2020

Cell Memory and Fate in Human Development

Anna Kresovich Yoney

CELL MEMORY AND FATE IN HUMAN DEVELOPMENT

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Anna Kresovich Yoney
June 2020

CELL MEMORY AND FATE IN HUMAN DEVELOPMENT

Anna Kresovich Yoney, Ph.D.
The Rockefeller University 2020

Development of a single cell, the fertilized egg, into an entire organism is a fascinating example of biological organization that gives rise to many forms of multicellular life. Cell-cell signaling is central to development, particularly in organisms such as vertebrates, where embryonic cells become progressively restricted in their potential through a hierarchy of decision points guided by the signals they receive. All cells contain the same genetic information in the form of DNA, and we now know that the DNA is modified epigenetically to impart distinct functions to each adult cell type. However, we are just beginning to uncover how the epigenome is modified over the course of development and how signaling pathways might direct these modifications.

One of the earliest developmental decision points in bilaterian embryos, which includes vertebrates, occurs during gastrulation when cells are directed into one of three primary germ layers. Model systems, including frog, fish, chicken, and mouse, have been very useful for uncovering the identity and function of the key signaling networks driving vertebrate gastrulation. However, how these signals function together in time to modify cellular behavior has been difficult to tease apart in embryos. With the isolation of human embryonic stem cells (hESCs), many aspects of development can now be reconstituted and manipulated in culture, thus allowing us for the first time to investigate the early stages of our own development. We can now also address how mechanisms proposed from model systems function dynamically or in unanticipated ways.

In this work, I used hESCs and a stem cell-based model of gastrulation (human gastruloids) to disentangle the function and timing of two key signaling pathways, namely Wnt and Activin/Nodal. From studies in model organisms both pathways are known to be required for initiating gastrulation and for the formation of mesoderm and endoderm, but it has been proposed that the level of Activin/Nodal signaling ultimately determines germ layer identity. In order to test these predictions in a model of human gastrulation, I first developed hESC lines with fluorescent reporters of Activin/Nodal signaling that facilitate quantitative measurements of the cellular response in real-time. With these lines I was able to establish that hESCs display a dose-dependent response to Activin ligands at the level of the intracellular effector Smad2. However, this response is transient, and the cells do not adopt germ layer identities even when the pathway response is transiently saturated. By manipulating the relative timing of Wnt and Activin ligands, I demonstrated that hESCs can record Wnt signals without differentiating and that this signaling memory makes cells competent to form mesoderm and endoderm in response to subsequently supplied Activin. My observations do not eliminate the co-requirement for Wnt and Activin in germ layer differentiation, which has been emphasized by previous studies of mouse and human ESCs, but rather they highlight the fact that instructive Wnt signals occur temporally up-stream of Activin signals to induce complete germ layer patterning.

My initial efforts to uncover the mechanism of Wnt signaling memory identified a role for the bromodomain and extra-terminal domain (BET) family of epigenetic readers, which facilitate transcription through interactions with acetylated residues of histones and other proteins and are thought to promote tumor formation in the context of cancer. I showed that inhibition of BET

protein function selectively eliminates mesoderm and endoderm in human gastruloids and produces phenotypes in frog embryos consistent with disruption of early Wnt signaling. The concept that Wnt signaling primes cells to differentiate without directing them to commit to a particular fate defines a new aspect to how this signaling pathway functions. These observations also establish an experimental context to further investigate the link between developmental signaling pathways and the epigenetic landscape.

As I hope to demonstrate throughout this thesis, hESCs are a powerful means to uncover both conserved and potentially human-specific mechanisms of vertebrate development. In the context of Wnt and Activin signaling my results highlight the importance of the temporal interaction between Wnt and Activin signaling in patterning the vertebrate embryo and suggests an evolutionarily conversed mechanism of Wnt signaling memory that could explain how signals are integrated to direct cellular behavior in both development and disease.

This thesis is dedicated to my family, especially to my mother, Veda, my father, Jim, my sister, Karla, and my aunts, Lisa and Diane

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors, Professors Eric Siggia and Ali Brivanlou. I came to graduate school with a plan to continue working in physics and biology, which is how I ended up doing a rotation project with Eric and eventually joining his group. What I did not realize is that I would become a developmental biologist and also a full member of Ali's group through their close collaboration and shared appreciation for the beauty and source of wonderment that is the developing embryo. It was not always easy working with two advisors, especially ones with such different personalities as Ali and Eric, but I learned from both of them invaluable lessons about science and life, especially the power of thinking creatively, without fear and limitations. One of the advantages of being co-advised by Eric and Ali was the opportunity to work with a group of people from diverse scientific backgrounds and perspectives, who turned out to be friends as much as colleagues, including Mijo Simunovic, Jakob Metzger, Fred Etoc, and Francesco Piccolo. I would like to especially thank Fred, who supported me with his ideas and enthusiasm through the most uncertain times of my research. I was very fortunate to have the opportunity to work with Professor Lucy Bai from Penn State, who came to our group for a sabbatical in my final year. Although her sabbatical was interrupted by the COVID-19 pandemic, her efforts have pushed the current research beyond what I was able to present here. I am greatly indebted to Tomo Haremaki, who provided invaluable support for the frog experiments and to Shu Li for her technical support. To other past and present members of the Brivanlou and Siggia groups, including Lauren Gerber, Iain Martyn, Zeeshan Ozair, Szilvia Galgoczi, Tatiane Kanno, Riccardo De Santis, Aryeh Warmflash, Albert Ruzo, Christian Markopoulos, Gist Croft, Alessia Deglincerti, Lauren Pietila, Joanna Krzyspiak, Tien Phan-Everson, Rohan Soman, Tiago Laundos, Ariel Waisman, Caroline Lara, and Adam Marks, thank you for making the lab such an inspiring, unique, and enjoyable place to do science.

In addition to Eric and Ali, I was fortunate to interact with several faculty members at Rockefeller and in the NYC scientific community, who I have come to deeply admire. My committee members, Professors Fred Cross and Vanessa Ruta, were a great source of support and wisdom throughout my Ph.D. At every committee meeting they gave me criticism and guidance, which often prompted me to undertake key experiments in the following year and which was critical to me having made it to this point. I had the pleasure of interacting with Professor Kat Hadjantonakis and her group at Memorial Sloan Kettering, who are dedicated to pushing the boundary of what is possible to observe and study in developing mouse embryos. I would also like to thank Professor Danwei Huangfu from Memorial Sloan Kettering for kindly accepting to serve as my external examiner.

I received immense support on the administrative side from Melanie Lee, Peter Ingrassia, Jean-Marx Santel, Adam Souza, and Michael Heke, who made critical tasks happen daily behind the scenes. I would additionally like to thank Melanie for making the Physics Center such a welcoming place for the younger generations. Thank you to Dean Sid Strickland and Associate Dean Emily Harms, and everyone on the administrative staff, especially Cris Rosario, Kristen Cullen, Stephanie Fernandez, and Marta Delgado, for supporting the entire graduate program and maintaining such a pleasant and friendly environment. I spent a great deal of time in the Bio-Imaging Resource Center, which is headed by Alison North and is and has been supported by wonderful staff members including, Pablo Ariel, Kaye Thomas, Tao Tong, Christina Pyrgaki, and Carlos Rico.

I feel extremely fortunate to have had the opportunity to attend the Marine Biological Laboratory's embryology course at Woods hole, MA in 2018, which was made possible by the financial support of the Marine Biological Laboratory through the Burroughs Wellcome Fund, the David Rockefeller Graduate Program, and my advisors, Eric and Ali. The experience, which was an intense 6-week exploration of experimental embryology set against the backdrop of summer on Cape Cod, turned out to be one of the best times of my personal and scientific life. The history of biology, especially of embryology, permeates so deeply at Woods Hole that it is impossible not to be swept up in it and, as far as I can tell, not to be forever influenced by it. Throughout the course, I was privileged to work alongside so many great people, whom I now consider as friends and colleagues, Laurel Yohe, Melvin Bonilla, Andrea Attardi, Anneke Kakebeen, Marla Tharp, Catherine May, Darcy Mishkind, Boris Tezak, Max Courgeon, Bruno Moretti, Paul Sanchez, Stefania Gutierrez, Ashley Rasys, Jack Allen, Shiri Kult, Jay Smith, Kate Nesbit, Lily Tang, Shinuo Weng, Aastha Garde, Sandra Edwards, Andrew Fraser, and my roommate, Martyna Lukoseviciute. I would like to thank the course directors, Professors Dave Sherwood and Rich Schneider, and the teaching assistants, Amber Rock, Hannah Rosenblatt, and Chris Pineda, who dedicated so much time and energy to make it all possible. I am also very grateful to time and effort of the module instructors and assistants, especially Professors Richard Harland, Claudio Stern, Andrea Streit, Nipam Patel, and then postdoctoral scientist John Young.

I would not have arrived at Rockefeller as prepared as I was for graduate school if it had not been for the training and guidance of my former physics professor and first scientific mentor, Professor Hanna Salman. I worked with Hanna during my undergraduate studies when he was opening his lab in the Department of Physics at the University of Pittsburgh. With his support I received much recognition for my research efforts, including the Halliday-Resnick Award for Excellence in Undergraduate Research and a National Science Foundation Graduate Research Fellowship, which I held as a graduate student. Hanna introduced me to Israel, to New York City, to Rockefeller, and to the beauty of doing experiments. We have had many conversations on science, politics, and life over the past ten years, and I am very grateful for our continued friendship.

I would like to thank my family and friends inside and outside of science, all of who have made my life so enjoyable and filled with love. I would especially like to thank my friend and former partner, Christian Zierhut, for his tremendous love and support over the course of my Ph.D. I believe that he critically read and commented on every document of significance, which I wrote during that time and for which I am greatly indebted. I had the sincere pleasure to get to know New York and Vienna with him over those years.

This thesis was written and defended (through the power of technology) in the midst of the COVID-19 pandemic. It is a strange and surreal time that emphasizes how connected we are as a global society. The pandemic has brought to the forefront many of the grave and interrelated problems underlying our modern society, including climate change, wealth and social inequality, and the inability of governments, under the control of rampant capitalism and dictatorships, to provide fundamental support for their citizens. Our society is deeply indebted to those people, whose work is essential and must continue during this period of great risk. Only time will tell how much loss we will suffer and how we will rebuild as a society.

STATEMENT OF CONTRIBUTIONS

Parts of this thesis were modified from a published manuscript that was co-written with Ali Brivanlou and Eric Siggia. The homology donor constructs for the RUES2-RFP-SMAD1, RUES2-mCit-SMAD2, and RUES2-GFP-SMAD4 cell lines were generated by Albert Ruzo. Christoph Kirst wrote the StemCellTracker software, which was implemented for image analysis. RNA- and ATAC-sequencing data analysis was carried out in collaboration with Tom Carroll and Lucy Bai. Tom contributed Figure 2.9C, and Lucy contributed Figure 3.14. The *Xenopus* experiments were done with help from Tomo Haremaki.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTE ON GENE NAMES AND SYMBOLS

I refer to several species throughout this thesis, each of which has its own conventions for how to write the names and symbols of genes and gene products. Wherever possible I tried to use the correct format for each species. In parts of the text where I refer to gene or gene products generally I use a format, which has some but not too much capitalization and which is closest to the conventions used for mouse. For example, I use *Lefty1* and Lefty1 to refer generally to the left-right determination factor 1 gene and gene product, respectively. When symbols are short acronyms, I use all capital letters as in BMP, which is the symbol for bone morphogenetic protein.

CHAPTER 1. INTRODUCTION

Gastrulation in vertebrates and in all bilaterians, transforms embryonic cells into three primary germ layers (endoderm, mesoderm, and ectoderm), which are lineage-restricted cell states that give rise to distinct tissues in the adult. Gastrulation also establishes the anterior-posterior and dorsal-ventral body axes, marking it as a critical patterning event in the early embryo. The signaling pathways involved in vertebrate gastrulation are found in all extant metazoan phyla, suggesting that they were present in the last common ancestor of all animals and that their evolution facilitated the elaboration of animal body plans (Adamska et al., 2007; Nakanishi, Sogabe, & Degnan, 2014; Nichols, Dirks, Pearse, & King, 2006). In this chapter I review some classical embryological concepts, which provide a framework for thinking about developmental processes. I describe the molecular architecture of two key signaling pathways, namely the transforming growth factor β (TGFβ) and Wnt pathways that direct vertebrate gastrulation. I also describe early patterning events in vertebrate embryos with an emphasis on frog and mouse, which are the model organisms that inspired much of the hypotheses driving my thesis research. Lastly, I present a stem cell-based system that our group developed to model and investigate human gastrulation, termed gastruloids.

1.1 Classical embryological concepts

Throughout much of the last decade, developmental biologists have been particularly fascinated with the mechanisms of development in vertebrate embryos. This fascination stems from the fact that in addition to sharing features that are characteristic of our own embryonic development, vertebrate embryos display a striking ability to regulate embryogenesis through inductive signaling. The most famous and influential demonstration of inductive signaling came from the organizer experiments of Hans Spemann and Hilde Mangold in 1924. They found that transplanted tissue from of one amphibian embryo could generate a complete secondary body axis in a host embryo (Spemann & Mangold, 1924). The organizer tissue provides signals that pattern the host mesoderm and that induce neural tissue from a region of the host that would normally give rise to skin. *Regulative development* and the *inductive signals* that guide it contrast with *mosaic development* in which regions of the embryo are pre-specified, for example by localized *cytoplasmic determinants*, as to what they will become. Despite historical attempts to classify embryos based on their primary mode of development, embryos are neither entirely mosaic nor regulative (Lawrence & Levine, 2006). Nonetheless, vertebrate embryos have been the focus for dissecting the nature of inductive signals.

Inductive signals themselves can be classified based on their role in tissue differentiation, that is, whether they are simply *permissive* for differentiation to occur or whether they are *instructive* with respect to the fate that the responding tissue adopts (Figure 1.1A) (Slack, 1991). Clear examples of permissive signals have yet to be identified in early development, but could include extra-cellular matrix proteins that allow cells to adopt one of many potential adherent cell fates (Gilbert, 2000). Based experiments presented in this thesis, I will argue that Wnt signals could be considered as permissive signals in human embryonic stem cell differentiation (hESCs). Instructive signals, of which there are numerous examples at all stages of development, can be subdivided based on the number of fates that they can induce in cells of the responding tissue. An *appositional signal* induces a single fate, typically in a neighboring region, whereas a *morphogen* induces two or more fates in addition to the default fate, which cells adopt in the absence of signal. The term *morphogen* was originally coined by Alan Turing to describe any

signal that gives rise to form (Turing, 1952). Although Turing used the concept of a morphogen in his reaction-diffusion model of pattern generation, it was through the work of Lewis Wolpert and Francis Crick that the term became synonymous with a diffusing substance (Sharpe, 2019). Wolpert proposed the concept of a morphogen gradient as one way to solve the problem of spatial patterning or regionalization of tissues during development (Figure 1.1B) (Wolpert, 1969). It was not the only solution he proposed to the iconic French Flag Problem, but it was the one that stuck and found some experimental grounding in the following decades (Sharpe, 2019). Despite the evidence for signaling gradients and threshold responses that have emerged, the significance and generality of morphogen gradients in establishing patterning *in vivo* is still an open and heavily debated question (Gurdon & Bourillot, 2001; Harland & Gerhart, 1997; Heasman, 2006; Smith, 2009). With the advent of stem cell culture and new experimental techniques, it may be possible to definitively test long held ideas about how inductive signals, including morphogens, function to pattern embryonic tissues and at least to provide a more dynamic picture of inductive processes.

In addition to defining distinct modes of development and types of inductive signals, embryologists of the last century also defined several related concepts to explain the processes they observed (Slack, 1991). Embryonic cells have a *potency*, which defines the number of cell types they can generate if put into the correct environment. As development proceeds, cell fates are refined through a hierarchy of decision points leading to the formation of the specialized cell types and a reduction of potency. In regulative contexts, cells *commit* to a new state based on the signals they receive, and they take on the appropriate molecular and physiological characteristics of that state through the process of *differentiation*. Cells that are reversibly committed to form a specific structure are *specified*, whereas those that are irreversibly committed are *determined*. The ability of cells to respond to an inductive signal defines their *competency*. Some of these classical concepts, namely *potency* and *differentiation*, have translated directly to stem cell research. With increasing efforts to generate stem cell-based models that recapitulate *in vivo* structures and in combination with new experimental technologies, e.g. single-cell RNA sequencing, ATAC-sequencing, etc., we can investigate and define or modify more of these classical concepts in modern molecular terms. In return, they are useful to guide and contextualize findings from *in vitro* reconstituted systems.

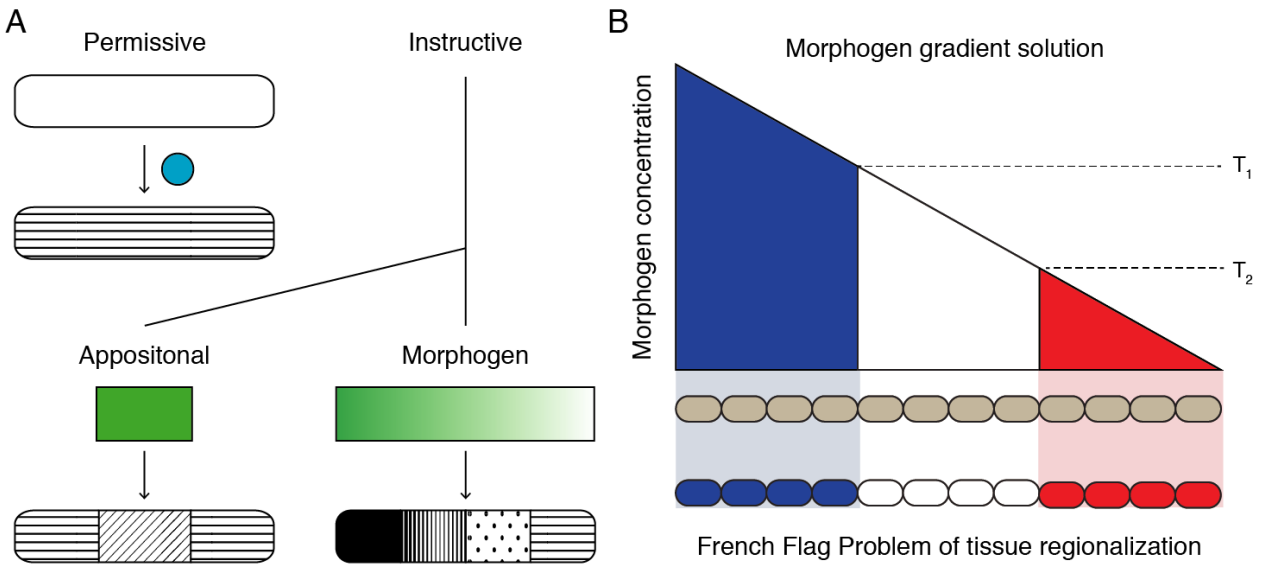


Figure 1.1 Types of inductive signals and their roles in tissue patterning
(A) Permissive signals (blue circle) are required for tissue differentiation, but they cannot influence which fate is selected. Instructive signals direct the responding tissue to adopt a specific fate over other possible alternative fates. Instructive signals can be appositional (solid green box) or distributed in a morphogen gradient (graded green box). Figure and definitions adapted from (Slack, 1991). **(B)** Lewis Wolpert proposed the morphogen gradient model as one solution to the problem of tissue regionalization, which he and Waddington (one year earlier) famously depicted as the French Flag Problem. In this model, cells exposed to high levels of morphogen (concentration > T_1) adopt the blue fate. Cells exposed to low levels (concentration < T_2) adopt the default red fate, and cells exposed to intermediate levels (concentration between T_1 and T_2) adopt the white fate.

1.2 Signaling pathways in gastrulation

Secreted proteins and the signaling pathways they regulate carry the task of instructing dynamic and coordinated cell differentiation during vertebrate gastrulation. Among those secreted proteins are the TGFβ superfamily members, BMP and Activin/Nodal, and Wnt, as well as their respective inhibitors.

1.2.1 TGFβ signaling

TGFβ signaling produces different responses, which are highly dependent on cell type and biological context. In many scenarios, TGFβ elicits two seemingly opposing responses. For example, in human embryonic stem cells (hESCs), TGFβ signaling can maintain pluripotency as well as promote differentiation. Although the basic pathway architecture is relatively simple and was correctly inferred several decades ago, resolving the context-dependent nature of the TGFβ response continues to be a major challenge due to the complexity of the interactions among the canonical signaling effectors and additional nuclear factors (Massagué, 2012).

Secreted proteins of the TGFβ superfamily signal through a complex of type I and type II serine/threonine kinase receptors (Figure 1.2). Upon ligand binding, type II receptors phosphorylate and activate type I receptors, which in turn phosphorylate and activate Smad transcription factors on a C-terminal motif. Receptor-regulated Smads (R-Smads) from a complex with a co-Smad, Smad4, and translocate to the nucleus, where the complex interacts with additional cofactors to activate or repress target gene expression, including the expression of key transcription factors that regulate cellular identity during development. Target genes also include pathway components, such as intra- and extracellular inhibitors, that further tune the response in both a cell autonomous and non-autonomous manner.

The mammalian genome contains over 30 TGFβ superfamily ligands belonging to a number of distinct protein families, which can be divided into two major classes based on the branch of the pathway through which they signal (Massagué, 2012). BMPs signal through R-Smad1, 5, and 8, and TGFβs, Activins, and Nodals signal through R-Smad2 and 3 (Figure 1.2). Each branch has a distinct set of type I receptors that activate the R-Smads and interact with different subsets of ligands. Different type II receptors also interact with different subsets of ligands but they can be shared across both braches of the pathway. Classification of ligands based on their signaling branch does not always reflect the phylogenetic relationships based on sequence, as in the case for Nodals, which are more closely related by sequence to BMPs (Pang, Ryan, Baxevanis, & Martindale, 2011). However, this classification is useful with regards to cellular response, because R-Smads and their binding partners are primarily responsible for the distinct transcriptional responses attributed to each branch of the pathway (Massague, 2005). In addition to secreted activators, each branch of the pathway can be regulated by secreted inhibitors that bind to ligands of one or both braches. TGFβ activator/inhibitor pairs have been proposed to establish tissue patterning through Turing reaction-diffusion mechanisms and through long-range morphogen gradients, which provide positional information (Müller et al., 2012; Smith, 2009).

All R-Smads and Smad4 contain two globular domains, termed MH1 and MH2, separated by a flexible linker region. MH1 functions as a DNA-binding domain. MH2 serves as a versatile hub for mediating interactions with between SMADs and other proteins in the cytoplasm and nucleus. The linker region, which contains multiple phosphorylation sites for intracellular

serine/threonine kinases, provides an additional node for pathway regulation. For example, phosphorylation of linker sites by glycogen synthase kinase 3 (GSK3) leads to the recruitment of E3 ubiquitin ligases and subsequent proteasome-mediated Smad degradation (Figure 1.2). The region of the MH1 domain that directly contacts DNA is conserved at the sequence level in all R-Smads and Smad4 and is, therefore, unlikely to provide much target gene specificity to the two signaling branches (Shi et al., 1998). Thus, the prevailing notion is that the interaction of Smads with cofactors via the MH2 domain provides branch- or other context-dependent transcriptional responses. Inhibitory Smads (I-Smads), Smad6 and 7, lack the MH1 domain but contain the MH2 domain, which interacts with R-Smads and the type I receptors. Through these interactions, I-Smads block R-Smad activation and can induce receptor degradation.

As discussed below, Nodal of which there is a single gene in the mammalian genome is the key activator of Smad2/3 in early mammalian development. Most stem cell protocols rely on commercially available recombinant Activin A rather than Nodal, because of its higher activity and the fact that Activin A can induce developmentally relevant, functional cell types from pluripotent stem cells (D'Amour et al., 2005; Kubo et al., 2004; McLean et al., 2007). In most contexts Activin and Nodal are thought to elicit similar transcriptional responses (Shen, 2007). However, Nodal relies on an additional EGF-CFC family co-receptor, Crypto/Cryptic, to signal, which makes it susceptible to inhibition by the secreted proteins Lefty1 and 2 (Figure 1.2) In my thesis research, I used Activin A, referred to throughout this work as Activin, as the exogenous source of Smad2 and 3 activating ligands, because more active forms of Nodal are not yet commercially available (Fuerer, Nostro, & Constam, 2014).

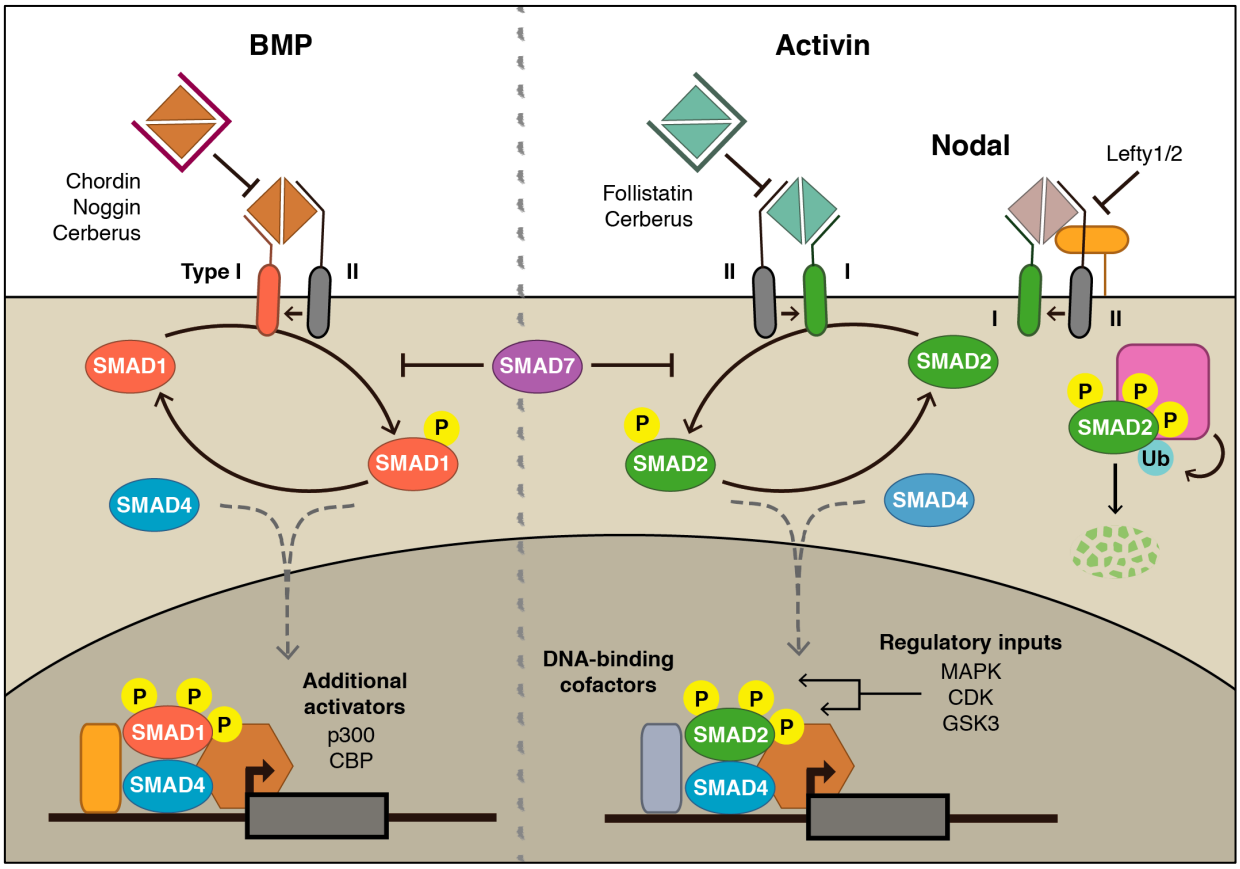


Figure 1.2 Schematic of the TGFβ signaling pathway
BMP and Activin signaling represent the two braches of the TGFβ pathway that are defined by the set of R-Smads, which are activated during signaling, here represented by Smad1 and Smad2. Ligand binding induces phosphorylation of type I receptors by type II receptors. Type I receptors subsequently phosphorylate R-Smads on their C-terminal tail. R-Smads from both branches form a complex with Smad4 and with other cofactors in the nucleus, including cell type-specific transcription factors, histone modifying enzymes, and transcription promoting complexes, to regulate target gene expression. Pathway target genes include among other pathway components I-Smad6 and 7 and secreted inhibitors that block signaling by ligands of one (e.g. Chordin, Noggin, and Follistatin) or both branches of the pathway (e.g. Cerberus). Nodal, which also signals through Smad2, requires an additional co-receptor to assemble the active type I/II receptor complex. The inhibitors Lefty1 and Lefty2 block Nodal activity by binding to Nodal ligands as well as to the co-receptor Crypto/Cryptic (C. Chen & Shen, 2004). Pathway activity of either branch can also be regulated at the level of Smad linker domain phosphorylation, which is carried out by kinases that themselves are regulated by numerous inputs. Linker phosphorylation leads to recruitment of E3 ubiquitin ligases and proteasome-mediated Smad degradation. Linker phosphorylation has also been proposed to regulate Smad subcellular localization and transcriptional activity (Massague, 2005).

