figure 4.7D). This result strongly suggests the involvement of secreted inhibitors in WNT-mediated self-organization and PS formation.

Figure 4.7 | E-CAD does not explain long term WNT patterning

(A) High density wild type and E-CAD$^{-/-}$ micropatterns fixed and stained for indicated markers after 48 h of WNT3A stimulation. (B) Time-course of LEF1 and BRA expression in high density E-CAD$^{-/-}$ micropatterns. Note that as time progresses the signalling response is gradually excluded from the colony center. (C) 12 h WNT3A response measured by BRA in high high density (HHD) wild type and E-CAD-/- micropatterns. HHD micropatterns are left to grow for an additional 36 h beyond that of HD micropatterns before stimulating with WNT3A. One sees that even at this extreme the knockdown of E-CAD allows a WNT3A response into the center of the micropattern. (D) High density E-CAD$^{-/-}$ micropatterned colonies stimulated with 6 µM CHIR or WNT3A and fixed and stained for BRA after 48 h. The dramatic difference between the two indicates that extra-cellular regulation of the WNT3A pathway may be a dominant factor, since CHIR is a small molecule and acts intercellularly, skipping extracellular regulation.


A
WT
E-CAD-/-
48 h DAPI/SOX17/BRA
B
E-CAD-/- 48 h
E-CAD-/- 12 h
E-CAD-/- 24 h
SOX2/LEF1/BRA
C
HHD stimulated with
WNT3A for 12 h
WT
E-CAD-/-
DAPI/BRA
D
CHIR
WNT3A
48 h DAPI/BRA

WNT induces WNT and NODAL inhibitors

To further test this hypothesis, and identify the relevant inhibitors, we focused on WNT inhibitors whose loss-of-function leads to early gastrulation defects phenotypes in the mouse. These include SFRP1, SFRP2, DKK1, and DKK3[144]. Additionally, as we showed in Chapter 2 that WNT activates NODAL, and these the two ligands have been shown to act synergistically to induce mesendodermal genes[145], we also included LEFTY1, LEFTY2, and CER1 on our list. qPCR was used to assess the induction of these inhibitors when cells were treated with WNT3A alone, or WNT3A+SB to distinguish direct versus indirect induction. After 12 and 24 hours of stimulation, expression of SFRP1, SFRP2, and DKK3 remain relatively unchanged regardless of WNT3A or WNT3A+SB treatment (Figure 4.8). DKK1 expression, however, was highly up-regulated in response to WNT3A. Similar to what has been previously reported[145], this appears to depend on synergy between the WNT and NODAL pathways, since DKK1 induction is lower in WNT3A+SB conditions. A stronger dependence on SMAD2 signalling was observed for the WNT3A induction of CER1 expression. Finally, the expression of LEFTY1 and LEFTY2 depend even more on NODAL signalling since they are also down-regulated in WNT3A+SB and cannot be activated with WNT3A alone. Thus DKK1 and CER1 emerged as the leading candidates involved in WNT self-organization.

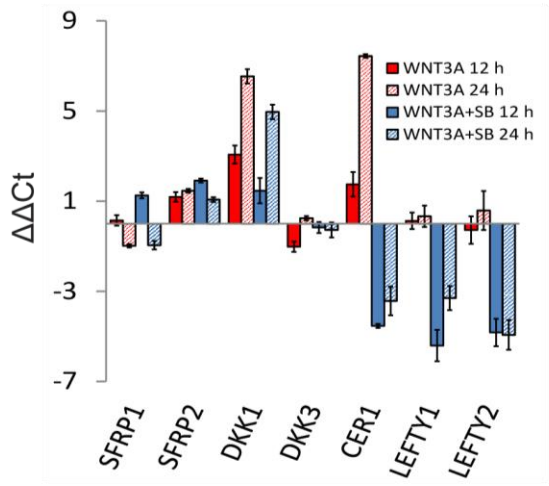


Figure 4.8 | WNT3A also induces WNT and NODAL inhibitors

qPCR of secreted inhibitors of interest upon 12 and 24 h of WNT3A or WNT3A+SB stimulation in micropatterned colonies. Note that the NODAL inhibitors are all severely downregulated when one inhibits the NODAL receptor with SB. (Note also that Conditioned Media also has endogenous ACTIVIN/NODAL activity that would contribute to the effect seen with SB even though no additional ACTIVIN was added).

DKK1 controls the size of the PS

To test whether DKK1 and CER1 are required for the late time WNT pattern, we used CRISPR/Cas9 to generate DKK1 and CER1 knockouts. Each clonal line was stimulated with WNT3A and compared to wild type. After 12 hours of stimulation, DKK1$^{-/-}$ colonies showed no difference with the control (Figure 4.9A). After 48 hours of treatment, however, the size of the PS was dramatically increased when compared to the wild type and E-CAD$^{-/-}$ lines, with only a small center of SOX2+ undifferentiated cells remaining (Figure 4.9B and C). This result, which was confirmed in two additional clonal DKK1$^{-/-}$ lines (Figure 4.12A), demonstrates that DKK1 activity is required for WNT-mediated patterning, and is consistent with a reaction-diffusion model.

To further test this interpretation and see whether human DKK1 protein can protect cells from WNT ligand at a distance in a non-cell autonomous manner, we created a clonal RUES2 cell line that expresses human DKK1 tagged with V5 epitope under the control of a doxycycline promoter (Figure 4.10A). When these RUES2-DOX:DKK1-V5 cells are seeded sparsely into E-CAD$^{-/-}$ micropatterns and made to express DKK1, we can see that they can block BRA expression in cells up to ~5 cell lengths away from them, thus demonstrating that human DKK1 can act as a long-range diffusible WNT inhibitor (Figure 4.10B and C).

To test for epistasis between WNT inhibition at early times mediated by E-CAD, and at late times by DKK1, we generated a double E-CAD$^{-/-}$DKK1$^{-/-}$ knockout line. In response to WNT3A stimulation for 48 hours, all cells in the micropatterned colonies differentiated, with no SOX2+ cells left in the center (Figure 4.9B and C). This suggests that DKK1 and E-CAD are the two major players among the collection of WNT inhibitors that block differentiation in our micropatterns.

Comparison of the expanded PS in the single and double knockouts with wild type and classification of cell types (see Methods and Figure 5.2) also established that both the total number and the ratio of BRA to SOX17 cells changed (Figure 4.11A). Whereas the proportion BRA+/SOX17- cells in RUES2 wild type line was ~10%, in DKK1$^{-/-}$ cells it doubled to 20%, and in E-CAD$^{-/-}$DKK1$^{-/-}$ it tripled to ~30%. The fraction of cells that express both BRA and SOX17 also greatly increases in the mutant lines. This suggests that in addition to determining the size of the PS, DKK1 may also be involved in the segregation of mesodermal and endodermal fates.

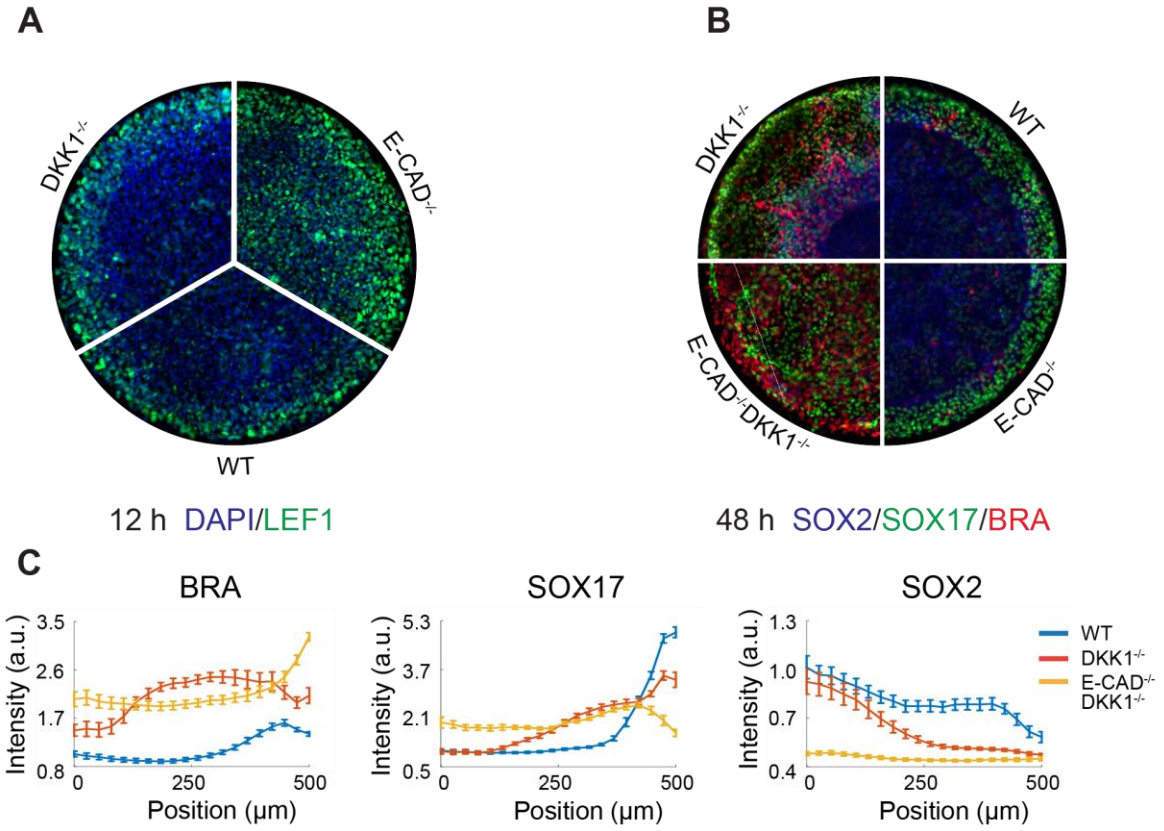
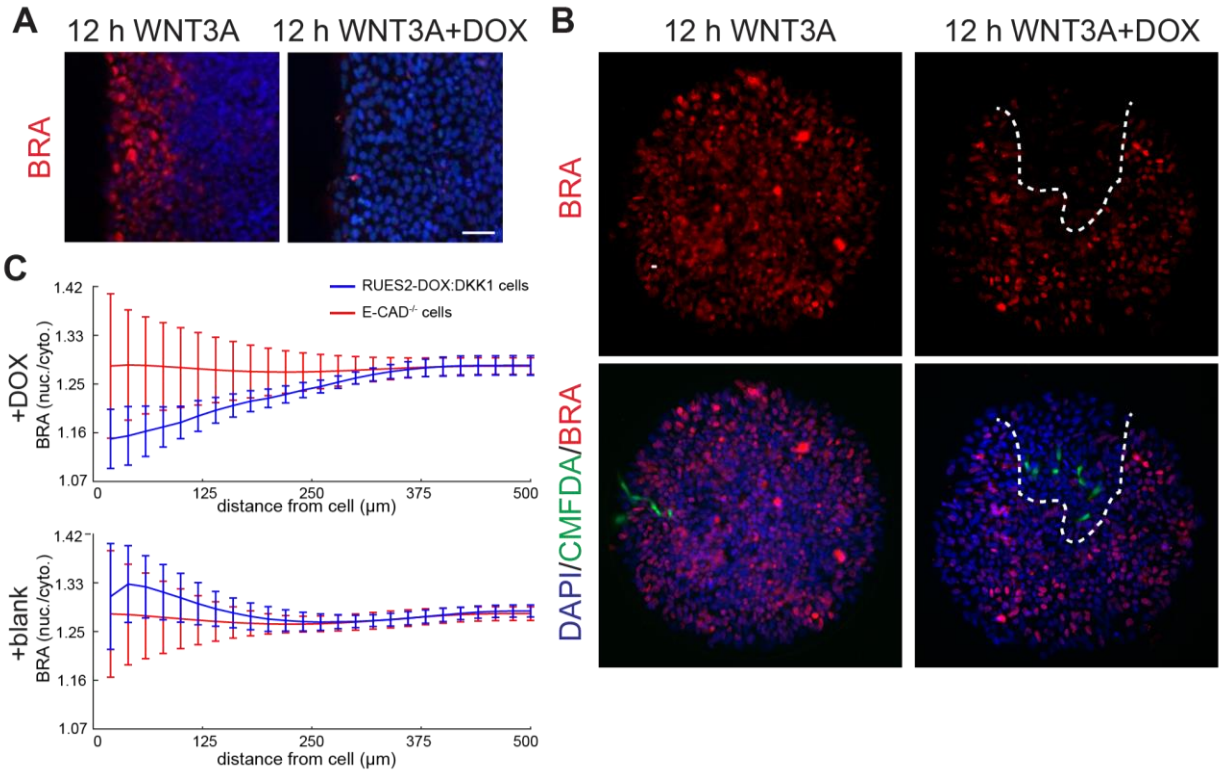


Figure 4.9 | DKK1 controls spatial extent of WNT3A patterning

(A) Comparison of 12 h WNT3A response in high density micropatterns of DKK1$^{-/-}$ cells vs E-CAD $^{-/-}$ cells vs wild type cells. As expected (since no DKK1 is transcribed in pluripotency) DKK1$^{-/-}$ micropatterns resemble wild type micropatterns and are less sensitive to WNT3A than E-CAD$^{-/-}$ cells. (B) After 48 h of WNT3A stimulation on high density micropatterns, DKK1$^{-/-}$ cells show a dramatic increase in WNT3A patterned region compared with E-CAD$^{-/-}$ and wild type cells. DKK1 and E-CAD also act synergistically, as loss of both genes cause complete loss in the edge restriction of WNT3A patterning. (C) Quantification of IF for pluripotency marker SOX2 and differentiation markers SOX17 and BRA in DKK1$^{-/-}$, E-CAD$^{-/-}$, and DKK1$^{-/-}$E-CAD$^{-/-}$ cells (*n*=20 colonies per condition) following B.

Figure 4.10 | DKK1 acts as a long-range diffusible WNT inhibitor

(A) Edge of a high density 1000 μm diameter RUES2-DOX:DKK1-V5 micropattern. Cells were either given DOX or blank media for 12 h and then stimulated with WNT3A for a further 12 h (again with or without DOX continuing the pre-treatment). Cells were then stained for BRA. Scale bar, 50 μm. (B) 500 μm diameter E-CAD$^{-/-}$ micropatterns seeded with 1% CMFDA cell tracker marked RUES2-DOX:DKK1-V5 cells (green). Cells were either given DOX or blank media for 12 h immediately after RI removal, and then stimulated with WNT3A for a further 12 h (again with or without DOX continuing the pre-treatment). Micropatterns were then fixed and stained for DAPI and BRA. In the DOX induced micropatterns one can see a BRA exclusion zone around the DKK1 expressing cells (marked roughly with dashed line). (C) Quantification of (B). 5 DOX-induced and 5 blank stimulated micropatterns were segmented and all cells were first classified as either RUES2-DOX:DKK1-V5 (green) or E-CAD$^{-/-}$ cells. For each cell the nuclear to cytoplasmic ratio of BRA in all neighbouring cells within radius r (x-axis) was calculated and averaged. These averages were then averaged across all cells of their cell type (50 for RUES2-DOX:DKK1-V5, and 5000 for E-CAD$^{-/-}$) and plotted as shown with error bars indicating the standard deviation. One can see that in the DOX induced micropatterns the cells neighbouring the RUES2-DOX:DKK1-V5 cells show lower BRA than the average, and this effect is half-maximal at a distance of ~130 μm, showing long-range action. This shielding does not happen when one looks at just neighbours of E-CAD$^{-/-}$cells, or in the blank stimulated micropatterns.


A
12 h WNT3A
12 h WNT3A+DOX
BRA
B
12 h WNT3A
12 h WNT3A+DOX
BRA
DAPI/CMFDA/BRA
C
RUES2-DOX:DKK1 cells
E-CAD-/- cells
+DOX
BRA (nuc./cyto.)
1.42
1.33
1.25
1.16
1.07
0
125
250
375
500
distance from cell (μm)
+blank
BRA (nuc./cyto.)
distance from cell (μm)

CER1 biases mesoderm versus endoderm fate decision

When stimulated with WNT3A, CER1$^{-/-}$ cells did not show any change in the size of the PS domain in comparison to wild type (Figure 4.11B). There was, however, a significant shift in the proportion of mesodermal versus endodermal fates. Unlike the DKK1$^{-/-}$ or E-CAD$^{-/-}$DKK1$^{-/-}$ cells, this time the shift was towards greater endoderm, with almost all differentiated cells expressing SOX17 and none expressing BRA (Figure 4.11A). This represents a similar phenotype to the WNT3A+ACTIVIN treatment (Figure 2.3A), and prompted us to investigate the status of NODAL/ACTIVIN signalling. We find that SMAD2 signalling in the CER1$^{-/-}$ knockout line is significantly enhanced (Figure 4.11C and D), penetrating farther into the colony from the edge and with a higher nuclear to cytoplasmic ratio. It is known in mouse that CER1 inhibits BMP and NODAL signalling but not WNT signalling[146] (at odds with other vertebrates systems where its tri-functionality motivates its name[144]). For human however it is only known that CER1 inhibits NODAL and a subset of BMP ligands, with a verdict on WNT inhibition still awaiting[147,148]. Since the size of the PS remains unchanged while SMAD2 signalling increases along with the proportion of endodermal cells in the CER1$^{-/-}$ colonies, our results suggest that in hESCs CER1 acts primarily as a NODAL inhibitor rather than a WNT inhibitor.

As it was previously shown in the mouse that the most dramatic CER1 phenotype is when it is doubly knocked out with LEFTY1[149], we also generated CER1$^{-/-}$LEFTY1$^{-/-}$ and LEFTY1$^{-/-}$ clonal cell lines. We detected no difference in the WNT response phenotype between wild type and LEFTY1$^{-/-}$ cells, and no difference between CER1$^{-/-}$ and CER1$^{-/-}$LEFTY1$^{-/-}$ cells (Figure 4.12B and C). In order to check for all other players identified in our RNA-seq and qPCR results, we generated DKK3$^{-/-}$ and SFRP1$^{-/-}$SFRP2$^{-/-}$ clonal cell lines. None of these lines displayed any phenotypic difference

when compared to wild type (Figure 4.12D-F). We conclude that DKK1 and CER1 are the major secreted inhibitors that control WNT patterning in our model system.

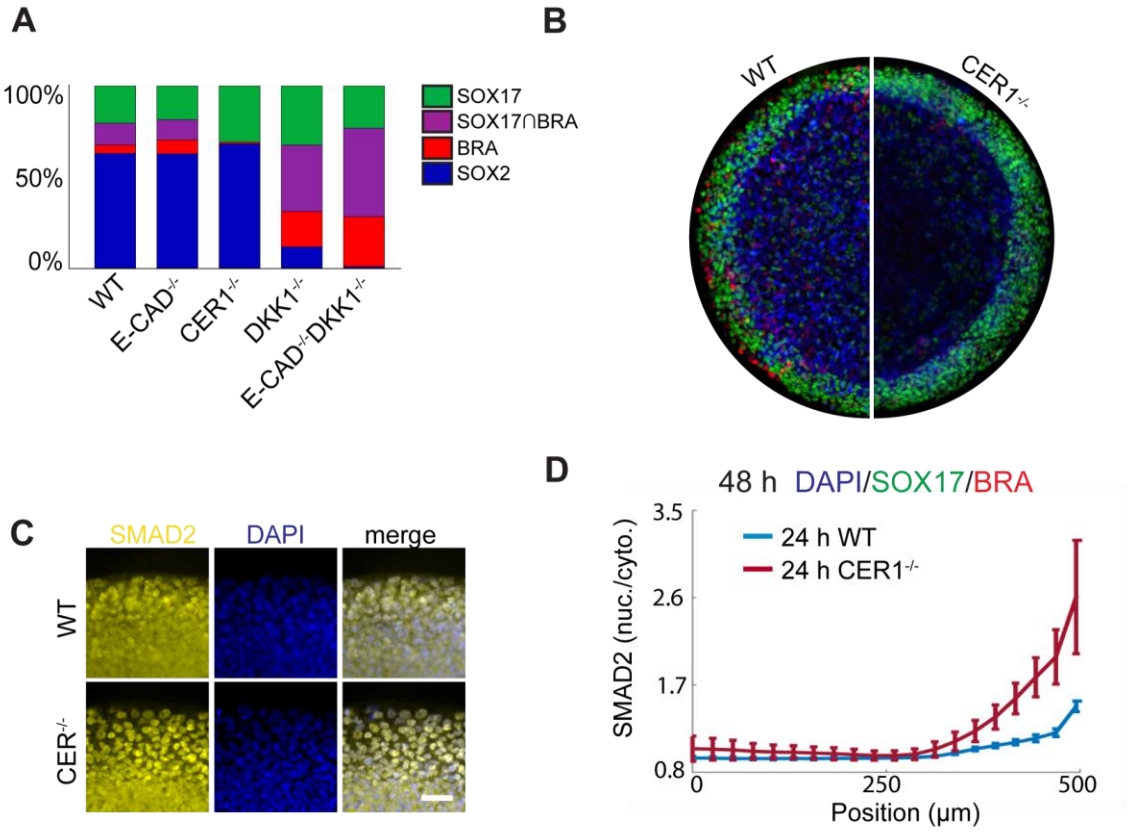


Figure 4.11 | CER1 influences the mesoderm vs endoderm fate decision

(A) Classification of cells from wild type, CER1$^{-/-}$, DKK1$^{-/-}$, E-CAD$^{-/-}$, and DKK1$^{-/-}$E-CAD$^{-/-}$ micropatterns into 4 different subpopulations. Classification was performed by fitting single cell SOX17, BRA, and SOX2 levels to a Gaussian mixture model (See Methods and Figure 6.1). Note that in addition to increasing the spatial extent of WNT3A patterning, DKK1 also influences the proportion of differentiated cells that commit to either mesoderm (BRA) or endoderm (SOX17), with significantly more cells expressing BRA when DKK1 is knocked out. Note also the decline in BRA cells in the CER1$^{-/-}$ micropatterns compared to the wild type. (B) Comparison of wild type and CER1$^{-/-}$ micropatterns after 48 h of WNT3A stimulation. Notice the higher number of BRA cells in the wild type. (C) SMAD2 levels and the ring of activity are increased in CER1$^{-/-}$ cells compared to wild type (micropatterns stimulated with WNT3A, fixed and stained after 48 h). Scale bar, 50 μm. (D) Quantification of F, *n*=20 colonies per condition.

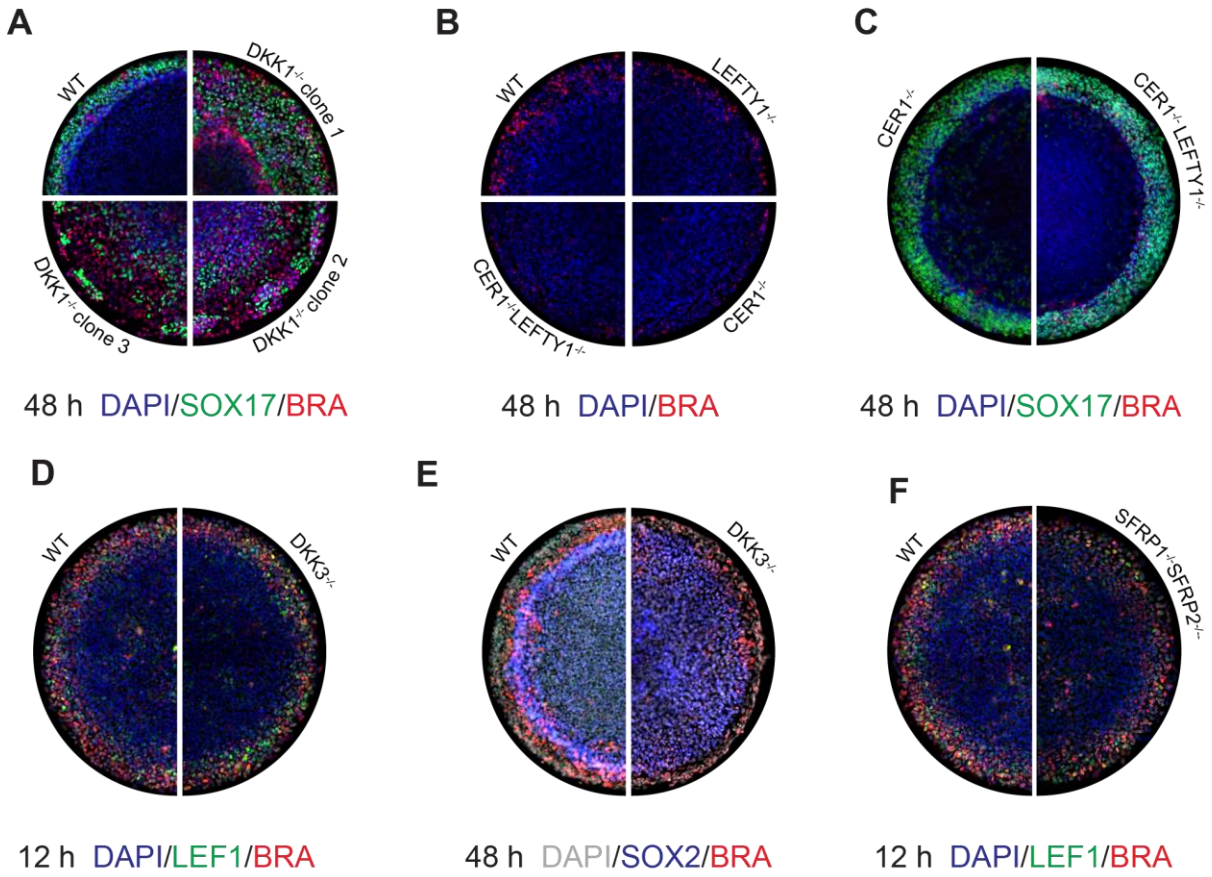


Figure 4.12 | Other WNT inhibitor CRISPR knockout lines and clones

(A) Other DKK1$^{-/-}$ clonal lines with different frameshift mutations also exhibit the same micropattern phenotype when stimulated with WNT3A at high density and fixed and stained after 48 hours. (B) No discernible difference in SOX17 expression at 48 h between CER1$^{-/-}$ and CER1$^{-/-}$LEFTY1$^{-/-}$ micropatterns under WNT3A stimulation. (C) LEFTY1$^{-/-}$ micropatterns show no discernible difference with wild type micropatterns in number of BRA cells. However, both CER1$^{-/-}$ and CER1$^{-/-}$LEFTY1$^{-/-}$ (with a different CER1 frameshift mutation) show similar phenotype in having fewer BRA cells. Thus CER1$^{-/-}$ and not LEFTY1$^{-/-}$ is the main NODAL inhibitor during WNT induced patterning. (D) and (E): no discernible difference at 12 h or 48 h between wild type and DKK3$^{-/-}$ micropatterns under WNT3A stimulation. (F) No discernible difference at 12 h between wild type and SFRP1$^{-/-}$SFRP2$^{-/-}$ micropatterns under WNT3A stimulation.

An edge to center WNT/EMT wave

The size of the PS region in the DKK1$^{-/-}$ cell line at 48 hours is intermediate between the smaller wild type PS region and the fully converted PS region of the double E-CAD$^{-/-}$DKK1$^{-/-}$ cell line. Given that in both RUES2 and DKK1$^{-/-}$ colonies WNT signalling begins at the edge (Figure 4.9A), an important and relevant question is whether the 48 hour result is at steady-state, or if given more time the PS would eventually expand inward and consume the entire colony. To address this question we fixed and stained wild type and DKK1$^{-/-}$ micropatterns at 12, 24, 48, and 72 hours. We find that while differentiation starts similarly for both, the wild type micropatterns seems to reach a steady state by 24-48 hours while differentiation and EMT in the DKK1$^{-/-}$ colonies continues to proceeds inwards, eventually almost consuming the entire micropattern by 72 hours (Figure 4.13A). This is consistent with a wave of WNT differentiation proceeding from the outer edge to the center.

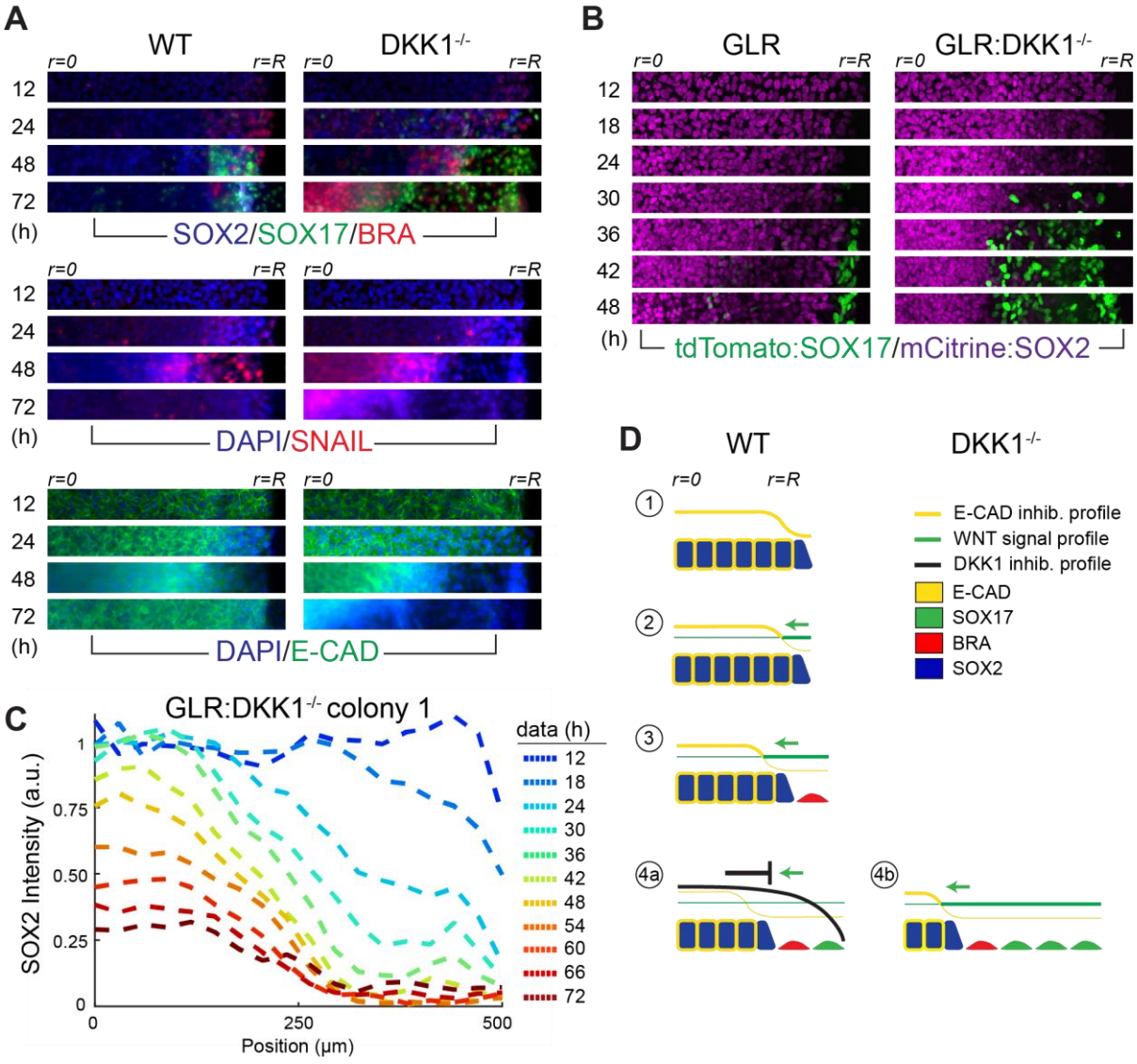
To confirm the existence of this wave and study it further, we knocked out DKK1 in the RUES2-GLR (Germ Layer Reporter) cell line used in Chapter 3, where ectoderm (SOX2), mesoderm (BRA), and endoderm (SOX17) germ layers are tagged with 3 separate fluorescent markers. This enabled us to evaluate the change in differentiation and fate acquisition in the same micropattern across different times. Stimulation of the control RUES2-GLR with WNT3A leads to a downregulation of SOX2 and an upregulation of SOX17 that begins at the edge but stops a few cell layers inward, maintaining the PS at the periphery. RUES2-GLR colonies in which the DKK1$^{-/-}$ mutation is introduced, however, display a wave of progressive downregulation of SOX2 and upregulation of SOX17, which, as in the stained time-course, begins at the outer edge and does not stop (Figure 4.13B and C).

What is the mechanism for such a wave? Given that E-CAD expression recedes as the differentiation front advances (Figure 4.13A) we posit that the wave results from a positive EMT feedback loop. WNT3A is known to induce EMT and to down-regulate E-CAD through the transcription factor SNAIL[150,151]. As illustrated in Figure 4.13D, WNT3A first down-regulates E-CAD in WNT susceptible edge cells and causes them to go through EMT. In so doing, these cells destabilize their E-CAD junctions with their neighbours. This leads to a domino like propagation of EMT from cell to cell via shared cell contacts. If the differentiated cells were induced to secrete a diffusing and thus long-ranged WNT inhibitor, i.e. DKK1, then it will accumulate in the center more than the edges and the wave would be expected to halt[71]. This is how the wild type cells achieve the observed steady state we, and why the DKK1$^{-/-}$and GLR:DKK1$^{-/-}$ cells fail to do so.

In addition to these dynamics, results from Chapter 2 showed that WNT ligand also upregulates WNT production in hESCs (Figure 2.1C). To test if this endogenous WNT signalling contributed to the dynamics we compared wild type and DKK1$^{-/-}$ micropatterns with and without IWP2, the small molecule used previously that blocks all WNT ligand secretion. We find no significant differences (Figure 4.14A), most likely due to the fact that we were already stimulating our micropatterns with a high dose if WNT in the media and thus are in a saturated regime where endogenous WNT does not make any significant contribution to the dynamics. We also tested the involvement of ACTIVIN/NODAL signalling (which has a baseline activity in our media) in this wave by comparing DKK1$^{-/-}$ micropatterns stimulated either with WNT3A or WNT3A+SB. Consistent with other studies of EMT[152], our wave stops when we block ACTIVIN/NODAL signalling with SB (Figure 4.14B).

Figure 4.13 | Patterning via a WNT/EMT wave

(A) Radial sections of WNT3A stimulated high density wild type or DKK1$^{-/-}$ micropatterns fixed and stained for the indicated markers at the indicated times. Interior of each colony (r=0) is on the left of each section, the edge of the colony (r=R) is on the right. (B) Time-lapse radial sections of WNT3A stimulated high density GLR micropattern and GLR:DKK1$^{-/-}$ micropattern. Ten micropatterns for each condition were imaged in the same session, and the examples shown here are representative. One notes that SOX17 turns on slightly earlier in the GLR:DKK1$^{-/-}$ micropatterns than in the GLR micropatterns, and that, as with the immunostaining data, a wave of SOX2 downregulation and differentiation starts on the periphery. This wave halts in the GLR line, but continues proceeding inward in the GLR:DKK1$^{-/-}$ micropatterns. (C) Quantification of single-cell SOX2 expression in the same GLR:DKK1$^{-/-}$ micropattern shown in (B). (D) Qualitative model of WNT/EMT wave spreading and stabilization. ①: prior to WNT3A stimulation E-CAD creates a bias so that only cells on the immediate periphery are sensitive to WNT ligand. ②: Application of WNT3A ligand results in only boundary cells responding and differentiating. ③: As these cells undergo EMT, they lose E-CAD junctions and expose interior cells, enabling them now to respond to the WNT ligand. ④: If checked by secreted DKK1 from the differentiating cells, however, the boundary cells become protected from WNT ligand and the wave stops, as illustrated in (a); if left unchecked, this cycle will enable a wave of differentiation to travel progressively across the colony from outside to inside, as illustrated in (b).


A
WT
DKK1-/-
r=0
r=R
12
24
48
72
(h)
SOX2/SOX17/BRA
DAPI/SNAIL
DAPI/E-CAD
B
GLR
GLR:DKK1-/-
18
30
36
42
tdTomato:SOX17/mCitrine:SOX2
C
GLR:DKK1-/- colony 1
SOX2 Intensity (a.u.)
Position (μm)
data (h)
54
60
66
D
E-CAD inhib. profile
WNT signal profile
DKK1 inhib. profile
E-CAD
SOX17
BRA
SOX2
4a
4b

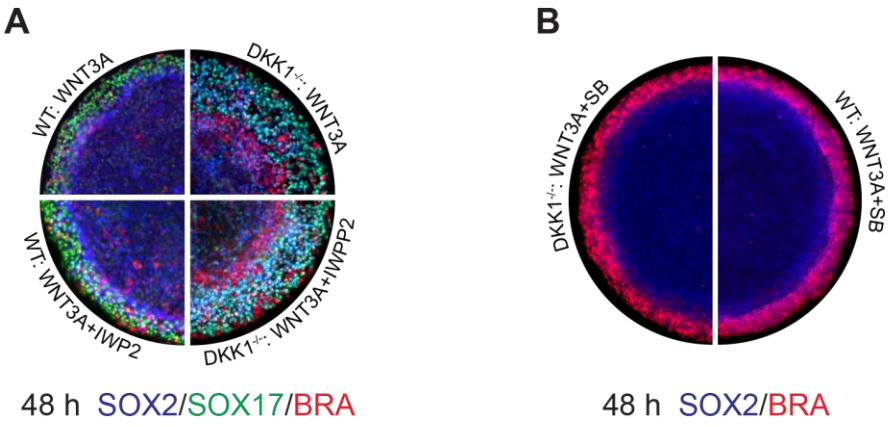


Figure 4.14 | Endogenous WNT signalling has little effect on patterning

(A) Wild type and DKK1$^{-/-}$ high density micropatterns stimulated with either WNT3A or WNT3A+IWP2 for 48 h and stained for BRA, SOX2, and SOX17. No significant differences between wild type IWP2 and non-IWP2 or DKK1$^{-/-}$ IWP2 and non-IWP2 stimulated colonies were observed. (B) Wild type and DKK1$^{-/-}$ high density micropatterns stimulated with WNT3A+SB for 48 h and stained for BRA and SOX2.

A quantitative dynamic model

The spatial pattern of WNT signalling in our colonies is defined by inhibition from E-CAD at the earliest times and DKK1 at later times. These inhibitors operate very differently in space: E-CAD bridges adjacent cells, while DKK1 diffuses across the colony and also leaks out at the edges[71]. Downstream of WNT, NODAL and CER1 are produced and together with WNT generate mesoderm and endoderm fates. To fully unravel the complexity involved in this process we formulated a quantitative dynamic model.

A good model will use a portion of the data to fit parameters and then make testable informative predictions about the remainder of the data, and do this with as few variables as possible. With these criteria in mind we define a 2D partial differential equation (PDE) model where the intracellular WNT signal, $W(r,t)$, is normalized to [0,1] and where a simple Michaelis-Menten system of equations links it to DKK1 and E-CAD. One advantage of our formulation is that it is separable, so we can fit the DKK1 specific parameters to the E-CAD$^{-/-}$ data and vice versa (Figure 4.15A). Since we cannot directly measure WNT levels we use the immunofluorescence data for LEF1 at 12 hours and the percentage loss of SOX2 at 48 hours as surrogates, after normalizing to [0,1]. For the full list of equations, descriptions of the variables, and initial and boundary conditions, the reader is directed to the Supplemental Methods.