1.2.2 Wnt signaling

Wnt pathway regulation is complex and many biochemical aspects are still being elucidated (MacDonald, Tamai, & He, 2009). However, it is clear that Wnt-activated receptors can signal via several intracellular pathways that include the canonical β-catenin-dependent pathway and the noncanonical planar cell polarity (PCP) pathway. The PCP pathway, which establishes cell- and tissue-level polarity in the plane of an epithelium, is critical for developmental processes including the coordination of morphogenetic movements during gastrulation (Butler & Wallingford, 2017). Here, I will focus on the canonical pathway, because of its role in stem cell differentiation and tissue patterning mediated by transcriptional and epigenetic changes, which are central themes of my thesis work.

Canonical Wnt signaling functions by relieving the repression of its intracellular effector, β-catenin (Figure 1.3). In the absence of Wnt ligands ("off" state), cytoplasmic β-catenin is bound by a destruction complex that includes two scaffolding proteins, Axin and adenomatous polyposis coli (APC), and two serine/threonine kinases, casein kinase 1 (CK1) and GSK3. CK1 and GSK3 phosphorylate β-catenin targeting it for recognition by the E3 ubiquitin ligase β-TrCP and subsequent proteasome-mediated degradation. In the "on" state, Wnt ligands interact with seven-pass transmembrane Frizzled receptors and low-density lipoprotein receptor-related protein (LRP) 5 or 6. The bound receptors recruit the scaffolding protein, Dishevelled, which results in recruitment of the destruction complex and phosphorylation of LRP by GSK3 and CK1. Phosphorylated LRP inhibits GSK3 from targeting β-catenin and, thus, results in β-catenin stabilization. Stabilized β-catenin accumulates in the cytoplasm and subsequently moves into the nucleus where it interacts with TCF/LEF transcription factors. In most contexts, this interaction results in the displacement of the co-repressor Groucho and transcriptional activation. In addition to TCF/LEF, β-catenin interacts with other transcription factors, histone modifying enzymes, and transcription promoting complexes to regulate target gene expression.

Wnt signaling can be regulated by secreted inhibitors, such as Dickkopf (Dkk) family proteins and Secreted Frizzled-related proteins (sFRPs), which inhibit Wnt signaling by binding to LRP receptors or to Wnt ligands, respectively. In addition to extracellular regulation, the responsiveness of a cell to Wnt stimulation is tuned by the abundance and activity of the pathway components. For example, Wnt signaling activates *Axin2* expression in most cell types thereby forming a negative feedback loop to dampen the Wnt response. Axin stability and activity is also regulated by post-translational modifications, including phosphorylation, poly(ADP-ribosyl)ation and ubiquitination. The abundance of cytoplasmic β-catenin itself can be modulated through incorporation and release from adherens junctions, where it plays a second critical role as a structural component to link cadherins to the actin cytoskeleton.

Similar to TGFβ family ligands, Wnts and their inhibitors have been proposed to function as Turing activator/inhibitor pairs (Sick, Reinker, Timmer, & Schlake, 2006). However, the view that Wnts act as long-range morphogens is complicated by the fact that they are hydrophobic and cannot freely move about in the extracellular space (Clevers, Loh, & Nusse, 2014). The hydrophobic nature of Wnt ligands is due to a covalent lipid modification, which is added to one or more amino acid residues by the transmembrane palmitoyltransferase Porcupine (Figure 1.3). This lipid modification is added to Wnt ligands in the endoplasmic reticulum and is required for them to be trafficked to the plasma membrane and secreted. Although several mechanisms have

been proposed to explain how Wnts move long distances between cells, including through formation of lipoprotein particles, recent experiments in flies engineered with membrane-tethered Wingless, the main *Drosophila* Wnt, have challenged the concept that Wnts must move at all to pattern developing tissues (Alexandre, Baena-Lopez, & Vincent, 2013). Rather these experiments, along with others in *Xenopus* and those in hESCs presented in the following chapters, suggest that cells possess mechanisms to record developmental signals, which allow them to integrate multiple signaling inputs across space and time even as the availability of ligands changes.

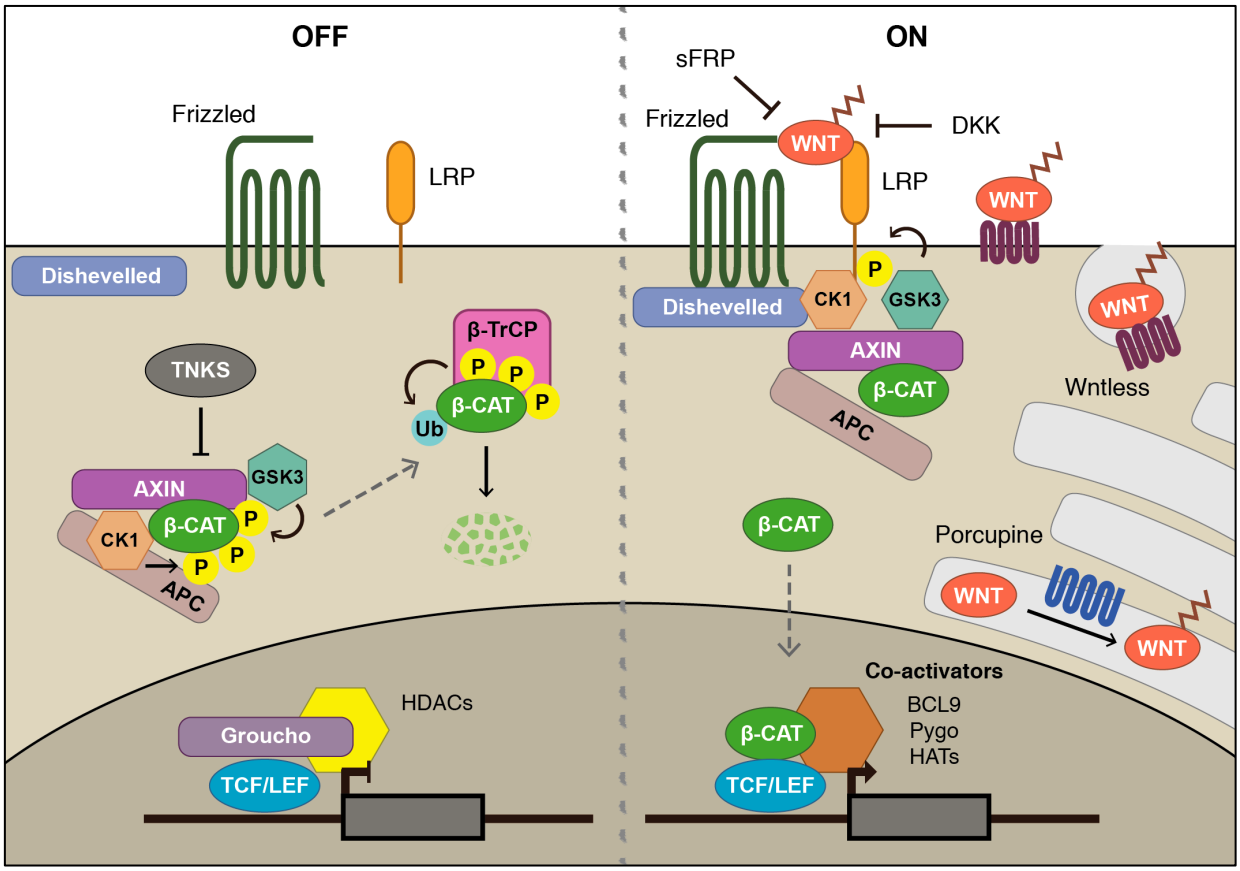


Figure 1.3 Schematic of the canonical Wnt signaling pathway
In the absence of Wnt signals, β-catenin is continually phosphorylated by CK1 and GSK3 ("off" state). Phosphorylation targets β-catenin (β-CAT) for ubiquitination by β-TrCP and proteasome-mediated degradation. GSK3 and CK1 are part of the β-catenin destruction complex, which also includes the scaffolding proteins, Axin and APC. The activity and stability of the destruction complex is regulated by post-translational modifications, including poly(ADP-ribosyl)ation of Axin, which is added by Tankyrase (TNKS). poly(ADP-ribosyl)ation leads to Axin ubiquitination and degradation, thus increasing the sensitivity of cells to Wnt stimulation. In the "on" state, Wnt ligands bind Frizzled and LRP5/6 receptors, which results in the recruitment of Dishevelled, phosphorylation of LRP, and recruitment of the destruction complex. LRP phosphorylation blocks GSK3-mediated phosphorylation of β-catenin. β-catenin accumulates in the cytoplasm and subsequently moves into the nucleus where it interacts with DNA-bound TCF/LEF. Interaction of β-catenin with TCF/LEF displaces Groucho, which represses transcription in part through the recruitment of histone deacetylases (HDACs). Activation of Wnt target genes is dependent on the interaction of β-catenin with other co-factors (e.g. BCL9 and Pygo), histone modifying enzymes (e.g. histone acetyltransferases, HATs), and transcription promoting complexes. Wnts are hydrophobic due to lipid modifications added in the endoplasmic reticulum by the transmembrane palmitoyltransferase Porcupine. These lipid modifications are recognized by a second transmembrane protein, Wntless, which binds Wnts and transports them to the plasma membrane.

1.2.3 Pathway interactions

The TGFβ and Wnt pathways can be regulated by shared kinases, most notably GSK3, which phosphorylates β-catenin and the Smads leading to their degradation. Wnt signaling inhibits GSK3 activity and could, therefore, indirectly modulate TGFβ signaling as well (Demagny, Araki, & De Robertis, 2014). Studies in mouse and human embryonic stem cells have revealed that β-catenin and Smad2/4 complexes co-regulate the expression of key transcription factors that define mesoendodermal lineages (Funa et al., 2015; Singh et al., 2012; Wang et al., 2017). This joint regulation may involve co-binding of these effectors to overlapping sites within regulatory DNA as well as other mechanisms of co-regulation that operate over long genomic distances, e.g. enhancer-promoter looping (Estarás, Benner, & Jones, 2015; Funa et al., 2015). In classical embryological terms the interactions of Wnt and Nodal signaling remain ill defined. For example, whether one signal is permissive versus instructive with respect to mesoderm and endoderm specification and patterning is not known (Figure 1.1A). Likewise, whether there is a temporal order to the function of Wnt and Nodal has not been possible to dissect through traditional genetic manipulations in embryos.

1.3 Gastrulation in vertebrates

During vertebrate gastrulation, embryonic cells are specified into one of the three primary germ layers in a spatially and temporally ordered sequence. *Endoderm* forms the digestive and respiratory tracts and associated organs (e.g. the lungs, liver, and pancreas). *Mesoderm* forms the skeleton, muscles, the circulatory system, and most of the reproductive system, excluding germ cells, which are not derived from the primary germ layers, and *ectoderm* forms the skin, sensory organs, and the nervous system. Amphibian embryos, primarily of the frog *Xenopus laevis*, have been used for decades to study vertebrate development. The mouse embryo is the best-studied model of mammalian development, but the ability to culture embryos outside of the uterus for extended periods of time is still limited. Therefore, most of our knowledge about mouse and, hence, mammalian gastrulation is based on genetic perturbations followed by analysis at fixed time points. These experiments and the classical cutting and pasting experiments in amphibians offered critical identification of key molecules and developmental principles but limited insight into the dynamic processes underlying germ layer patterning.

1.3.1 *Xenopus laevis:* the vertebrate posterchild

Immediately after fertilization the frog embryo contains two axes: the animal-vegetal axis, which is setup in the egg by localized determinants, and the dorsal-ventral axis, which is defined relative to the site of sperm entry and which depends on the re-localization of some of those determinants. The first twelve rounds of cell division occur without growth and with very little zygotic transcription or cell-cell signaling (Heasman, 2006). Following these cleavage divisions the mid-blastula embryo is comprised of a ball of cells surrounding a fluid-filled cavity. At this stage, the embryo can already be divided into three regions based on what fates the cells give rise to during gastrulation. That is, although cells are still pluripotent, they are already regionally specified into the three germ layers (Heasman, Wylie, Hausen, & Smith, 1984). The end of the twelfth cell cycle marks the mid-blastula transition (MBT), at which point global zygotic transcription and cell-cell signaling become active and set the stage for gastrulation.

The first visible sign of gastrulation occurs during the 15th cell cycle with the formation of the blastopore lip, which is the organizer region identified by Spemann and Mangold. The blastopore

forms on the dorsal side of the embryo and marks the site of invagination of cells from the surface to the interior of the embryo. Cells of the vegetable hemisphere, which give rise to the endoderm, move first through the blastopore, followed by the cells of the equatorial or marginal zone, which give rise to mesoderm. Cells of the animal region, which give rise to ectoderm, remain on the surface of the embryo. Surprisingly, many of the secreted molecules identified from the organizer region, turned out to be inhibitors of inductive signaling pathways. These inhibitors generate neural tissue from the overlying ectoderm by protecting it from signals that induce epidermis, mesoderm, and endoderm, a mechanism known as the default model of neural induction that was articulated by Brivanlou and colleagues (Muñoz-Sanjuán & Brivanlou, 2002). BMP signaling on the ventral side induces epidermis. Signals that induce endoderm, primarily those of the TGFβ family that activate Smad2/3, are initially localized to the vegetal region by the action of maternal mRNAs. At later stages, these signals are propagated to the marginal zone to induce and pattern mesodermal fates.

During the late 1960s, Pieter Nieuwkoop began a series of explant experiments that ultimately led to the discovery of Activin as one of those mesoderm inducing signals (Nieuwkoop, 1969; Smith, Price, Van Nimmen, & Huylebroeck, 1990; van den Eijnden-Van Raaij et al., 1990). When presented to mid-blastula animal cap explants, Activin was sufficient to induce different types of mesoderm based on its concentration and was, thus, qualified as a morphogen (Green & Smith, 1990; Wilson & Melton, 1994). When the Activin receptors and the Smad pathway were characterized in the 1990s, it was shown that micro-injection of different amounts of synthetic mRNAs encoding Smad2 into animal caps, recapitulated the dose-dependent effects of Activin, independently confirming that the activation of different thresholds of the pathway was sufficient for patterning mesoderm fates (Shimizu & Gurdon, 1999). Despite ongoing debates about the role of morphogen gradients in embryos, the proposal that Activin signaling functions in a dose-dependent manner to pattern tissues is a powerful one that has propagated throughout studies of vertebrate gastrulation, including those in mouse and mammalian embryonic stem cells.

1.3.2 Getting warmer: moving from amphibians to mammals

Several key differences exist in the early development of amphibians and mammals. In amniotes, which include mammals, birds, and reptiles, zygotic transcription begins after a few cell divisions before any overt morphological changes or cellular differentiation has occurred (Jukam, Shariati, & Skotheim, 2017). Cell divisions are relatively slow and accompanied by growth, and inductive signaling drives early patterning events up to and including gastrulation. Embryonic cells comprise a simple epithelium surrounded by extra-embryonic tissues that line fluid-filled cavities surrounding the developing embryo. Both embryonic and extra-embryonic tissues are involved in inductive interactions. In birds and primates, including humans, the embryonic epithelium, or epiblast, is disc-shaped. In mouse, the epiblast is cup-shaped, and, therefore, a proximal-distal axis exists that does not exist in other amniotes.

Gastrulation is initiated with the formation of the primitive streak on the future posterior side of the epiblast. Cells of the endoderm and mesoderm undergo an epithelial-to-mesenchymal transition (EMT), ingress through the primitive streak, and migrate anteriorly and laterally between the overlying epiblast, which gives rise to ectoderm, and the extra-embryonic endoderm. Prior to gastrulation in the frog embryo territories of prospective endoderm and mesoderm are physically distinct. In the mouse embryo progenitors of these germ layers are in

close proximity to one another, and the mechanisms that guide lineage segregation in this context are still unclear (Arnold & Robertson, 2009). However, it is thought that at least in the anterior primitive streak a bi-potent intermediate exists, which give can rise to both definitive endoderm and axial mesoderm fates. In frogs gastrulation beings around 9 hours post fertilization at 23 °C and is complete by 14 hours. Gastrulation in mouse begins around 6 days post conception and is complete by 7.5 days, and in humans it begins around 13 to 14 days post conception and may last several days (M. A. Hill, 2020). At the conclusion of gastrulation in all vertebrates, the dorsal-ventral and anterior-posterior axes are established and the three germ layers are ordered such that ectoderm lies on the dorsal side, endoderm on the ventral side, and mesoderm in between.

As in frog, patterning of the germ layers in mouse requires the action of inductive signals, and their inhibitors. Inhibitors of Nodal and Wnt signaling are first expressed prior to gastrulation in the extra-embryonic endoderm. These inhibitors promote anterior neural fate in the adjacent epiblast by restricting inductive signals to the future posterior side, which are already expressed in the posterior extra-embryonic endoderm. The formation of the primitive streak in the proximal-posterior region is driven by a feedback loop between Nodal, BMP, and Wnt signaling (Ben-Haim et al., 2006). Nodal signaling from the epiblast activates *BMP4* expression in the overlying extra-embryonic trophectoderm. BMP signaling in turn induces *Wnt3* expression. Nodal signaling maintains its own expression in the epiblast prior to gastrulation through a Smad-FoxH1 auto-regulatory enhancer in the first intron of the *Nodal* gene (Norris, Brennan, Bikoff, & Robertson, 2002). During gastrulation Wnt signaling further activates and maintains *Nodal* expression through a second TCF/LEF-dependent 5' enhancer (Ben-Haim et al., 2006). Mesendodermal derivatives of the anterior primitive streak, which meet the criteria of an organizer, secrete Wnt, Nodal, and BMP inhibitors, which are necessary to maintain the neural identity of the ectoderm in the anterior region and to establish patterning within the primitive streak.

In frog embryos multiple TGFβ signals activate the Smad2/3 branch of the pathway and potentially function to induce and pattern the endoderm and mesoderm. These include Vg1, Activins, and Xenopus nodal-related factors (Xnrs). In early mouse development, Nodal is the primary Smad2/3 activating ligand, and *Nodal* homozygous mutant or *Smad2*;*Smad3* double-homozygous mutant embryos fail to gastrulate (F. L. Conlon et al., 1994; Dunn, Vincent, Oxburgh, Robertson, & Bikoff, 2004). Elegant genetic experiments that reduced Nodal activity by eliminating combinations of *Smad2* and *Smad3* alleles or the *Nodal* 5' enhancer supported the idea that Nodal functions as a morphogen in mouse gastrulation (Dunn et al., 2004; S. D. Vincent, Dunn, Hayashi, Norris, & Robertson, 2003). Removal of signaling effectors does not directly indicate a requirement for Nodal gradients, however, and interactions between Nodal, BMP, and Wnt beyond initiation of the primitive streak, which nonetheless may be critical for patterning, have been difficult to determine through genetic loss-of-function studies. Differentiation happens concomitant with tissue rearrangements, implying an additional source of signal variation in terms of levels and duration as gastrulation proceeds. Thus, many questions as to how patterning is established within the mammalian primitive streak remain unresolved.

1.4 Stem cell-based models of gastrulation

With the isolation of mouse and human ESCs and the derivation of induced pluripotent cells, all of which can be propagated in culture and differentiated into derivatives of the three embryonic germ layers, many aspects of mammalian development can now be reconstituted and manipulated in culture with more control than was possible with traditional model systems (Evans & Kaufman, 1981; Takahashi & Yamanaka, 2006; Thomson et al., 1998). Additionally, hESC-based models can be used to uncover conserved and potentially human-specific features of development.

1.4.1 Micropatterned hESC culture

In standard culture conditions, hESCs grow as monolayer, epithelial colonies. By providing combinations of extrinsic signals cells can be directed to differentiate into mature cell types (e.g. insulin-secreting β cells or dopaminergic neurons). Current differentiation protocols often produce heterogeneous cell populations that lack spatial organization, and on-going research efforts are aimed at eliminating this heterogeneity for clinical applications. In order to model many aspects of development, hESC culture systems that display reproducible, spatial organization of multiple cell types are required. Towards this aim, our group utilizes a platform for culturing hESCs as colonies with a defined two-dimensional geometry in the range of 100 - 1000 μm, which are produced by coating patterned surfaces with extracellular matrix proteins to which cells adhere (Deglincerti et al., 2016; Warmflash, Sorre, Etoc, Siggia, & Brivanlou, 2014). The pre-gastrulation human embryo is a disc-shaped epithelium with a radius of 100 - 200 μm, which continues to expand as gastrulation proceeds (M. A. Hill, 2020). Therefore, micropatterned culture facilitates analysis of hundreds of nominally identical hESCs colonies over biologically relevant size ranges (Figure 1.4A–B).

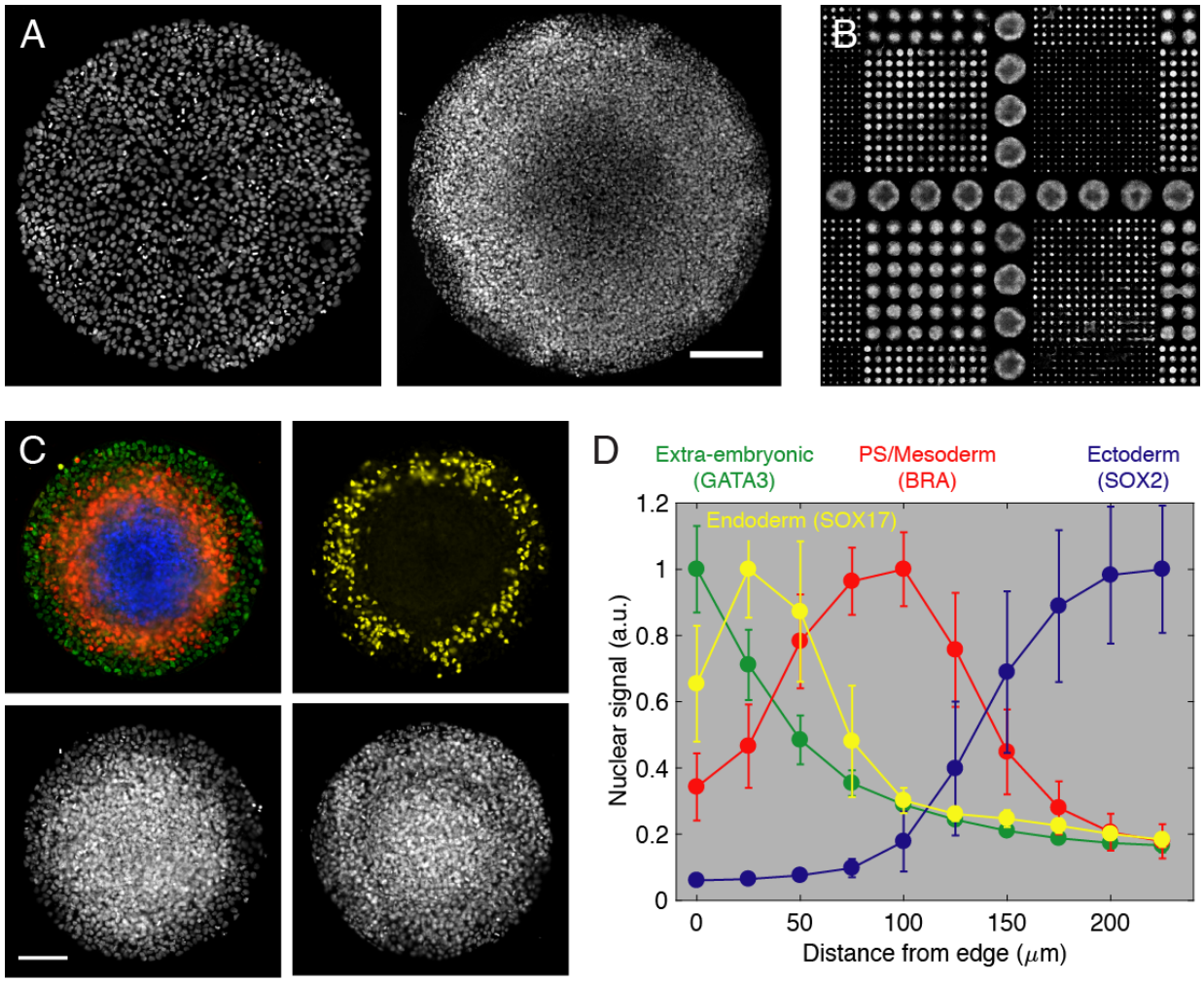


Figure 1.4 Human gastruloids from micropatterned cell culture
(A) Micropatterned culture enforces confinement of colony area, which results in increasing cell density and robust self-organization of the germ layers. At early time points, the colony is a uniform monolayer of cells (left). After stimulation with BMP4 (50 ng/mL, 48 h) the density of cells continually increases and cells pile up in the region of the putative primitive streak (right). Colony radii, 500 µm; scale bar, 200 µm. **(B)** DAPI image of an entire micropatterned coverslip containing hundreds of colonies of various sizes (radii, 40 - 500 µm). **(C)** Following differentiation with BMP4, transcription factors delineating each fate were detected by immunofluorescence (top, left). Extra-embryonic cells express GATA3 (green). Cells within the primitive streak (PS) and mesoderm express Brachyury (BRA, red). Ectoderm cells continue to express SOX2 (blue), but down-regulate NANOG and OCT4. In an independent experiment, SOX17 (yellow), which marks endoderm cells, was also analyzed marker SOX17 (top, right). DAPI, which labels all nuclei, is shown in gray scale for each colony (bottom, left and right). Colony radii, 250 µm; scale bar, 100 µm. **(D)** Colonies are grouped by radius and the average per cell median marker intensity is determined as a function of distance from the colony edge. Error bars represent the standard deviation over colonies on a single coverslip.

1.4.2 Human gastruloids

We initially used the micropatterned platform to address to what extent the signaling pathways directing gastrulation in other vertebrates may also be involved in human gastrulation. Using culture medium previously established to maintain pluripotency under standard conditions, hESCs grown on micropatterned substrates likewise maintain pluripotency as indicated by the co-expression of the transcription factors OCT4, NANOG, and SOX2 (Boyer et al., 2005). In response to BMP4, we have previously shown that micropatterned hESCs self-organize to induce and pattern embryonic and extra-embryonic germ layers (Warmflash et al., 2014). Ectoderm is specified at the colony center, extra-embryonic tissue at the edge, and mesoderm and endoderm in between, thus, reflecting the dorsal-ventral ordering of the germ layers *in vivo* (Figure 1.4C–D). We also observed molecular and morphological signatures of the primitive streak, such as the expression of EMT markers and migration of cells underneath the epithelial layer (Martyn, Siggia, & Brivanlou, 2019b; Warmflash et al., 2014). Consistent with the observations in mouse, BMP signaling induces *WNT3* expression, and Wnt signaling induces *NODAL* expression (Martyn, Kanno, Ruzo, Siggia, & Brivanlou, 2018). Application of Wnt ligands (in the form of the commercially available mouse Wnt3a) can also induce the emergence of a primitive streak and the organization of embryonic germ layers in micropatterned hESC colonies (Martyn et al., 2018; Martyn, Brivanlou, & Siggia, 2019a). Therefore, we termed this model a *human gastruloid* as it recapitulates many of the early events of mammalian gastrulation.

The third pathway, namely Activin/Nodal signaling, which makes up the signaling feedback loop in mouse, has an intriguing and seemingly contradictory role in hESCs. Activation of the Smad2/3 branch of the TGFβ pathway is required for hESC pluripotency maintenance (James, Levine, Besser, & Hemmati-Brivanlou, 2005; Vallier, Alexander, & Pedersen, 2005). However, as discussed in the previous sections, signaling through Smad2/3 is also necessary for fate specification and patterning of the vertebrate embryo. Consistent with this latter role, the Activin/Nodal pathway is necessary for self-organization in human gastruloids as inhibition of signaling blocks induction of embryonic mesoderm and endoderm by BMP and Wnt (Figure 1.5A) (Martyn et al., 2018; Warmflash et al., 2014). However, consistent with its role in pluripotency, Activin stimulation alone is not sufficient to induce stable differentiation of micropatterned colonies (Figure 1.5B) (Yoney et al., 2018). Despite these findings, the mechanisms by which Activin/Nodal signaling function in mammalian gastrulation remain perplexing, and a number of key questions are still unresolved. Among them are: how can a single pathway carry two opposing functions before and after the onset of gastrulation? Is the remainder of the transcriptional feedback loop identified in mouse conserved in human gastruloids, i.e. does Activin/Nodal signaling induce *BPM4* expression to close the loop? In addition to the transcriptional connections among the pathways, are there combinatorial signaling inputs that instruct cell fate decisions or does patterning rely primarily on the morphogen activity of Activin/Nodal?