Figure 4.15C shows the quality of the fits to the immunofluorescence data at 12 and 48 hours in the two knockout lines that we consider quite acceptable (Figures 6.2 and 6.4). Our model with no additional adjustments is then able to predict and reproduce the 72 hour data in Figure 4.15C as well as the EMT wave (in the DKK1$^{-/-}$ background with endogenous E-CAD), which is most

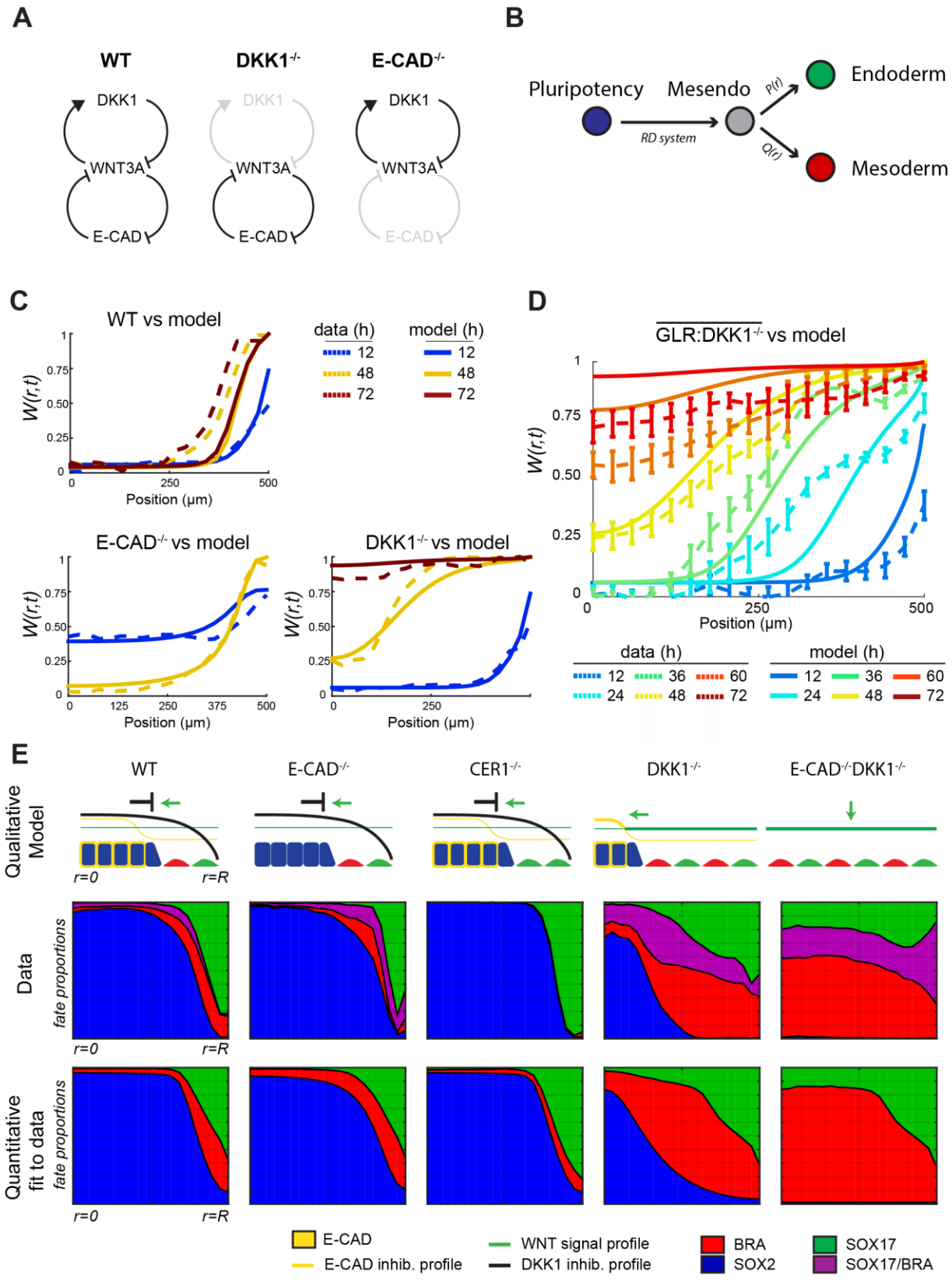
visible in the GLR:DKK1$^{-/-}$ cell line (Figure 4.15D). This live data is somewhat variable due to density variation and phototoxicity (see Appendix) but the prediction of the shape of the wave front and where it should be at a specific time closely follows the observations. The collapse of the wave after it reaches a radius of about 200 μm is also predicted and the model agrees with both the live data and the immunofluorescence data in Figure 4.13A.

The model makes explicit and mathematically precise that the reciprocal inhibition between WNT and E-CAD (Figure 4.15A) can give rise to bistability, and how, in the absence of the long-range secreted inhibitor DKK1, the bistability resolves by an inward propagating wave that eliminates the epithelial state in favour of the PS mesenchyme. The model situates the WNT system within a general class of problems where waves result from bistability. The interest for development, elaborated in discussion, is that waves propagate information faster than diffusion, and the proposed mechanism is generic and largely parameter independent.

Having captured the WNT response dynamics for the different cell lines, we can further model whether the differentiated cells become endoderm or mesoderm. To do so we assume that NODAL favours endoderm over mesoderm[131], and we describe this as a branching probability from an intermediate mesendoderm state (Figure 4.15B), since we lack more quantitative data about the genetic network for mesendoderm specification. As a surrogate for NODAL signalling we used measured SMAD2 profiles at 24 hours for each cell line (Figure 4.11C and Figure 6.6 in Appendix). Figure 4.15E shows the comparison of model and data. The conversion of mesoderm to endoderm due to up regulation of NODAL is clearly visible in the comparison of CER1$^{-/-}$ with wild type or E-CAD$^{-/-}$.

Figure 4.15 | A quantitative model of WNT patterning dynamics in the PS

(A) Gene regulatory network of WNT3A, E-CAD, and DKK1 used in the model. Note how the DKK1$^{-/-}$ and E-CAD$^{-/-}$ cell lines can be used to simplify this network and fit the subcomponents separately. (B) Simple model of mesoderm versus endoderm fate decision. The reaction-diffusion system determines the probability that a cell at a given radius in a given background commits to differentiate by going from the pluripotent state to the intermediate mesdendo state. Once there, the probability *P(r),* which is a linearly rescaled function of radial nuclear SMAD2 profile, is used to determine the probability of cell going to endoderm versus mesoderm. (C) Comparison of the simulated PDE model to the WT, DKK1$^{-/-}$, and E-CAD$^{-/-}$ data. The fit was made using the 12 h LEF1 response (Figure 2D), the 48 h differentiated cell response (i.e. 1–SOX2 population) shown in Figure 4.16A, and the 72 h differentiated cell response (again 1–SOX2 population) shown in Figure 4.16B. The remaining data plotted are model predictions. (D) WNT/EMT wave in the DKK1$^{-/-}$ micropatterns. Solid lines are the model predictions, dashed lines and error bars represent the average and standard deviation of 10 continuously imaged GLR:DKK1$^{-/-}$ micropatterns, respectively. (E) Qualitative diagram of WNT patterning dynamics for each genetic background, and fit of quantitative model to the data. For the data, single cells from the 48 h micropatterns of each genetic background were classified into 4 different subpopulations according to a Gaussian mixture model based on SOX17, BRA, and SOX2 levels (see Methods). Classified cells were then further subdivided according to their radial position and the population was normalized, allowing us thus to obtain the probability of belonging to a specific fate at a specific radii given a specific genetic background.


A
WT
DKK1-/-
E-CAD-/-
DKK1
WNT3A
E-CAD
B
Pluripotency
Mesendo
RD system
P(r)
Q(r)
Endoderm
Mesoderm
C
WT vs model
data (h)
model (h)
12
48
72
W(r,t)
Position (μm)
E-CAD-/- vs model
DKK1-/- vs model
D
GLR:DKK1-/- vs model
12
24
36
48
60
72
E
WT
E-CAD-/-
CER1-/-
DKK1-/-
E-CAD-/-DKK1-/-
Qualitative Model
Data
fate proportions
Quantitative fit to data
r=0
r=R
E-CAD
E-CAD inhib. profile
WNT signal profile
DKK1 inhib. profile
BRA
SOX2
SOX17
SOX17/BRA

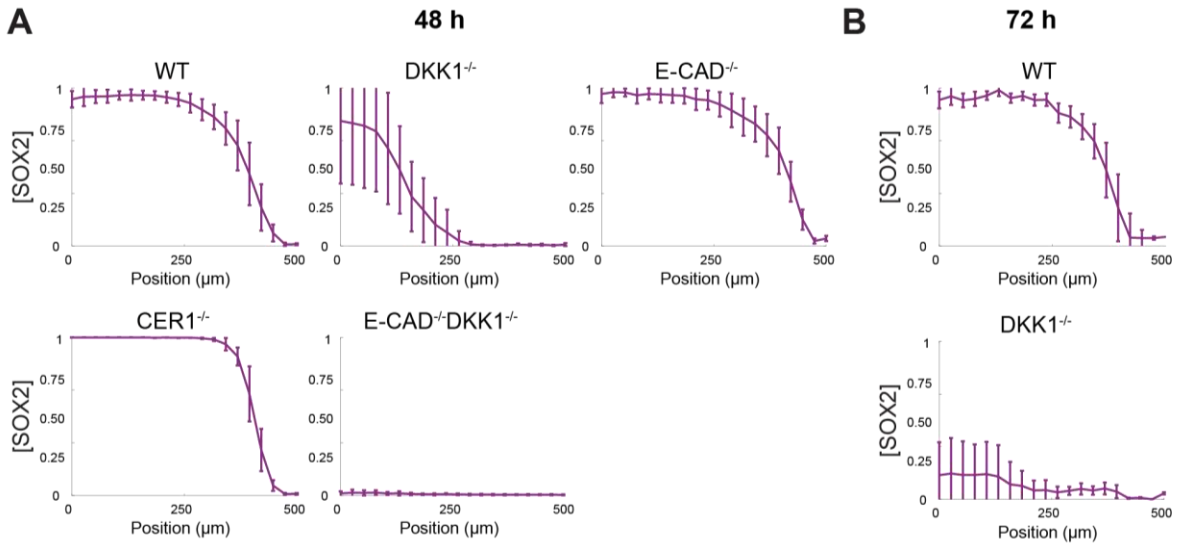


Figure 4.16 | Data used to fit model

(A) Proportion of cells classified as SOX2+ at 48 h plotted as a function of radius. Error bars represent the standard deviation of *n=20* colonies. (B) Proportion of cells from micropatterns in Figure 4.13A classified as SOX2+ at 72 h plotted as a function of radius. Error bars represent the standard deviation of *n=20* colonies.

Discussion

In this chapter I have unraveled the molecular mechanism underlying the WNT-mediated self-organization in our hESC model of human PS. This mechanism includes several elements. The first element is the "pre-pattern" of WNT sensitivity. Unlike the situation with BMP4 this pre-pattern is not imposed by polarity of receptor localization, but rather by β-CAT mechano-sensation via E-CAD, and the cytoskeleton, which is governed by tissue geometry. Given the importance of this pre-pattern for WNT patterning in a model human epiblast, how applicable might such pre-patterns be to gastrulation more generally? Gastrulation takes many forms across all animals, but a common theme is the correlation of WNT signalling with invagination[153] (such as for example through the PS in amniotes or through a blastopore in lower orders). Invagination necessarily involves a breaking of geometric symmetry and a concentration of mechanical forces in an area of high curvature, and it is tempting to think that β-CAT mechanosensation is also involved here in either pre-patterning or reinforcing a WNT signal. Other notable examples of β-CAT mechanosensation as a critical element for WNT patterning include are found *Drosophila* gastrulation[154], *Xenopus* embryonic explants[154], and avian follicles[155].

The second element in the WNT patterning process is the Turing-like activator-inhibitor pair of WNT and DKK1. In a manner similar to BMP4 which in species-specific manner induces NOGGIN, the WNT ligand also directly induces the expression of its own inhibitor: DKK1. In the mouse the anterior visceral endoderm (AVE) on the opposite side of the embryo from the PS is traditionally thought of as the major source of DKK1, with this inhibitor being produced in a WNT independent manner[31,156]. The expression of DKK1 by the AVE makes disentangling any intrinsic contribution

of the epiblast to the position and control of the PS in the embryo difficult. Given that our synthetic epiblast achieves a stable pattern via the induction of DKK1 by WNT without external sources, however, we might expect to see DKK1 produced by cells in the PS. This is exactly what recent single cell and FACS sorted RNA-seq data shows[157,158]. Similarly, results from studies in rabbit gastrulation show that DKK1 is also expressed in cells in the PS, especially in epiblast cells undergoing EMT[159]. While expression of activator and inhibitor on opposite sides of an embryo (such as in PS and AVE) seems logical and necessary, the Turing mechanism requires production of the inhibitor where the activator is highest, since the inhibitor must also spread rapidly to confine the activator.

We discovered that the same is true for the CER1, the third major element in the WNT patterning process. Despite also being thought of as an AVE product, single cell RNA-seq studies show that CER1 is also expressed in the mouse and rabbit PS[157,159]. In our human model we find that CER1 acts as a NODAL inhibitor and controls the balance between mesodermal and endodermal fates emerging from the back of the EMT wave. Thus E-CAD, DKK1, and CER1 are the primary determinants of WNT-induced PS patterning in artificial human epiblasts.

Our knockouts of CER1 and DKK1 in hESCs also reveal possible species specific differences between human, mouse, chick, and rabbit. For example, a difference between human and other studied species is that the knockout of CER1, or even the double knockout of CER1 and LEFTY1, does not lead to an expansion of the PS in our model system. This is surprising since in mouse it leads to multiple streaks[149] and in the chick NODAL repression by Cerberus is thought to be the dominant repressor of ectopic streaks[46]. Another difference is that the mouse Dkk1 knockout

does not broaden the PS as defined by Bra expression, though it does broaden the expression of a synthetic Wnt reporter[39]. This could suggest that PS formation in human is dominated more by WNT signals whereas in mouse the PS and Bra require other signals (such as NODAL) as well. Indeed, recent experiments with mouse epiblast-like cells (EpiLCs) on micropatterns[75] suggest that there is a stronger dependence on NODAL signalling to achieve a PS-like region and Bra positive cells than in hESCs. We hope that future comparative experiments will shed light on these apparent species specific differences, and will also test for possible confounding *in vitro* versus *in vivo* differences as well.

Our results for the DKK1 and CER1 knockouts have implications not just for the spatial control of the streak, but also for mesoderm versus endoderm cell fate decisions within the streak. We found a higher ratio of mesoderm to endoderm in DKK1$^{-/-}$ cells, which is consistent with the reduction in the mouse of anterior endoderm in favour of mesoderm following the analogous knockout[39]. Conversely, removal of CER1 favors endoderm over mesoderm in our system, with up regulation of nuclear SMAD2 compared with wild type. Although less well understood than the appearance of multiple streaks, the elimination of CER1 and LEFTY in the mouse embryo also results in a marked conversion of BRA cells to more NODAL regulated fates, such as anterior endoderm and axial mesoderm[149]. Based on these observations, as well as the result that WNT induces both DKK1 and NODAL in hESCs, we speculate that mammalian anterior endoderm is specified by a transient WNT3 pulse terminated by self-induced DKK1 and followed by up regulation of NODAL signalling. In fact, a transient application of WNT followed by its removal and the addition of ACTIVIN is precisely the most efficient and commonly used endoderm differentiation protocol used for hESCs[70,160–162].

One of the more intriguing results to emerge from our investigation was the presence of an EMT wave from the edge of our micropatterns towards the center. This wave is generated by downregulation of E-CAD by EMT and the negative regulation of WNT signalling by E-CAD, but also requires ACTIVIN/NODAL activity since SB treatment blocks the wave. As DKK1 acts as a negative feedback to halt the wave, the wave is much longer and more apparent in the DKK1$^{-/-}$ micropatterns than wild type micropatterns. This is not to say, however, that wild type cells could not undergo such a wave *in vivo*. For example, removal of DKK1 expressing cells from the front, as happens in migration of the mesenchymal cells out of the PS *in vivo,* would allow the wave to continue progressing. This cannot take place in our system since cells are confined to the surface of the micropattern. We note that although it is harder to cleanly isolate, in mouse and chick embryos there is evidence for cooperative EMT in maintaining and extending the PS[163,164].

Traveling waves of activity have been seen in several other developmental contexts, but it is important to distinguish propagation of a front separating two states with phase waves, such as in somitogenesis[165]. Waves with a propagating front have been seen in several other developmental contexts, such as calcium waves following fertilization or in large embryos that presumably function to synchronize tissues[166,167]. A wave of mitotic activity in frog extracts has also recently been observed and linked to bistability in the CDK system[168]. Why might a wave be useful for patterning? A wave is rationalized as way to spread information more rapidly than diffusion in a large system, and the hundred micron scale disk-shaped epiblast in rabbit and humans may require such a solution. Patterning via a wave may be a widespread mechanism for tissue patterning since it only requires a bistable system and some means for the favoured state to spread between cells.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

A major conclusion of this work is that the role and control of the Wnt pathway in PS formation in our human model is widely conserved with that of model organisms. Not only is it the case for example that Wnt is necessary and sufficient for the induction of human PS and patterning of mesoderm and endoderm as is with nearly every other studied model organism, but the same exact Wnt ligand and signalling hierarchy network are shared between mouse and human. Also shared between mouse and human is the same major secreted inhibitor that controls Wnt signalling. More remarkably, the degree of conservation of the role of the Wnt pathway in amniote PS patterning is such that a gastruloid human organizer is able to induce a secondary axis when grafted into chick, a species separated by hundreds of millions of years of evolution and closer to dinosaurs than to us. Conceptually speaking, this conclusion ties ourselves and our species to the vast picture of Wnt signalling in early gastrulation we have gleaned from decades of studies in model organisms. Technically speaking, this conclusion supports and offers validation for future comparative approaches that use hESCs and model organisms as a means to study early human gastrulation. Immediate experiments that suggest themselves are further mapping of the transcriptomic profiles of gastruloid cells to cells in gastrulating mouse embryos, and functionally mapping other gastruloid components to those of the gastrulating chick embryo via traditional embryological cut-and-paste experiments.

A more thought-provoking conclusion of this work is that it was not enough just to consider the individual molecules and genes involved in our Wnt patterning process; geometry and mechanical forces were emergent irreducible qualities that directly shaped and controlled the

final outcome. Our work thus emphasizes the need for an equal emphasis on the "morpho" aspects with the "genetic" aspects in future morphogenetic studies. In our particular case of the human PS, a structure whose purpose in development is first and foremost to break spatial and cell type symmetries, an important question is to what degree are naturally arising morphological asymmetries (such as edges versus centers or high or low mechanical stresses) exploited by signalling pathways and genes to break large scale symmetries? Are there universal mechanisms or classes for this, and, as in the case of the Wnt wave, can they be reduced to simple universal models? We anticipate that answers to such questions will continue to be most fruitfully pursued in synthetic in vitro systems, such as our hESC micropatterned PS model where complex and developmentally relevant patterning dynamics can be followed, quantified, and ultimately deconstructed.

MATERIALS AND METHODS

Ethics Statement for Chick-Human Chimeras

This work was conducted according to protocols approved by the Tri-Institutional Stem Cell Initiative Embryonic Stem Cell Research Oversight (Tri-SCI ESCRO) Committee, an independent committee charged with oversight of research with human pluripotent stem cells and embryos to ensure conformance with University policies, and guidelines from the U.S. National Academy of Sciences (NAS) and the International Society for Stem Cell Research (ISSCR). The Tri-SCI ESCRO is composed of members with scientific and bioethical expertise. The ESCRO review of these protocols was conducted prior to the May 2016 issuance of the ISSCR Guidelines, but the review addressed the issues of growth and culture of human-chick chimeras and in vitro culture of embryo-like structures and anticipated the ISSCR Guidelines (specifically Recommendations 2.1.3 and 2.1.5, which are pertinent to this study). As part of these protocols, the human cells transplants were limited to <10% compared to host animal at any given stage, and no chicken-human chimeras were allowed to hatch. Additionally, the researchers considered that the self-organized structures that arose from the experiments lacked human organismal potential due to their in vitro culture without the necessary non-embryonic tissues or support that is present in vivo. The ESCRO Committee also reviewed and approved the NIH grants HD080699 and GM101653 that funded this study, and approved the initial derivation of the RUES2 cell line which is listed in the NIH Human Embryonic Stem Cell Registry.

Cell Culture of Human Embryonic Stem Cells

The RUES2 cell line was derived in our lab and has been described previously[60,93]. Cells were grown in HUESM medium (Table 6.1) conditioned by mouse embryonic fibroblasts (MEF-CM) and supplemented with 20ng/mL bFGF. Mycoplasma testing was carried out at 2 month intervals. For maintenance, cells were grown on GelTrex (Invitrogen) coated tissue culture dishes (BD Biosciences, 1:40 dilution). The dishes were coated overnight at 4°C and then incubated at 37°C for at least 20 minutes before the cells were seeded on the surface. Cells were passaged using Gentle Cell Dissociation Reagent (Stem Cell Technologies 07174).