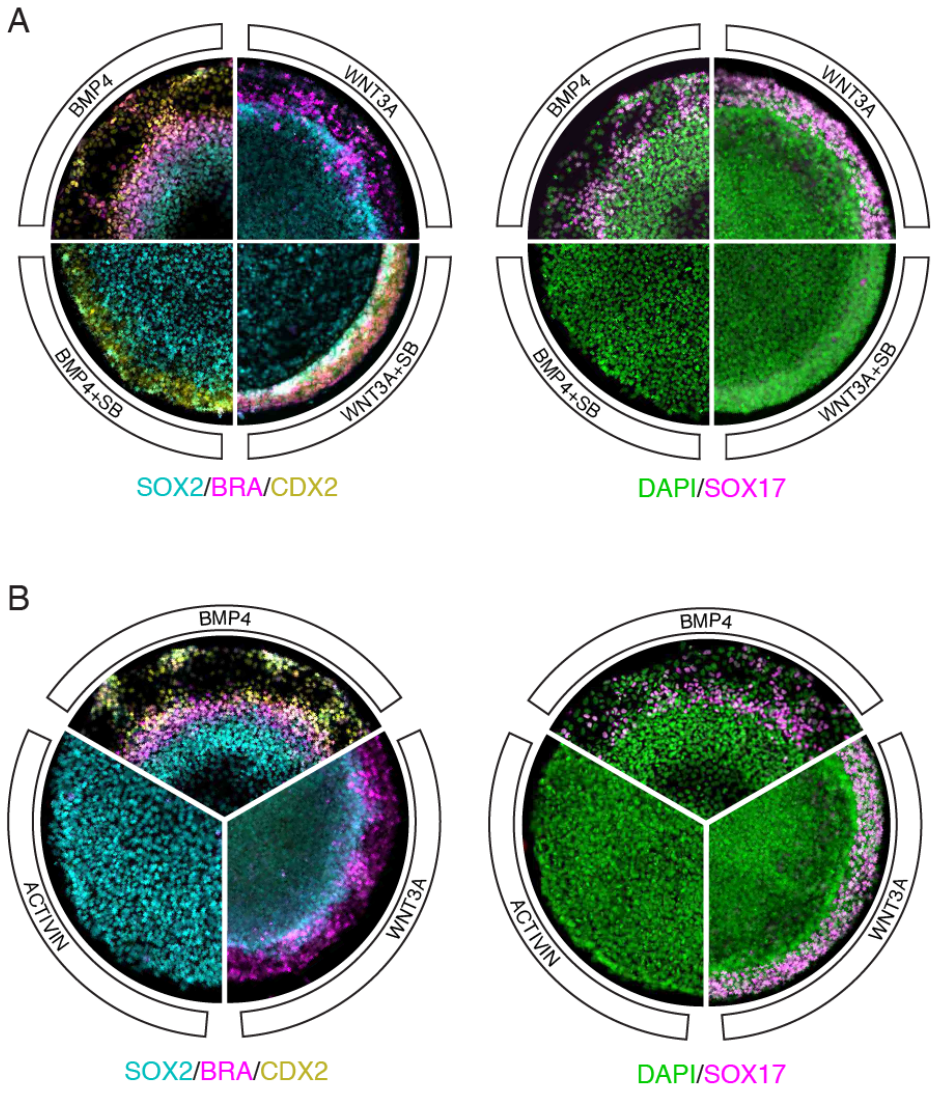


Figure 1.5 Activin modifies gastruloid patterning but it cannot induce it
(A) WNT3A-stimulated colonies lack the extra-embryonic layer, marked by CDX2 (yellow, left), which is induced by high BMP signaling. Instead endoderm, marked by SOX17 (magenta, right), is induced at the outermost edge. The primitive streak and mesoderm, which express BRA (magenta, left), separate the endoderm from the ectoderm layer, which expresses SOX2 (cyan, left). The Activin/Nodal inhibitor SB-431542 (SB, 10 μM) eliminates embryonic endoderm and mesoderm, as cells in the WNT3A+SB condition that continue to express low levels of BRA also express CDX2, which marks them as extra-embryonic mesoderm (Martyn, Siggia, & Brivanlou, 2019b). **(B)** Human gastruloids stimulated with BMP4 or WNT3A generate radially organized germ layers, whereas colonies stimulated with Activin remain pluripotent and continue to express SOX2 as well as OCT4 and NANOG. Colonies were analyzed by immunofluorescence 48 h after stimulation in (A) and (B). DAPI (green, right). BMP4 (50 ng/mL); WNT3A (200 ng/mL); Activin (100 ng/mL).

1.4.3 Mechanisms of gastruloid patterning

As expected from vertebrate model systems, human gastruloid patterning depends on the induction of BMP, Wnt, and Activin/Nodal inhibitors (Etoc et al., 2016; Martyn, Brivanlou, & Siggia, 2019a; Warmflash et al., 2014). Interestingly, the BMP inhibitor, Noggin, is a direct target of BMP signaling in hESCs but not in the equivalent mouse epiblast-like stem cells (Etoc et al., 2016). Mouse embryos that are double-homozygous mutants for *Noggin* and a second BMP inhibitor, *Chordin*, show severe anterior defects, including loss of head structures (Bachiller et al., 2000). However, neither of the single homozygous mutants display early stage defects, suggesting that Noggin and Chordin have redundant activities during mouse gastrulation. In contrast, removal of *noggin (NOG)* alone is sufficient to eliminate the ectoderm territory of human gastruloids (Etoc et al., 2016). Thus, inhibitors of inductive signaling protect anterior fates in both vertebrate embryos and human gastruloids, but their transcriptional regulation and functional redundancy appear to differ between species.

In addition to the inhibitor-based mechanism motivated by embryology, we identified a second mechanism, which restricts the early response to BMP or Activin to the colony edge independent of inhibitor production. This phenomenon results from the cell density-dependent localization of the TGFβ receptors to the basal-lateral surface of cells, i.e. below tight junctions, where they are shielded from ligands in the overlying medium (Etoc et al., 2016). Cells at the colony edge that have an exposed lateral surface or can possibly extend basal-lateral projections respond to BMP and Activin at all densities (Figure 1.6). Although Wnt stimulation has a edge-dependent response, as well, this results primarily from spatial differences in the incorporation of β-catenin into adherens junctions and less from the differential localization of Wnt receptors (Martyn, Brivanlou, & Siggia, 2019a). Basal-lateral BMP receptor restriction was recently found to play a role in patterning the early mouse embryo, demonstrating the *in vivo* relevance of this geometry-based mechanism (Z. Zhang, Zwick, Loew, Grimley, & Ramanathan, 2019), and both receptor restriction and inhibitor production are critical to robustly pattern the human gastruloid (Etoc et al., 2016).

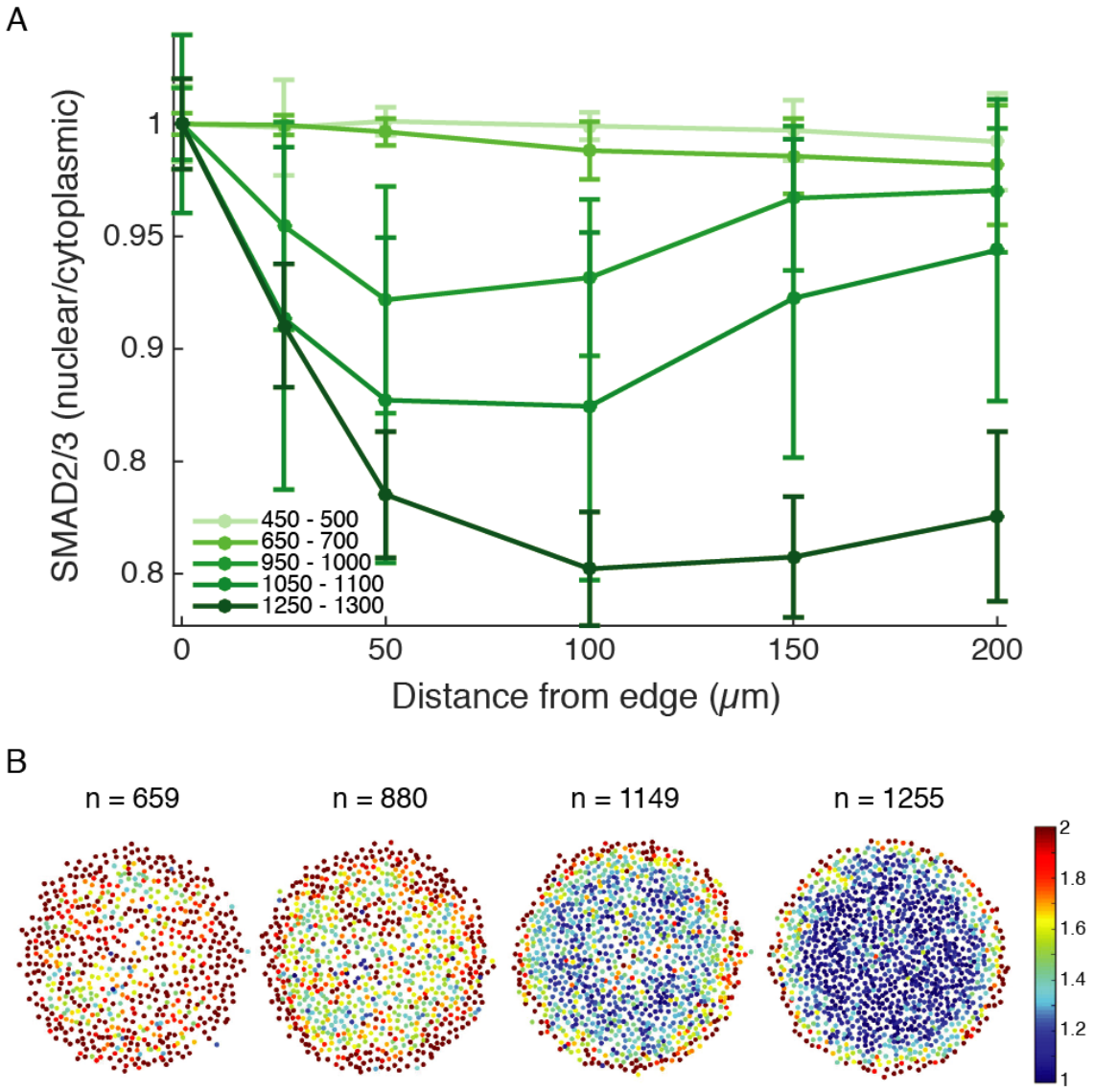


Figure 1.6 Density-dependent restriction of the Activin response in hESC colonies
(A) Cells in low-density colonies respond to Activin independent of their radial position. At high density only cells near the colony border respond. The response to Activin (5 ng/mL) was quantified by the nuclear-to-cytoplasmic ratio of SMAD2/3 immunofluorescence at 1 h after ligand presentation. The ratio, which has been shown to be an accurate readout for pathway activity, was further normalized to the value at the colony edge, because nuclear staining intensity and hence the ratio varies with cell density independent of stimulation. The average ratio as a function of distance from the edge is shown for colonies of radius 250 μm grouped by cell number. Error bars represent the standard deviation. **(B)** Sample colonies from the groups shown in (A). Color scale represents the SMAD2/3 nuclear-to-cytoplasmic ratio without the additional normalization. The number of cells (*n*) identified in each colony by automated image analysis is indicated.

CHAPTER 2. TGFβ SIGNALING IN PLURIPOTENCY

Simulation of micropatterned hESC colonies with exogenous BMP and Wnt ligands proved to an informative approach to dissecting the role of these pathways in germ layer patterning and led to the establishment of the human gastruloid model (Etoc et al., 2016; Martyn et al., 2018; Warmflash et al., 2014). However, in the case of Activin/Nodal signaling, pathway stimulation failed to induce patterning, despite the fact that it is key to initiating primitive streak formation in the mouse. The pluripotent status of hESCs is dependent on SMAD2/3 activation, which is maintained by all standard culture media (James et al., 2005; Singh et al., 2012; Vallier et al., 2005). However, I observed that cells display increased levels of nuclear SMAD2/3 when further stimulated with Activin, leading to an apparent discrepancy between this response and the absence of a change in fate. In this chapter, I describe our efforts to reconcile this discrepancy, which necessitated the development of more quantitative reporters of TGFβ signaling. Through imaging-based studies of hESC reporter lines engineered by CRISPR/Cas9-mediated tagging of SMAD1, SMAD2, or SMAD4, we demonstrated that each branch of the TGFβ pathway has distinct signaling dynamics. In response to Activin, SMAD2 displays a dramatic but transient nuclear accumulation, which stands in sharp contrast to the stable BMP4-induced SMAD1 response. Activin stimulation induces transient meso- and endodermal (referred to as mesendodermal) gene transcription, which correlates with SMAD2 dynamics. This induction, however, is not sustained and cells revert back to pluripotency at later times. These findings support a revision of the notion that the morphogen activity of Activin/Nodal is sufficient to pattern germ layer derivatives from hESCs and instead highlight the importance of a combination of inductive signals in this process.

2.1 Signaling dynamics

It has been proposed that different pathway activation dynamics can lead to different cellular outcomes, a classic but currently-studied example of this being receptor tyrosine kinases that signal through MAPK/ERK to regulate diverse cellular processes from yeast to mammals (P. Conlon, Gelin-Licht, Ganesan, Zhang, & Levchenko, 2016; la Cova, Townley, Regot, & Greenwald, 2017; Marshall, 1995). We initially set out to characterize TGFβ signaling dynamics in hESCs, and specifically to determine how the increase in nuclear SMAD2/3 in response to Activin evolves in time.

2.1.1 Generation of R-SMAD reporter hESC lines

Standard immunofluorescence techniques are inefficient for making highly quantitative measurements with high temporal resolution. Therefore, we used CRISPR/Cas9 homology donor-mediated genome engineering on RUES2, a well-characterized hESC line previously derived by the Brivanlou group, to fluorescently tag the N-terminus of the endogenous receptor-associated, R-SMAD1, with tagRFP (RUES2-RFP-SMAD1) and R-SMAD2 with mCitrine (RUES2-mCit-SMAD2) (Figure 2.1 and Figure 2.2). As activation of the pathway leads to the binding of SMAD1 and 2 to the co-SMAD, SMAD4, we also analyzed a GFP-tagged SMAD4 line (RUES2-GFP-SMAD4) (Nemashkalo, Ruzo, Heemskerk, & Warmflash, 2017). Each line was additionally transfected with ePiggyBac transposable elements carrying a nuclear marker (H2B-mCitrine or H2B-mCherry) in order to analyze the response of individual cells. N-terminal SMAD fusion proteins were shown to function similarly to endogenous proteins in biochemical and cell-based assays (Schmierer & Hill, 2005). Additionally, the SMAD response dynamics

measured with our reporter lines, matches the behavior of the endogenous proteins as measured by immunofluorescence and western blotting, discussed in more detail in the following sections.

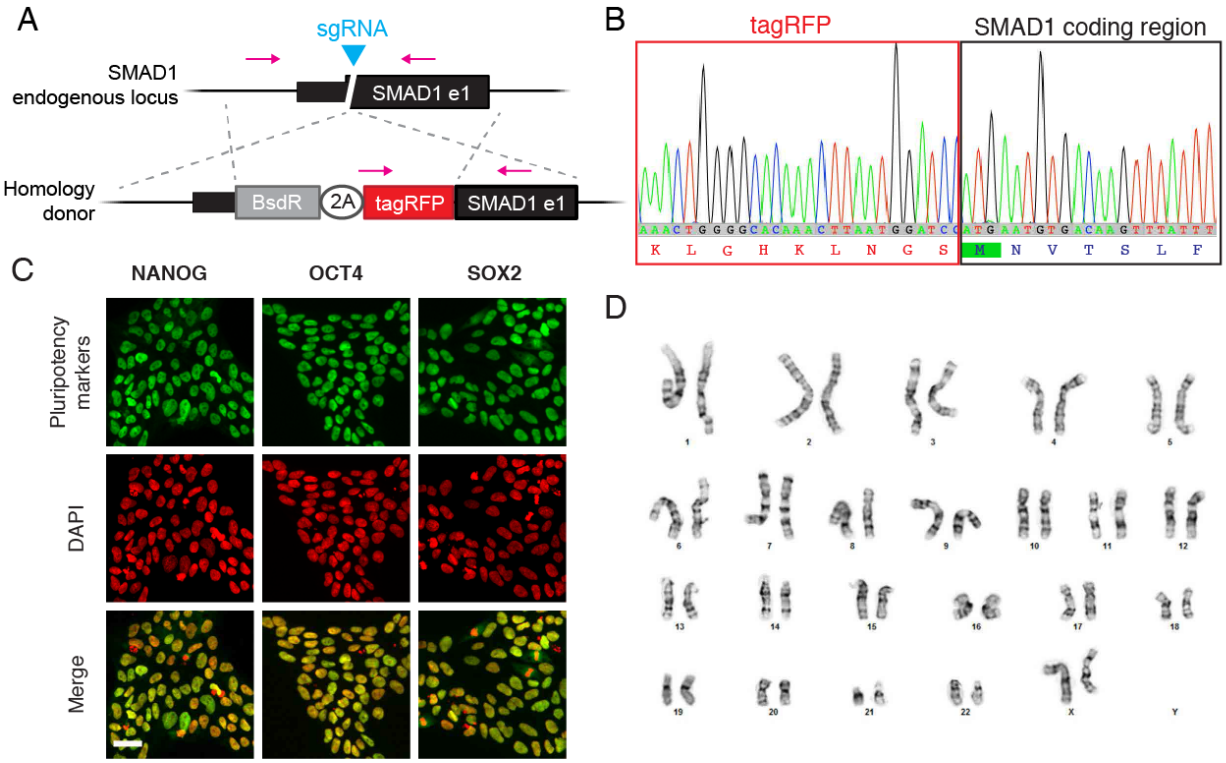


Figure 2.1 RUES2-RFP-SMAD1 reporter line generation
(A) Schematic of the modification of the endogenous SMAD1 locus via CRISPR/Cas9 genome engineering. A blasticidin resistance gene (BsdR) is separated from the tagRFP-SMAD1 fusion by a T2A self-cleaving peptide. The SMAD1 coding sequence (e1) and the upstream untranslated regions are not affected by the modification. sgRNA protospacer sequence: 5'-GCAGCACTAGTTATACTCCT-3'. **(B)** Sequence analysis using loci specific primers (magenta arrows in A) demonstrated that both of the SMAD1 alleles had been modified. The trace shows the modified allele sequence. **(C)** RUES2-RFP-SMAD1 cells maintain expression of pluripotency markers under standard growth conditions. Separate samples were prepared for each marker and immunofluorescence staining was carried out using Alexa Fluor 647-conjgated secondary antibodies to avoid signal from the reporter constructs: NANOG, OCT4, and SOX2 (green) and DAPI (red). Scale bar, 50 μm. **(D)** RUES2-RFP-SMAD1 has a normal XX karyotype.

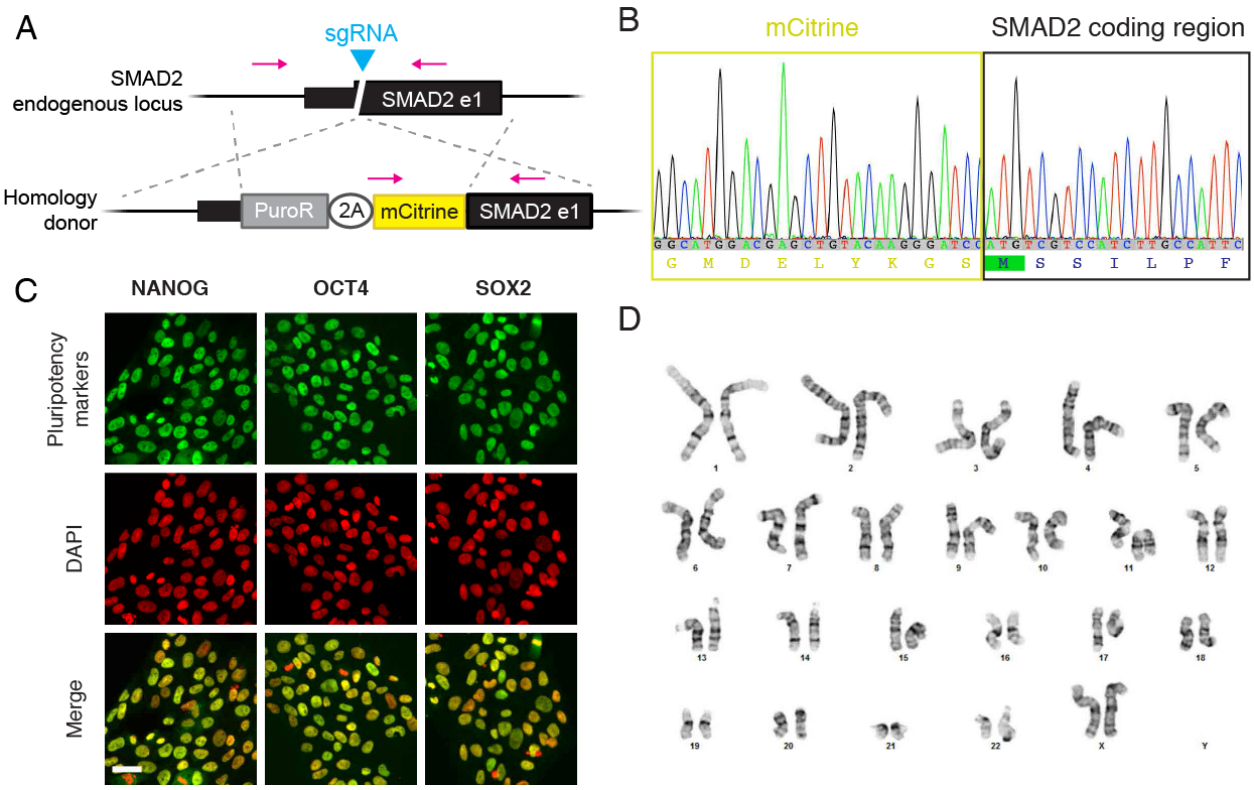


Figure 2.2 RUES2-mCit-SMAD2 reporter line generation
(A) Schematic of the modification of the endogenous SMAD2 locus via CRISPR/Cas9 genome engineering. A puromycin resistance gene (PuroR) is separated from the mCitrine-SMAD2 fusion by a T2A self-cleaving peptide. The SMAD2 coding sequence (e1) and the upstream untranslated regions are not affected by the modification. sgRNA protospacer sequence: 5'-GGACGACATGTTCTTACCAA-3'. **(B)** Sequence analysis using loci specific primers (magenta arrows in A) demonstrated that only one of the SMAD2 alleles had been targeted. The trace shows the modified allele sequence. **(C)** RUES2-mCit-SMAD2 cells maintain expression of pluripotency markers under standard growth conditions. Separate samples were prepared for each marker and immunofluorescence staining was carried out using Alexa Fluor 647-conjgated secondary antibodies to avoid signal from the reporter constructs: NANOG, OCT4, and SOX2 (green) and DAPI (red). Scale bar, 50 µm. **(D)** RUES2-mCit-SMAD2 has a normal XX karyotype.

2.1.2 Analysis of R-SMAD reporter lines on micropatterned colonies

In our previous studies we used mouse embryonic fibroblast conditioned medium (MEF-CM), which was one of the first media developed to maintain hESCs. However, it contains high levels of serum and poorly defined levels of SMAD2/3 activating ligands, both of which can complicate experimental interpretation. Therefore, we analyzed TGFβ signaling in RUES2 cells grown on micropatterned colonies in chemically-defined TeSR-E7 (E7) medium, which is a version of serum-free E8 medium that lacks any TGFβ activity (G. Chen et al., 2011). In response to BMP4, we detected an increase in the SMAD1 nuclear signal that was stable over 12 hours (Figure 2.3A). The SMAD1 response was observed at the colony edge, consistent with our previous immunofluorescence results and our discovery that the TGFβ receptors are exposed to added ligands only at the border of dense colonies (Etoc et al., 2016; Warmflash et al., 2014). However, in contrast to our previous results in which we could detect a response as early as 1 hour after BMP4 presentation, RFP-SMAD1 only shows a measureable nuclear accumulation after several hours. This difference most likely results from a higher sensitivity of immunofluorescence over the endogenous reporter in this context. In low-density cell culture, the results of which are presented below, we are able to detect a rapid response to BMP, suggesting that the reporter can accumulate on similarly fast timescales. In response to Activin, SMAD2 responded only transiently: a pulse of nuclear translocation over the first 1-2 hours was followed by a gradual decrease over the next 6 hours (Figure 2.3B). However, the long-term SMAD2 nuclear level did not return completely to the pre-stimulus level. As in the case of SMAD1, the SMAD2 response was highest at the colony edge, consistent with our immunofluorescence data (Figure 1.6)

According to the textbook picture, the R-SMADs localize to the cytoplasm in the absence of signaling. In the case of the RFP-SMAD1 reporter line a faint nuclear signal was detected prior to ligand presentation, but no cytoplasmic signal was measured above the background (Figure 2.4A–B). However, in the case of the mCitrine-SMAD2 reporter line an enrichment of the mCitrine signal in the cytoplasm prior to ligand presentation was detected, which decreased concomitantly with an increase in the nuclear signal following Activin presentation (Figure 2.4C–D). The difference in the cytoplasmic accumulation of the reporters could result from differences in the rate of R-Smad degradation in the unstimulated state, as it has been reported that Smad1 but not Smad2 can be targeted for ubiquitination and degradation in the absence of signaling by the E3 ubiquitin ligase, Smurf1 (H. Zhu, Kavsak, Abdollah, Wrana, & Thomsen, 1999). Because we could not detect a cytoplasmic signal, the RFP-SMAD1 response was quantified in individual cells as the median nuclear signal divided by the median H2B signal, which normalizes for intensity changes that result from focal drift (Figure 2.3C). The mCitrine-SMAD2 response was quantified in individual cells as the nuclear-to-cytoplasmic ratio, which was previously shown to be an accurate readout for pathway activity (Figure 2.3D) (Warmflash et al., 2012). This method of normalization conveniently corrects for focal drift and bleaching of the reporter fluorescence, which would affect estimates of the signal adaptation.

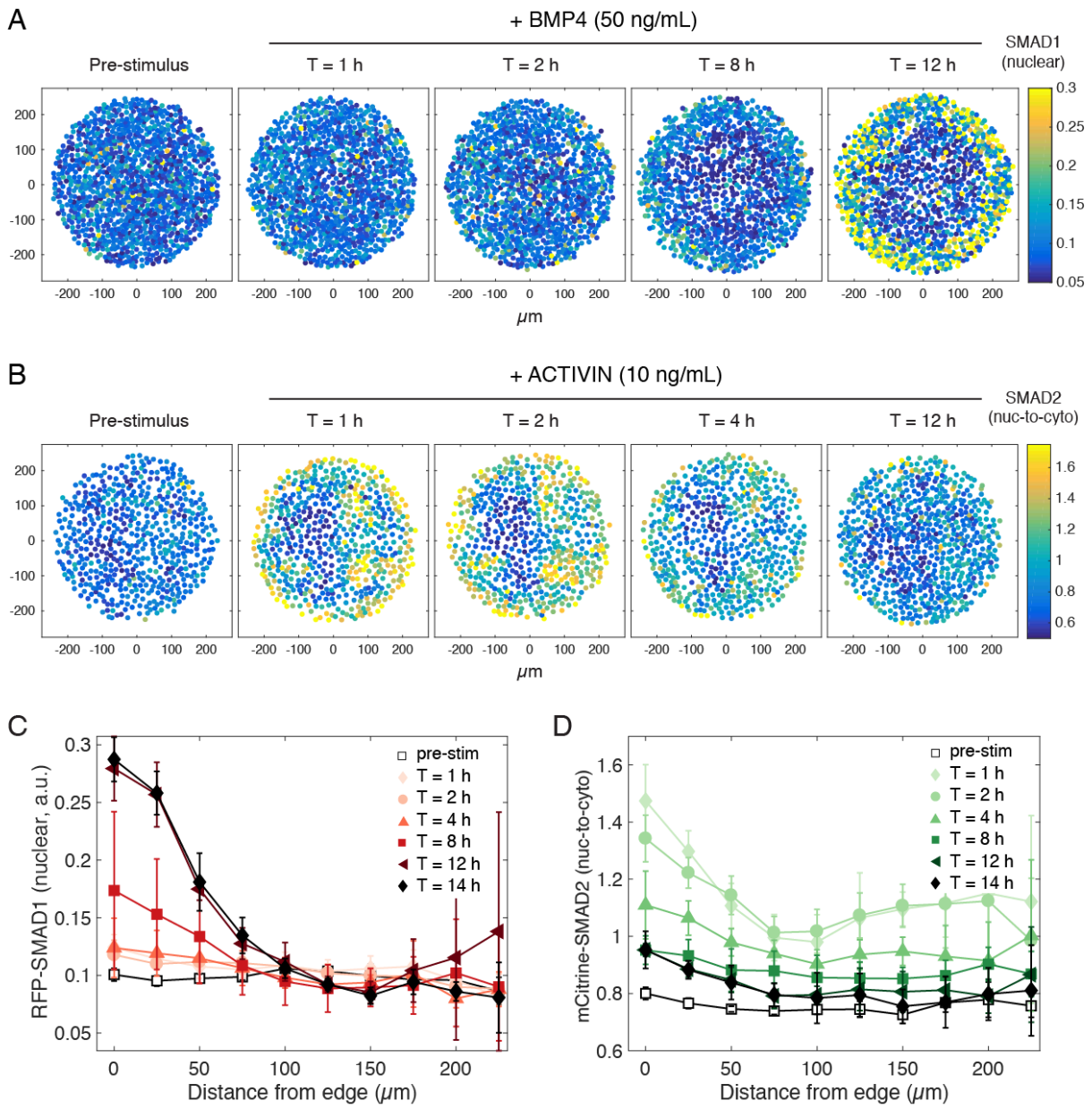


Figure 2.3 BMP and Activin signal with different temporal dynamics in hESCs
Micropatterned colonies with **(A)** RUES2-RFP-SMAD1 in E7 stimulated with BMP4 (50 ng/mL) and **(B)** RUES2-mCit-SMAD2 in E7 stimulated with Activin (10 ng/mL). **(C)** The average RFP-SMAD1 nuclear signal was analyzed as a function of radial position at different time points following BMP4 stimulation. The nuclear RFP-SMAD1 signal analyzed in individual cells was normalized to the median H2B-mCitrine signal. **(D)** The average mCitrine-SMAD2 nuclear-to-cytoplasmic ratio was analyzed as a function of radial position within the colony at different time points following Activin stimulation. The single-cell nuclear mCitrine-SMAD2 single was normalized to the single-cell cytoplasmic signal. Error bars in (C) and (D) represent the standard deviation over colonies from one experiment (n = 5 colonies).

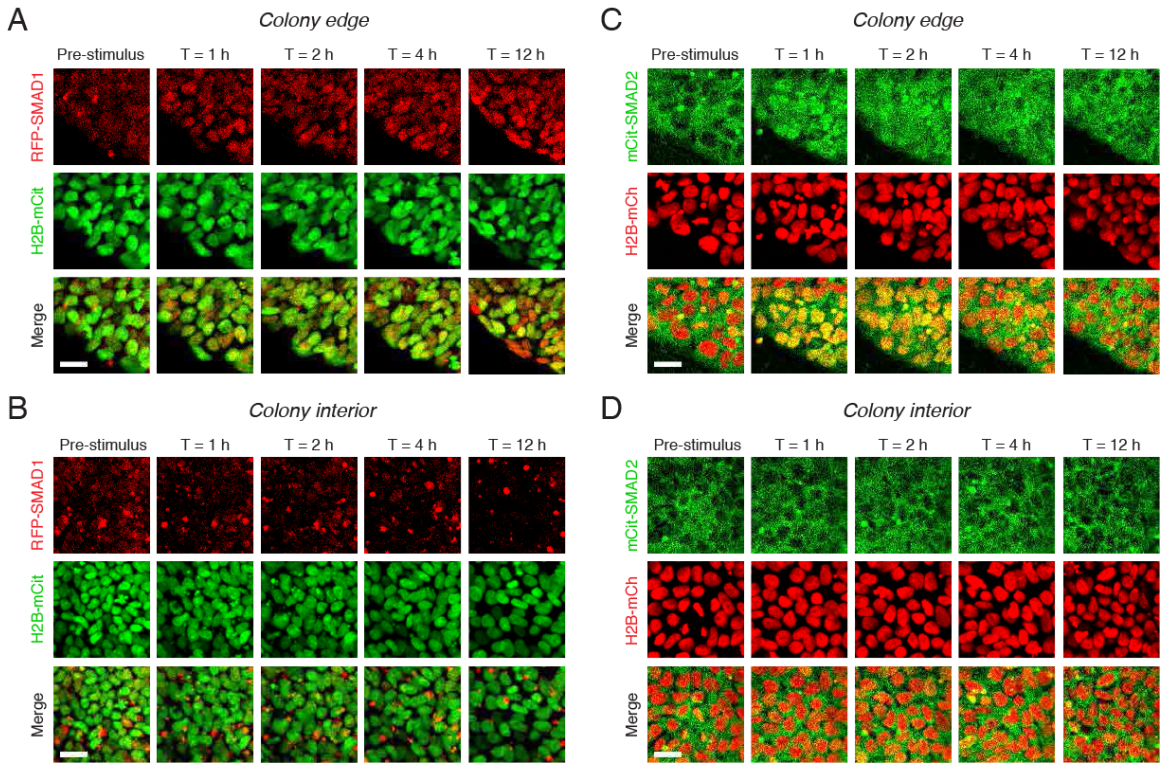


Figure 2.4 R-SMAD reporter lines grown on micropatterned colonies
RUES2-RFP-SMAD1 grown on micropatterned colonies in E7 and stimulated with BMP4 (50 ng/mL). Images show the response of cells **(A)** at the colony edge and **(B)** in the colony center as a function of time following BMP4 presentation. The intensity range was adjusted to the same minimum and maximum values in all images in both (A) and (B). The bright RFP puncta in (B) do not completely overlap with nuclei, so we did not consider that as true SMAD1 signal. Scale bars, 25 μm. **(C)** and **(D)** Same as in (A) and (B) for RUES2-mCit-SMAD2 in E7 stimulated with Activin (10 ng/mL).

2.1.3 Analysis of SMAD reporter lines in low-density culture

In order to determine the cell-autonomous TGFβ responses, we performed similar experiments on dissociated reporter cells. As observed at the edge of micropatterned colonies, the SMAD1 response to BMP was stable (Figure 2.5A–B), whereas the SMAD2 response to Activin was transient with near-complete adaptation to pre-stimulus levels (Figure 2.5C–D). SMAD4 followed the dynamics of the corresponding R-SMADs (Figure 2.5E–F). The stable and adaptive response of SMAD4 to BMP and Activin, respectively, is consistent with recent experiments that also utilized this reporter line (Heemskerk et al., 2019; Nemashkalo et al., 2017). The population average response curves presented in Figure 2.5 result from individual cells responding with similar dynamics, which is evident in the population-level histograms and in individual cell traces (Figure 2.6). Thus, in response to their activating ligands, the two branches of the TGFβ pathway display distinct dynamics of signal transduction at the level of both the R-SMAD and the co-SMAD. This demonstrates that modifying signals do not influence the dynamics of the initial response in micropatterned colonies, apart from the receptor localization, which restricts the response spatially.

Figure 2.5 Analysis of the SMAD response in low-density cell culture
Reporter cells were plated at low-density to analyze the cell-autonomous response. **(A)** RFP-SMAD1 response to BMP4 (5 ng/mL). **(B)** Quantification of the RFP-SMAD1 nuclear signal as a function of time and BMP4 concentration. The solid lines represent the average response for BMP4 0 ng/mL (light red), 0.5 ng/mL (medium red), and 5 ng/mL (dark red), and the dashed lines represent the standard deviation for –BMP4 (0 ng/mL, light red) and BMP4 (5 ng/mL, dark red) conditions. Similar results were obtained in two independent experiments. **(C)** mCitrine-SMAD2 response to Activin (10 ng/mL). **(D)** Quantification of the mCitrine-SMAD2 nuclear-to-cytoplasmic ratio as a function of time and Activin concentration. The solid lines represent the average response for Activin 0 ng/mL - 10 ng/mL, with darker shades of green indicating higher concentrations. The dashed lines represent the population standard deviation for –Activin (0 ng/mL, lightest green) and Activin (10 ng/mL, darkest green) conditions. Data were collectively obtained from three independent experiments. **(E)** GFP-SMAD4 response to Activin (10 ng/mL), BMP4 (10 ng/mL), or cells that were left unstimulated (E7). **(F)** Quantification of the GFP-SMAD4 nuclear-to-cytoplasmic ratio as a function of time in E7 (blue), E7 + BMP4 (10 ng/mL, red), or E7 + ACTIVIN (10 ng/mL, green). Solid lines represent the average response at each time point. In all experiments, ligands were added at T = 0 h, and images were acquired every 10 minutes. The average response was calculated for n > 200 cells per time point. Scale bars, 25 μm.

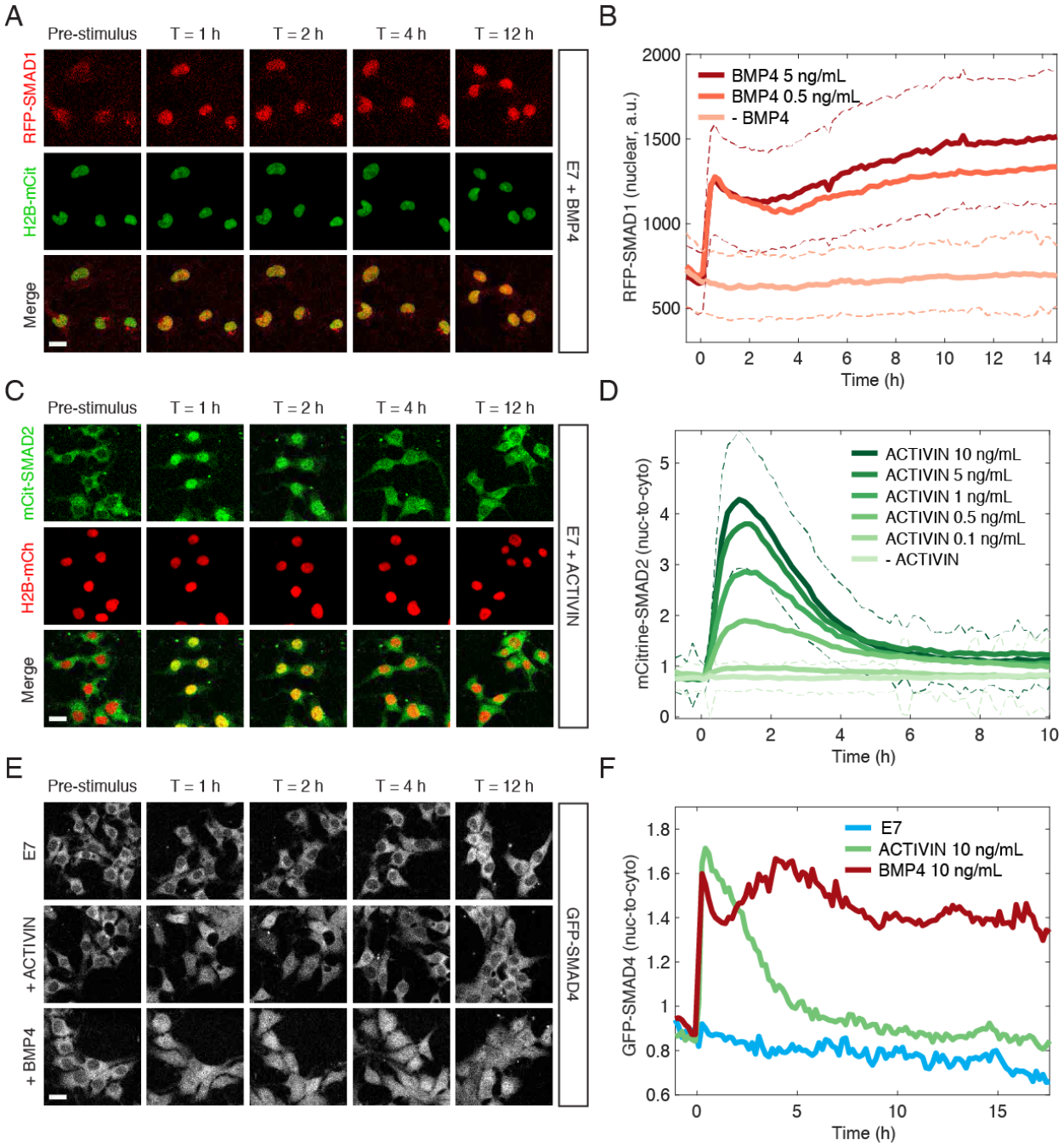
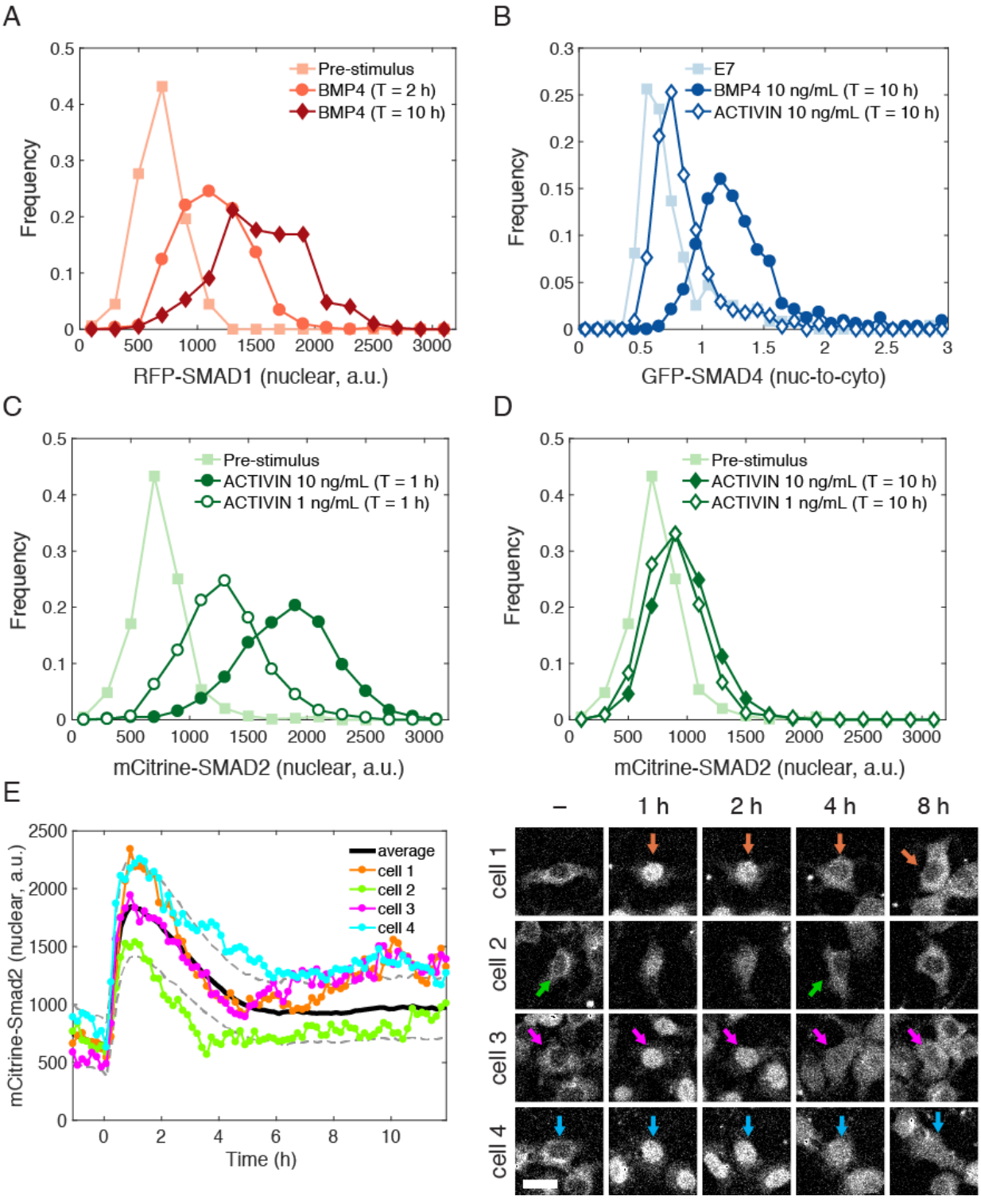

A
Pre-stimulus
T = 1 h
T = 2 h
T = 4 h
T = 12 h
RFP-SMAD1
H2B-mCit
Merge
E7 + BMP4
B
BMP4 5 ng/mL
BMP4 0.5 ng/mL
- BMP4
RFP-SMAD1 (nuclear, a.u.)
Time (h)
C
mCit-SMAD2
H2B-mCh
Merge
E7 + ACTIVIN
D
ACTIVIN 10 ng/mL
ACTIVIN 5 ng/mL
ACTIVIN 1 ng/mL
ACTIVIN 0.5 ng/mL
ACTIVIN 0.1 ng/mL
- ACTIVIN
mCitrine-SMAD2 (nuc-to-cyto)
Time (h)
E
E7
+ ACTIVIN
+ BMP4
GFP-SMAD4
F
E7
ACTIVIN 10 ng/mL
BMP4 10 ng/mL
GFP-SMAD4 (nuc-to-cyto)
Time (h)

Figure 2.6 Individual cell responses reflect population-level dynamics
(A) Histograms of the RFP-SMAD1 nuclear signal quantified in individual cells at different time points following stimulation with BMP4 (5 ng/mL) indicate that the nuclear levels across the population shift up. **(B)** Histograms of the GFP-SMAD4 nuclear-to-cytoplasmic ratio quantified in individual cells in E7, E7 + BMP4 (10 ng/mL), and E7 + Activin (10 ng/mL) show the different levels of nuclear signal increase across the population at T = 10 h after stimulation. Histograms of the mCitrine-SMAD2 nuclear signal following stimulation with Activin (10 ng/mL or 1 ng/mL) at **(C)** T = 1 h and **(D)** T = 10 h indicate that the nuclear levels across the population shift up and down. **(E)** Single-cell traces of the SMAD2 response following stimulation with Activin (10 ng/mL) capture the cell-to-cell variability but demonstrate that individual cells respond with dynamics that are similar to the population average response (solid black line). Gray dashed lines represent the standard deviation. Images from the indicated time points or prior to stimulation (–) show the analyzed cells labeled with arrows colored according to the plot. Scale bar, 25 μm.


A
Pre-stimulus
BMP4 (T = 2 h)
BMP4 (T = 10 h)
Frequency
RFP-SMAD1 (nuclear, a.u.)
B
E7
BMP4 10 ng/mL (T = 10 h)
ACTIVIN 10 ng/mL (T = 10 h)
Frequency
GFP-SMAD4 (nuc-to-cyto)
C
Pre-stimulus
ACTIVIN 10 ng/mL (T = 1 h)
ACTIVIN 1 ng/mL (T = 1 h)
Frequency
mCitrine-SMAD2 (nuclear, a.u.)
D
Pre-stimulus
ACTIVIN 10 ng/mL (T = 10 h)
ACTIVIN 1 ng/mL (T = 10 h)
Frequency
mCitrine-SMAD2 (nuclear, a.u.)
E
average
cell 1
cell 2
cell 3
cell 4
mCitrine-Smad2 (nuclear, a.u.)
Time (h)
–
1 h
2 h
4 h
8 h
cell 1
cell 2
cell 3
cell 4

2.1.4 Activin dose-dependent response

One of the major aims driving my thesis work was to determine whether Activin functions as morphogen in hESCs. An important first step to addressing this question was to determine if and over which concentration range cells are capable of differentially responding to Activin stimulation. In low-density culture, we found that the SMAD2 peak response displays a strong sigmoidal dependence on Activin concentration (Figure 2.5D and Figure 2.7A). However, the post-stimulation baseline (average ratio at T > 8 h after Activin addition) above a threshold concentration of 0.5 ng/mL and the time scale of the transient response were not dependent on the dose (Figure 2.7A–B). Adaption in SMAD2 does not result from loss of Activin activity, as culture medium recovered from cells that were incubated with Activin for 12 hours, elicited a similar SMAD2 response when presented to unstimulated cells (Figure 2.8A). Instead, Activin signaling appears to induce a negative feedback mechanism to keep SMAD2 nuclear levels at a fixed, relatively low level independent of the external ligand concentration, which is also supported by the observation that stimulation with a second larger dose of Activin was dramatically attenuated (Figure 2.8B). Thus, hESCs detect a step-increase in Activin concentration, which is reflected in the transient nuclear accumulation of SMAD2. As SMAD2 is a transcription factor, this short-term response could be translated to dose-dependent cell fate acquisition.

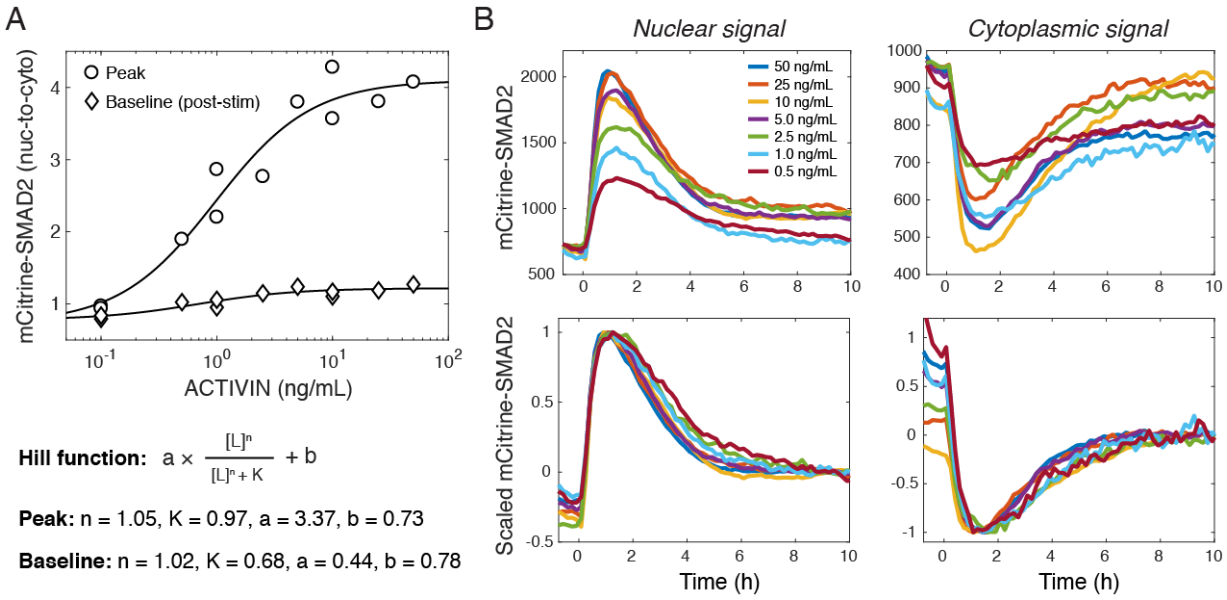


Figure 2.7 Activin/SMAD2 response characteristics
(A) The SMAD2 peak (circles) and post-stimulation baseline (diamonds) nuclear-to-cytoplasmic ratio plotted as a function of Activin concentration. The post-stimulation baseline ratio was calculated as the average ratio for T > 8 h post-stimulation. Solid lines are fits to the Hill function of the form shown, where L is the ligand concentration, n is the Hill coefficient, K is the inflection point, and a and b are constants. The values of the parameters for each fit are also indicated. **(B)** The average mCitrine-SMAD2 nuclear signal (top left) and cytoplasmic signal (top right) as a function of time and Activin concentration. The scaled signals (bottom left and right) were normalized by subtracting the tail SMAD2 fluorescence signal (average response at T > 8 h) and dividing by the peak signal, which is the maximum in the case of the nuclear signal and the minimum in the case of the cytoplasmic signal. Scaling collapses the curves indicating that the time scale of the transient response is similar at each concentration.

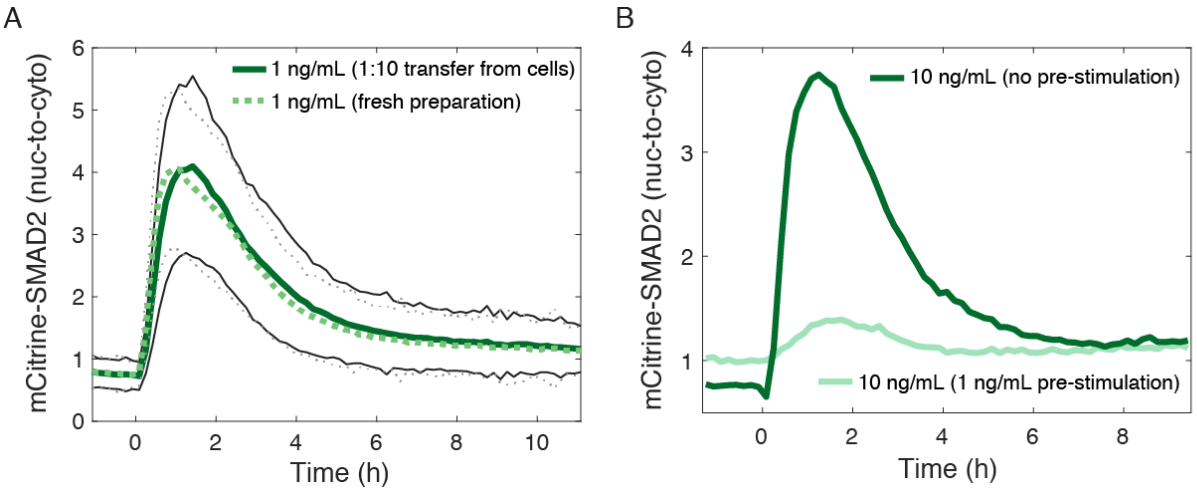


Figure 2.8 Activin induces a negative feedback to limit SMAD2 nuclear accumulation
(A) The mCitrine-SMAD2 response to Activin diluted 1:10 from cells incubated for 12 h with 10 ng/mL (solid green line) elicits a similar response as medium freshly prepared with Activin (dashed green line) indicating that ligand activity is not lost over the time course of SMAD2 adaptation. Solid black lines represent the standard deviation for the 1:10 transfer response and dotted grays lines represent the standard deviation for the response to the fresh preparation. **(B)** The response to Activin (10 ng/mL, dark green line) is dramatically attenuated if cells are stimulated with a first dose of Activin (1 ng/mL, light green line). The pre-stimulation was carried out for approximately 12 h and the background level of 1 ng/mL was maintained when cells were presented with the second dose of 10 ng/mL.

2.2 Fate outcomes

We have previously shown that BMP4 signaling induces a sustained transcriptional response leading to gastruloid differentiation (Etoc et al., 2016; Warmflash et al., 2014). This is consistent with the stable nature of SMAD1 signaling presented above. The adaptive behavior of SMAD2 signaling prompted us to ask whether the short SMAD2 signaling peak was sufficient to elicit a transcriptional response and fate changes in hESCs exposed to Activin.

2.2.1 Transcriptional response to Activin

Given that the cell-autonomous SMAD2 response resembles that observed at the edge of micropatterned colonies, we used low-density culture to facilitate analysis of the corresponding transcriptional response. We performed bulk total RNA-sequencing at several time points following Activin stimulation (Figure 2.9A). 3,529 genes showed a change in expression level of at least two-fold during the experimental time course, which is consistent the number of genes identified by ChIP-sequencing to have nearby SMAD2/3 binding sites (Brown et al., 2011; Tsankov et al., 2015). The differentially expressed genes fell into three distinct groups (complete lists available as supplementary tables with Yoney et al., 2018). The first, which consisted of the majority of transcripts (2,956), peaked at 2.5 hours and declined at later time points (Figure 2.9A, magenta box). This group matched the timing of the transient SMAD2 response and it included key regulators of mesendodermal differentiation, such as EOMES, HHEX, GATA2, and GATA3 (Loh et al., 2014; Teo et al., 2011). The second group, which consisted of 452

transcripts, showed stable induction (Figure 2.9A, orange box). This group included genes expressed during pluripotency, such as NANOG, NODAL, LEFTY1, LEFTY2 and SMAD7 (Sato et al., 2003). Finally, the third group, which consisted of 121 transcripts, represented genes that were stably or transiently down regulated upon Activin presentation and included genes that are involved in signaling pathways not previously associated with pluripotency or differentiation, such as insulin signaling and cAMP response (Figure 2.9A, gray box). Although, all differentially expressed genes as defined by our criteria may not be direct targets of SMAD2, these results nonetheless suggest that cells transiently activate differentiation in response to Activin. Examination of the signaling hierarchy involved in gastruloid self-organization, that going from BMP to Wnt to Nodal signaling, revealed the presence of feedback loops at all three levels. Activin induced the expression of its own ligands and inhibitors as well as those of the BMP and Wnt pathway (supplementary tables provided with Yoney et al., 2018). However, despite the increased expression of ligands and inhibitors, the overall threshold of signaling is presumably not sufficient to induce mesendodermal fates from either the BMP or Wnt pathway downstream of Activin signaling.

We took two additional approaches to evaluate the transient and stable transcriptional responses following Activin stimulation. First, we performed motif enrichment analysis on our gene groups. We selected 10 transcription factors that regulate primitive streak and mesendodermal differentiation: MIXL1, LEF1, BRA, GATA6, FOXH1, FOXA1, FOXA2, GSC, SOX17, and EOMES, and asked if their binding sites are enriched in the promoter region of the genes belonging to each of the dynamic groups. In support of our hypothesis, the motifs were significantly enriched only in the transiently expressed genes of group 1 and not within group 2 or 3 (Figure 2.9B). This suggests that the gene regulatory network activated during the peak of SMAD2 signaling is associated with mesendodermal differentiation. We additionally compared our gene groups with tissue specific genes identified in isolated endoderm, mesoderm, and ectoderm/epiblast tissue from E7.5 mouse embryos (Lu et al., 2018). Although all groups contained some significant enrichment of genes from one or more of the mouse germ layers, group 1 displayed the most significant enrichment of endodermal genes (Figure 2.9C). Overall our data demonstrate that during the peak of SMAD2 nuclear accumulation, hESCs are *en route* for differentiation. However, the mesendodermal differentiation program is not maintained and cells return to pluripotency.