Micropatterned Cell Culture

Two methods to make micropatterns were used. In the first, they were simply purchased from CYTOO. In the second, they were made from scratch in-house. The need for the second method was a result of a 6 month manufacturing failure by CYTOO during the second year of my studies. All of the results obtained with Method 2 were repeated with Method 1 once it was certain that this manufacturing failure was fixed and not likely to repeat.

Method 1: Pre-patterned micropatterned chips were purchased from CYTOO (Arena A, EMB, or 500μm). The coverslips were first coated with 10 ug/ml laminin 521 (Biolamina) diluted in PBS++ for 3 hours at 37°C. The coverslips were then prepared for cell seeding with 3 washes of PBS++.

Method 2: Micropatterned coverslips were made according to a new protocol I devised that significantly reduced operating costs. First, 22x22 mm #1 coverslips were spin-coated with a thin

layer of PDMS (RTV615A Momentive) and left to set overnight. They were then coated with 5µg/ml laminin 521 (Biolamina) diluted in PBS with calcium and magnesium (PBS++) for 2 hours at 37°C. After 2 washes with PBS++, coverslips were placed under a positive feature UV Quartz Mask (Applied Image Inc) in a home-made UV oven. Laminin not protected by the features in the mask was burned off by 10 minutes of deep UV application (185 nm wavelength). Coverslips were then removed, washed twice more with PBS++, and then left at 4°C overnight in 1% F127-Pluronic (Sigma) solution in PBS++. The now patterned coverslips were used within 1 week of fabrication, and are ready for cells after 3 washes with PBS++

For both methods: Cells were dissociated from growth plates with StemPro Accutase (Life Technologies) for 7 minutes. Cells were washed once with growth media, washed again with PBS, and then re-suspended in growth media with 10 µM ROCK-inhibitor Y-27632 (Abcam). Coverslips were placed in 35 mm tissue culture plastic dishes, and $1x10^6$ cells in 2 mL of media were used for each coverslip. After 1 h ROCK-inhibitor was removed and replaced with standard growth media supplemented with Pen Strep (Life Technologies). Concentrations of ligands and small molecules used were: 100 ng/mL WNT3A, 100 ng/mL Activin-A, 10 µM SB, 6 µM CHIR, 50 ng/mL BMP4, 0.5 µM cytochalasin B, 10 µM blebbistatin.

Transwell Cell Culture

We used Costar Transwell 24-well plates with 0.4 µm pore-sized clear polycarbonate membrane inserts (Fisher Scientific 07-200-147). Membranes were coated with 10 µg/mL of laminin 521 (Biolamina) diluted in PBS++ for 3 h, followed by washing 3 times with PBS++. Single cells were

collected and seeded as per micropattern protocol. To image the membrane the transwell was removed from the multi-well plate after fixing and staining and placed on top of a coverslip.

Generation of knockout lines

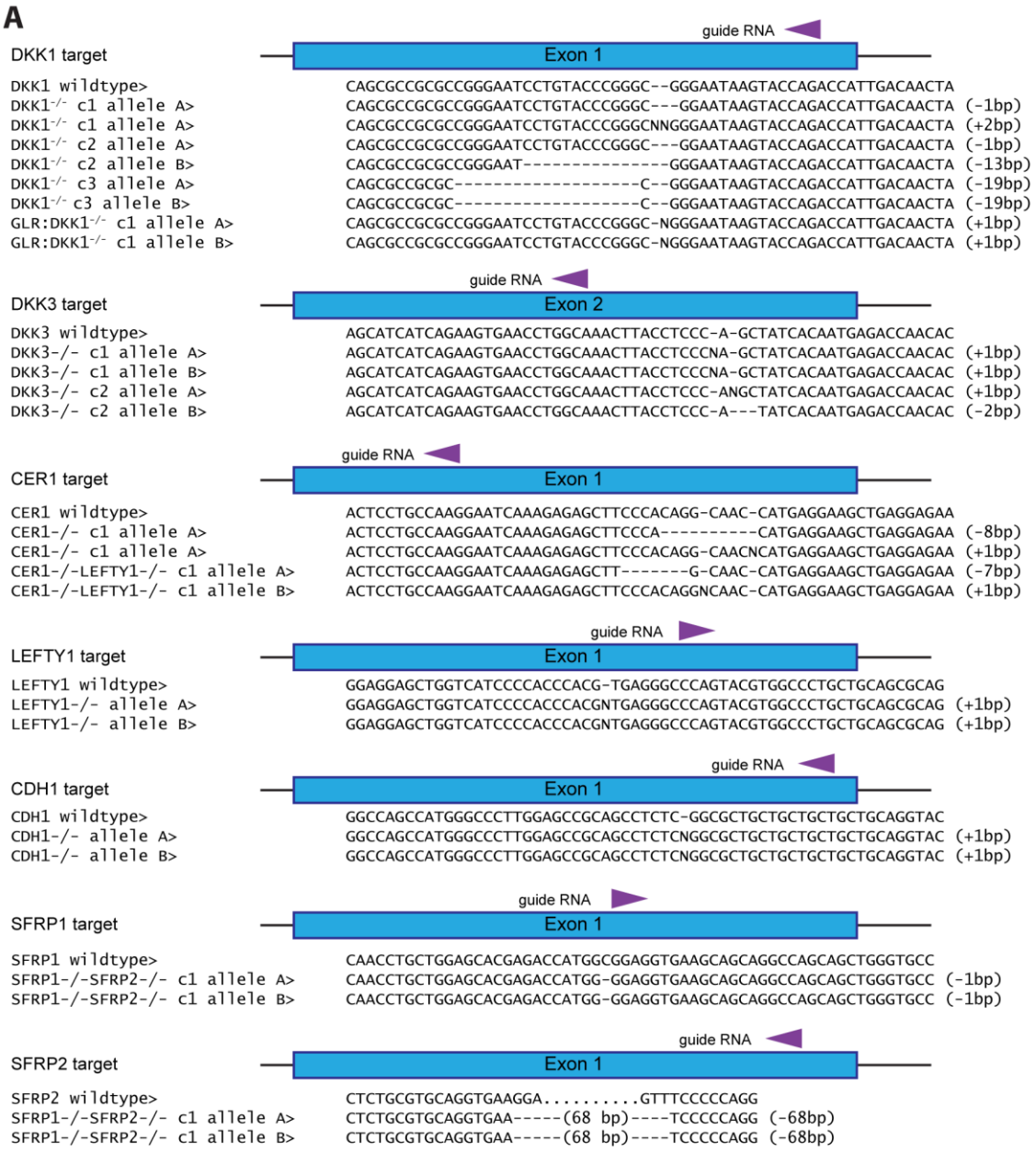
The CRISPR/Cas9 system was used to generate all knockout lines. Three rules of reproducibility and quality control were applied: (i) at least two independent clones for each gene clonal lines were isolated and studied in parallel; (ii) lack of off-target effects was assessed by qPCR; and (ii) their ability to maintain the pluripotent state as assessed by expression of NANOG, OCT4, and SOX2. For the E-CAD, DKK1, and CER1 knockouts one clone of each was also assessed for chromosomal integrity by karyotyping. The sgRNA target for each gene is listed in Table 6.2. A list of all knockout lines is given in Table 6.3. The sgRNAs were cloned into a pX330 plasmid[169] that we modified to co-express a puromycin-2A-EGFP cassette, as this strategy gave us a higher percentage of successfully targeted clones. Transfection was carried out using the B-016 setting of a Nucleofector II instrument and using the Cell Line Nucleofector Kit L (Lonza). Transfected cells were immediately seeded in ROCK-inhibitor on GelTrex coated culture dishes, and puromycin was added after 24 h for 24 h. Cells were then passaged as single cells using Accutase (Stem Cell Technologies) and sparsely seeded to facilitate picking individual clones. Clones were handpicked with a 20uL pipette tip and once expanded genomic DNA from each clone was extracted with the DNeasy Blood & Tissue kit (Qiagen). The locus for each targeted gene was then PCR-amplified using the primers listed in Table 6.4, and submitted for Sanger sequencing. The resulting chromatograms for each clone were decomposed using the TIDE webtool[170] http://tide.nki.nl. Only clones that showed a high probability for both alleles of the gene of

interest having a missense mutation leading to a premature stop codon were kept. Examples of these mutations are shown in Figure 5.1A. As a check on the integrity of our most critical knockout cell lines, the E-CAD$^{-/-}$ clone 1, DKK1$^{-/-}$ clone 1, and CER1$^{-/-}$ clone 1 cell lines were additionally sent for karyotyping and were found to be karyotypically normal (Figure 5.1B). In the case of E-CAD we note that although there are 4 recorded isoforms (https://www.ncbi.nlm.nih.gov/gene/999) our sgRNA target is after the start codon of each and is present in all of these isoforms. As an additional check to ensure that our CRISPR/Cas9 mutation results in a complete knockdown instead of a hypomorph, we note that the E-CAD antibody we use to check for E-CAD presence using immunofluorescence (Figure 4.5A) was produced by immunization with a synthetic peptide corresponding to the sequence surrounding Pro780 of human E-cadherin[171]. This corresponds to Exon 15, which is near the C-terminus of the final protein and is shared between all the E-CAD isoforms present in our cells (Figure 4.3F). Immunofluorescence tests with an antibody from a different supplier that targets a different conserved region (Abcam #ab40772 produced by immunization with synthetic peptide within human E CAD) aa 600-700[172]) also gave the same result (Figure 4.4H).

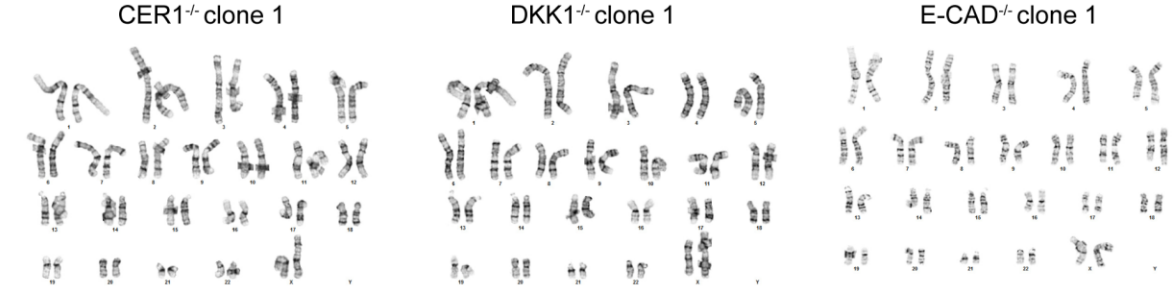
Figure 5.1 | Confirmation of CRISPR knockouts

(A) Sequences of the mutated alleles in the CRISPR/Cas9 knockout cell lines. (B) The E-CAD$^{-/-}$ clone 1, DKK1$^{-/-}$ clone 1, and CER1$^{-/-}$ clone 1 cell lines were karyotypically normal.

A


guide RNA
DKK1 target
Exon 1
DKK1 wildtype> CAGCGCCGCGCCGGGAATCCTGTACCCGGGC--GGGAATAAGTACCAGACCATTGACAACTA
DKK1-/- c1 allele A> CAGCGCCGCGCCGGGAATCCTGTACCCGGGC---GGAATAAGTACCAGACCATTGACAACTA (-1bp)
DKK1-/- c1 allele A> CAGCGCCGCGCCGGGAATCCTGTACCCGGGCNNGGGAATAAGTACCAGACCATTGACAACTA (+2bp)
DKK1-/- c2 allele A> CAGCGCCGCGCCGGGAATCCTGTACCCGGGC---GGAATAAGTACCAGACCATTGACAACTA (-1bp)
DKK1-/- c2 allele B> CAGCGCCGCGCCGGGAAT---------------GGGAATAAGTACCAGACCATTGACAACTA (-13bp)
DKK1-/- c3 allele A> CAGCGCCGCGC--------------------C--GGGAATAAGTACCAGACCATTGACAACTA (-19bp)
DKK1-/- c3 allele B> CAGCGCCGCGC--------------------C--GGGAATAAGTACCAGACCATTGACAACTA (-19bp)
GLR:DKK1-/- c1 allele A> CAGCGCCGCGCCGGGAATCCTGTACCCGGGC-NGGGAATAAGTACCAGACCATTGACAACTA (+1bp)
GLR:DKK1-/- c1 allele B> CAGCGCCGCGCCGGGAATCCTGTACCCGGGC-NGGGAATAAGTACCAGACCATTGACAACTA (+1bp)
guide RNA
DKK3 target
Exon 2
DKK3 wildtype> AGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCC-A-GCTATCACAATGAGACCAACAC
DKK3-/- c1 allele A> AGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCNA-GCTATCACAATGAGACCAACAC (+1bp)
DKK3-/- c1 allele B> AGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCNA-GCTATCACAATGAGACCAACAC (+1bp)
DKK3-/- c2 allele A> AGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCC-ANGCTATCACAATGAGACCAACAC (+1bp)
DKK3-/- c2 allele B> AGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCC-A---TATCACAATGAGACCAACAC (-2bp)
guide RNA
CER1 target
Exon 1
CER1 wildtype> ACTCCTGCCAAGGAATCAAAGAGAGCTTCCCACAGG-CAAC-CATGAGGAAGCTGAGGAGAA
CER1-/- c1 allele A> ACTCCTGCCAAGGAATCAAAGAGAGCTTCCCA----------CATGAGGAAGCTGAGGAGAA (-8bp)
CER1-/- c1 allele A> ACTCCTGCCAAGGAATCAAAGAGAGCTTCCCACAGG-CAACNCATGAGGAAGCTGAGGAGAA (+1bp)
CER1-/-LEFTY1-/- c1 allele A> ACTCCTGCCAAGGAATCAAAGAGAGCTT-------G-CAAC-CATGAGGAAGCTGAGGAGAA (-7bp)
CER1-/-LEFTY1-/- c1 allele B> ACTCCTGCCAAGGAATCAAAGAGAGCTTCCCACAGGNCAAC-CATGAGGAAGCTGAGGAGAA (+1bp)
guide RNA
LEFTY1 target
Exon 1
LEFTY1 wildtype> GGAGGAGCTGGTCATCCCCACCCACG-TGAGGGCCCAGTACGTGGCCCTGCTGCAGCGCAG
LEFTY1-/- allele A> GGAGGAGCTGGTCATCCCCACCCACGNTGAGGGCCCAGTACGTGGCCCTGCTGCAGCGCAG (+1bp)
LEFTY1-/- allele B> GGAGGAGCTGGTCATCCCCACCCACGNTGAGGGCCCAGTACGTGGCCCTGCTGCAGCGCAG (+1bp)
guide RNA
CDH1 target
Exon 1
CDH1 wildtype> GGCCAGCCATGGGCCCTTGGAGCCGCAGCCTCTC-GGCGCTGCTGCTGCTGCTGCAGGTAC
CDH1-/- allele A> GGCCAGCCATGGGCCCTTGGAGCCGCAGCCTCTCNGGCGCTGCTGCTGCTGCTGCAGGTAC (+1bp)
CDH1-/- allele B> GGCCAGCCATGGGCCCTTGGAGCCGCAGCCTCTCNGGCGCTGCTGCTGCTGCTGCAGGTAC (+1bp)
guide RNA
SFRP1 target
Exon 1
SFRP1 wildtype> CAACCTGCTGGAGCACGAGACCATGGCGGAGGTGAAGCAGCAGGCCAGCAGCTGGGTGCC
SFRP1-/-SFRP2-/- c1 allele A> CAACCTGCTGGAGCACGAGACCATGG-GGAGGTGAAGCAGCAGGCCAGCAGCTGGGTGCC (-1bp)
SFRP1-/-SFRP2-/- c1 allele B> CAACCTGCTGGAGCACGAGACCATGG-GGAGGTGAAGCAGCAGGCCAGCAGCTGGGTGCC (-1bp)
guide RNA
SFRP2 target
Exon 1
SFRP2 wildtype> CTCTGCGTGCAGGTGAAGGA.....GTTTCCCCCAGG
SFRP1-/-SFRP2-/- c1 allele A> CTCTGCGTGCAGGTGAA-----(68 bp)----TCCCCCAGG (-68bp)
SFRP1-/-SFRP2-/- c1 allele B> CTCTGCGTGCAGGTGAA-----(68 bp)----TCCCCCAGG (-68bp)

B


CER1-/- clone 1
DKK1-/- clone 1
E-CAD-/- clone 1

Generation of E-CAD insertion and DKK1 inducible cell lines

Full length E-CAD and E-CAD lacking β-CAT binding domain were amplified via PCR from hE-cadherin-pcDNA3 and hE-cadherin/Δβ-catenin-pcDNA3 plasmids (Addgene plasmids #45769 and #45772) and inserted downstream of a pCAG promoter and puromycin resistance cassette flanked by 1kb homology arms for the AAVS1 safe harbour locus. This plasmid and TALENS targeting the AAVS1 site were then nucleofected into $1x10^6$ pluripotent E-CAD$^{-/-}$ cells using the B-016 setting on an Amaxa Nucleofector II (Lonza). Nucleofected cells were then plated as per maintenance conditions, but supplemented with 10 μM ROCK-inhibitor. Selection for both puromycin commenced after 2 days, and ROCK-inhibitor was maintained until colonies reached adequate size (typically 8-16 cells per colony). To derive pure clones, individual colonies were picked in an IVF hood with a 20 μL pipette tip and seeded into separate wells with growth media and ROCK-inhibitor. Once successfully established, each clone was assayed functionally for brightness and homogeneity of the overexpressed E-CAD. The RUES2-DOX:DKK1-V5 cell line was obtained in an identical fashion, except for the use of a TRE promotor instead of a pCAG promoter and an rTA-T2A-puromycin element instead of only puromycin. The V5 tag was attached to the C terminal end of DKK1, and DKK1 itself was obtained from a commercially available cDNA plasmid (Thermofisher Scientific # MHS6278-202801665).

Generation of KikGR photo-convertible hESC cell Line

pCAG:KikGR was a gift from Anna-Katerina Hadjantonakis[173] (Addgene plasmid # 32608). The KikGR protein from this plasmid was amplified with forward primer 5-ATTGGATCCCGGATGGTGAGTGTGATTACATCAGAA-3 and reverse primer 5-

TATGCGGCCGCCGGTTACTTGGCCAGCCTTG-3 and, using BamHI and NotI cloning sites, was inserted into an ePiggyBac plasmid[93] with a pCAG promoter and puromycin resistance cassette. This plasmid, along with a plasmid carrying the piggybac transposase and another ePiggyBac plasmid carrying a H2B-RFP657 fluorescent protein and blasticidin resistance, were nucleofected into $1x10^6$ pluripotent RUES2 cells using the B-016 setting on an Amaxa Nucleofector II (Lonza). Nucleofected cells were then plated as per maintenance conditions, but supplemented with 10 μM ROCK-inhibitor. Selection for both puromycin and blasticidin commenced after 2 days, and ROCK-inhibitor was maintained until colonies reached adequate size (typically 8-16 cells per colony). To derive pure clones, individual colonies were picked in an IVF hood with a 20uL pipette tip and seeded into separate wells with growth media and ROCK-inhibitor. Once successfully established, each clone was assayed functionally for brightness and homogeneity of the KikGR and H2B-RFP647 fluorescent proteins. Each clone was also assayed functionally for its ability to recapitulate the self-organization in micropatterns when stimulated with BMP4. Three successful clones were selected, and one was used for subsequent studies.