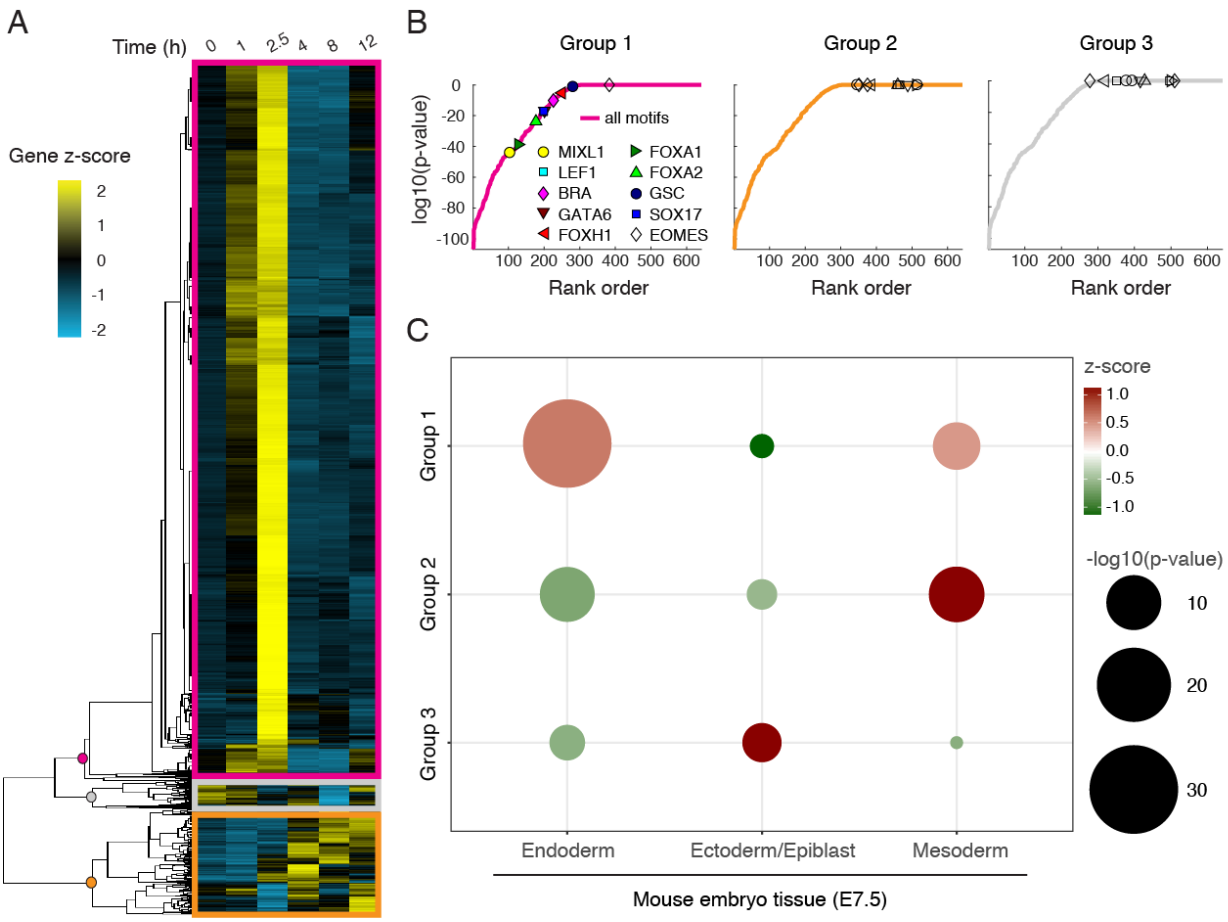


Figure 2.9 Activin signaling results in stable and transient transcriptional responses
(A) Hierarchical clustering of RNA-sequencing time course data for genes showing a fold-change > 2 in response to Activin (10 ng/mL) relative to the pre-stimulus level (T = 0 h). The gene z-score at each time point was calculated by subtracting the average and dividing by the standard deviation of the normalized read counts across all time points. **(B)** Motif enrichment analysis of the gene groups obtained from hierarchical clustering shown in (A). The solid lines represent all motifs analyzed and ordered by the p-value of their enrichment corrected for multiple tests. Symbols indicate select motifs that are associated with primitive streak and mesendodermal differentiation: corrected p-value < 10^{-1} (filled symbols) and corrected p-value > 10^{-1} (empty symbols). **(C)** Comparison of the genes in each of the groups identified in (A) to genes specifically enriched in isolated endoderm, mesoderm, and ectoderm/epiblast tissue from E7.5 mouse embryos (Lu et al., 2018). The size of the circles indicates the –log10(p-value) of the enrichment and the color scale indicates the z-score of the log2(observed/expected) ratio across each group.

2.2.2 *SMAD3-/-* RUES2 lines

In the mouse embryo, Smad3 is dispensable for early development as demonstrated by the fact that Smad3 knockout mice make it to adulthood (Datto et al., 1999; X. Yang et al., 1999; Y. Zhu, Richardson, Parada, & Graff, 1998). In the absence of Smad2 in the epiblast, Smad3 can mediate some mesoderm induction during gastrulation. However, more anterior mesendodermal lineages are completely eliminated and the embryos fail at gastrulation suggesting a critical role for Smad2 in this process (Dunn et al., 2004; S. D. Vincent et al., 2003). In order to decipher whether the transient SMAD2 response is sufficient to drive the transcriptional program downstream of Activin presentation in hESCs, we generated two independent RUES2 *SMAD3* homozygous knockout lines (RUES2-SMAD3-/-) using CRISPR-Cas9 genome editing to introduce a frameshift mutation into the first coding exon shared by all SMAD3 isoforms. These lines maintained expression of pluripotency markers NANOG, OCT4, and SOX2, which is consistent with the previous finding that SMAD2, but not SMAD3, regulates NANOG expression to promote pluripotency in hESCs and mouse epiblast stem cells (Figure 2.10A) (Sakaki-Yumoto, Liu, Ramalho-Santos, Yoshida, & Derynck, 2013). In response to Activin, RUES2-SMAD3-/- cells displayed a transcriptional response identical to the parental RUES2 line for both pluripotency- and mesendoderm-associated Activin target genes (Figure 2.10B–C). Although our results cannot rule out possible redundancy between SMAD2 and SMAD3, we conclude that SMAD3 is not required for maintenance of pluripotency or the transcriptional response dynamics following Activin presentation.

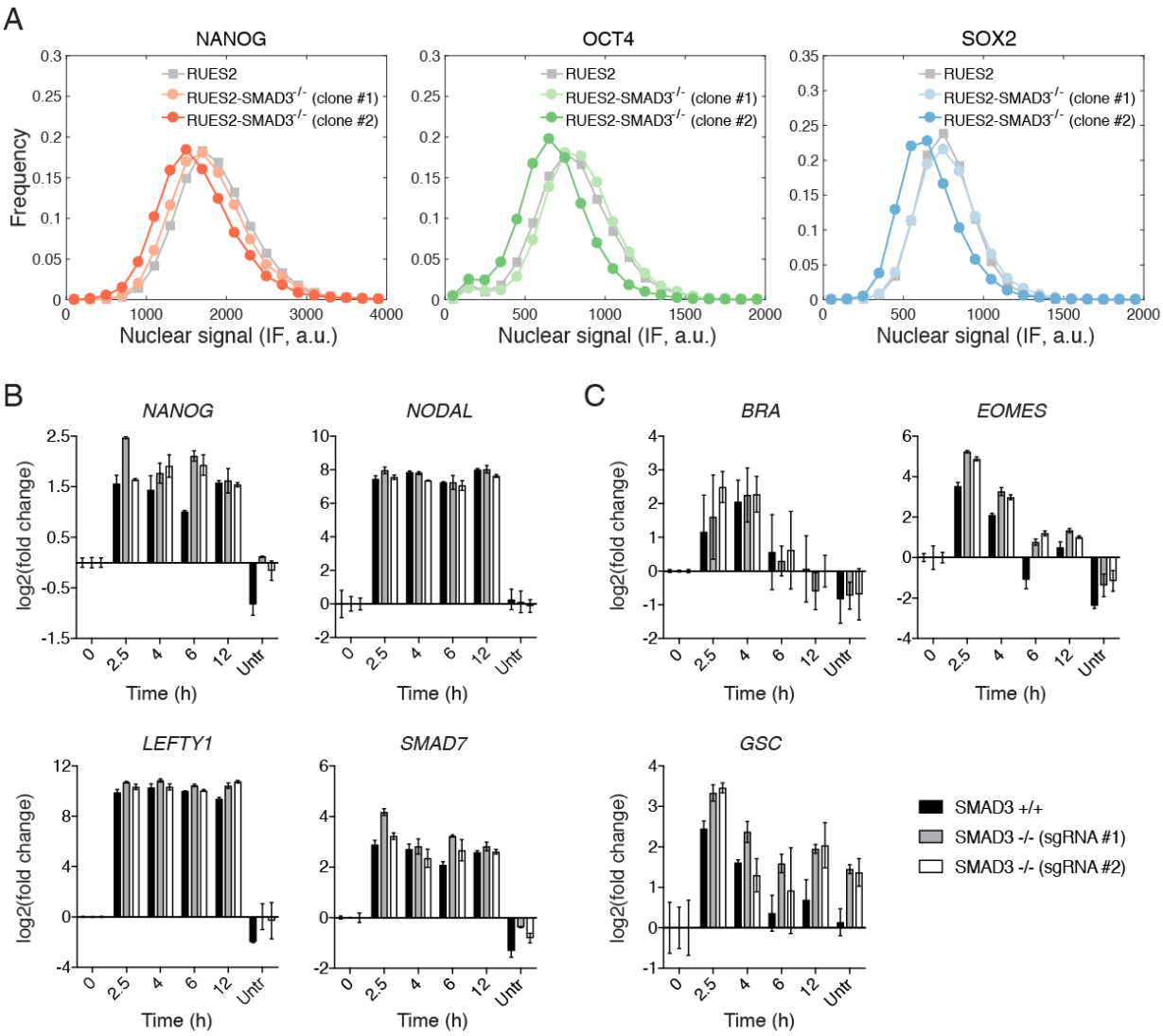


Figure 2.10 SMAD3 is not required for the transcriptional response to Activin in hESCs
(A) RUES2-SMAD3-/- cells maintain expression of the pluripotency markers under standard growth conditions in MEF-CM. Cells were fixed and analyzed by immunofluorescence for NANOG, OCT4, and SOX2 expression. Histograms show the nuclear signal quantified in single cells (n > 10,000 cells per line). **(B–C)** RT-PCR analysis of (B) pluripotency and (C) and mesendoderm-associated Activin/Nodal target genes following presentation of Activin (10 ng/mL) in the parental RUES2 line (black bars) and the RUES2-SMAD3-/- lines (gray and white bars). An additional sample was collected that was left untreated for the 12 h time course (untr). Expression in each sample was normalized to GAPDH and then to the pre-stimulus level (T = 0 h). Data represents the mean ± S.D. for n = 3 technical replicates. Similar results were obtained for the parental line in three independent experiments.

2.2.3 Implications for pluripotency

As presented above the post-stimulation SMAD2 baseline remains elevated relative to the pre-stimulus level following adaptation, and it is independent of exogenous Activin over a wide range of concentrations. We asked if the increase in the SMAD2 baseline after the peak response is required for the maintenance of pluripotency. In order to address this question, RUES2-mCit-SMAD2 cells were treated with SB-431542 (SB), which inhibits the activity of the type I Activin/Nodal receptors, 8 hours after Activin stimulation and analyzed for their ability to maintain pluripotency. SB treatment led to a decrease in SMAD2 nuclear-to-cytoplasmic levels back to the unstimulated baseline (Figure 2.11A). As observed previously, presentation of SB led to a loss of pluripotency, as indicated by the loss of NANOG expression in RUES2-mCit-SMAD2 (Figure 2.11B) (James et al., 2005; Vallier et al., 2005; Xu et al., 2008). Analysis of the parental RUES2 line confirmed the loss of pluripotency under the same experimental conditions (Figure 2.11C). We conclude that the elevated baseline following adaptation of the SMAD2 response is dependent on continued activation of the TGFβ receptors and is responsible for maintaining the pluripotency program long-term.

The fact that the SMAD2 post-stimulation baseline is the same regardless of Activin concentration above 0.5 ng/mL, suggests that pluripotency is insensitive to graded ligand levels above this threshold. To test this hypothesis, we treated cells with different levels of Activin and compared the expression of NANOG, OCT4, and SOX2 after 2 days of stimulation. Expression of all three markers was similar at three different concentrations of exogenous Activin (1, 10, and 100 ng/mL) and expression of NANOG and OCT4 was elevated relative to the –Activin condition (Figure 2.12A–B). To test the response to complete inhibition of SMAD2 signaling we presented SB for the same amount of time. Pluripotency was not maintained under these conditions as demonstrated by down-regulation of NANOG and OCT4 below the –Activin condition (Figure 2.12A–B). In order to compare immunofluorescence data across biological replicates, which may differ in absolute intensity values, we computed the Kolmogorov-Smirnov (KS) distance between each marker's cumulative distribution function (CDF) measured under different conditions to a reference CDF, in this case the CDF measured under the 1 ng/mL Activin condition. In biological replicates using RUES2 or a second hESC line, RUES1, the distributions of NANOG and OCT4 levels under the 100 or 10 ng/mL Activin conditions are reproducibly close to the distributions measured under the 1 ng/mL Activin condition (small KS distance, Figure 2.12C). The distributions measured under the –Activin and SB conditions are reproducibly far from the distributions measured under the 1 ng/mL Activin condition (large KS distance, Figure 2.12C). This confirms the insensitivity of the pluripotent state to graded Activin and demonstrates the quantitative reproducibility of our findings.

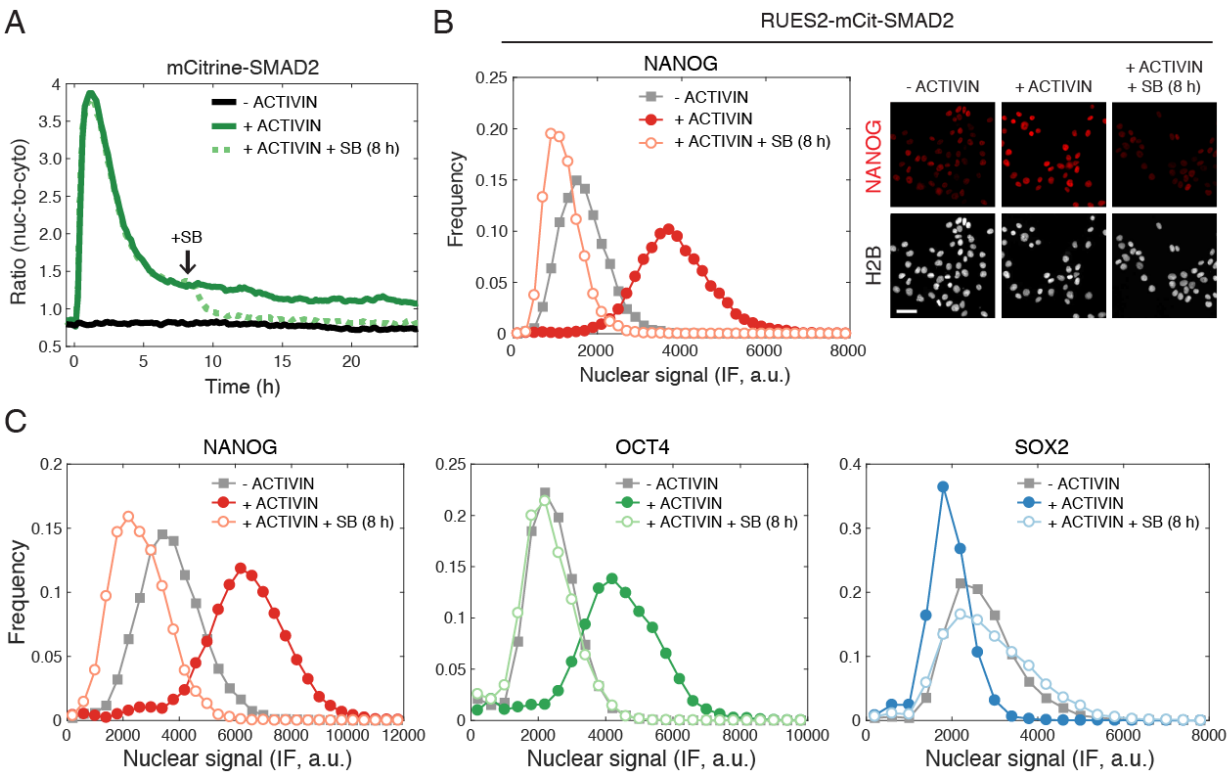


Figure 2.11 Post-adaptation SMAD2 levels maintains pluripotency
(A) Quantification of the mCitrine-SMAD2 nuclear-to-cytoplasmic ratio following treatment with Activin (10 ng/mL added at T = 0 h, solid green line). Following the transient SMAD2 response, SB (10 µM) was added to one of the samples (dotted green line, added at T = 8 h). A third sample was left untreated in E7 (–Activin) for the duration of the experiment (solid black line). Images were acquired every 10 minutes. Lines represent the average response (n > 200 cells analyzed per time point). **(B)** The samples in (A) were fixed 24 h after Activin presentation and analyzed for NANOG expression by immunofluorescence. Scale bar, 50 µm. **(C)** Analysis of the parental RUES2 line under the same experimental conditions shown in (A) and (B). Histograms in (A) and (C) show the nuclear signal quantified in single cells (n > 5,000 cells per condition).

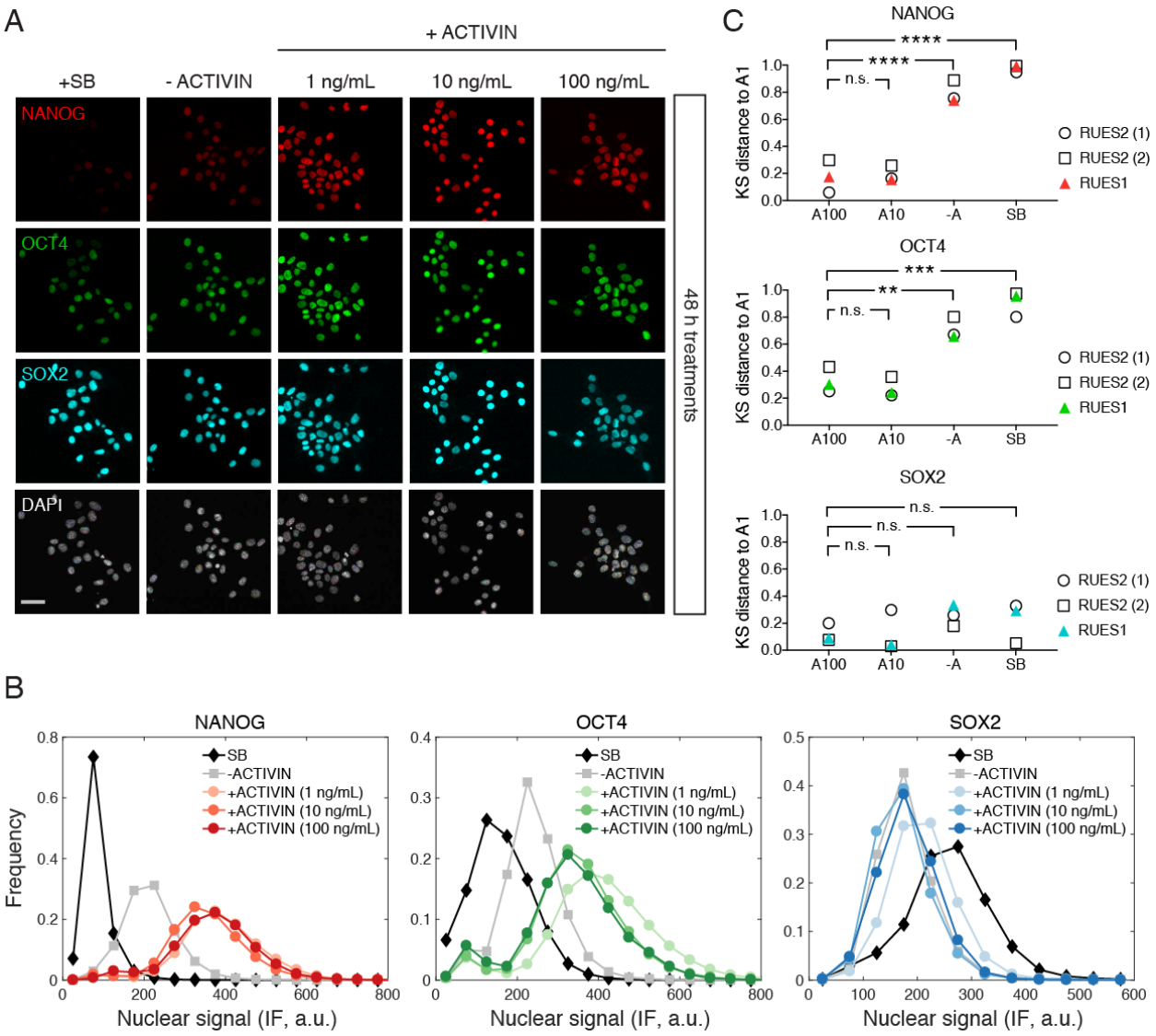


Figure 2.12 Pluripotency is insensitive to the dose of exogenous Activin
(A–B) RUES2 cells were cultured in E7 with different levels of Activin (0, 1, 10, 100 ng/mL) or SB (10 µM) for 2 days at which point cells were fixed and analyzed by immunofluorescence. (A) Images: NANOG (red), OCT4 (green), SOX2 (cyan), DAPI (gray). Scale bar, 50 µm. (B) Histograms showing the nuclear IF signal quantified in single cells (n > 5,000 cells per condition). **(C)** Kolmogorov-Smirnov (KS) distance of the cumulative probability distribution (CDF) of each marker to the reference CDF (1 ng/mL ACTIVIN condition) for independent experiments in RUES2 and in RUES1. n.s., not significant, **P< 0.01, ***P<0.001, ****P<0.0001, ANOVA.

2.3 Discussion

These results represent one of the first highly quantitative characterizations of Activin/SMAD2 signaling in hESCs. According to our estimates, most differentiation protocols that currently utilize Activin, apply concentrations (50 - 100 ng/mL) that are well above pathway saturation at the level of SMAD2. Therefore, it is possible that protocols can be re-optimized using our response curves, or conversely that these reporter lines can be used to estimate SMAD responses in other experimental conditions. Our results also demonstrate that similar to cells of the *Xenopus* animal cap, human embryonic cells are also capable sensing and responding to changes in exogenous ligand levels, and in this respect, Activin meets this definition of a morphogen. In contrast to Xenopus, there is only a transient induction of differentiation that happens at the peak of SMAD2 nuclear levels. This transient differentiation in response to Activin alone was not expected from previous studies in hESCs, which focused on the role of SMAD2/3 in pluripotency maintenance. Whether this differentiation is functionally meaningful and whether it could be stabilized simply by forcing SMAD2 to remain in the nucleus is a potential question for future studies.

A second, related question that is raised by these experiments is: what is the mechanism of SMAD2 adaption? One possibility is that Activin signaling induces receptor internalization and that the receptors are not replenished in the presence of excess ligands. This type of mechanism has been reported for adaptive TGFβ signaling in a human keratinocyte cell line (Vizán et al., 2013). Our RNA-sequencing data following Activin stimulation provide candidate genes, which remain elevated after SMAD2 adaptation and which could, therefore, negatively regulate the response at later times, the most likely candidates from this analysis being BAMBI and SMAD7. BAMBI is a surface-bound decoy receptor for Activin and BMP ligands, and SMAD7 is an inhibitory SMAD, and thus, intracellular, which has been reported to negatively regulate signaling through both branches of the pathway by several mechanisms (Massague, 2005). We found that following one round of ligand stimulation, cells are refractory to a second larger dose of Activin, which suggests that regardless of the specific mechanism of adaptation, it has an excess buffering capacity.

CHAPTER 3. TEMPORAL INTERACTIONS BETWEEN WNT AND ACTIVIN

A major advancement in the stem cell field was made with the discovery of protocols to induce endoderm and its derivatives, including clinically-relevant insulin-producing β cells (D'Amour et al., 2005; McLean et al., 2007; Pagliuca et al., 2014). The initial protocols, which generated definitive endoderm through a primitive-streak/mesendoderm intermediate, relied on the application of Activin or Nodal ligands in combination with PI3K/Akt inhibition, either through low serum medium or small molecule inhibitors. Follow-up studies recognized the cascade of interactions that link PI3K inhibition ultimately to GSK3β inhibition and, hence, to Wnt activation. Although there may be additional roles for PI3K modulation in hESCs, the field has largely moved towards mesendoderm protocols that utilize purified recombinant Wnt protein or small molecules that directly inhibit GSK3β (Loh et al., 2014; Singh et al., 2012; Sumi, Tsuneyoshi, Nakatsuji, & Suemori, 2008). We have demonstrated that Activin alone cannot drive stable differentiation of human gastruloids. However, Wnt can lead to differentiation and self-organization of mesoderm and endoderm in a SMAD2/3-dependent manner (Martyn et al., 2018). Furthermore, co-presentation of Activin with Wnt to micropatterned hESCs leads to formation of a primitive streak region, which expresses the organizer-specific marker Goosecoid (GSC) and which can induce a secondary axis when grafted into chick embryos (Martyn et al., 2018). Given these observations and evidence in the mouse that Wnt is operating directly up-stream of increased Nodal signaling, we set out to address how hESCs with a history of Wnt signaling respond to Activin stimulation.

In this chapter, I present a simple protocol using low-density cell culture that allowed us to dissect the temporal interaction between Wnt and Activin signals. Despite the existence of multiple transcriptional feedbacks between the different inductive signals, our results clearly demonstrate that with respect to fate acquisition Wnt signaling acts temporally upstream of Activin. Cells appear to maintain a memory of earlier Wnt signals that alone does not induce a particular fate. However, subsequent presentation of Activin results in dose-dependent mesendoderm induction. Inspired by recent work in *Xenopus*, we used our two-step protocol to screen for epigenetic modulators of Wnt memory establishment and uncovered a potentially conserved role for the Bromodomain and extra-terminal (BET) family of epigenetic readers in vertebrate germ layer patterning.

3.1 Establishment of Wnt signaling memory

In micropatterned cell culture robust patterning is dependent on the reproducible, high cell density that results from confinement of colony area (Etoc et al., 2016). In low-density cell culture secreted ligands appear to be less efficiently shared between cells, which facilitates precise control of their timing through exogenous proteins and small molecule inhibitors. We initially used this latter method to dissect the temporal requirements for Wnt and Activin signals in mesendoderm induction. Our results are applicable to recent efforts to model the dynamic patterning of human gastruloids, demonstrating the power of combining these approaches to dissect developmental mechanisms (Chhabra, Liu, Goh, Kong, & Warmflash, 2019).

3.1.1 Wnt priming in low-density cell culture

To address how prior Wnt stimulation affects the response of cells to Activin, low-density cultures were stimulated with or without recombinant Wnt3a (referred to as Wnt) for 24 hours, washed, and then stimulated with or without Activin for an additional 24 hours (Figure 3.1A–D).

Surprisingly, when cultures were treated with Wnt in the absence of Activin, cells remained pluripotent, as indicated by the co-expression of the pluripotency makers, NANOG, OCT4, and SOX2, and the absence of differentiation markers, BRA, EOMES, and GSC. As discussed below, we did detect induction of *BRA* and *EOMES* at the RNA level. However, this expression decreased after cells were returned to E7, which we took as further evidence that cells had maintained or returned to pluripotency following one day of Wnt stimulation. If cells were exposed to Wnt followed by Activin, mesendodermal fates were robustly induced, as indicated by expression of BRA, EOMES, and GSC. KS distance analysis using the Activin only treated cells as the reference condition demonstrated the quantitative reproducibility of our observations in RUES1 and RUES2 (Figure 3.1E–F). We obtained similar results if cells were stimulated with a GSK3 inhibitor, CHIR99021 (CHIR), instead of recombinant Wnt protein (Figure 3.2A). Similar results were also obtained in the RUES2-SMAD3-/- line, suggesting that SMAD3 is not required for stable mesendoderm induction in response to Wnt and Activin (Figure 3.2B). Typically, 1-2 days of Wnt stimulation or GSK3 inhibition is sufficient do induce BRA expression (Martyn et al., 2018; Martyn, Brivanlou, & Siggia, 2019a). We attribute the differences between our results presented here and prior studies by our group and others to the use of medium without Smad2/3 activating ligands as well as the use of low-density culture conditions, which likely further limits differentiation by limiting the action of endogenously produced ligands, including Nodal.

In order to address the possibility that cells merely require Wnt and Activin signals in close succession, we treated cells with Activin for 24 hours, washed, and then immediately stimulated with Wnt for an additional 24 h. We did not observe mesendoderm induction using the reverse order, Activin followed by Wnt, or with the presentation of either signal alone (Figure 3.2C–D). Therefore, we termed the protocol in which Wnt is presented first, *Wnt priming*, because it does not induce overt differentiation but is required to change the response of cells to subsequent Activin stimulation. In agreement with the morphogen patterning models from frog and mouse, differentiation following Wnt priming was Activin concentration dependent, as demonstrated by the induction of BRA at low Activin and the induction of BRA, EOMES and GSC at high Activin concentrations (Figure 3.3).

Figure 3.1 Wnt priming unveils Activin-dependent mesendoderm differentiation
(A–B) hESCs were cultured for one day in E7 with or without Wnt3a (100 ng/mL, top bar). On the second day cells were washed to remove Wnt and cultured in E7 with or without Activin (10 ng/mL, bottom bar). After the second day cells were fixed and analyzed by immunofluorescence for expression of the pluripotency markers. **(A)** Images: NANOG (red), OCT4 (green), SOX2 (cyan), DAPI (gray). Scale bar, 50 μm. **(B)** Histograms showing the nuclear signal quantified in single cells (n > 5,000 cells per condition). **(C–D)** Analysis of mesendoderm marker expression in the same experimental setup as that shown in (A–B). **(C)** Images: BRA (red), EOMES (green), GSC (cyan), DAPI (gray). Scale bar, 50 μm. **(D)** Histograms showing the nuclear immunofluorescence signal quantified in single cells (n > 5,000 cells per condition). **(E–F)** Kolmogorov-Smirnov (KS) distance of the cumulative probability distribution (CDF) of each marker to the reference CDF (-/ACT) for independent experiments in RUES2 and RUES1 for the **(E)** pluripotency **(F)** and mesendoderm marker sets. n.s. (or comparison not shown), not significant, *P<0.05, **P< 0.01, ***P<0.001, ANOVA.