Transplantation of Human Organizer into Chick Host

A more detailed protocol for chick-human organizer grafts can be found in the "Methods in Molecular Biology: Chimera Methods" handbook, soon to be published by Springer Nature. An abbreviated version is as follows. Fertilized White Leghorn chicken eggs were obtained from Charles River Laboratories and incubated at 37-38°C and 50% of humidity and staged according to Hamburger and Hamilton[117]. Chick embryos were then removed from the egg and set up in Early-Chick culture[130], with Pannett-Compton saline solution as final wash and residual liquid in

the culture. hESCs were grown on EMB Cytoo coverslips. All other culture details remained the same. Once grown to the indicated time with the indicated stimulation conditions, 500 µm diameter colonies were peeled off whole with tungsten needles (Fine Science Tools). These colonies were washed twice with Pannett-Compton solution to remove culture growth factors and ligands. Colonies were then moved to chick embryos and grafted into the marginal zone between the area opaca and area pellucida approximately 90° away from the site of primitive streak initiation, following the example of a typical Hensen's node graft[174]. The grafted embryos were then returned to the incubator to develop and were imaged live one day after and were ultimately fixed between 24-48 hours post-graft. Due to background from the agar mount and chick auto-fluorescence, unfortunately only the SOX17-tdTomato marker had high enough signal to noise to be imaged live. In all steps Pen/Strep was used to minimize the chance for bacterial contamination.

hESC immunofluorescence

Cultures were fixed in 4% paraformaldehyde for 20 minutes, washed twice with PBS, and then blocked and permeabilized with 3% donkey serum and 0.1% Triton X-100 in PBS for 30 minutes. Cultures were incubated overnight with primary antibodies in this blocking buffer at 4°C (for primary antibodies and dilutions, see Table 6.4), washed 3 times with PBS+0.1% Tween-20 for 30 minutes each, and then incubated with secondary donkey antibodies (Alexa 488, Alexa 555, Alexa 647) and DAPI for 30 minutes before a final washing with PBS and mounting onto glass slides for imaging.

Chick immunofluorescence

Embryos were fixed in 4% paraformaldehyde in PBS for 1 hour at room temperature or overnight at 4°C. They were then washed 3 times with PBST (PBS+0.5% Triton X-100) for 1 hour each on nutator and blocked and permeabilized with 3% donkey serum, 1% bovine serum albumin in PBST for 2 hours, also at room temperature. Next, they were incubated overnight with anti-SOX2 antibody (R&D AF2018) diluted in blocking buffer at 4°C. The next day embryos were washed 3 times with PBST for 1 hour each on a nutator and then incubated with secondary donkey antibody Alexa-594, anti-human nuclear antigen (Novus Biologicals NBP2-34525AF647), and DAPI overnight. Embryos were washed times with PBST for 1 hour each and mounted in glass slides with fluoromount to image.

Chick *in situ* hybridization

Chicken *SOX3* probe was kindly provided by F.M. Vieceli and the whole mount *in situ* hybridization was performed using previously described procedures[175]. Briefly, the embryos were fixed overnight in 4% paraformaldehyde in PBS 24-48 hours after the grafting. The embryos were then washed 3 times with PBS+0.1% Tween-20, and then dehydrated through a methanol series (25% methanol/PBS, 50% methanol/PBS, 75% methanol/PBS, 100% methanol), and rehydrated (100% methanol, 75% methanol/PBS, 50% methanol/PBS, 25% methanol/PBS PBS), 15 minutes each step at room temperature. Next, the embryos were incubated with Proteinase K 10µg/ml for 5 minutes, rinsed twice in PBS+0.1% Tween-20, incubated in 2mg/ml glycine in PBS+0.1% Tween-20, washed 2 times in PBS+0.1% Tween-20 for 5 minutes each and post-fix for 20 minutes in 4% paraformaldehyde +0.2% glutaraldehyde in PBS. The embryos were then

hybridized at 70° C using antisense RNA chicken *SOX3, OTX2, HOXB1,* or *GBX2* probe labeled with digoxigenin-11-UTP. The probe was localized using AP-conjugated antibodies and the signal was developed with BM-Purple.

hESC RNA-Seq and qPCR data

RNA was collected in Trizol at indicated time points from either mircopatterned colonies or from small un-patterned colonies. Total RNA was purified using the RNeasy mini kit (Qiagen). qPCR was performed as described previously[71] and primer sequences are listed in Table 6.5.

RNA-seq data is from a previously published data set[71], and all raw data are available from the GEO database, accession number GSE77057.

Cell-tracking with photo-convertible line

RUES2-KikGR-RFP657-H2B cells were plated onto micropatterned CYTOO chips instead of home-made chips in order to accommodate the 19.5x19.5 mm spaced CYTOO chip holder. Immediately after stimulation with BMP, WNT3A, or WNT3A+SB, each chip was sequentially loaded into the CYTOO chip holder, placed on the microscope, photo-converted, washed with PBS, and then returned to the culture dish. Photo-conversion was carried out on a custom-built spinning-disk confocal Inverted Zeiss Axiovert 200 microscope with a Photonics Instruments Digital Mosaic system with a 405 nm laser. Regions of Interest (ROIs) for photo-conversion were programmed with custom Matlab code, and then loaded into the Metamorph software used to operate the microscope. Individual colonies were found, aligned with the ROI, and had their stage position stored. Using a custom written Metamorph script, each colony was sequentially imaged with

GFP and RFP filters, exposed to 3162 ms of 405 nm light from the laser, and then imaged again to check for complete photo-conversion. Once photo-converted, each CYTOO chip was returned to its native 35 mm dish and placed in an incubator. For tracking these cells at later times, the micropatterned chips were taken out of the incubator and sequentially re-loaded in the CYTOO holder and imaged with the afore-mentioned Lecia SP8 confocal microscope. They were then washed and returned to the incubator. For tracking cells with the RUES2-KikGR-RFP657-H2B cell line, segmentation was carried as for fixed cells, except here we used the H2B-RFP647 fluorescence signal instead of a DAPI signal as the nuclear marker. We then trained an Ilastik classifier to binarize cells as photo-converted or unconverted, and binned the converted cells into a radial histogram. The plots in Figure 2.7D represent the average of n=5 colonies.

Microscopy and Image Analysis

Fixed cell images were acquired with either a Zeiss Axio Observer and a 20x/0.8 numerical aperture (NA) lens, or with a Leica SP8 inverted confocal microscope with a 40×/1.1-NA water-immersion objective. Image analysis and stitching was performed with ImageJ and custom Matlab routines. Images were first background subtracted and normalized and then stitched on a colony by colony basis. Scenes for background subtraction and normalization were acquired in the spaces between colonies where no cells were present. For segmentation of individual cells, we first used Ilastik classification to separate foreground from background. The classifier was trained for each experiment on the DAPI images of 4 randomly chosen stitched colonies from that experiment. Once foreground and background were obtained, the DAPI channel was then filtered with a median and h-max filter and subtracted against a gradient of the image in order

to identify the nuclei centers. These centers were then used as seeds for a watershed, against which the background mask was applied to obtain the final segmentation. Using this segmentation mask we then obtained average intensities for each cell of the nuclear markers in the other channels. For radial plots, the intensity of IF signal for each marker was normalized to the DAPI intensity, and these corrected single-cell expressions were then radially binned and averaged. The final radial profile represents the average of the indicated number of colonies. For clustering and classification, single-cell intensities were log transformed and then clusters were fit with a Gaussian mixture model (see Figure 6.2 for example). For the images used in Figure 3.5, these were also deconvolved with Autoquant software and analysed in Imaris. In these images the notochord-like feature was identified by a combination of manual and Ilastik classification based on DAPI morphology, and cells belonging to this structure were segmented and false-coloured with the assistance of custom Python 3D segmentation software written by J. Metzger.

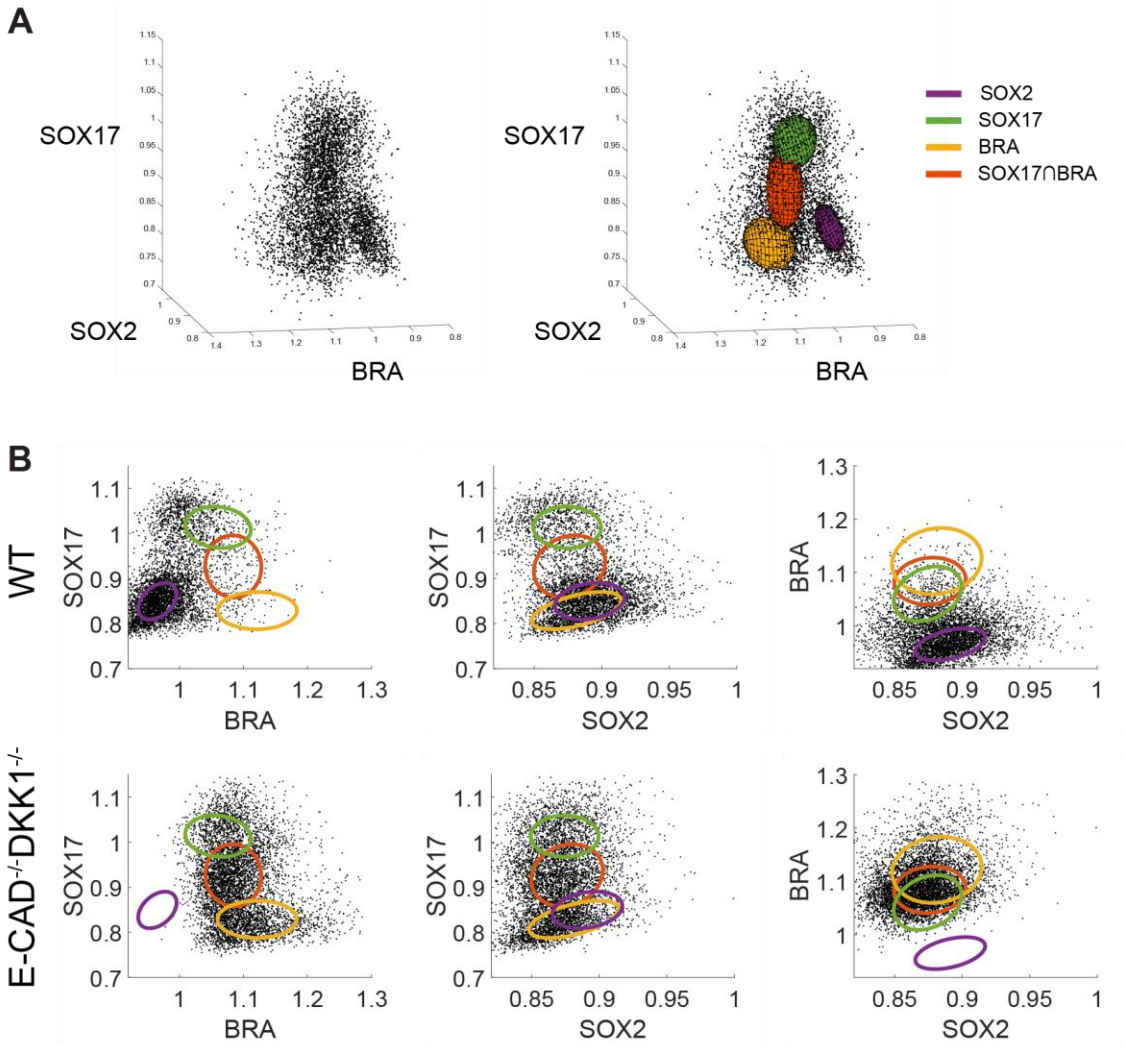


Figure 5.2 | Fate clustering and classification

(A) Scatterplots of combined cells from wild type, DKK1$^{-/-}$, E-CAD$^{-/-}$, E-CAD$^{-/-}$DKK1$^{-/-}$, and CER1$^{-/-}$ 48 h micropatterns data showing clustering into the Gaussian mixture model used for Figure 14.11A. (B) Example scatterplots of cells from wild type and E-CAD$^{-/-}$DKK1$^{-/-}$ micropatterns projected into 2D for better visualization of the clustering.

RUES2-GLR time-lapse imaging

RUES2-GLR cells were dissociated to single cells from growth plates with StemPro Accutase (Life Technologies), washed, and then re-suspended in MEF-CM with 10 µM ROCK-inhibitor Y-27632 (Abcam). CYTOO micropatterned chips were placed in 35mm tissue culture plastic dishes, and $8x10^5$ cells in 2mL of media were added to each coverslip. After 1 hour ROCK-inhibitor was removed and replaced with standard MEF-CM, supplemented with Pen/Strep (Life Technologies), and incubated overnight. The following morning the micropatterned coverslip was carefully removed from the dish and placed in a coverslip holder (CYTOOchambers from CYTOO), to which 1 mL of MEF-CM+Pen/Strep+50ng/mL BMP4 was added to induce differentiation. Immediately after media addition, the holder was transferred to a spinning-disk confocal microscope (CellVoyager CV1000, Yokogawa), in which fluorescent images were acquired every 30 minutes for 2 days. Multichannel time-lapse movies were generated from the raw images using ImageJ analysis software.

Mathematical Modelling See Appendix for details of the mathematical modelling.

APPENDIX

Supplementary Tables

Table 6.1 | Composition of Harvard University Embryonic Stem Cell Medium (HUESM)

HUESM Media (500mL)
245 mL DMEM w/ L-Glutamine
100 mL Knockout serum replacement
10 mL B27 supplement w/o Vitamin A
5 ml GlutaMAX
5 mL Non-essential amino acids
900 µL β-mercaptoethanol

Table 6.2 | sgRNA designs

Gene Symbol	**Forward primer**	**Reverse Primer**
CER1	caccgcttcctcatggttgcctgt	aaacacaggcaaccatgaggaagc
DKK1	caccgtctggtacttattcccgcc	aaacggcgggaataagtaccagac
DKK3	caccgtctcattgtgatagctggg	aaaccccagctatcacaatgagac
E-CAD	caccgcagcagcagcagcgccgag	aaacctcggcgctgctgctgctgc
LEFTY1	caccggtcatccccacccacgtga	aaactcacgtgggtggggatgacc
SFRP1	caccggagcacgagaccatggcgg	aaacccgccatggtctcgtgctcc
SFRP2	caccggtggtcgctgctagcgagg	aaaccctcgctagcagcgaccacc

Table 6.3 | Knockout lines

Target	Separate clones
RUES2:DKK1$^{-/-}$	3
RUES2-GLR:DKK1$^{-/-}$	2
RUES2:DKK3$^{-/-}$	2
RUES2:E-CAD$^{-/-}$	2
RUES2:CER1$^{-/-}$	1
RUES2:LEFTY1$^{-/-}$	1
RUES2:SFRP1$^{-/-}$	1
RUES2:SFRP2$^{-/-}$	1
RUES2:SFRP1$^{-/-}$SFRP2$^{-/-}$	1
RUES2:LEFTY1$^{-/-}$CER1$^{-/-}$	1

Table 6.4 | Target locus sequencing designs

Gene Symbol	Forward primer	Reverse Primer
CER1	tgctctttaagccccagaca	gcccaaggaccaaatctgta
DKK1	tttgttgtctccctcccaag	ctgtccacctctccaaaacc
DKK3	acccacctcccagagagatt	gcaggaagcctagggttaaa
E-CAD	tagagggtcaccgcgtctat	acaccccacccccctacac
LEFTY1	cctccaccccaggctataa	tggcagtacgaaaacagacc
SFRP1	agcctccggagtcagtgc	ttaggaatcacgtgcacagc
SFRP2	gcacccagcgaagagagc	ctggtacaactcgggaggag

Table 6.5 | hESC antibody information

Antigen	Antibody	Dilution
Active β-CATENIN	Millipore 05-665	1:400
BRACHYURY	R&D Systems AF-2085	1:300
CDX2	Abcam Ab-15258	1:50
COLLAGEN IV	Abcam Ab-6586	1:100
E-CADHERIN	Cell Signalling 3195	1:200
EOMES	Abcam Ab-23345	1:100
FOXA2	SCBT 6554	1:200
FOXF1	R&D Systems AF-4798	1:200
FZD7	R&D Systems MAB1981	1:200
GOOSECOID	R&D AF-4086	1:100
HAND1	R&D Systems AF-3168	1:100
ISL1	Abcam Ab-109517	1:200
LEF1	Cell Signalling 2230	1:200
LRP6	R&D Systems 1505	1:100
NANOG	R&D Systems AF-1997	1:200
N-CADHERIN	BioLegend 350802	1:200
OCT4	BD 611203	1:400
OTX2	SCBT 30659	1:200
PITX2	Abcam Ab-55599	1:100
SMAD2	BD bioscience 610842	1:100
SNAIL	R&D Systems AF-3639	1:200
SOX17	R&D Systems AF-1924	1:200
SOX2	Cell Signalling 3579	1:200
TBX6	R&D Systems AF-4744	1:100
GATA3	Thermofisher MA1-028	1:100
GATA6	Cell Signalling 5851	1:400
ZO1	Invitrogen 617300	1:200

Table 6.6 | qPCR Primer designs

Gene Symbol	Forward primer	Reverse Primer
AXIN2	ctccccaccttgaatgaaga	tggctggtgcaaagacatag
BMP4	tccacagcactggtcttgag	gggatgttctccagatgttctt
BRACHYURY	catgcaggtgagttgtcagaa	gctgtgacaggtacccaacc
CER1	gccatgaagtacattgggaga	cacagccttcgtgggttatag
CHORDIN	ccagccaggaggacacac	gtgcccacgttcaggaag
DKK1	ccttggatgggtattccaga	cctgaggcacagtctgatga
DKK3	agagcctgatggagccttg	ggcttgcacacatacaccag
LEF1	agatcaccccacctcttgg	atgagggatgccagttgtgt
LEFTY1	ctgcacaccctggacctt	atcccctgcaggtcaatgta
LEFTY2	cctggacctcagggactatg	atcccctgcaggtcaatgta
MSGN1	ggagaagctcaggatgagga	gtctgtgagttccccgatgt
NODAL	agacatcatccgcagccta	caaaagcaaacgtccagttct
NOGGIN	gaagctgcggaggaagttac	tacagcacggggcagaat
SFRP1	gctggagcacgagaccat	tggcagttcttgttgagca
SFRP2	gctagcagcgaccacctc	tttttgcaggcttcacatacc
SFRP5	tgcagtcaaattcatgttctcc	gccccgtagaagaaagggta
WNT3	ctcgctggctacccaattt	gagcccagagatgtgtactgc

Mathematical Modelling

1 Overview

In this supplement we present a minimal model that fits the micropattern WNT3A phenotype data for each CRISPR knockout cell line. The four key features of this data that the model captures are:

1. The geometric edge bias mediated by E-CAD.
2. The edge restriction of the WNT response by the secreted inhibitor DKK1.
3. The edge-to-interior EMT and WNT response wave, as especially observed in the DKK1$^{-/-}$ line.
4. The bifurcation of WNT differentiated cells to mesoderm versus endoderm depending on the relative levels of NODAL signal received.

To capture the first three features we need a model with two spatial diffusion terms: (1) to describe the spreading of DKK1 and its loss from the colony edge, and (2) to describe the implicit cell-cell communication via E-CAD that is responsible for the EMT wave as explained in the main text. As will be shown, this part of the model will give us the WNT signal received by cells at a given radius at a given time and can be used to determine the proportion of differentiated cells at 48 h after stimulation. As will also be shown, this part of the model consists of 10 parameters, 4 of which can be independently fitted to the E-CAD$^{-/-}$ phenotype, 5 of which can be fitted independently to the DKK1$^{-/-}$ phenotype, and 1 which can be estimated from the E-CAD$^{-/-}$DKK1$^{-/-}$ cell line.

To capture the last feature (the bifurcation of differentiated cells into either mesoderm or endoderm), ideally one would use detailed data of the gene regulatory networks for mesendoderm formation and mesoderm and endoderm bifurcation. However, these are not known to the required level of detail, and so we instead simply use the NODAL signalling level as defined by nuclear SMAD2 to determine fates, with cells with higher nuclear SMAD2 being more likely to be directed to endoderm rather than mesoderm.

2 PDE System

In the first part of the model we need to track three quantities: the WNT response, the level of DKK1, and the level of E-CAD. We let $W(r,t)$ represent the WNT response, we let $I(r,t)$ represent DKK1, and we let $E(r,t)$ represent E-CAD. $W(r,t)$ is the only quantity that we measure experimentally and that we use to fit the model (measured as nuclear LEF1 at 12 h and at late times, the relative loss of SOX2 i.e., 1-SOX2 after rescaling). Since all of the relevant experiments were conducted on 1000μm diameter disc micropatterns, this 2D geometry is the only geometry we consider in our model as well (though it can be easily generalized to other geometries).

Initial conditions: We assume that $W(r,t)$and $I(r,t)$ are off everywhere and that $E(r,t)$ is uniformly on. For the boundary condition on $W(r,t)$, we take this variable to be cell intrinsic, thus there is no diffusion of $W(r,t)$.

Boundary conditions: For the boundary condition on $I(r,t)$, we assume that it behaves like the BMP secreted inhibitor NOGGIN in the modelling of Etoc et al[71]. Thus DKK1 can freely diffuse over the colony and is quickly lost to the media on the edge, giving the boundary condition $I(R,t)=0$. Finally, for $E(r,t)$, since the E-CAD state of a cell also depends on its neighbours, we allow for a simple diffusion-like coupling with strength D_E. We do not initially know the scale of this dependence, whether it just immediate neighbours or if the coupling is long range. Thus D_E is one of the key parameters to be fitted, and, as will be shown, it has a direct effect on the speed of the EMT wave. For the boundary condition, we have experimentally determined that cells on the periphery of the micropatterns have a reduced inhibitory E-CAD effect and are thus more sensitive to WNT3A ligand (see Chapter 4 and especially Figure 4.5). We incorporate this observation by setting $E(R,t)=0$.

To model the interaction between the WNT response, DKK1, and E-CAD, simple Michaelis–Menten dynamics with activation and inhibition between the three species as shown in Figure 4.15 suffice. Mechanistically, we justify these interactions as follows: the inhibition of WNT by DKK1 is at the receptor level and is well known[144]; the activation of DKK1 by WNT is at the transcription level and was shown in Figure 4.8; the inhibition of E-CAD by WNT is also at the transcription level, though it operates more indirectly by WNT first turning on SNAIL and working with FGF[31] (which is always present and constant in our media); and there is extensive literature on the interaction of E-CAD with β-CAT and the effect this has on WNT signalling[176]. For this last interaction, we choose to model it as $E+W\rightarrow E$, i.e. akin to simple enzymatic degradation instead of a more complicated sequestration/release mechanism that might be modelled as $E+$

$W \rightleftharpoons C, C \rightarrow \emptyset$. We make this choice because although there is some evidence for this latter picture[177], much of the dynamic details of how E-CAD affects WNT signalling remain to be worked out. Thus in lieu of a more detailed mechanism, we use simple enzymatic degradation as a reasonable phenomenological approximation of the system, and we find that we achieve a better, more robust fit with it than with the sequestration/release model, even though the latter has more variables.

Since the scales on the three dynamical variables depend on imaging conditions and are thus arbitrary, we can absorb certain constants into their definitions without affecting the generality of the model. We set the maximum rate of synthesis of E-CAD to be equal to ν_E, the rate of natural E-CAD decay, so that in pluripotency the E-CAD level in each cell is 1. Likewise, we set the maximum rate of the WNT response to be equal to its degradation rate ν_W so that the WNT response varies from 0 to 1. We also assume, based on our evidence of direct induction of DKK1 by WNT (Figure 4.8) that this WNT response rate is a good approximation for the rate of DKK1 synthesis. The coefficient of W in equation (2) can be chosen to equal the degradation rate of W in (1) by adjusting the scale of I so we do not require a DKK1 specific synthesis rate. We do though let DKK1 have its own degradation rate ν_I.

Combining all of the above we obtain:

$$\frac{dW(r,t)}{dt} = \nu_W \frac{\theta_w^n}{\theta_W^n + I^n} - kEW - \nu_W W \qquad (1)$$

$$\frac{dI(r,t)}{dt} = D_I \nabla^2 I + \nu_W W - \nu_I I \qquad (2)$$

$$\frac{dE(r,t)}{dt} = D_E \nabla^2 E + \nu_E \left(\frac{\theta_E^m}{\theta_E^m + W^m} - E \right) \qquad (3)$$

Note that if we are purely interested in exploring the possible dynamics we could simplify this system further, rescaling the time for example so that $t = \tau * t_0$ with $t_0 = 1/\nu_W$. However, as we wish to use this model to fit our data, we need a match of the timescales between model and experiment, and hence we need to fit ν_W. Fortunately, we can get help in estimating its value since in the E-CAD$^{-/-}$DKK1$^{-/-}$ cell line the PDE system reduces to the one parameter equation:

$$\frac{dW(r,t)}{dt} = \nu_W(1 - W) \qquad (4)$$

A fit of the qPCR measured LEF1 response in small, unpatterned colonies of E-CAD$^{-/-}$DKK1$^{-/-}$ cells at various times results in a value of $\nu_W = 0.24/h$ (Figure 6.1). The upper and lower 95% confidence bounds are 0.07 and $0.40/h$. Since this estimate is based on mRNA data and so represents an upper limit, we restrict our search of possible ν_W values to be between 0.07 and $0.24/h$. We find that the final best fit value is $\nu_W = 0.11/h$, which means that $W(r,t)$ would reaches its half-maximal value ~6 hours after stimulation if the inhibition of E-CAD and DKK1 are not included.

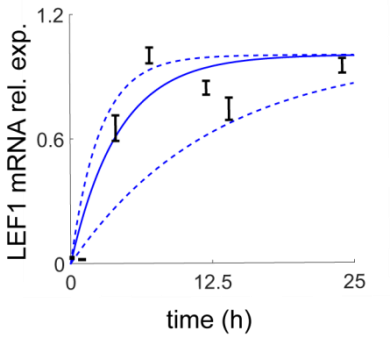


Figure 6.1 | Fit to Equation 4

Solid line indicates best fit of LEF1 qPCR timecourse in unpatterned E-CAD$^{-/-}$DKK1$^{-/-}$ cells to Equation 4. Dashed lines indicate 95% confidence interval in the fit. $\nu_W = 0.24\ (0.07, 0.40)$. Error bars represent the standard deviation of three biological replicates.

2.1 Fitting DKK1 and WNT interaction with E-CAD$^{-/-}$ cell line

In the E-CAD$^{-/-}$ cell line the system reduces to just equations (1) and (2) with $E(r,\tau)=0$. This leaves just 4 parameters to be fitted: n, θ_w, D_I and ν_I. This parameter space is small enough that we can perform a simple exhaustive grid search to find the global optima where the model WNT response curve at 12 and 48 h matches the data at 12 hours (measured by LEF1) and 48 hours (measured by BRA or SOX17 in Figure 4.13, which is also equivalent to 1-SOX2) as calculated by a least squares fit. Where possible, this grid search was centered on biologically relevant values for variables, or from previous similar simulations[71].

Figure 6.2 shows the best fit to the data. The Hill parameter n can adjust the slope of the curve, but the effects of the other parameters are less obvious. To understand these effects we do two things: (1) we perturb each parameter individually and see the change to the fit, and (2), we let the remaining parameters adapt to the perturbed parameter to see if and how the perturbation can be corrected. Doing this we can see for example that doubling D_I means that more cells on the boundary respond to WNT. Allowing the other parameters to vary while holding the doubled D_I constant, we can also see that decreasing ν_I can counteract this perturbation. This makes sense as increased D_I means more inhibitor is lost and decreasing ν_I counteracts this by making the inhibitor degrade less. The opposite is also true, since halving D_I means fewer cells on the boundary respond to WNT and this effect can be counteracted by increasing ν_I. For the WNT response threshold θ_w, we find that doubling it leads to less inhibitor and more differentiation, and that this can be counteracted by lowering the DKK1 degradation rate. Halving θ_w can conversely be counteracted by increasing ν_I. Taken together, these results imply that there is

some redundancy between certain parameters in our model. This is not surprising as although we formulated our model to be minimal in the number of parameters, we also based it directly on the known biology and variables that could be easily perturbed. A more abstract model with a reduced number of variables might give more independent parameter estimates, but it would be harder to explore and relate to the fundamental biology at hand.

Perturbing the E-CAD$^{-/-}$ fit

	D_I (mm^2h^{-1})	v_I (h^{-1})	θ_W
(1)	0.0024	0.0017	0.3000
(2)	0.0046	0.0017	0.3000
(3)	0.0046	0.0009	0.2073
(4)	0.0012	0.0017	0.3000
(5)	0.0012	0.0034	0.3628
(6)	0.0024	0.0850	0.3000
(7)	0.0022	0.0850	0.1893
(8)	0.0024	0	0.3000
(9)	0.0025	0	0.3045
(10)	0.0024	0.0017	0.5442
(11)	0.0013	0.0017	0.5442
(12)	0.0024	0.0017	0.1361
(13)	0.0050	0.0066	0.1361

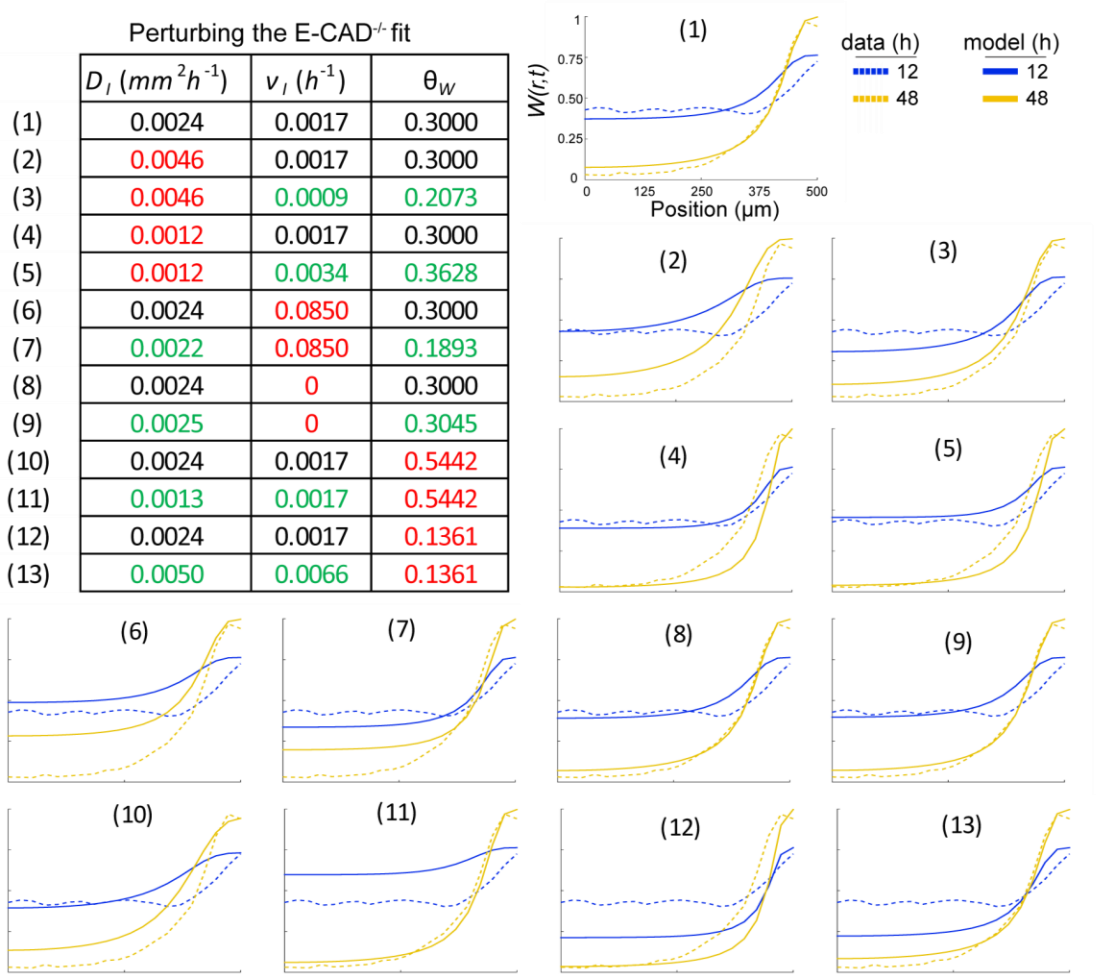


Figure 6.2 | Fit of DKK1 specific parameters with E-CAD$^{-/-}$ cell line

Values in black in the table are the best fit parameters, values in red are the parameters specifically perturbed, and values in green are the new best fit parameters based on the perturbed parameter.

2.2 Fitting E-CAD and WNT interaction with DKK1$^{-/-}$ cell line

In the DKK1$^{-/-}$ cell line the system reduces to just equations (1) and (3), with $I(r,\tau)=0$. As can be seen when solving for $W(r,\tau)$ in equilibrium (see Equation 5), this system undergoes a super-critical pitchfork or "cusp" bifurcation to two different WNT states (Figure 6.2) depending on the choice of just 3 parameters: m, θ_E, and k.

$$\frac{1}{W}=\frac{k}{\nu_W}*\frac{\theta_E^m}{\theta_E^m+W^m}+1 \qquad (5)$$

More interestingly, it can be shown that, depending on initial conditions or boundary conditions, this system can admit traveling wave solutions that switch from the higher equilibrium state to the lower equilibrium state[178,179]. In 1D these waves travel with a unique shape and have a speed largely dependent on the coupling constant (here D_E) and time scale of E-CAD synthesis and degradation (v_E). Since in our case we have the boundary condition $E(R,t)=0$, then we can have the situation whereby WNT stimulation directs the periphery cells to a WNT "high" equilibrium state that is more stable than the WNT "low" equilibrium state that their more central neighbours are in. This creates an unstable boundary that can only be resolved by a travelling wave solution. By adjusting m, θ_E, k, v_E and D_E ,we can moderate the speed and shape of the wave to fit the profile of WNT responsive cells in the DKK1$^{-/-}$ micropatterns at 12 hours (LEF1) and 48 hours (BRA and SOX17, as measured in Figure 4.13). As for the previous fit, this parameter space is sufficiently small enough that we can perform a grid search to find the global optima.

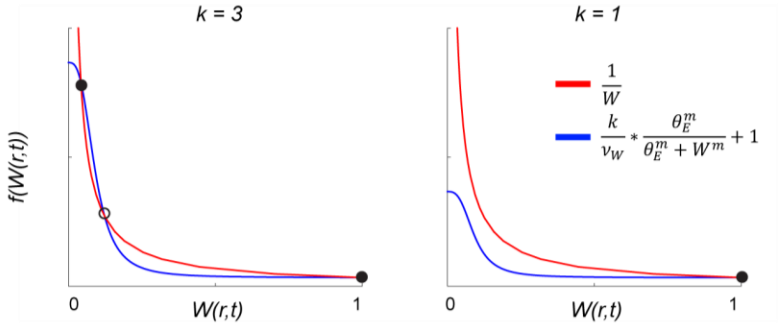


Figure 6.3 | Plot of cusp bifurcation from Equation 5 with different values of *k*

Solid circles mark stable points, open circles mark unstable points.

The best fit is shown in Figure 6.4. As for the DKK1 fit, the Hill parameter m can adjust the slope of the curve ($m = 3$ in our fit), and to learn the effect of the other parameters we again use selected perturbations. As one might expect, increasing the coupling constant D_E increases the speed of the WNT response wave while decreasing it slows it down. This parameter has the largest effect on wave speed, though changing ν_E, the effective turnover rate of E-CAD, also can modify speed since increasing it makes cells next in line in the wave respond faster once the wave reaches them. This can be noted in the perturbation analysis, as changing either D_E or ν_E leads to a change in the other variable to compensate for it. As one expects from Equation 5, θ_E and k are linked, so perturbing one leads to a correction from the other. Due to the non-linearity of the bifurcation they control they are also quite sensitive to perturbation, so just halving or doubling one leads to a collapse of the wave and one dominant state (i.e. either E-CAD is too strong and cells resist the WNT signal, or E-CAD is too weak and even a small WNT signal can downregulate it anywhere).

Perturbing the DKK1$^{-/-}$ fit

	$D_E\ (mm^2\,h^{-1})$	$v_E\ (h^{-1})$	θ_E	k
(1)	0.0043	8.0402	0.09005	2.4184
(2)	0.0086	8.0402	0.09005	2.4184
(3)	0.0086	30.0788	0.0272	9.0474
(4)	0.0021	8.0402	0.09005	2.4184
(5)	0.0021	67.0019	0.0367	6.0619
(6)	0.0043	8	0.09005	2.4184
(7)	0.0058	8	0.0272	9.0474
(8)	0.0043	4	0.09005	2.4184
(9)	0.0058	4	0.1650	1.2217
(10)	0.0043	8.0402	0.1801	20.1534
(11)	0.0332	6.0619	0.1801	1.2217
(12)	0.0043	8.0402	0.0450	20.1534
(13)	0.0045	0.0149	0.0450	6.0619
(14)	0.0043	8.0402	0.09005	4.8368
(15)	0.0548	100	0.0496	4.8368
(16)	0.0043	8.0402	0.09005	1.2092
(17)	0.0905	100	0.1650	1.2092

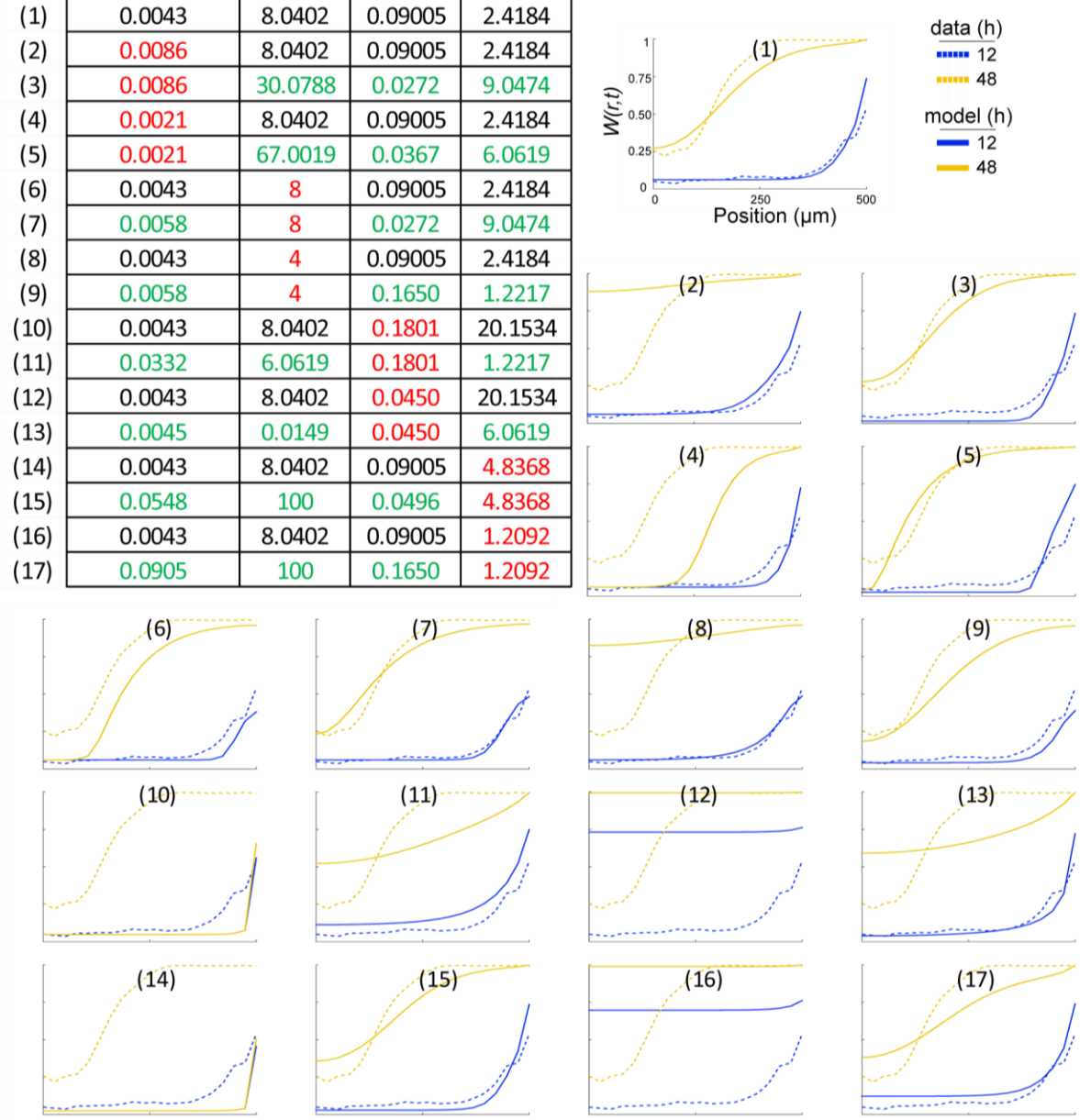


Figure 6.4 | Fit of E-CAD specific parameters with DKK1$^{-/-}$ cell line

Values in black in the table are the best fit parameters, values in red are the parameters specifically perturbed, and values in green are the new best fit parameters based on the perturbed parameter.