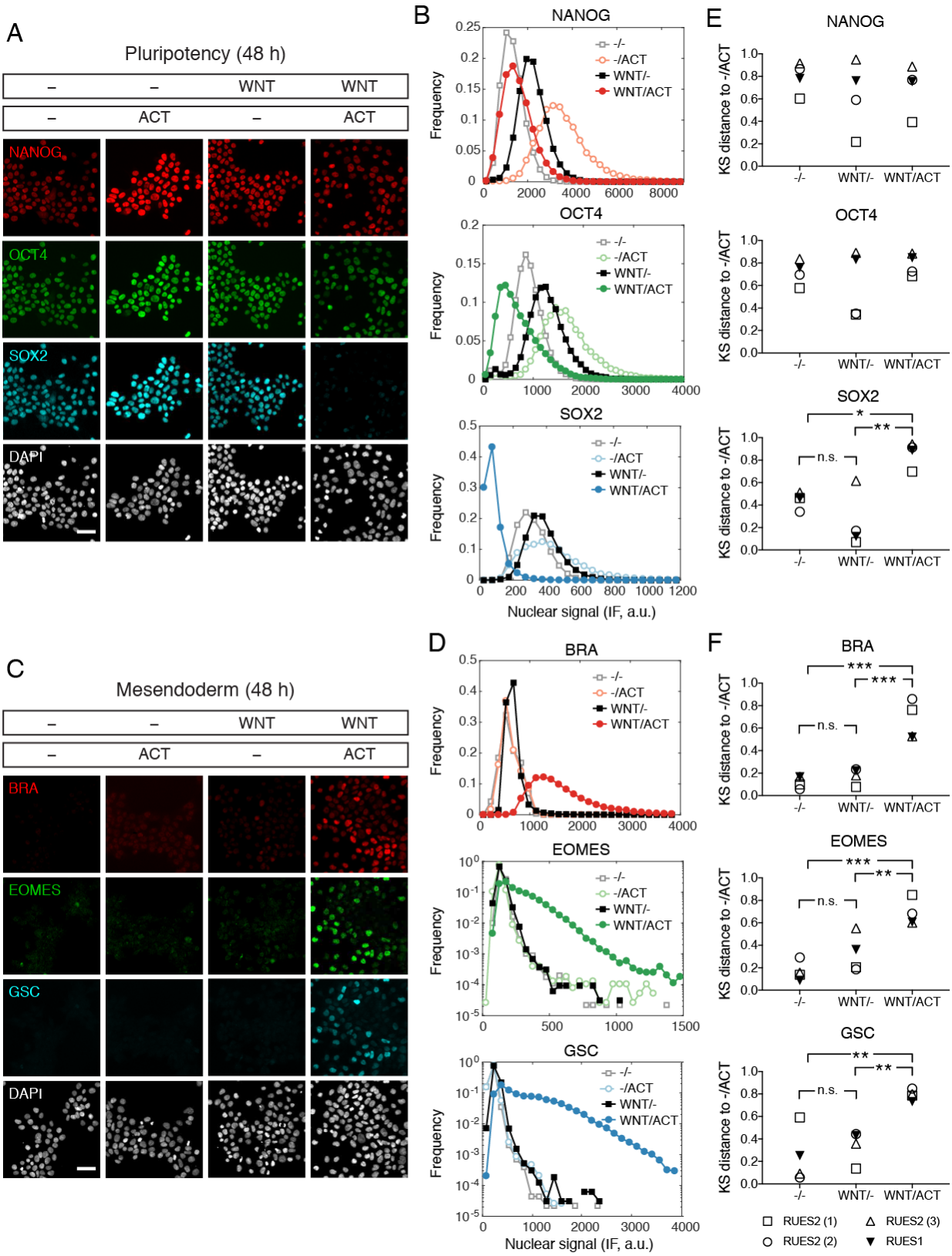
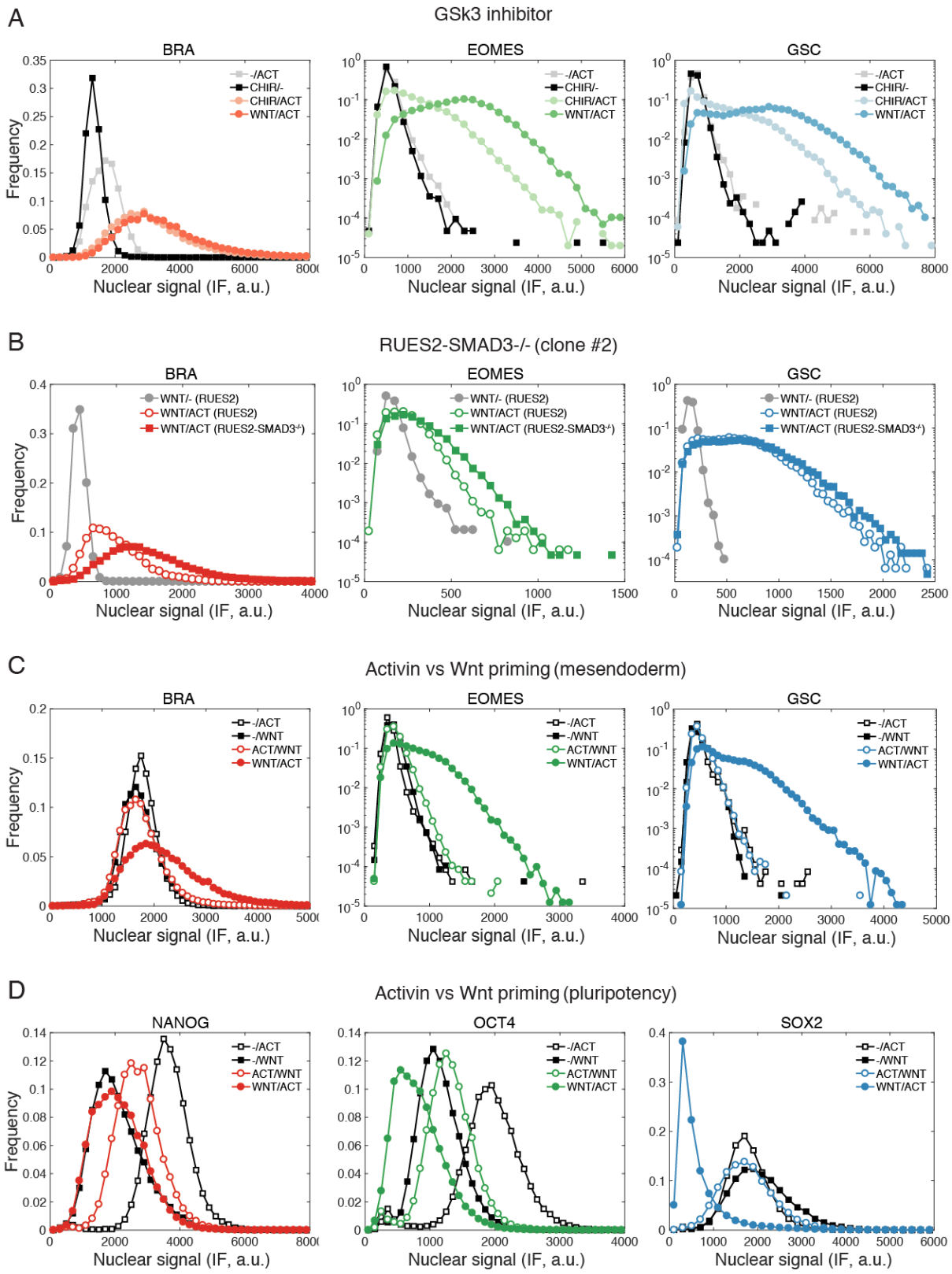

A
Pluripotency (48 h)
–
–
WNT
WNT
–
ACT
–
ACT
NANOG
OCT4
SOX2
DAPI
B
NANOG
OCT4
SOX2
Frequency
-/-
-/ACT
WNT/-
WNT/ACT
Nuclear signal (IF, a.u.)
E
NANOG
OCT4
SOX2
KS distance to -/ACT
n.s.
*
**
C
Mesendoderm (48 h)
BRA
EOMES
GSC
DAPI
D
BRA
EOMES
GSC
F
***
RUES2 (1)
RUES2 (2)
RUES2 (3)
RUES1

Figure 3.2 Wnt priming control experiments
(A) hESCs were cultured for one day in E7 with or without CHIR99021 (CHIR, 2.5 μM). On the second day cells were washed to remove CHIR and cultured in E7 with or without Activin (10 ng/mL). After the second day cells were fixed and analyzed by immunofluorescence (IF) for expression of the mesendodermal markers, BRA, EOMES, and GSC. The results of Wnt priming were obtained in parallel for direct comparison. Histograms represent the nuclear signal quantified in single cells (n > 5,000 cells per condition). **(B)** Mesendoderm is induced in both RUES2 and RUES2-SMAD3-/- cells in response to Wnt priming followed by Activin stimulation. Experimental setup and analysis are the same as that described in Figure 3.3. **(C–D)** Wnt priming followed by Activin stimulation results in **(C)** mesendoderm induction and **(D)** loss of pluripotency, as indicated by changes in maker expression. The reverse order, i.e. Activin priming, or the presentation of either signal alone is not sufficient to induce changes in fate.


A
GSk3 inhibitor
BRA
EOMES
GSC
-/ACT
CHIR/-
CHIR/ACT
WNT/ACT
Frequency
Nuclear signal (IF, a.u.)
B
RUES2-SMAD3-/- (clone #2)
WNT/- (RUES2)
WNT/ACT (RUES2)
WNT/ACT (RUES2-SMAD3-/-)
C
Activin vs Wnt priming (mesendoderm)
-/WNT
ACT/WNT
D
Activin vs Wnt priming (pluripotency)
NANOG
OCT4
SOX2

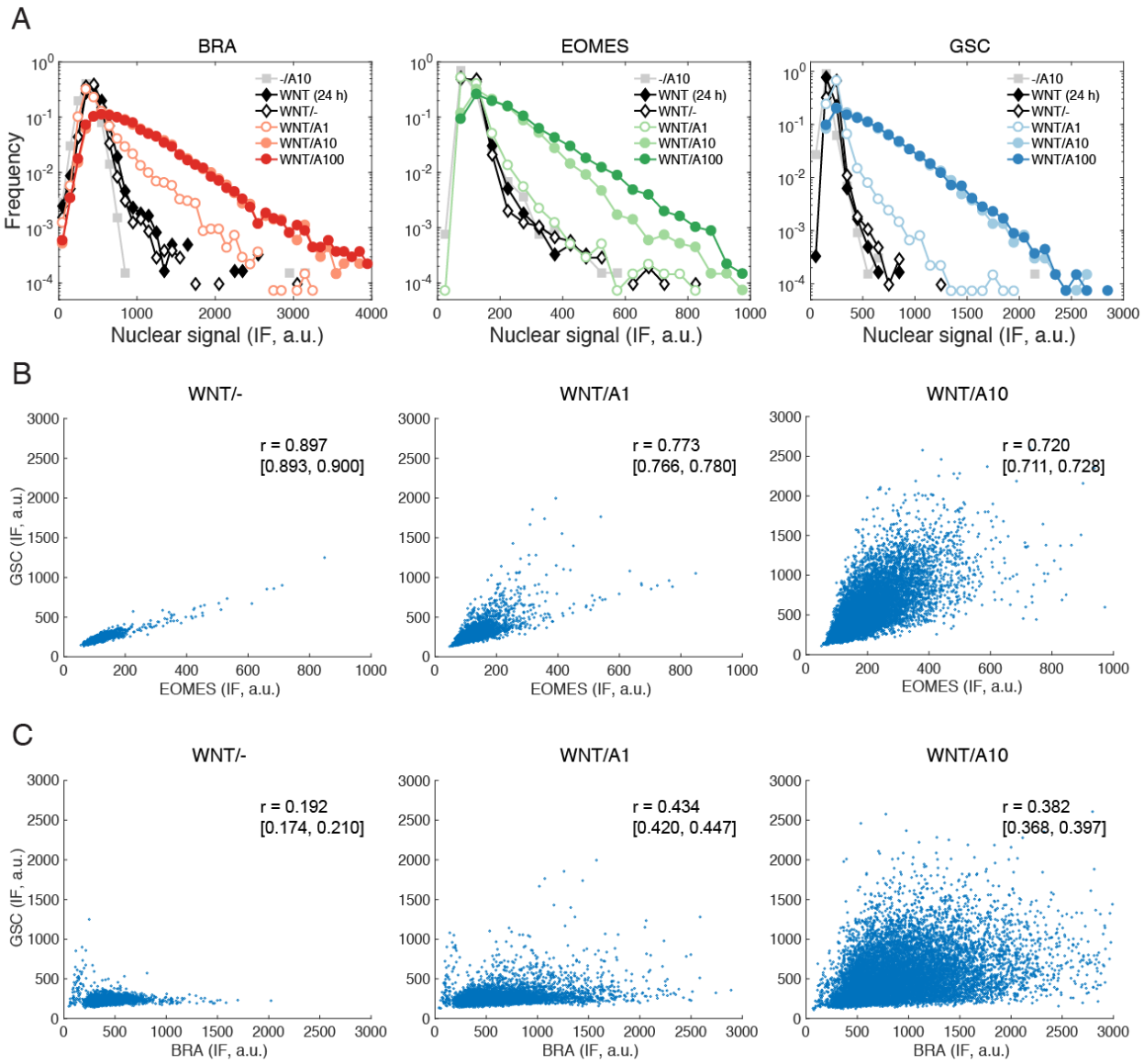


Figure 3.3 Activin functions as a morphogen following Wnt priming
(A) Histograms showing immunofluorescence data from samples treated with different doses of Activin (1, 10, 100 ng/mL) for 12 h following Wnt priming. The data for Wnt (24 h) shows that all markers were poorly expressed prior to Activin addition and that their distributions are identical to cells in pluripotency conditions (-/ACT10, 36 h) or cells that were left untreated with Activin for the additional 12 h (WNT/-). **(B)** Scatter plot of the single-cell EOMES and GSC nuclear signal from the data shown in (A). EOMES and GSC expression increases in WNT/ACT10 condition and expression is positively correlated. **(C)** Scatter plot of the single-cell BRA and GSC nuclear signal from the data shown in (A). Expression is not well correlated. Similar numbers of cells were analyzed in all conditions in (B–C), but the data points cluster on top of one another in the WNT/- and WNT/ACT1 conditions. r, Pearson correlation coefficient with the 95% confidence interval shown in brackets.

3.1.2 Insights into the mechanism of Wnt priming

In order to address the mechanism of Wnt priming, we first asked whether the transcriptional effector of canonical Wnt signaling, β-catenin, is involved in this process. To this end, we treated cells with the small molecule inhibitor endo-IWR-1 throughout the two-day Wnt and Activin protocol (Figure 3.1). endo-IWR-1 stabilizes Axin and, thus, the β-catenin destruction complex through inhibition of Tankyrase (B. Chen et al., 2009; Huang et al., 2009). In the presence of endo-IWR-1 mesendoderm maker expression was eliminated and pluripotency marker expression was maintained (Figure 3.4A–B). Although the mechanism of endo-IWR-1 inhibition is indirect, these results nonetheless suggest a requirement for β-catenin in the priming step. We then asked if endogenously produced Wnt is required for induction of mesendoderm, particularly during the Activin stimulation phase when exogenous Wnt has been removed. To address this we blocked endogenous Wnt secretion using a small molecule inhibitor of Porcupine, IWP-2, which was again maintained throughout the two-day protocol (B. Chen et al., 2009). We found that addition of IWP-2 does not affect Activin-dependent mesendoderm differentiation and loss of pluripotency following Wnt priming (Figure 3.4A–B). However, IWP-2 does block Wnt-dependent mesendoderm differentiation downstream of BMP4, indicating that the molecule is effective at this concentration (Figure 3.4C–D). We attempted to assess the stability of the Wnt-primed state by incubation of cells in neutral medium (E7) prior to Activin stimulation. We observed an Activin-dependent induction of mesendoderm markers after one day in neutral medium following Wnt priming (Figure 3.5). However, over this time course the culture density continues to increase and endogenous SMAD2/3 activating ligands begin to take effect, which can be seen in the increase in mesendodermal markers in the absence of added Activin (Figure 3.5). Therefore, an alternative experimental setup is needed to unambiguously probe this aspect of Wnt priming.

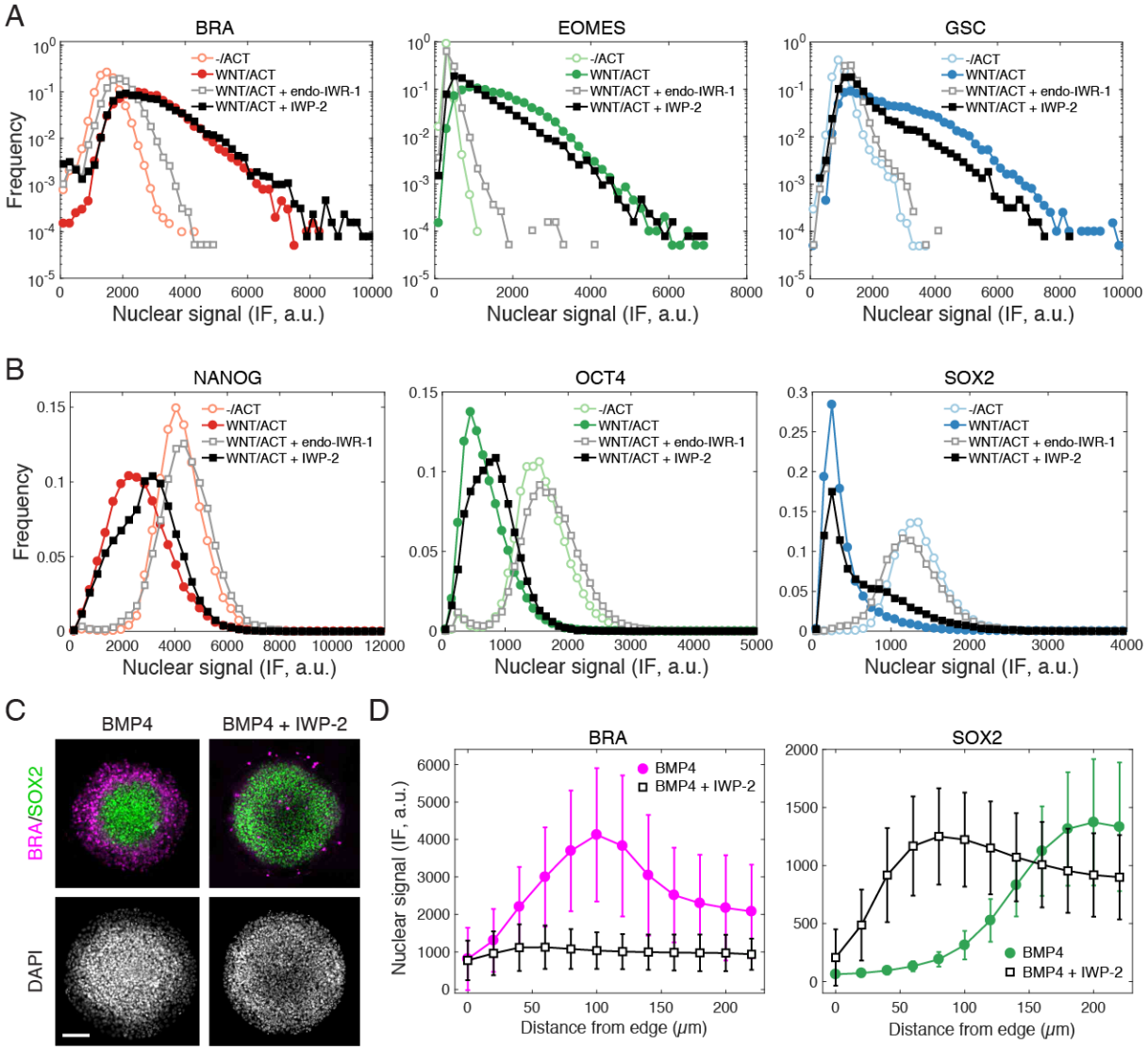


Figure 3.4 Wnt signaling memory is established through the canonical pathway
(A–B) Cells were cultured for one day with Wnt (100 ng/mL), washed, and then stimulated with Activin (10 ng/mL). In order to block Wnt signaling through β-catenin cells were treated with endo-IWR-1 (1 μM), which was added with Wnt on day 1 and again with Activin on day 2. In order to block Wnt secretion, cells were similarly treated with IWP-2 (1 μM). Cells were fixed and analyzed by immunofluorescence (IF) for **(A)** mesendoderm and **(B)** pluripotency markers. Histograms show the nuclear IF signal quantified in single cells (n > 5,000 cells per condition). Cells cultured without Wnt, corresponding to the pluripotency condition (-/ACT), represent the background signal. **(B–C)** Micropatterned colonies treated with BMP4 (50 ng/mL) or BMP4 + IWP-2 (1 μM) for 42 hours. The colonies were fixed and analyzed by immunofluorescence. (B) Images: BRA (magenta), SOX2 (green), DAPI (gray). Scale bar, 100 μm. (C) Quantification of the mean nuclear fluorescence as a function of radial position from the colony edge. Error bars represent the standard deviation across n = 17 (BMP4, filled circles) and n = 12 (BMP4 + IWP-2, open squares) colonies from one experiment.

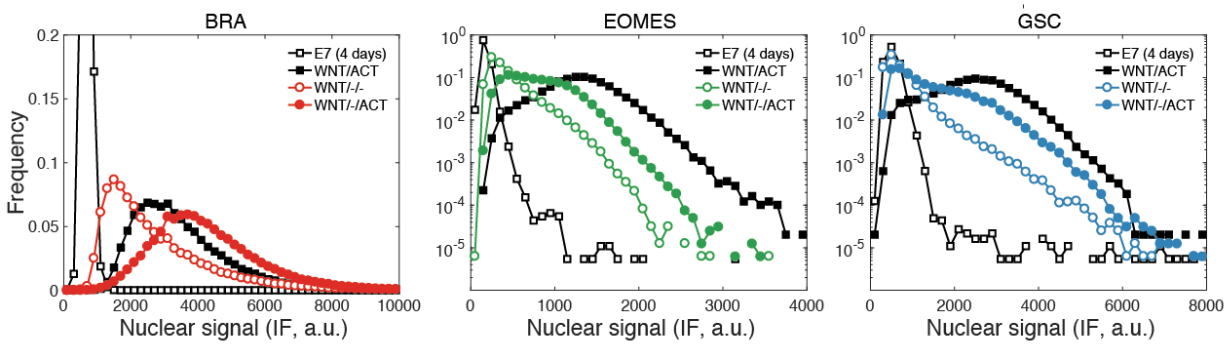


Figure 3.5 Wnt-primed state is stable in neutral medium
Cells were cultured for one day with Wnt (100 ng/mL), washed, and then incubated for one day in neutral medium (E7) prior to Activin (10 ng/mL) stimulation. Cells stimulated with Activin (solid circles) showed a greater increase in mesendoderm marker expression compared to cells kept in neutral medium following Wnt priming (open circles). Marker expression in all Wnt stimulated samples increased above the background signal in cells cultured in neutral medium only (open squares). Marker levels in cells primed with Wnt immediately followed by Activin stimulation were obtained in parallel for comparison (black squares).

Since we have shown that the SMAD response dynamics can be stable or transient depending on the branch being activated, we next asked if Wnt priming affects the dynamics of SMAD2 signal transduction. The transient response of SMAD2 and SMAD4 and the elevation in the baseline post-stimulation was the same whether or not cells were previously exposed to Wnt (Figure 3.6A). However, transcription of mesendodermal genes was stabilized in response to Activin following Wnt priming consistent with the increased expression at the protein level presented above (Figure 3.6B). As in the case of pluripotency, stable fate acquisition required on-going SMAD2 signaling in the elevated baseline as treatment with SB, 8 hours after Activin stimulation, eliminated mesendodermal differentiation at 24 h (Figure 3.7A). In contrast, and further arguing for a mechanism of Wnt priming that is temporally up-stream of Activin, addition of SB during Wnt priming did not block mesendoderm differentiation (Figure 3.7B). Taken together, our results suggest the presence of an unexpected *Wnt signaling memory* in cells that is mediated via β-catenin and is established prior to and is required for the morphogen activity of Activin.

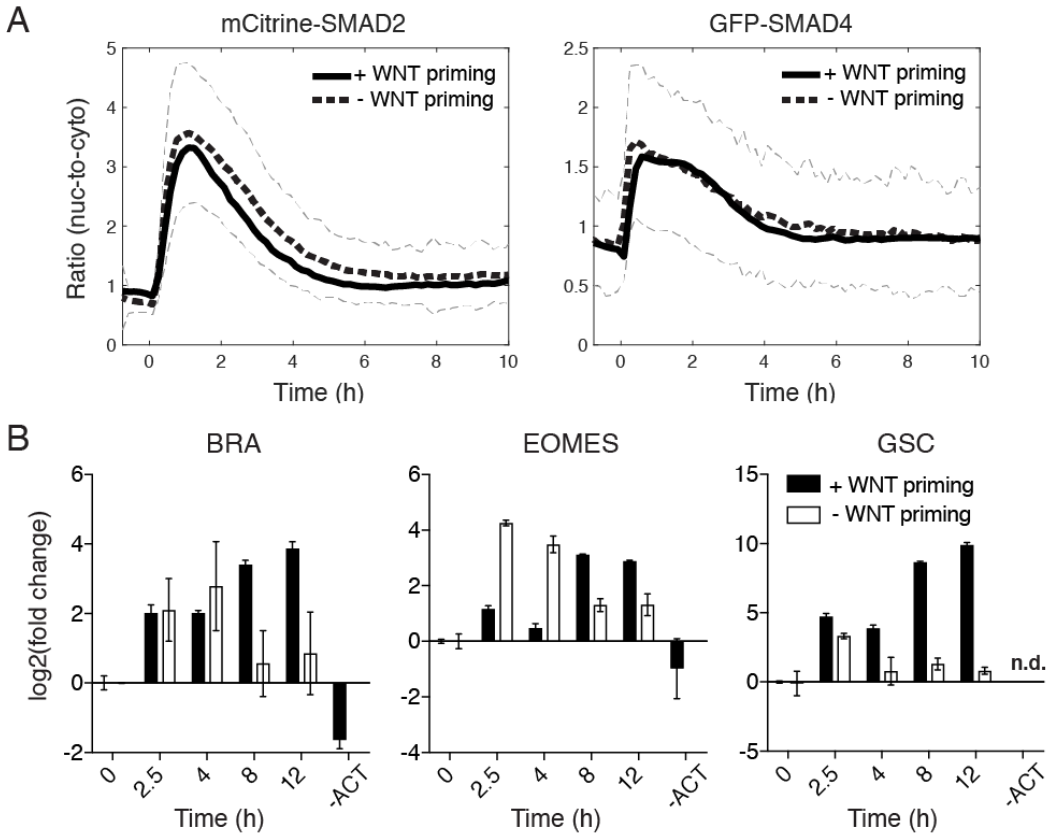


Figure 3.6 Adaptive SMAD signaling leads to stable mesendoderm gene expression
(A) mCitrine-SMAD2 (left) and GFP-SMAD4 (right) response to Activin (10 ng/mL) with or without Wnt priming (black solid and dashed lines, respectively). Activin was added at T = 0 h, which corresponds to 24 h of Wnt stimulation. Images were acquired every 10 minutes. The black solid and dashed lines represent the average response at each time point and the gray dashed lines represent the population standard deviation of the SMAD2 response without Wnt priming (n > 200 cells per time point). Similar results were obtained in two independent experiments. **(B)** Transcriptional response of mesendoderm genes to Activin with or without WNT priming (solid and open bars, respectively). An additional sample was collected that was treated with Wnt and left untreated with Activin (-ACT). Expression in each sample was normalized to GAPDH and then to the level prior to Activin addition (T = 0 h). Data represents the mean ± S.D. for n = 3 technical replicates. n.d., not detected. Similar results were obtained in two independent experiments.

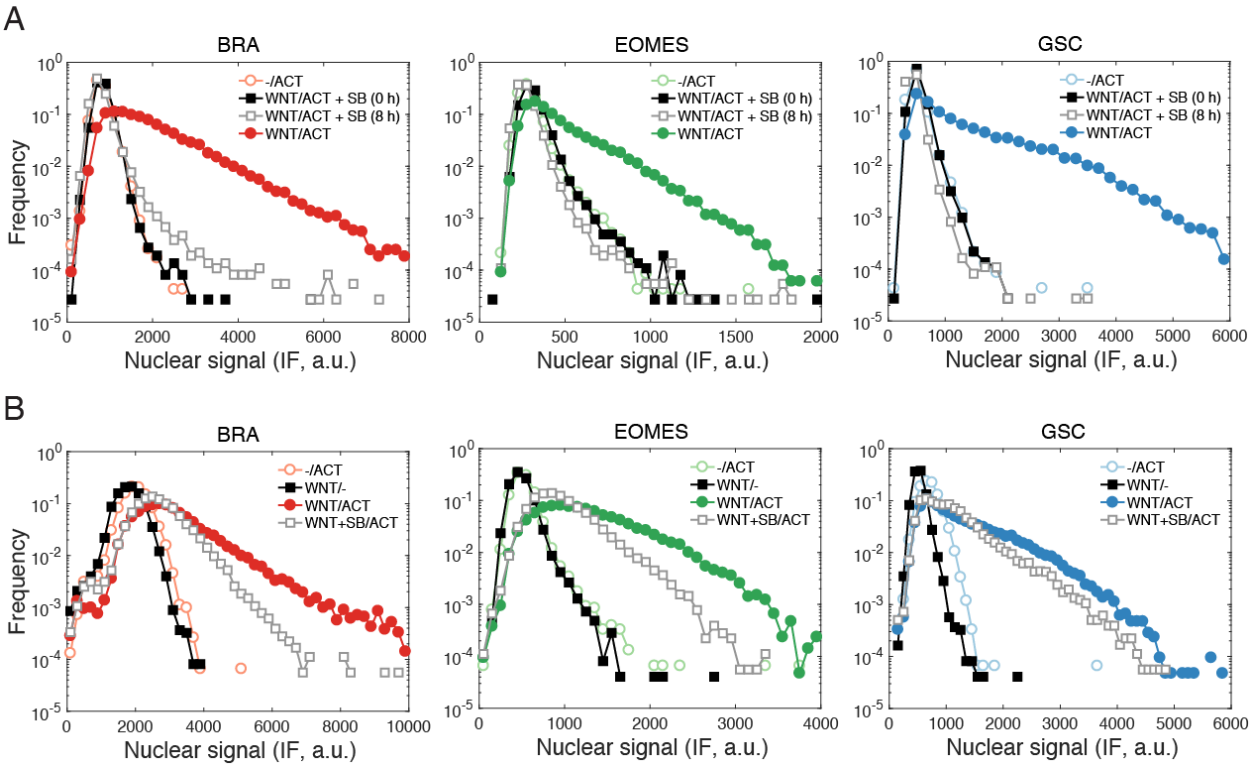


Figure 3.7 Activin/SMAD signaling is dispensable for Wnt priming
(A) Cells were cultured for one day with Wnt (100 ng/mL). On the second day they were washed to remove Wnt and treated with Activin (10 ng/mL) or Activin followed by SB (10 μM) addition at the indicated times. Cells were fixed after 24 h of Activin stimulation (48 h total) and analyzed for BRA, EOMES, and GSC expression by immunofluorescence (IF). Marker expression is eliminated when SB is added at T = 0 (black squares) or T = 8 h (open gray squares) relative to the start of Activin stimulation. **(B)** Cells were cultured for one day with Wnt (100 ng/mL) either on its own or in combination with SB (10 μM) in order to block Activin signaling during Wnt priming. On the second day cells were washed to remove Wnt and treated with Activin (10 ng/mL). Cells were fixed at after 24 h of Activin stimulation (48 h total) and analyzed for BRA, EOMES, and GSC expression by immunofluorescence. Addition of SB during WNT priming (open gray squares) only moderately reduces marker expression relative to the standard treatment (WNT/ACT, filled circles). In (A) and (B) histograms show the nuclear IF signal quantified in individual cells (n > 5,000 cells per condition). Cells cultured without Wnt (-/ACT, open circles) represent background signal levels.

3.2 Future directions: mechanism of Wnt priming and memory

One clear future direction for this work is to further address the mechanism of Wnt priming and memory, because of the implications it has for patterning in the embryo as well as our understanding of Wnt signaling in general. Three hypotheses that we could consider are: Wnt/β-catenin signaling (1) induces a co-factor, which is maintained after the Wnt signal is removed either through protein stability or a self-sustaining positive feedback, (2) represses an inhibitory factor that blocks mesendoderm differentiation in pluripotent cells, and (3) mediates epigenetic changes, e.g. histone modifications, that make mesendodermal genes stably accessible to subsequent Activin/Smad2-mediated transcriptional activation. Similar to Smad2 signaling dynamics, Wnt/β-catenin signaling also displays a transient dose-dependent response in pluripotent hESCs (Massey et al., 2019). At the ligand concentrations we use for priming (50 - 100 ng/mL) nuclear β-catenin levels adapt on a timescale of 10 - 15 hours, which is consistent with the minimum time required for Wnt priming (Figure 3.8). Although these observations and those described in the previous section establish important timescales for Wnt priming and memory. They do not obviously favor one hypothesis over the other, and both mechanisms could play a role in establishing the Wnt-primed state.

Additionally, Wnt adaptation is not complete at the concentrations used here (Massey et al., 2019). Therefore, it is also possible that low levels of nuclear β-catenin remain in the nucleus after ligand removal and interact with SMAD2/3 to drive mesendodermal differentiation. This mechanism would be consistent with recent studies, which concluded that β-catenin and Smad2/3 interact to regulate mesendodermal gene expression (Funa et al., 2015; Singh et al., 2012; Wang et al., 2017). However, many β-catenin and Smad2/3 binding sites within key genes do not overlap indicating that other indirect interactions between these pathways are possible (Estarás et al., 2015). Nevertheless, the possibility that β-catenin is retained in nucleus should be addressed in future experiments, for example, through the use of small molecule inhibitors that disrupt β-catenin/TCF interactions or genome engineering of inducible protein degradation systems (Holland, Fachinetti, Han, & Cleveland, 2012; Yan, Li, & An, 2017).

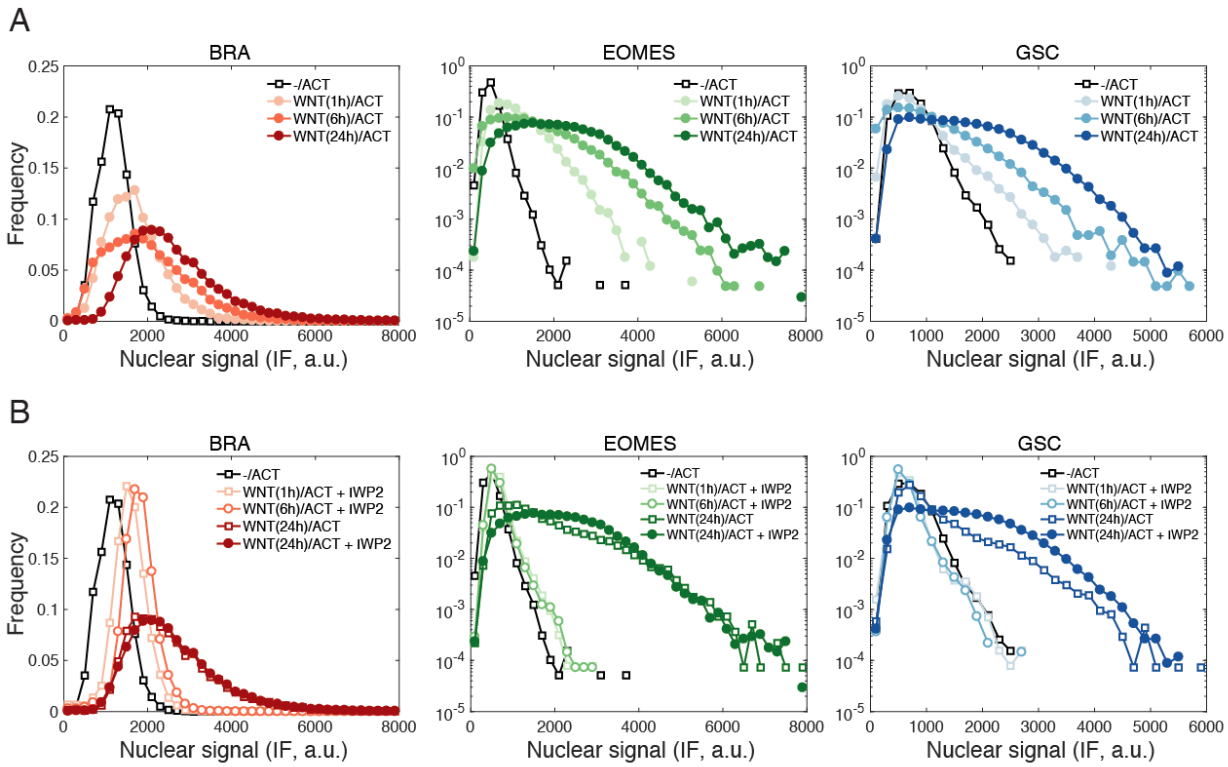


Figure 3.8 Priming is established after several hours of Wnt stimulation
(A) Cells were cultured for different amounts of time with Wnt (100 ng/mL), washed, and then stimulated for 24 h with Activin (10 ng/mL). Immunofluorescence analysis of mesendoderm marker expression suggests that as little as 1 h of stimulation appears to prime cells. **(B)** Similar experiments carried out with IWP-2 (1 μM) to block secretion of endogenously produced ligands indicates that priming takes between 6 - 24 h, as 1 and 6 h priming in the presence of IWP-2 is no longer sufficient to produce induction of BRA, EOMES, and GSC. Histograms show the nuclear IF signal quantified in single cells (n > 5,000 cells per condition). Cells cultured without Wnt, corresponding to the pluripotency condition (-/ACT), represent the background signal.

3.2.1 Epigenetic modifiers of mesendoderm differentiation

In the nucleus β-catenin interacts with many proteins through both its 12 Armadillo repeats (e.g. LEF/TCF) and its C-terminal domain (e.g. histone acetyltransferases CBP and p300, components of the mediator complex) (Valenta, Hausmann, & Basler, 2012). In the context of pre-gastrulation *Xenopus* development, Blythe and colleagues found that β-catenin recruits the H3R8 methyltransferase Prmt2 to methylate histone residues at the promoters of key dorsal genes, thereby poising them for activation at MBT. Surprisingly, knockdown of β-catenin after the 32-cell stage does not disrupt this activation, which occurs 6 cell cycles later (Blythe, Cha, Tadjuidje, Heasman, & Klein, 2010). Inspired by these experiments, which indicate that cells maintain a memory of earlier Wnt signaling at level of their chromatin architecture, we initially followed up on hypothesis (3) and investigated a possible role for epigenetic changes mediated by Wnt signaling in our system.

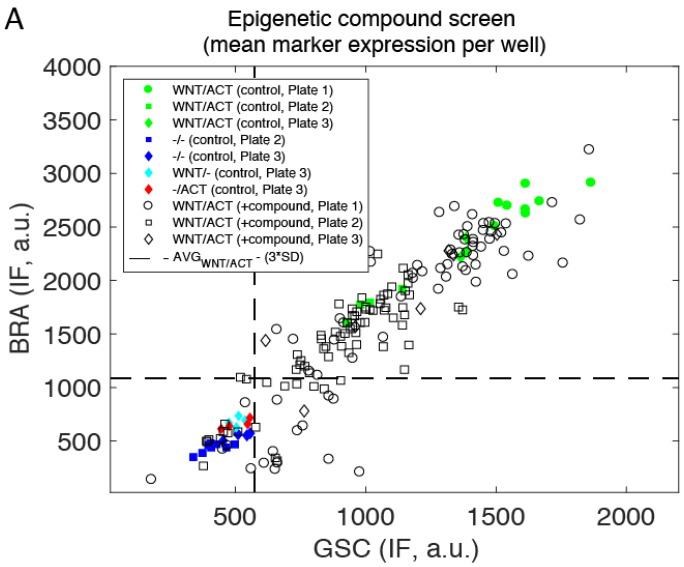
Rather than focus on the model described in *Xenopus*, we chose an unbiased approached and took advantage of a commercially-available epigenetic compound library (Selleck Chemicals) to screen for molecules that disrupt the ability of Wnt priming to change the response of the cells to subsequently supplied Activin. In our setup, the compounds were added with Wnt on the first day and both Wnt and the compounds were removed by washing before the addition of Activin on the second day. Several compounds were identified that eliminated mesendoderm differentiation, including 6 out of 8 compounds in the library that target the BET family of bromodomain-containing epigenetic readers as well as a generic bromodomain inhibitor (Figure 3.9). These compounds were validated by additional experiments, which demonstrated that over the concentration range at which BET inhibition (BETi) blocks mesendoderm induction, pluripotency was maintained (Figure 3.10). In the top hits, we also identified a Tankyrase inhibitor, NVP-TNKS656, which inhibits Wnt signaling in a manner similar to endo-IWR-1 and, thus, serves as an internal control for our screen setup (Huang et al., 2009; Shultz et al., 2013). The library did contain 5 Prmt inhibitors. However, none of them reduced marker expression, as would be expected if the mechanism were similar to that proposed for β-catenin-dependent dorsal specification in *Xenopus*.

BET proteins of which there are 4 in mammals (BRD2, BRD3, BRD4, and testis-specific BRDT) recognize acetylation of lysine residues on proteins, primarily histone H3, which is associated with open chromatin and active transcription. BETs have gained increasing attention, because of their role in cancer and their status as potential as chemotherapy targets (Belkina & Denis, 2012). However, little is known about their functions in development, particularly in germ layer patterning. One possible mechanism of Wnt priming could involve Wnt-dependent acetylation that is then read about by BETs to promote later transcriptional activation. In this case, we might have expected inhibitors of histone acetyltransferases (HATs) to also be identified as hits. Of the 4 HAT inhibitors in the library all of them resulted in cell death, so we were unable to find further support for this potential mechanism. Although one histone deacetylase (HDAC) inhibitor was identified, there were a total of 35 in the library and we did not interpret this result as strong evidence against such a model.

There are additional caveats to supposing a specific mechanism from the screen results. Although the setup was intended to screen for compounds that block Wnt priming, the inhibitors could also work by blocking expression of a Smad2/3 co-factor (hypothesis 1) or by disrupting other transcriptional activation steps. These possibilities are currently being addressed through short-term inhibitor treatment. For example, rather than maintaining the inhibitors throughout the entire 24-hour Wnt priming step, we are applying them just prior to Activin stimulation and analyzing the effects on mesendoderm induction. If BETs are required for assembly or maintenance of transcriptional activation complexes but not co-factor induction, short-term treatment will presumably be effective at blocking Wnt priming.

Lastly, unlike histone methylation, which is generally thought to be stable in the absence of continual enzymatic activity, acetylation is thought to undergo rapid turnover. Therefore, the role of acetylation readers in mesendoderm induction may be distinct from the mechanism of Wnt signaling memory. Nonetheless, BETs could be interesting from the perspective of cell identity and may serve as an important molecular player in stabilizing new transcriptional programs as cells transition between states. In support of this view, recent experiments found that BET

function is required for maintenance of pluripotency in mESCs (Di Micco et al., 2014; Finley et al., 2018; Gonzales-Cope, Sidoli, Bhanu, Won, & Garcia, 2016; W. Liu et al., 2014; T. Wu, Pinto, Kamikawa, & Donohoe, 2015). However, BET inhibition through small molecules or *Brd4* knockout can be rescued through over-expression of key transcription factors that maintain pluripotency in the absence of BET activity (Finley et al., 2018). Interestingly, a similar compensation mechanism might underlie BET inhibitor resistance in cancer. In one study, resistance in leukaemia cells emerged from the ability of cells to up-regulate components of the Wnt and TGFβ pathway, which could thus increase activity of nuclear effectors of these pathways and provide stable transcriptional activation in the absence of BET function (Fong et al., 2015). These findings also suggest a critical link between BET activity and developmental signaling pathways in disease.

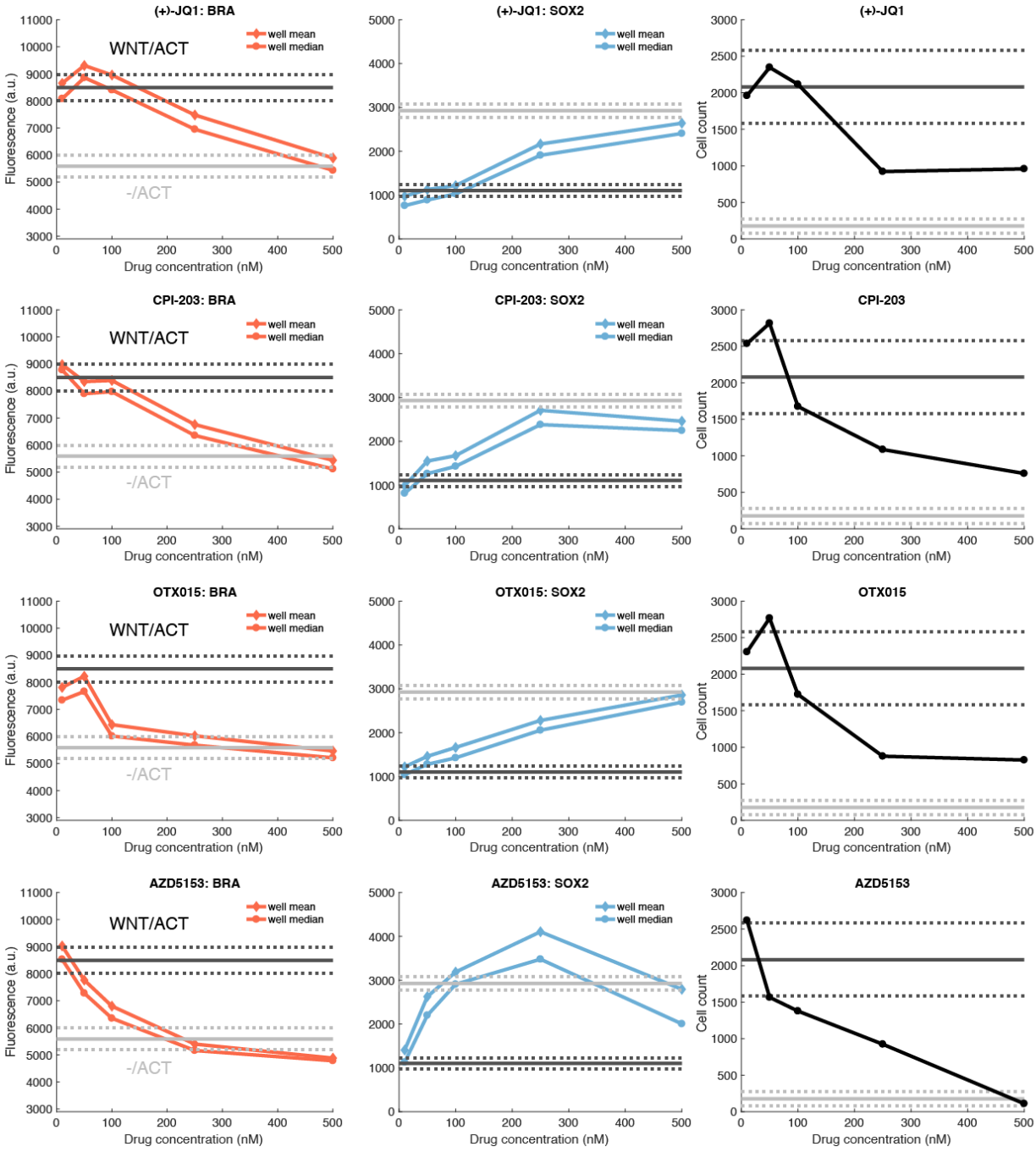


B

ID	Compound	Information
L02_F03*	SGI-1027	DNMT inhibitor (DNMT1, DNMT3A, and DNMT3B)
L01_C07	AZD1480	ATP-competitive JAK2 inhibitor
L02_A02	(+)-JQ1	**BET bromodomain** inhibitor
L02_B05	OTX015	**BET bromodomain** inhibitor (BRD2, BRD3, and BRD4; Phase 1)
L02_C03	NVP-TNKS656	highly potent, selective, tankyrase inhibitor (TNKS2)
L02_E02	I-BET-762	an inhibitor for **BET proteins**
L02_E04	CPI-203	**BET bromodomain** inhibitor with IC50 of 37 nM for BRD4.
L02_F04	MS436	a selective **BET bromodomain** inhibitor
L02_H02	Bromosporine	Bromosporine is a broad spectrum inhibitor for bromodomains
L02_H03	RG2833 (RGFP109)	HDAC inhibitor (HDAC1 and HDAC3)

Figure 3.9 Screen reveals BETs as regulators of mesendoderm differentiation in hESCs
(A) We carried out a small-scale screen in 96-well plates using 181 epigenetic-related compounds with known targets, most of which are inhibitors. The compounds (2 μM) were added with Wnt (50 ng/mL) on the first day, and both Wnt and the compounds were removed by washing before the addition of Activin (10 ng/mL) on the second day. Cells were fixed after 24 h of Activin stimulation and analyzed for BRA, EOMES, and GSC expression by immunofluorescence. Marker expression was quantified in individual cells, and the mean per cell BRA fluorescence in each well is shown plotted against the mean GSC fluorescence. Hits were identified as those compounds that reduced marker expression by 3 standard deviations (3*SD, dashed lines) below the fluorescence in positive or mesendoderm control wells (WNT/ACT, green symbols). This cutoff identified wells that had fluorescence levels similar to various negative control wells: -/- (dark blue symbols), WNT/- (light blue symbols), and -/ACT (red symbols). **(B)** Table of screen hits identified by the 3*SD and cell number cutoff. BET inhibitors are boxed. Bolded IDs indicate wells that had > 800 cells. The average number of cells in control wells was 1100 ± 620, and all hits had at least 100 cells. *SGI-1027 increased marker expression above WNT/ACT levels. However, cells had an abnormal nuclear morphology and the compound has not been further validated.

Figure 3.10 BET inhibition blocks mesendoderm induction following Wnt priming
BET inhibitors identified in the screen, three of which are shown (JQ1, CPI-203, and OTX015), and an additional inhibitor (AZD5153), were validated in their ability to block mesendoderm induction following Wnt and Activin stimulation. The inhibitors were added at the indicated concentrations with Wnt (50 ng/mL) on the first day and both Wnt and the compounds were removed by washing before the addition of Activin (10 ng/mL) on the second day. Cells were fixed after 24 h of Activin stimulation and analyzed for BRA (mesendoderm) and SOX2 (pluripotency) expression by immunofluorescence. Marker expression was quantified in individual cells in each well and the mean fluorescence per cell for each well is plotted as a function of inhibitor concentration. The mean and standard deviation of several control wells are shown as solid and dashed lines, respectively. Mesendoderm conditions (WNT/ACT, black lines) indicate the BRA+ and SOX2- levels, whereas pluripotency conditions (-/ACT, grey lines) indicate the BRA- and SOX2+ levels. Toxicity analysis in this context is complicated by the fact the cells in mesendoderm control conditions (WNT/ACT, black lines) proliferate more than those in pluripotency control conditions (-/ACT, grey lines). Comparison of cell number does not indicate a significant toxicity effect over the concentration range tested, as wells with compounds had at least as many cells as those maintained in pluripotency conditions with DMSO alone.


(+)-JQ1: BRA
(+)-JQ1: SOX2
(+)-JQ1
CPI-203: BRA
CPI-203: SOX2
CPI-203
OTX015: BRA
OTX015: SOX2
OTX015
AZD5153: BRA
AZD5153: SOX2
AZD5153
WNT/ACT
-/ACT
well mean
well median
Fluorescence (a.u.)
Cell count
Drug concentration (nM)

3.2.2 Conserved role for BETs in patterning the vertebrate embryo

Ultimately, we aim to apply the insight we gain from our reduced model system, i.e. low-density hESCs, back to the human gastruloid and potentially to vertebrate development in general. Towards this aim we tested the effects of BETi on human gastruloids. We found that BET inhibitors eliminate induction of mesoderm and endoderm but not extra-embryonic fates downstream of BMP4 (Figure 3.11 and 3.12). These results, which are consistent with our findings in low-density culture, suggest that BETi does not generically inhibit differentiation of hESCs but selectively blocks cell fates induced by Wnt and Activin/Nodal signaling. In order to address whether the role of BETs might be conserved in other vertebrates, we analyzed the effects of BETi on *Xenopus* embryos cultured with inhibitors from 2-cell to tailbud stage (Figure 3.13A). Consistent with disruption of early Wnt signaling and subsequent dorsal development, we found that drug treatment ventralized embryos to various degrees. Mild ventralization observed with CPI-203 treatment was evident in the reduced size of the head and other anterior structures in several embryos. Embryos treated with AZD5153 displayed severe ventralization, showing no signs of axial development.

In addition to Wnt signaling, dorsal-ventral patterning in *Xenopus*, particularly of the mesoderm, is also dependent on Activin/Nodal activity. One feature of the early animal cap experiments, which has since been down played in the Activin morphogen model, is that animal cap explants exhibit dorsal-ventral differences, which result from dorsal-specific nuclear localization of β-catenin by the 16-cell stage (Larabell et al., 1997; Rowning et al., 1997; Schneider, Steinbeisser, Warga, & Hausen, 1996). Only dorsal-half caps can induce all mesodermal fates in response to Activin, including organizer tissue, suggesting a critical link between Wnt and Activin signaling in this context, as well (Bolce, Hemmati-Brivanlou, Kushner, & Harland, 1992; Sokol & Melton, 1991). We found that animal caps that were uniformly dorsalized with LiCl treatment robustly induce *Xbra*, *Eomes*, and *Gsc* in response to Activin, which can be blocked by BETi (Figure 3.13B–C). Expression of dorsal β-catenin target, *Xnr3*, however, was not inhibited (J. Yang, Tan, Darken, Wilson, & Klein, 2002). Taken together, these results indicate a potentially conserved role for the BET family of epigenetic readers in vertebrate germ layer patterning mediated by Wnt and Activin/Nodal signaling. However, the molecular mechanisms by which these pathways interact and at what stage BETs play a role requires further investigation.

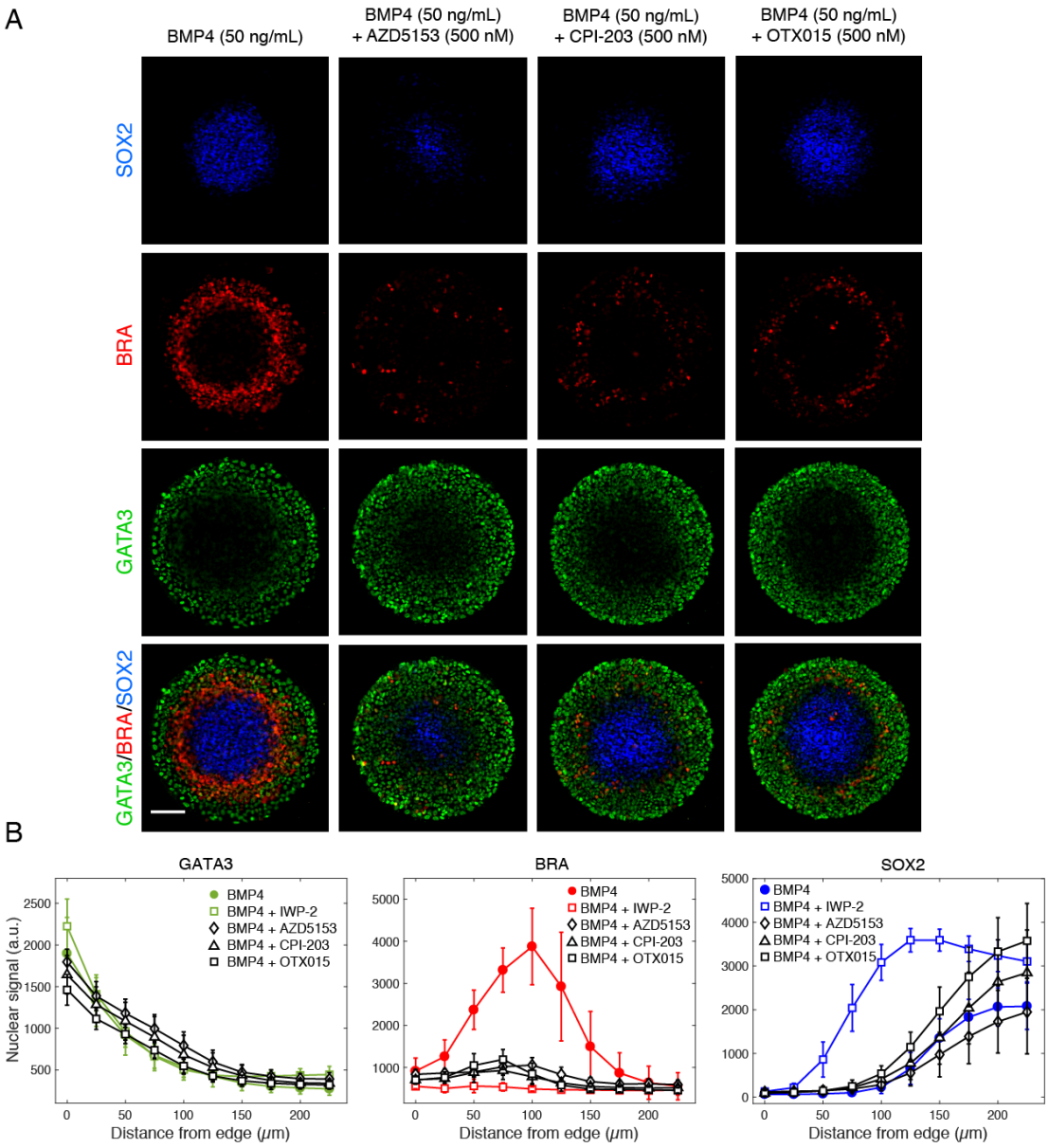


Figure 3.11 BET activity required to establish mesoderm in human gastruloids
(A) Micropatterned colonies were stimulated with BMP4 (50 ng/mL) for 48 h and analyzed by immunofluorescence for GATA3, BRA, and SOX2 expression. BET inhibitors drastically eliminate mesoderm, as indicated by the loss of BRA+ cells, but they do not block BMP4-induced GATA3+ extra-embryonic tissue. The reduction in the SOX2+ center population in some conditions may result from the loss of cells that produce TGFβ inhibitors. **(B)** Quantification of the average per cell median marker intensity as a function of the distance from the colony edge. Error bars represent the standard deviation over colonies on a single coverslip. The elimination of BRA+ cells by IWP-2 (1 μM) is shown for comparison. GATA3 induction by BMP4 is unchanged in the presence of IWP-2. However, the SOX2 domain is expanded suggesting that BETi operates distinctly from eliminating Wnt signals altogether in this system.