2.3 Applying fit to GLR:DKK1$^{-/-}$ time-lapse movie

As discussed in the main text, a further test of our fitted model was to apply it to our live-cell time-lapse GLR:DKK1$^{-/-}$ data. We found that though the model matches the data for the first 30 hours, beyond this there is a collapse of the inner SOX2 region that is quicker than what the model predicts (Figure 6.5). The agreement is a little better in a repeat of the experiment (Figure 6.5) but there is still a greater error between the model and these live cell data sets than the fixed antibody stained data we fit to. However, the virtual disappearance of the SOX2 domain in the colony center at 72 h is confirmed by the fixed data in Figure 4.13, which was not used in the fit. The difficulties in reproducing the movie data could entail the different dynamics of the live fluorescent reporter versus the LEF1 and the SOX2 stain, phototoxicity, or uncontrollable density differences.

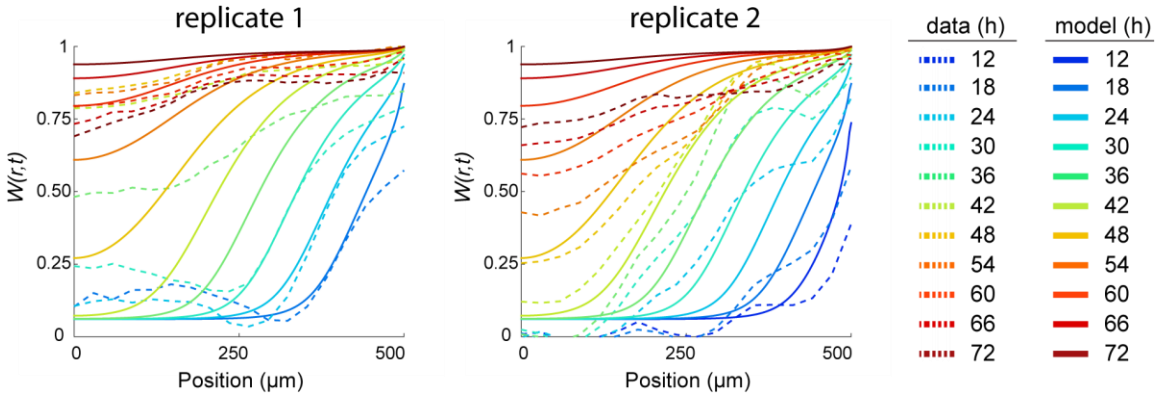


Figure 6.5 | Model comparison

Comparison of the previously fitted model to two replicates of continuously imaged GLR:DKK1$^{-/-}$ micropatterns, the first replicate is an average of 7 colonies, the second replicate is an average of 10 colonies.

3 Mesoderm versus endoderm fate bifurcation

Having obtained the radial profile of differentiated cells from Eqs 1-3 for each cell line, the next question is whether these cells will commit to mesoderm (BRA) or endoderm (SOX17). In the anterior mouse primitive streak, these are two mutually exclusive states that are influenced mainly by the relative levels of WNT versus NODAL signalling, with more WNT signalling leading to mesoderm and more NODAL signalling leading to endoderm. The effect of NODAL signalling on endoderm fate choice was also shown in Figure 2.3. Here we do not attempt to model the mechanism for this bistability or capture the dynamics. Instead, we simply let the choice depend on the ratio of WNT to NODAL signalling, letting the differentiated cells go to endoderm with probability $P(r)$ and going to mesoderm with probability $Q(r) = 1 - P(r)$. Since we are focused only on the fraction of cells at each radial position that differentiated, and in our model a differentiated cell is a cell that responded to the WNT signal, we assume that the WNT level is approximately the same across differentiated cells and thus a constant. NODAL signalling however is not the same for all cells, so we use measured nuclear SMAD2 levels for each cell line to approximate it (Figure 4.11 and Figure 6.6). Note that we examine the SMAD2 levels at 24 h rather than 48 h, as 48 h is too late and the fate decision has already been made. Thus our probabilities are

$$P(r) = \frac{NODAL(r)}{WNT(r) + NODAL(r)} = \frac{SMAD2_i(r) - b_1}{b_2 + (SMAD2_i(r) - b_1)} \quad (6)$$

$$Q(r) = \frac{WNT(r)}{WNT(r) + NODAL(r)} = 1 - P(r) \quad (7)$$

This effectively amounts to just a rescaled version of the measured SMAD2 profiles for each of our cell lines.

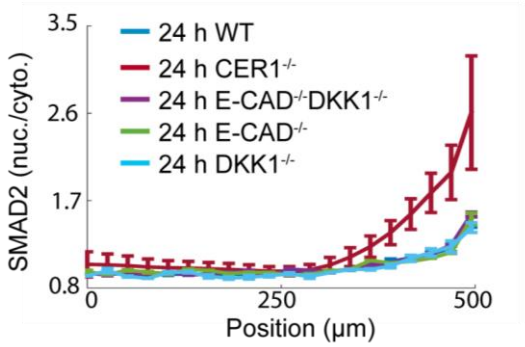


Figure 6.6 | Nuclear SMAD2

Radial profiles of nuclear SMAD2 in the cell lines modelled, n=20 colonies per condition.

4 Comparison of full model with data

Putting together the PDE system and the fate bifurcation parts of the model and comparing it to the measured SOX2, BRA, and SOX17 radial profiles at 48 h for each cell line, we see that our model captures the data very well (Figure 4.15E). One obvious missing feature, however, is the dual SOX17 and BRA positive cell population. These are cells that express both markers and have not yet committed to one fate versus the other (though we expect they will as time progresses) and are found in most significant numbers in the DKK1$^{-/-}$ and E-CAD$^{-/-}$DKK1$^{-/-}$ cell lines. If forced to apportion this population to either mesoderm or endoderm, one can see that classifying them as mesoderm would give the better fit to the model. Interestingly the major exception to this is at the colony edges in the double knockout E-CAD$^{-/-}$DKK1$^{-/-}$ cell line, where doing so would lead the model to underestimate the number of BRA cells counted in the imaging. To further investigate this large discrepancy between the model and the data, we re-examined our E-CAD$^{-/-}$ DKK1$^{-/-}$ micropattern data. We found that on the colony edge (and only on the colony edge) BRA cells sometimes lumped together in discrete clusters (Figure 6.7). As these clusters were

immediately surrounded by a monolayer primarily composed of SOX17 cells, instead of the salt-and-pepper BRA or SOX17 pattern found further in the interior of the colony, we suppose that these clusters formed due to cell movement, i.e. cells were initially patterned in a stochastic manner but the BRA cells then moved and coalesced together while the SOX17 cells either stayed in place or underwent unbiased migration. That this only takes place on the edge of the E-CAD$^{-/-}$ DKK1$^{-/-}$ micropatterns could be explained by noting that it is in this regime that cells are the least restricted from moving, both by E-CAD and by existence of neighbours. The clusters are often quite 3 dimensional, and since our analysis involves just 2D segmentation on epifluorescence images, classification of cells in these clusters may be biased and introduce errors such as overcounting the number of SOX17/BRA cells. Migration was also not included in the model, and introducing a biased outward migration term for BRA cells for example might lead to an improved (though more complex) fit.

Overall, our model faithfully captures the E-CAD and DKK1 patterning dynamics as observed in our micropatterns and live-reporter lines, and in the most striking case where it does not, it still proves informative, pointing to a segregation and clustering of mesoderm versus endoderm cells that may be due to cell migration and merits further investigation.

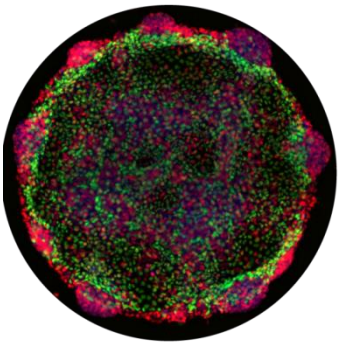
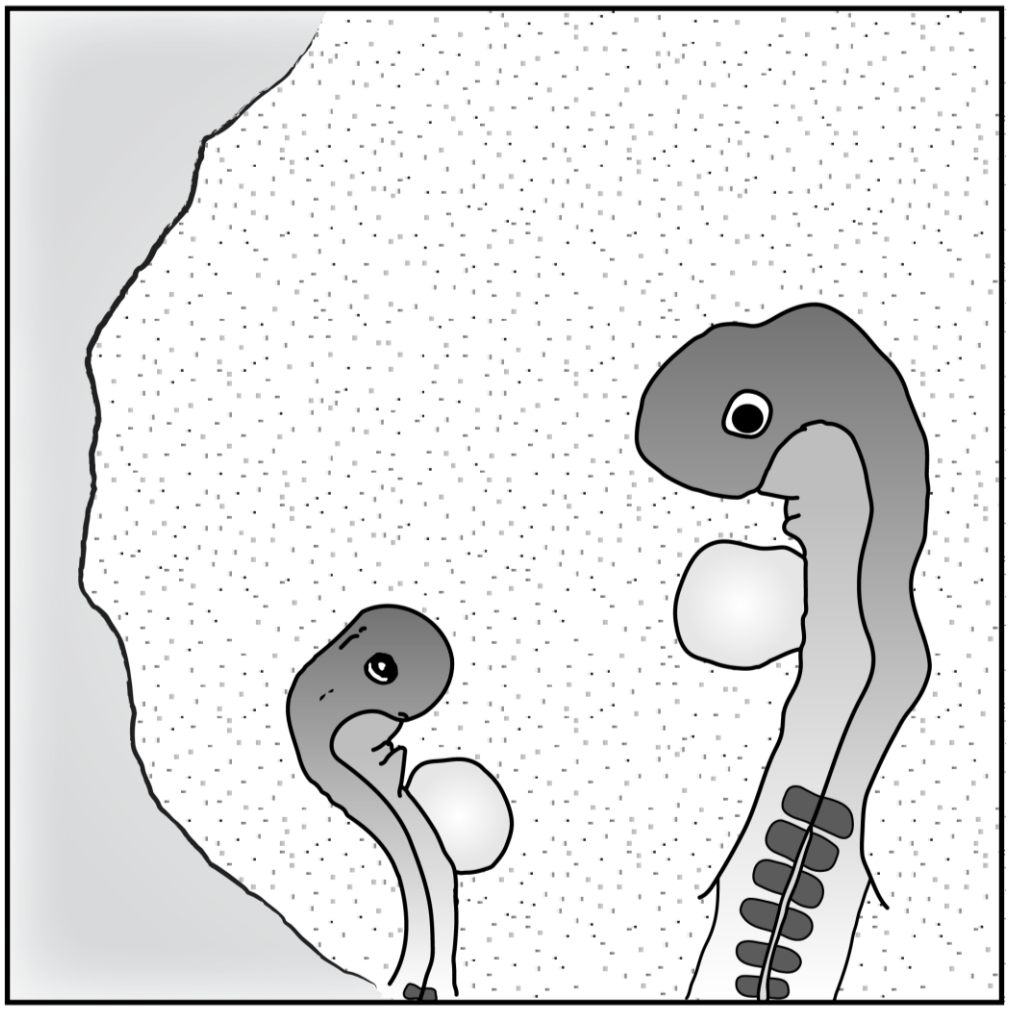


Figure 6.7 | BRA clusters on periphery

Example of an E-CAD$^{-/-}$DKK1$^{-/-}$ micropattern with BRA clusters on the periphery. The center remains a salt-and-pepper mix of SOX17 and BRA cells.

POSTSCRIPT



"Try and think of this as a wnt-wnt situation."

REFERENCES

1. Nusse, R. & Varmus, H. E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31,** 99–109 (1982).

2. Nüsslein-Volhard, C. & Wieschaus, E. Mutations affecting segment number and polarity in Drosophila. *Nature* **287,** 795–801 (1980).

3. Sharma, R. P. & Chopra, V. L. Effect of the Wingless (wg1) mutation on wing and haltere development in Drosophila melanogaster. *Dev. Biol.* **48,** 461–465 (1976).

4. Rijsewijk, F. *et al.* The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* **50,** 649–657 (1987).

5. Ruiz-Trillo, I., Roger, A. J., Burger, G., Gray, M. W. & Lang, B. F. A phylogenomic investigation into the origin of metazoa. *Mol. Biol. Evol.* **25,** 664–672 (2008).

6. Schierwater, B. *et al.* Concatenated analysis sheds light on early metazoan evolution and fuels a modern 'urmetazoon' hypothesis. *PLoS Biol.* **7,** e20 (2009).

7. Sebe-Pedros, A., de Mendoza, A., Lang, B. F., Degnan, B. M. & Ruiz-Trillo, I. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan Capsaspora owczarzaki. *Mol. Biol. Evol.* **28,** 1241–1254 (2011).

8. King, N. *et al.* The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. *Nature* **451,** 783–788 (2008).

9. Willert, K. & Nusse, R. Wnt proteins. *Cold Spring Harb. Perspect. Biol.* **4,** a007864 (2012).

10. Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423,** 448–452 (2003).

11. Takada, R. *et al.* Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. *Dev. Cell* **11,** 791–801 (2006).

12. Rios-Esteves, J. & Resh, M. D. Stearoyl CoA desaturase is required to produce active, lipid-modified Wnt proteins. *Cell Rep.* **4,** 1072–1081 (2013).

13. Rios-Esteves, J., Haugen, B. & Resh, M. D. Identification of key residues and regions important for porcupine-mediated Wnt acylation. *J. Biol. Chem.* **289,** 17009–17019 (2014).

14. Nusse, R. & Clevers, H. Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169,** 985–999 (2017).

15. Farin, H. F. *et al.* Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530,** 340–343 (2016).

16. van Amerongen, R. Alternative Wnt pathways and receptors. *Cold Spring Harb. Perspect. Biol.* **4,** (2012).

17. Petersen, C. P. & Reddien, P. W. Wnt signaling and the polarity of the primary body axis. *Cell* **139,** 1056–68 (2009).

18. Wodarz, A. & Nusse, R. Mechanisms of Wnt signaling in development. *Annu. Rev. Cell Dev. Biol.* **14,** 59–88 (1998).

19. He, X., Semenov, M., Tamai, K. & Zeng, X. LDL receptor-related proteins 5 and 6 in Wnt/β-catenin signaling: arrows point the way. *Development* **131,** 1663–1677 (2004).

20. Stamos, J. L. & Weis, W. I. The β-catenin destruction complex. *Cold Spring Harb. Perspect. Biol.* **5,** a007898 (2013).

21. Cadigan, K. M. & Waterman, M. L. TCF/LEFs and Wnt Signaling in the Nucleus. *Cold Spring Harb Symp Quant Biol* **4,** 1–22 (2012).

22. MacDonald, B. T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev. Cell* **17,** 9–26 (2009).

23. Nelson, W. J. & Nusse, R. Convergence of Wnt, β-catenin, and cadherin pathways. *Science* **303,** 1483–1487 (2004).

24. Heuberger, J. & Birchmeier, W. Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. *Cold Spring Harb. Perspect. Biol.* **2,** 1–25 (2010).

25. Valenta, T., Hausmann, G. & Basler, K. The many faces and functions of β-catenin. *EMBO J.* **31,** 2714–36 (2012).

26. Wieschaus, E., Nüsslein-Volhard, C. & Jurgens, G. Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster : III. Zygotic loci on the X-chromosome and fourth chromosome. *Wilhelm Roux's Arch. Dev. Biol.* **193,** 296–307 (1984).

27. Ozawa, M., Baribault, H. & Kemler, R. The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J.* **8,** 1711–7 (1989).

28. Martin, B. L. & Kimelman, D. Wnt signaling and the evolution of embryonic posterior development. *Curr. Biol.* **19,** R215-9 (2009).

29. Houston, D. W. & Wylie, C. The role of Wnts in gastrulation. in *Gastrulation: from cells to embryo* (ed. Stern, C. D.) 521–537 (Cold Spring Harbor Press, 2004).

30. 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).

31. 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).

32. Tam, P. P. L. & Loebel, D. a F. Gene function in mouse embryogenesis: get set for gastrulation. *Nat. Rev. Genet.* **8,** 368–81 (2007).

33. Liu, P. *et al.* Requirement for Wnt3 in vertebrate axis formation. *Nat. Genet.* **22,** 361–365 (1999).

34. Mohamed, O. A., Clarke, H. J. & Dufort, D. β-catenin signaling marks the prospective site of primitive streak formation in the mouse embryo. *Dev. Dyn.* **231,** 416–424 (2004).

35. Mukhopadhyay, M. *et al.* Dickkopf1 Is Required for Embryonic Head Induction and Limb Morphogenesis in the Mouse. *Dev. Cell* **1,** 423–434 (2001).

36. Glinka, A. *et al.* Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* **391,** 357–362 (1998).

37. Haegel, H. *et al.* Lack of β-catenin affects mouse development at gastrulation. *Development* **121,** 3529–3537 (1995).

38. Huelsken, J. *et al.* Requirement for beta-catenin in anterior-posterior axis formation in mice. *J. Cell Biol.* **148,** 567–578 (2000).

39. Lewis, S. L. *et al.* Dkk1 and Wnt3 interact to control head morphogenesis in the mouse. *Development* **135,** 1791–1801 (2008).

40. Popperl, H. *et al.* Misexpression of Cwnt8C in the mouse induces an ectopic embryonic axis and causes a truncation of the anterior neuroectoderm. *Development* **124,** 2997–3005 (1997).

41. Merrill, B. J. *et al.* Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development* **131,** 263–274 (2004).

42. Ishikawa, T., Tamai, Y., Li, Q., Oshima, M. & Taketo, M. M. Requirement for tumor suppressor Apc in the morphogenesis of anterior and ventral mouse embryo. *Dev. Biol.* **253,** 230–246 (2003).

43. Skromne, I. & Stern, C. D. Interactions between Wnt and Vg1 signalling pathways initiate primitive streak formation in the chick embryo. *Development* **128,** 2915–2927 (2001).

44. Shah, S. B. *et al.* Misexpression of chick Vg1 in the marginal zone induces primitive streak formation. *Development* **124,** 5127–5138 (1997).

45. Pfeffer, P. L., De Robertis, E. M. & Izpisua-Belmonte, J. C. Crescent, a novel chick gene encoding a Frizzled-like cysteine-rich domain, is expressed in anterior regions during early embryogenesis. *Int. J. Dev. Biol.* **41,** 449–458 (1997).

46. Bertocchini, F. & Stern, C. D. The hypoblast of the chick embryo positions the primitive streak by antagonizing nodal signaling. *Dev. Cell* **3,** 735–744 (2002).

47. Stern, C. D. & Downs, K. M. The hypoblast (visceral endoderm): an evo-devo perspective. *Development* **139,** 1059–1069 (2012).

48. Deglincerti, A. *et al.* Self-organization of the in vitro attached human embryo. *Nature* **533,** 251–254 (2016).

49. Shahbazi, M. N. *et al.* Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* **18,** 700–708 (2016).

50. Fogarty, N. M. E. *et al.* Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* **550,** 67–73 (2017).

51. Hyun, I., Wilkerson, A. & Johnston, J. Embryology policy: Revisit the 14-day rule. *Nature* **533,** 169–171 (2016).

52. International Society of Stem Cell Research. *Guidelines for Stem Cell Research and Clinical Translation*. *International Society for Stem Cell Research* (2016).

53. Gasser, R. F., Cork, R. J., Stillwell, B. J. & McWilliams, D. T. Rebirth of human embryology. *Dev. Dyn.* **243,** 621–628 (2014).

54. O'Rahilly, R. & Müller, F. *Developmental Stages in Human Embryos*. (Carnegie Institution of Washington, 1987).

55. O'Rahilly, R. & Muller, F. Developmental stages in human embryos: revised and new measurements. *Cells. Tissues. Organs* **192,** 73–84 (2010).

56. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282,** 1145–1147 (1998).

57. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4,** 487–492 (2009).

58. Itskovitz-Eldor, J. *et al.* Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol. Med.* **6,** 88–95 (2000).

59. Wesselschmidt, R. L. The teratoma assay: an in vivo assessment of pluripotency. *Methods Mol. Biol.* **767,** 231–241 (2011).

60. James, D., Noggle, S. A., Swigut, T. & Brivanlou, A. H. Contribution of human embryonic stem cells to mouse blastocysts. *Dev. Biol.* **295,** 90–102 (2006).

61. Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10,** 771–785 (2012).

62. McCauley, H. A. & Wells, J. M. Pluripotent stem cell-derived organoids: using principles of developmental biology to grow human tissues in a dish. *Development* **144,** 958–962 (2017).

63. Spence, J. R. *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470,** 105–109 (2011).

64. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459,** 262–265 (2009).

65. Dye, B. R. *et al.* In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* **4,** (2015).

66. Takasato, M. *et al.* Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **526,** 564–568 (2015).

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

68. Simunovic, M. & Brivanlou, A. H. Embryoids, organoids and gastruloids: new approaches to understanding embryogenesis. *Development* **144,** 976–985 (2017).

69. 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,** 847–54 (2014).

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

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

72. Brennan, J. *et al.* Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411,** 965–969 (2001).

73. Zorn, A. M. & Wells, J. M. Vertebrate Endoderm Development and Organ Formation. *Annu. Rev. Cell Dev. Biol.* **25,** 221–251 (2009).

74. Faial, T. *et al.* Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic

stem cells. *Development* **142,** 2121–35 (2015).

75. Morgani, S. M., Metzger, J. J., Nichols, J., Siggia, E. D. & Hadjantonakis, A.-K. Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. *Elife* **7,** e32839 (2018).

76. Firulli, A. B., McFadden, D. G., Lin, Q., Srivastava, D. & Olson, E. N. Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat. Genet.* **18,** 266–270 (1998).

77. Home, P. *et al.* Genetic redundancy of GATA factors in the extraembryonic trophoblast lineage ensures the progression of preimplantation and postimplantation mammalian development. *Development* **144,** 876–888 (2017).

78. Ralston, A. *et al.* Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137,** 395–403 (2010).

79. Manaia, A., Lemarchandel, V., Klaine, M., Romeo, P. & Godin, I. Lmo2 and GATA-3 associated expression in intraembryonic hemogenic sites. *Development* **127,** 643–653 (2000).

80. Beck, F., Erler, T., Russell, A. & James, R. Expression of Cdx-2 in the Mouse Embryo and Placenta: Possible Role in Patterning of the Extra-Embryonic Membranes. *Dev. Dyn.* **204,** 219–227 (1995).

81. Sherwood, R. I., Maehr, R., Mazzoni, E. O. & Melton, D. A. Wnt signaling specifies and patterns intestinal endoderm. *Mech. Dev.* **128,** 387–400 (2011).

82. Peterson, R. S. *et al.* The winged helix transcriptional activator HFH-8 is expressed in the mesoderm of the primitive streak stage of mouse embryos and its cellular derivatives. *Mech. Dev.* **69,** 53–69 (1997).

83. Mahlapuu, M., Ormestad, M., Enerback, S. & Carlsson, P. The forkhead transcription factor Foxf1 is required for differentiation of extra-embryonic and lateral plate mesoderm. *Development* **128,** 155–166 (2001).

84. Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R. & Grosveld, F. The transcription factor GATA6 is essential for early extraembryonic development. *Development* **126,** 723–732 (1999).

85. Cai, C. L. *et al.* Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev. Cell* **5,** 877–889 (2003).

86. Zhuang, S. *et al.* Expression of Isl1 during mouse development. *Gene Expr. Patterns* **13,** 407–412 (2013).

87. Chapman, D. L., Agulnik, I., Hancock, S., Silver, L. M. & Papaioannou, V. E. Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev. Biol.* **180,** 534–542 (1996).

88. Yoon, J. K., Moon, R. T. & Wold, B. The bHLH class protein pMesogenin1 can specify paraxial mesoderm phenotypes. *Dev. Biol.* **222,** 376–391 (2000).

89. Chalamalasetty, R. B. *et al.* Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation. *Development* **141,** 4285–4297 (2014).

90. Savory, J. G. A. *et al.* Cdx2 regulation of posterior development through non-Hox targets. *Development* **136,** 4099–4110 (2009).

91. Nowotschin, S. *et al.* Charting the emergent organotypic landscape of the mammalian gut endoderm at single-cell resolution. *bioRxiv* 471078 (2018). doi:10.1101/471078

92. Robb, L. & Tam, P. P. L. Gastrula organiser and embryonic patterning in the mouse. *Semin. Cell Dev. Biol.* **15,** 543–554 (2004).

93. Lacoste, A., Berenshteyn, F. & Brivanlou, A. H. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. *Cell Stem Cell* **5,** 332–342 (2009).

94. Parameswaran, M. & Tam, P. P. Regionalisation of cell fate and morphogenetic movement of the mesoderm during mouse gastrulation. *Dev. Genet.* **17,** 16–28 (1995).

95. Sutherland, A. E. Tissue morphodynamics shaping the early mouse embryo. *Semin. Cell Dev. Biol.* **55,** 89–98 (2016).

96. Viotti, M., Nowotschin, S. & Hadjantonakis, A.-K. SOX17 links gut endoderm morphogenesis and germ layer segregation. *Nat. Cell Biol.* **16,** 1146–1156 (2014).

97. Rivera-Perez, J. A. & Hadjantonakis, A.-K. The Dynamics of Morphogenesis in the Early Mouse Embryo. *Cold Spring Harb. Perspect. Biol.* **7,** (2014).

98. Hardy, K. M. *et al.* Non-canonical Wnt signaling through Wnt5a/b and a novel Wnt11 gene, Wnt11b, regulates cell migration during avian gastrulation. *Dev. Biol.* **320,** 391–401 (2008).

99. Yang, X., Chrisman, H. & Weijer, C. J. PDGF signalling controls the migration of mesoderm cells during chick gastrulation by regulating N-cadherin expression. *Development* **135,** 3521–3530 (2008).

100. Yang, X., Dormann, D., Munsterberg, A. E. & Weijer, C. J. Cell movement patterns during gastrulation in the chick are controlled by positive and negative chemotaxis mediated by FGF4 and FGF8. *Dev. Cell* **3,** 425–437 (2002).

101. Yue, Q., Wagstaff, L., Yang, X., Weijer, C. & Munsterberg, A. Wnt3a-mediated chemorepulsion controls movement patterns of cardiac progenitors and requires RhoA function. *Development* **135,** 1029–1037 (2008).

102. Sweetman, D., Wagstaff, L., Cooper, O., Weijer, C. & Munsterberg, A. The migration of paraxial and lateral plate mesoderm cells emerging from the late primitive streak is controlled by different Wnt signals. *BMC Dev. Biol.* **8,** 63 (2008).

103. Stankova, V., Tsikolia, N. & Viebahn, C. Rho kinase activity controls directional cell movements during primitive streak formation in the rabbit embryo. *Development* **142,** 92–98 (2015).

104. Nakamura, T. *et al.* Single-cell transcriptome of early embryos and cultured embryonic stem cells of cynomolgus monkeys. *Sci. data* **4,** 170067 (2017).

105. Spemann, H. & Mangold, H. Induction of embryonic primordia by implantation of organizers from a different species. *Int. J. Dev. Biol.* **45,** 13–38 (2001).

106. Oppenheimer, J. M. Transplantation experiments on developing teleosts (Fundulus and Perca). *J. Exp. Zool.* **72,** 409–437 (1936).

107. Waddington, C. H. Experiments on the Development of Chick and Duck Embryos, Cultivated in vitro. *Philos. Trans. R. Soc. London. Ser. B* **221,** 179 LP-230 (1932).

108. Beddington, R. S. Induction of a second neural axis by the mouse node. *Development* **120,** 613–620 (1994).

109. Kinder, S. J. *et al.* The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* **128,** 3623–3634 (2001).

110. Sharon, N., Mor, I., Golan-lev, T., Fainsod, A. & Benvenisty, N. Molecular and functional characterizations of gastrula organizer cells derived from human embryonic stem cells. *Stem Cells* **29,** 600–608 (2011).

111. Nishihara, A. & Hashimoto, C. Tail structure is formed when blastocoel roof contacts blastocoel floor in Xenopus laevis. *Dev. Growth Differ.* **56,** 214–222 (2014).

112. Le Douarin, N. M. & Jotereau, F. V. Tracing of cells of the avian thymus through embryonic life in interspecific chimeras. *J. Exp. Med.* **142,** 17–40 (1975).

113. Couly, G. F. & Le Douarin, N. M. Mapping of the early neural primordium in quail-chick chimeras. I. Developmental relationships between placodes, facial ectoderm, and prosencephalon. *Dev. Biol.* **110,** 422–439 (1985).

114. Douarin, N. Le, Dieterlen-Lièvre, F., Creuzet, S. & Teillet, M. A. Chapter 2 Quail-Chick

Transplantations. *Methods Cell Biol.* **87,** 19–58 (2008).

115. Stern, C. D. The chick: A great model system becomes even greater. *Dev. Cell* **8,** 9–17 (2005).

116. Mok, G. F., Alrefaei, A. F., McColl, J., Grocott, T. & Münsterberg, A. Chicken as a Developmental Model. *eLS* 1–8 (2015). doi:10.1002/9780470015902.a0021543

117. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88,** 49–92 (1951).

118. Goldstein, R. S. Transplantation of Human Embryonic Stem Cells and Derivatives to the Chick Embryo. in *Human Embryonic Stem Cell Protocols* (ed. Turksen, K.) 367–385 (Humana Press, 2010). doi:10.1007/978-1-60761-369-5_20

119. Lee, G. *et al.* Isolation and directed differentiation of neural crest stem cells derived from human embryonic stem cells. *Nat. Biotechnol.* **25,** 1468–1475 (2007).

120. Boulland, J.-L., Halasi, G., Kasumacic, N. & Glover, J. C. Xenotransplantation of Human Stem Cells into the Chicken Embryo. *J. Vis. Exp.* 11–15 (2010). doi:10.3791/2071

121. Glover, J. C., Boulland, J.-L., Halasi, G. & Kasumacic, N. Chimeric Animal Models in Human Stem Cell Biology. *ILAR J.* **51,** 62–73 (2010).

122. Li, W. *et al.* Engraftable neural crest stem cells derived from cynomolgus monkey embryonic stem cells. *Biomaterials* **39,** 75–84 (2015).

123. Fontaine-Pérus, J. *et al.* Mouse-chick chimera: a developmental model of murine neurogenic cells. *Development* **124,** 3025–3036 (1997).

124. Fontaine-Pérus, J. Mouse-Chick Chimera: An Experimental System for Study of Somite Development. in *Current Topics in Developmetal Biology* 269–300 (2000).

125. Fontaine-Pérus, J. & Chéraud, Y. Mouse-chick neural chimeras. *Int. J. Dev. Biol.* **49,** 349–353 (2005).

126. Jiang, X. *et al.* Isolation and Characterization of Neural Crest Stem Cells Derived From In Vitro–Differentiated Human Embryonic Stem Cells. *Stem Cells Dev.* **18,** 1059–1071 (2009).

127. Lee, H. *et al.* Directed Differentiation and Transplantation of Human Embryonic Stem Cell-Derived Motoneurons. *Stem Cells* **25,** 1931–1939 (2007).

128. Zhu, L., Belo, J. A., De Robertis, E. M. & Stern, C. D. Goosecoid Regulates the Neural Inducing Strength of the Mouse Node. *Dev. Biol.* **216,** 276–281 (1999).

129. Knoetgen, H., Teichmann, U., Wittler, L., Viebahn, C. & Kessel, M. Anterior neural induction by nodes from rabbits and mice. *Dev. Biol.* **225,** 370–80 (2000).

130. Chapman, S. C., Collignon, J., Schoenwolf, G. C. & Lumsden, A. Improved method for chick whole-embryo culture using a filter paper carrier. *Dev. Dyn.* **220,** 284–289 (2001).

131. Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D. & Brivanlou, A. H. Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature* **558,** 132–135 (2018).

132. Riggleman, B., Schedl, P. & Wieschaus, E. Spatial expression of the Drosophila segment polarity gene armadillo is posttranscriptionally regulated by wingless. *Cell* **63,** 549–560 (1990).

133. Valenta, T., Hausmann, G. & Basler, K. The many faces and functions of β-catenin. *EMBO J.* **31,** 2714–2736 (2012).

134. Cadigan, K. M. & Peifer, M. Wnt signaling from development to disease: insights from model systems. *Cold Spring Harb. Perspect. Biol.* **1,** a002881 (2009).

135. Estarás, C., Benner, C. & Jones, K. A. SMADs and YAP Compete to Control Elongation of β-Catenin: LEF-1-Recruited RNAPII during hESC Differentiation. *Mol. Cell* **58,** 780–793 (2014).

136. MacDonald, B. T. & He, X. Frizzled and LRP5/6 Receptors for Wnt/β-Catenin Signaling. in *Cold Spring Harbor Perspectives in Biology* (eds. Nusse, R., He, X. & van Amerongen, R.) **4,** 65–87 (Cold Spring Harbor Laboratory Press, 2013).

137. Fagotto, F., Funayama, N., Gluck, U. & Gumbiner, B. M. Binding to cadherins antagonizes the signaling activity of β-catenin during axis formation in Xenopus. *J. Cell Biol.* **132,** 1105–1114 (1996).

138. Libusova, L., Stemmler, M. P., Hierholzer, A., Schwarz, H. & Kemler, R. N-cadherin can structurally substitute for E-cadherin during intestinal development but leads to polyp formation. *Development* **137,** 2297–2305 (2010).

139. Rosowski, K. A., Mertz, A. F., Norcross, S., Dufresne, E. R. & Horsley, V. Edges of human embryonic stem cell colonies display distinct mechanical properties and differentiation potential. *Sci. Rep.* **5,** 14218 (2015).

140. Przybyla, L., Lakins, J. N. & Weaver, V. M. Tissue Mechanics Orchestrate Wnt-Dependent Human Embryonic Stem Cell Differentiation. *Cell Stem Cell* **19,** 462–475 (2016).

141. Kanellos, G. *et al.* ADF and Cofilin1 Control Actin Stress Fibers, Nuclear Integrity, and Cell Survival. *Cell Rep.* **13,** 1949–1964 (2015).

142. Gottardi, C. J., Wong, E. & Gumbiner, B. M. E-cadherin suppresses cellular transformation

by inhibiting β-catenin signaling in an adhesion-independent manner. *J. Cell Biol.* **153,** 1049–1059 (2001).

143. Li, D. *et al.* Integrated biochemical and mechanical signals regulate multifaceted human embryonic stem cell functions. *J. Cell Biol.* **191,** 631–644 (2010).

144. Cruciat, C.-M. & Niehrs, C. Secreted and Transmembrane Wnt Inhibitors and Activators. *Cold Spring Harb. Perspect. Biol.* **5,** a015081 (2013).

145. Funa, N. S. *et al.* β-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell* **16,** 639–652 (2015).

146. Belo, J. A. *et al.* Cerberus-like is a secreted BMP and nodal antagonist not essential for mouse development. *Genesis* **26,** 265–270 (2000).

147. Aykul, S., Ni, W., Mutatu, W. & Martinez-Hackert, E. Human cerberus prevents nodal-receptor binding, inhibits nodal signaling, and suppresses nodal-mediated phenotypes. *PLoS One* **10,** 1–23 (2015).

148. Aykul, S. & Martinez-Hackert, E. New Ligand Binding Function of Human Cerberus and Role of Proteolytic Processing in Regulating Ligand-Receptor Interactions and Antagonist Activity. *J. Mol. Biol.* **428,** 590–602 (2016).

149. Perea-Gomez, A. *et al.* Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Dev. Cell* **3,** 745–756 (2002).

150. Batlle, E. *et al.* The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2,** 84–89 (2000).

151. Cano, A. *et al.* The transcription factor Snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2,** 76–83 (2000).

152. Derynck, R., Muthusamy, B. P., Saeteurn, K. Y. & Biology, T. Signaling pathway cooperation in TGF-β-induced epithelial- mesenchymal transition. *Curr. Opin. Cell Biol.* **31,** 56–66 (2014).

153. Arendt, D. Comparative Aspects of Gastrulation. in *Gastrulation: From Cells to Embryo* (ed. Stern, C.) 679–693 (Cold Spring Harbor Laboratory Press, 2004).

154. Farge, E. Mechanotransduction in Development. in *Current Topics in Developmental Biology* (ed. Labouesse, M.) **95,** 243–265 (Elsevier Academic Press, 2011).

155. Shyer, A. E. *et al.* Emergent cellular self-organization and mechanosensation initiate follicle pattern in the avian skin. *Science* **357,** 811–815 (2017).

156. Kimura-Yoshida, C. *et al.* Canonical Wnt signaling and its antagonist regulate anterior-

posterior axis polarization by guiding cell migration in mouse visceral endoderm. *Dev. Cell* **9,** 639–50 (2005).

157. Peng, G. *et al.* Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo. *Dev. Cell* **36,** 681–697 (2016).

158. Simunovic, M. *et al.* Molecular mechanism of symmetry breaking in a 3D model of a human epiblast. *bioRxiv* 330704 (2018). doi:10.1101/330704

159. Idkowiak, J., Weisheit, G., Plitzner, J. & Viebahn, C. Hypoblast controls mesoderm generation and axial patterning in the gastrulating rabbit embryo. *Dev. Genes Evol.* **214,** 591–605 (2004).

160. D'Amour, K. A. *et al.* Production of pancreatic hormone–expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24,** 1392–1401 (2006).

161. Teo, A. K. K., Valdez, I. A., Dirice, E. & Kulkarni, R. N. Comparable generation of activin-induced definitive endoderm via additive Wnt or BMP signaling in absence of serum. *Stem Cell Reports* **3,** 5–14 (2014).

162. Yoney, A. *et al.* WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids. *Elife* **7,** e38279 (2018).

163. Voiculescu, O., Bodenstein, L., Lau, I. J. & Stern, C. D. Local cell interactions and self-amplifying individual cell ingression drive amniote gastrulation. *Elife* **3,** e01817 (2014).

164. Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M. & Sutherland, A. Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev. Dyn.* **241,** 270–83 (2012).

165. Aulehla, A. & Pourquie, O. Signaling Gradients during Paraxial Mesoderm Development. *Cold Spring Harb. Perspect. Biol.* **2,** (2010).

166. Jaffe, L. F. Calcium waves. *Philos. Trans. R. Soc. B Biol. Sci.* **363,** 1311–1317 (2008).

167. Vergassola, M., Deneke, V. E. & Di Talia, S. Mitotic waves in the early embryogenesis of Drosophila: Bistability traded for speed. *Proc. Natl. Acad. Sci. U. S. A.* **115,** E2165–E2174 (2018).

168. Chang, J. B. & Ferrell Jr, J. E. Mitotic trigger waves and the spatial coordination of the Xenopus cell cycle. *Nature* **500,** 603–607 (2013).

169. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/VCas Systems. *Science (80-.).* **339,** 819–823 (2013).

170. Brinkman, E. K., Chen, T., Amendola, M. & Van Steensel, B. Easy quantitative assessment

of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42,** (2014).

171. E-Cadherin (24E10) Rabbit mAb Datasheet. *Cell Signaling Technology* (2016).

172. E-Cadeherin (EP700Y) Rabbit mAb Datasheet. *Abcam*

173. Nowotschin, S. & Hadjantonakis, A.-K. Use of KikGR a photoconvertible green-to-red fluorescent protein for cell labeling and lineage analysis in ES cells and mouse embryos. *BMC Dev. Biol.* **9,** 49 (2009).

174. Psychoyos, D. & Finnell, R. Assay for Neural Induction in the Chick Embryo. *J. Vis. Exp.* 4–7 (2009). doi:10.3791/1027

175. Vieceli, F. M. *et al.* The transcription factor chicken Scratch2 is expressed in a subset of early postmitotic neural progenitors. *Gene Expr. Patterns* **13,** 189–196 (2013).

176. Fagotto, F. Looking beyond the Wnt pathway for the deep nature of β-catenin. *EMBO Rep.* **14,** 422–33 (2013).

177. Kam, Y. & Quaranta, V. Cadherin-bound β-catenin feeds into the Wnt pathway upon adherens junctions dissociation: Evidence for an intersection between β-catenin pools. *PLoS One* **4,** (2009).

178. Fisher, R. A. The Wave of Advance of Advantageous Genes. *Ann. Eugen.* **7,** 355–369 (1937).

179. Kolmogorov, A. M., Petrovskii, I. G. & Piskunov, N. S. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Sel. Work. AN Kolmogorov* **6,** 1–26 (1991).