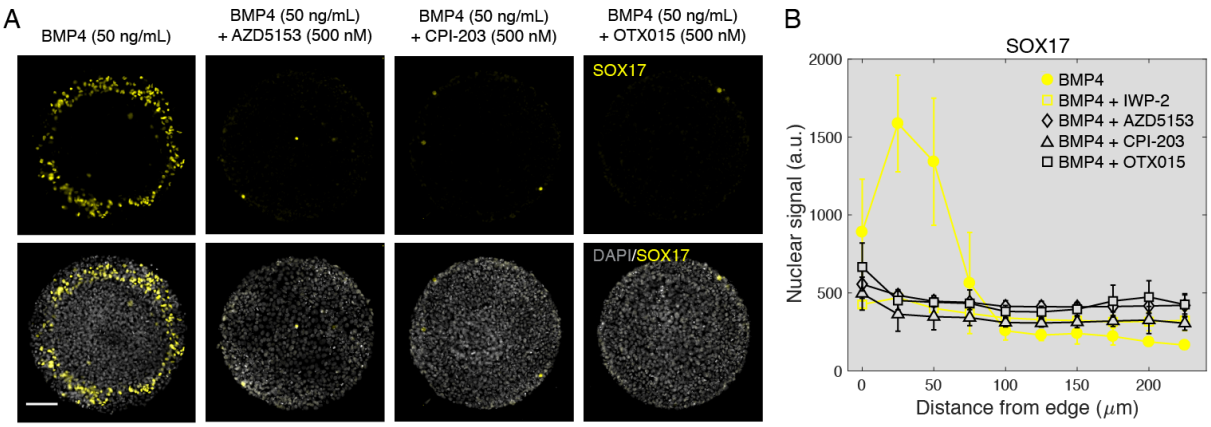
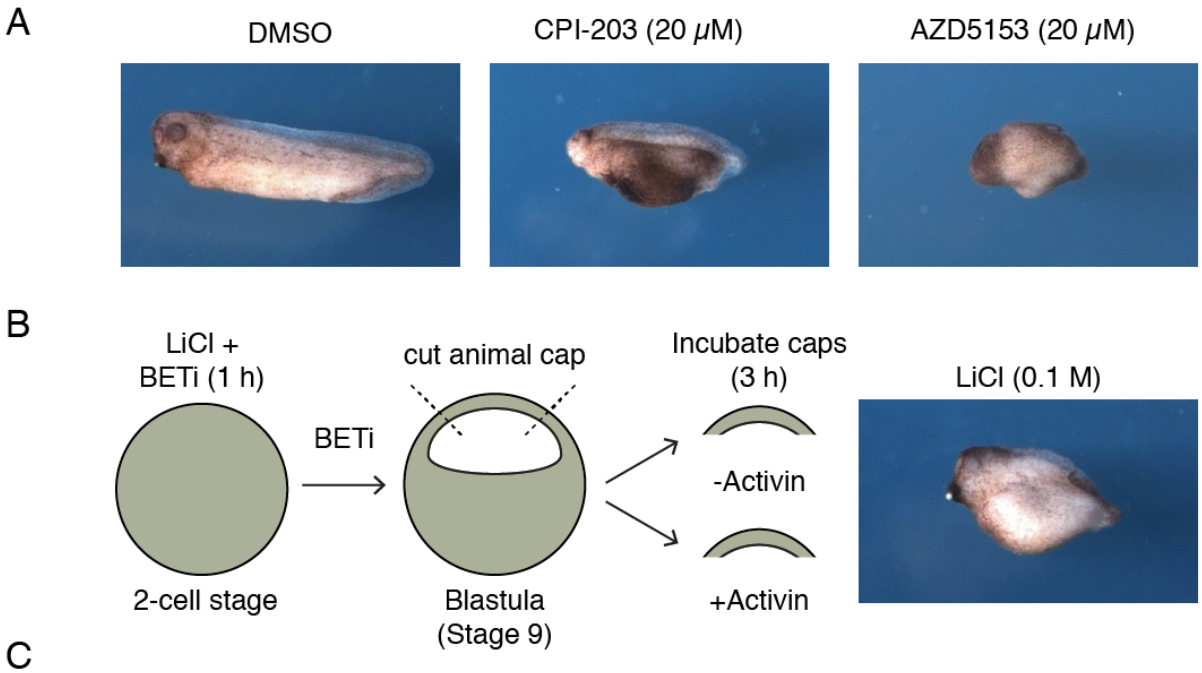


Figure 3.12 BET activity is required to establish endoderm in human gastruloids
(A) Micropatterned colonies were stimulated with BMP4 (50 ng/mL) for 48 h and analyzed by immunofluorescence for SOX17 expression. BET inhibitors drastically eliminate endoderm, as indicated by the loss of SOX17+ cells. **(B)** Quantification of the average per cell median marker intensity as a function of the distance from the colony edge. Error bars represent the standard deviation over colonies on a single coverslip. The elimination of SOX17+ cells by IWP-2 (1 μM) is shown for comparison.

Figure 3.13 BET inhibitors disrupt dorsal development in *Xenopus*
(A) *Xenopus laevis* embryos were cultured from the 2-cell to late tailbud stages with BET inhibitors or DMSO. DMSO-treated embryos show normal morphology. Embryos cultured with CPI-203 (20 μM) showed mild ventralization, as indicated by a reduction in dorsal anterior structures (4/7 embryos). Embryos cultured with AZD5153 (20 μM) showed severe ventralization, lacking any signs of an axis (9/9 embryos). **(B)** Following fertilization, embryos were dorsalized by treatment with LiCl (0.1 M for 1 h), which inhibits GSK3 and activates β-catenin throughout the embryo, as shown by the effects at the tailbud stage. Additionally, embryos were cultured with or without BET inhibitors (20 μM). At the blastula stage, animal caps were cut and cultured with or without Activin (0.5 ng/mL) for an additional 3 h, corresponding to gastrulation stages. Animal caps were frozen (8 per condition) and RNA was subsequently extracted. **(C)** RT-PCR analysis indicates that BET inhibitors block Activin-dependent induction of *Xbra*, *Eomes*, and *Gsc*, but not the induction of the dorsal β-catenin target, *Xnr3*. Expression in each sample was normalized to expression of the house keeping gene *odc* and then to the level in untreated caps. Data represents the mean ± S.D. for three technical replicates.


A
DMSO
CPI-203 (20 μM)
AZD5153 (20 μM)
B
LiCl +
BETi (1 h)
BETi
cut animal cap
Incubate caps
(3 h)
LiCl (0.1 M)
-Activin
+Activin
2-cell stage
Blastula
(Stage 9)
C
Xnr3
Eomes
Xbra
Gsc
Fold change
50.00
40.00
30.00
20.00
10.00
0.00
5.00
4.00
3.00
2.00
1.00
8.00
6.00
untr
LiCl
LiCl / ACT
LiCl + AZD
LiCl + AZD / ACT
LiCl + CPI
LiCl + CPI / ACT

3.2.3 Additional insights from ATAC- and RNA-sequencing data

We are attempting to address more generally the hypotheses enumerated at the beginning of this section through chromatin accessibility profiling using ATAC-sequencing (-seq) and bulk RNA-seq of hESCs at different stages of the Wnt and Activin protocol. In our initial analyses of chromatin accessibility, we identified a list of consensus peaks corresponding to open (nucleosome-depleted) regions of genome and found the peak intensity did not vary drastically across the conditions tested. In conditions corresponding to pluripotency, namely E7 and E7 + Activin, we observed peaks in the promoter and enhancer regions of key mesendodermal genes that overlap with Smad2/3 binding sites previously identified by ChIP-seq (Tsankov et al., 2015). The fact that these regions are already open in pluripotent hESCs is consistent with the ability to elicit transcription of these genes, albeit transient, in response to a pulse in SMAD2 signaling (Yoney et al., 2018). When we focused our analysis on regions corresponding to β-catenin binding, identified by ChIP-seq in Wnt-treated hESCs (Estarás et al., 2015), we found that these sites are also already open in pluripotent cells. However, we did observe a further increase in the ATAC-seq signal following 24 hours of Wnt stimulation that was reversed when Wnt was removed, and the cells were either stimulated with Activin or returned to E7 for an additional 24 hours (Figure 3.14). We took this reduction in ATAC signal as supporting the notion that once Wnt is washed out, β-catenin also leaves the nucleus. The BET inhibitor, CPI-203, did not block the increase in the ATAC signal at β-catenin peaks following 24 hours of Wnt stimulation. Instead, the ATAC signal in the Wnt + CPI condition actually increased relative to the E7 condition and the 24h Wnt condition (Figure 3.14). This result argues against the hypothesis that BETs are required to read out new Wnt-dependent acetylation that results in further opening of the chromatin.

From our previous RNA-seq data on Activin only treated cells, we identified a group of genes that showed transient induction and contained key regulators of mesendoderm differentiation (Yoney et al., 2018). With newly generated RNA-seq data from cells treated with Activin following Wnt priming, we observed that many of these mesendoderm-associated genes remain stably expressed in agreement with our analysis of BRA, EOMES, and GSC expression by RT-PCR (Figure 3.6). In order to identify potential co-factors that could regulate this stable expression, we carried out transcription factor binding motif analysis using the genomic sequences under ATAC-seq peaks within 10 kb of the start or end of each gene or within the introns. The most likely co-factor identified by this analysis was EOMES itself. EOMES has previously been proposed to interact with SMAD2/3 to regulate endoderm differentiation in mouse and human stem cells (Teo et al., 2011). In our conditions, EOMES expression increases at the RNA level following Wnt priming. Therefore, we plan to further investigate the possibility that EOMES mediates Wnt priming and memory through knockout and inducible expression studies. If EOMES is indeed the master regulator of the transition from pluripotency to mesendoderm fate in response to Wnt priming followed by Activin stimulation, we expect its knockout to block differentiation. On the other hand, forced expression should bypass the requirement for Wnt priming, as well as potentially modulate the timing of Wnt memory establishment and maintenance.

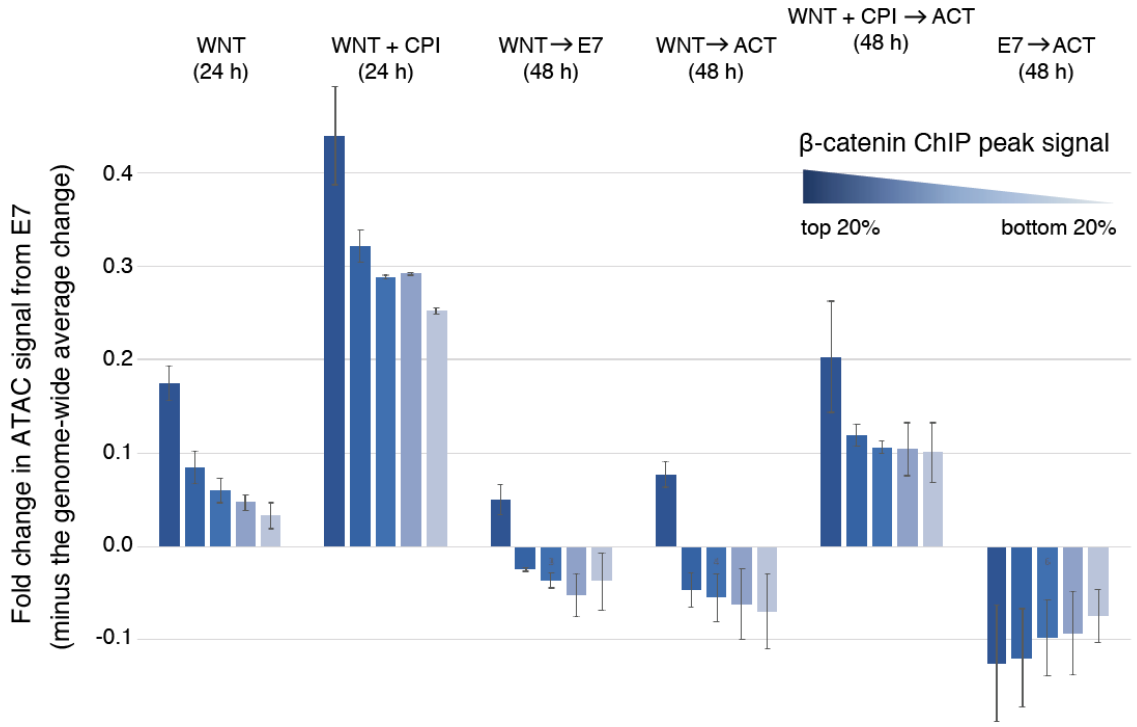


Figure 3.14 Analysis of ATAC-seq signal at β-catenin binding sites
β-catenin binding sites that were previously identified by ChIP-seq (Estarás et al., 2015) were grouped by signal intensity in increments of 20% of the total number of peaks, with the darkest blue bar for each condition corresponding to the top 20% of peaks and the lightest blue bar corresponding to the bottom 20%. For each condition the change in the average ATAC-seq signal within the different groups of β-catenin peaks was calculated relative to E7 (normalized to average genome wide change). This analysis revealed that Wnt stimulation (50 ng/mL, 24 h) increases the ATAC signal at β-catenin binding sites, which is further enhanced by CPI-203 (1 μM). When Wnt is removed the signal is reduced below the levels in E7 except for the strongest β-catenin peaks, which also show a reduction. Error bars represent the standard deviation across the peaks in each group.

CHAPTER 4. CONCLUSIONS

Stem cell-based systems offer the potential to study development with high spatial and temporal resolution, thus providing new insights into the dynamics of signaling, cell fate acquisition, and tissue patterning. In my thesis work I took advantage of hESCs, CRISPR/Cas9 genome engineering technology, and controlled cell culture methods to investigate these processes in the context of gastrulation. Using live reporters of SMAD1, SMAD2 and SMAD4, we found that BMP/SMAD1 signaling is stable, Activin/SMAD2 signaling is transient (or adaptive), and that co-SMAD4 follows the dynamics of the receptor-associated SMADs. In response to Activin, hESCs transiently induce mesendodermal gene expression, without generating mesendodermal fates, and the cells return to the state of pluripotency. However, the expression of Activin/Nodal inhibitors remained elevated and plausibly buffer the additional Activin signals, as expected from prior theory (Francois & Siggia, 2008). Pre-presentation of Wnt stabilizes transcription and fate acquisition without modifying SMAD2 dynamics. Thus, a transient signaling response is compatible with a stable cell fate change. Both pluripotency maintenance and mesendodermal fate acquisition are not affected by the loss of SMAD3, although our results cannot rule out possible redundancy between SMAD2 and SMAD3. Our data provide evidence for a previously undetected level of signal integration that implies the presence of a cellular memory. This memory operates not at the level of the SMAD2 signaling dynamics but by some other means to change the response of cells to Activin, some possible mechanisms of which are outlined in Chapter 3.

The Activin/Nodal branch of the TGFβ pathway has been shown to be necessary for maintenance of pluripotency, but not sufficient to induce mesendoderm in hESCs, except when influenced by modifiers of other signaling pathways (D'Amour et al., 2005; Funa et al., 2015; McLean et al., 2007; Singh et al., 2012). Our study shows that pre-stimulation of hESCs with Wnt, endows them with a memory that enables a graded response to subsequently applied Activin; simultaneous exposure to Wnt and Activin is not required. Furthermore, Wnt priming alone is not sufficient to elicit a stable fate change since in minimal media, without Activin, the cells maintain pluripotency. Interestingly, short Wnt activation combined with TGFβ inhibitors is used to generate neural crest or neuromesodermal progenitors from hESCs that go on to form presomitic mesoderm (Diaz-Cuadros et al., 2020; Funa et al., 2015). Our results combined with these findings indicate that Wnt may function as a permissive signal allowing cells to exit pluripotency but with the potential to take on several different fates depending on other instructive signals. This notion, thus, challenges the traditional view of Wnt as an instructive signal itself and establishes an important example of a permissive signal guiding early development (Slack, 1991). Furthermore, the ability of embryonic cells to record Wnt signals may be a broadly conserved and fundamental aspect of animal development, as shown by experiments in *Drosophila* in which the authors observed a morphogen effect for Dpp (a BMP homologue) after cells have lost contact with the source of Wnt (Alexandre et al., 2013).

In the context of vertebrate development, our findings force the reevaluation of the traditional literature regarding the sufficiency of the Activin/Nodal pathway for mesendodermal induction and patterning. A review of the evidence in model systems reveals that in all the experimental settings, at least some of the cells were either still under the influence, or had been previously exposed to Wnt signaling before or during Activin/Nodal signaling. For example, cells of the

Xenopus animal cap derived from the blastula stage embryo have been under maternal Wnt influence, hours before Activin is presented to the explants. This influence is the consequence of cortical rotation that occurs after sperm entry, and activates the Wnt pathway on the dorsal side of the embryo, as evidenced by the dorsal-specific nuclear localization of β-catenin by the 16-cell stage (Larabell et al., 1997; Rowning et al., 1997; Schneider et al., 1996). It is clear from the asymmetric elongation of the animal cap explants in response to Activin that the prior Wnt exposure sets up a dorsal-ventral pre-pattern that affects the response. It is also clear that the dorsal and ventral regions of the animal cap when separated respond differently to Activin (Bolce et al., 1992; Sokol & Melton, 1991). Although some mesoderm is induced in the ventral caps, only cells of the dorsal cap that have seen Wnt signals undergo Activin-mediated induction of GSC+ mesendoderm. These experiments demonstrate that both pathways are required for complete mesendodermal patterning without explicitly distinguishing their temporal relationship.

Our study revises this point of view by bringing a temporal order. We suggest that only cells primed by Wnt signals can respond to Activin to specify the full range of mesendodermal fates. Even in the extensively studied *Xenopus* embryo, it is still debated whether in the late blastula stage marginal zone, prior to the onset of gastrulation, an Activin/Nodal gradient already defines the medial to lateral mesendodermal fates or whether there is simply a bipartite division into dorsal organizer and ventral mesoderm (Harland & Gerhart, 1997; Smith, 2009). This debate also translates into whether Activin/Nodal is sufficient to pre-pattern the mesendoderm or whether complete patterning requires prior Wnt exposure present only in the dorsal part of the embryo. Controversies persist, since the embryo is rapidly developing, there are no live reporters of signaling, and fates are often assayed in early gastrulation rather than in the late blastula. In the mouse one of the earliest manifestation of the streak is the proximal-posterior expression of Wnt, which extends distally and anteriorly. By mid-streak stage, Nodal signaling is highest in the Node (Tam & Loebel, 2007). Thus, cells that leave the streak at various proximal-distal positions and Nodal levels, manifestly have been exposed to Wnt first (Tam & Behringer, 1997). This interpretation, however, does not eliminate the co-requirement for Nodal and Wnt, but rather suggests that instructive Wnt, required for mesendodermal differentiation, occurs temporally up-stream of Activin/Nodal signaling.

Lastly, as part of on-going work in which we have attempted to dissect the mechanism of Wnt priming and memory, we uncovered a role for the BET family of epigenetic readers in human mesendoderm induction and patterning, which is potentially conversed in frogs. The Wnt and Activin protocols developed here in combination with the human gastruloid system establish an experimental context in which to further investigate the mechanisms by which inductive signaling pathways intersect with the epigenetic landscape. These efforts will be important not just for our understanding of embryonic development but for also disease contexts, such as cancer, in which developmental mechanisms are frequently coopted to drive disease progression.

CHAPTER 5. MATERIALS AND METHODS

Table 5.1. Key reagents and resources

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (*Homo sapiens*, XX)	RUES2	US National Institutes of Health, human ESC registry	human ESC registry no. 0013; RRID:CVCL_VM29	Human embryonic stem cell line
cell line (*Homo sapiens*, XX)	RUES2-RFP-SMAD1	this paper		CRISPR/Cas9-modified human embryonic stem cell line
cell line (*Homo sapiens*, XX)	RUES2-mCit-SMAD2	this paper		CRISPR/Cas9-modified human embryonic stem cell line
cell line (*Homo sapiens*, XX)	RUES2-GFP-SMAD4	PMID: 28760810		CRISPR/Cas9-modified human embryonic stem cell line
cell line (*Homo sapiens*, XX)	RUES2-SMAD3-/- clone #1	this paper		CRISPR/Cas9-modified human embryonic stem cell line
cell line (*Homo sapiens*, XX)	RUES2-SMAD3-/- clone #2	this paper		CRISPR/Cas9-modified human embryonic stem cell line
cell line (*Homo sapiens*, XY)	RUES1	US National Institutes of Health, human ESC registry	human ESC registry no. 0012; RRID:CVCL_B809	Human embryonic stem cell line
antibody	anti-Brachyury (goat polyAb)	R&D Systems	Cat. #: AF2085; RRID:AB_2200235	IF (1:300)
antibody	anti-Brachyury (rabbit mAb)	R&D Systems	Cat. #: MAB20851	IF (1:200)
antibody	anti-Cdx2 (mouse mAb)	Abcam	Cat. #: ab15258; RRID:AB_2077042	IF (1:50)
antibody	anti-Eomes (mouse mAb)	R&D Systems	Cat. #: MAB6166; RRID:AB_10919889	IF (1:200)
antibody	anti-Lamin B1 (rabbit polyAb)	Proteintech	Cat. #: 12987-1-AP; RRID:AB_2136290	WB (1:2000)
antibody	anti-Goosecoid (goat polyAb)	R&D Systems	Cat. #: AF4086; RRID:AB_2114650	IF (1:100)
antibody	anti-Nanog (goat polyAb)	R&D Systems	Cat. #: AF1997; RRID:AB_355097	IF (1:200)
antibody	anti-Oct3/4 (mouse mAb)	BD Biosciences	Cat. #: 611203; RRID:AB_398737	IF (1:400)
antibody	anti-Smad2 (rabbit mAb)	Cell Signaling	Cat. #: 3122; RRID:AB_10697649	IF (1:200), WB (1:1000)
antibody	anti-Smad2/3 (mouse mAb)	BD Biosciences	Cat. #: 610842; RRID:AB_398161	IF (1:100)
antibody	anti-Sox17 (goat polyAb)	R&D Systems	Cat. #: AF1924; RRID:AB_355060	IF (1:200)
antibody	anti-Sox2 (rabbit mAb)	Cell Signaling	Cat. #: 3579; RRID:AB_2195767	IF (1:200)
peptide, recombinant protein	recombinant human/mouse/rat Activin A	R&D Systems	Cat. #: 338-AC/CF	

peptide, recombinant protein	recombinant human BMP4	R&D Systems	Cat. #: 314-BP	
peptide, recombinant protein	recombinant mouse Wnt3a	R&D Systems	Cat. #: 1324-WN	
peptide, recombinant protein	recombinant human Laminin-521	BioLamina		
chemical compound, drug	endo-IWR-1	Tocris	Cat. #: 3532	
chemical compound, drug	IWP-2	Stemgent	Cat. #: 04-0034	
chemical compound, drug	SB431542	Stemgent	Cat. #: 04-0010	
chemical compound, drug	Y-27632	Abcam	Cat. #: Ab120129	
chemical compound, drug	CHIR99021	EMD Millipore	Cat. #: 361559	
chemical compound, drug	AZD5153	Selleck Chemicals	Cat. #: S8344	
chemical compound, drug	CPI-203	Selleck Chemicals	Cat. #: S7304	
chemical compound, drug	(+)-JQ1	Selleck Chemicals	Cat. #: S7110	
chemical compound, drug	OTX015	Selleck Chemicals	Cat. #: S7360	
chemical compound, drug	Epigenetics compound library	Selleck Chemicals	Cat. #: L1900	All compounds were diluted to 2 μM in DMSO
software, algorithm	MATLAB	MathWorks	RRID:SCR_001622	
software, algorithm	StemCellTracker	https://github.com/ChristophKirst/StemCellTracker		Image segmentation and signal analysis
software, algorithm	R	https://www.r-project.org/		
software, algorithm	DESeq2 Bioconductor package	PMID: 25516281		Differential gene expression analysis
software, algorithm	PWMEnrich Bioconductor package	https://bioconductor.org/packages/release/bioc/html/PWMEnrich.html		Motif enrichment analysis
software, algorithm	GOseq Bioconductor package	PMID: 20132535		Statistical analysis of over/under represented categories
software, algorithm	AME	PMID: 25953851		Motif enrichment analysis
software, algorithm	Tide	PMID: 25300484		Sequence trace decomposition
software, algorithm	Cluster 3.0	PMID: 14871861		Clustering
software, algorithm	Java TreeView	PMID: 15180930		Data visualization

Human embryonic stem cell culture

Experiments were performed with the RUES2 hESC line (XX female; US National Institutes of Health, human ESC registry no. 0013) or CRISPR/Cas9 edited cell lines based on RUES2. Key experiments were repeated with the RUES1 hESC line (XY male; US National Institutes of Health, human ESC registry no. 0012). RUES1 and RUES2 have been authenticated by STR profiling and both tested negative for mycoplasma contamination. hESCs were grown in HUESM medium that was conditioned by mouse embryonic fibroblasts and supplemented with 20 ng/mL bFGF (MEF-CM). Cells were grown on tissue culture dishes (BD Biosciences, San Jose, CA) for maintenance and expansion at 37 °C and 5% CO_2. Dishes were coated overnight at 4 °C with Geltrex (Thermo Fisher Scientific, Waltham, MA) diluted 1:40 in DMEM/F12 and then incubated at 37 °C for at least 20 minutes before passaging. Cells were passaged as aggregates using Gentle Cell Dissociation Reagent (STEMCELL Technologies, Vancouver, Canada).

Micropatterned cell culture

Individual micropatterned coverslips (CYTOO, Grenoble, France) were washed one time with water for 5 minutes at room temperature (RT) to activate the surface according to the manufacturers recommendation. Coverslips were then coated at 37 °C for 2 hours with 20 µg/mL Laminin-521 (BioLamina, Sundbyberg, Sweden) in 0.5 mL PBS +Mg/+Ca. The laminin was then removed by serial dilutions without allowing the coverslip to dry (1:4 dilution in PBS –Mg/–Ca, six times). Chips were seeded immediately or stored overnight at 4 °C in 2 mL PBS –/– and seeded on the following day. Cell seeding was performed as follows. Cells growing in MEF-CM were washed once with PBS +/+ and dissociated to single cells with Accutase (STEMCELL Technologies). Cells were centrifuged and 600,000 cells were resuspended in 2 mL of MEF-CM supplemented with 10 µM Rock inhibitor (Y-27632, Abcam, Cambridge, MA). The cell suspension was then placed over the coverslip in a 35-mm tissue culture dish. The sample was left unperturbed for 10 minutes at RT in order to achieve homogeneous seeding of the cells throughout the chip before being moved to the incubator. After 2 hours, the medium was replaced with MEF-CM without Rock inhibitor (RI) and the cells were incubated overnight. On the following day, the medium was changed to MEF-CM with additional ligands, BMP4, Wnt3a, or Activin A (R&D Systems, Minneapolis, MN) with or without 10 µM SB431542 (Stemgent, Lexington, MA). For experiments in TeSR-E7 (STEMCELL Technologies), cells were moved to E7 medium the day after seeding and incubated for an additional 24 hours before adding ligands. Live imaging was carried out in E7 imaging medium, which was prepared with FluoroBrite DMEM (Thermo Fisher Scientific) according the published protocol for E8 (Beers et al., 2012). A step-by-step version of the micropatterned cell culture protocol with trouble-shooting tips is also available (Deglincerti et al., 2016).

Low-density cell culture

Optical-quality plastic tissue culture dishes or 24-well plates (ibidi, Martinsried, Germany) were coated with 10 µg/mL Laminin-521 in PBS +/+ for 2 hours at 37 °C or overnight at 4 °C. Single cells were collected as described for micropatterned cell culture and dishes were seeded with 50,000 cells resuspended in 1 mL E7 supplemented with 10 µM RI, which was added to 1 mL medium in each dish to bring the total volume to 2 mL. 24-well plates were seeded with 5,000 cells per well in 0.5 mL of medium, which was added to 0.5 mL to bring the total volume to 1 mL per well. The samples were incubated overnight. Plates were additionally placed inside an

empty tip box inside the incubator, and the bottom chamber of the tip box was filled with water. This improved the uniformity of the seeding, perhaps by limiting evaporation. Without this additional protection cells tended to accumulate at the well edges and in a small patch in the center. On the following day, the medium was changed to E7 supplemented with ligands or small molecules. For the 2-day protocol in which cells were switched from E7 +/– Wnt3a (100 ng/mL) to E7 +/– Activin A (10 ng/mL), the samples were washed with PBS +/+ before adding fresh medium. Live imaging was carried out in E7 imaging medium. RI (10 μM) was maintained throughout the duration of the experiment. Other small molecules were used at the following concentrations and were replaced every 24 hours: 2.5 μM CHIR99021 (EMD Millipore), 1 μM IWP-2 (Stemgent), 1 μM endo-IWR-1 (Tocris), and 10 μM SB431542 (Stemgent).

Compound screen

The screen was carried out using an epigenetic compound library (Selleck Chemicals). Optical-quality plastic 96-well plates (Greiner Bio-One) were coated with 5 μg/mL Laminin-521 in PBS +/+ overnight at 4 °C. Single cells were collected as described for micropatterned cell culture and the plates were seeded with 5,000 cells per well in 100 μL E7 supplemented with 10 μM RI, which was added to 100 μL to bring the total volume well 200 μL per well. The plates were incubated overnight inside empty tip boxes inside the incubator, and the bottom chamber of the tip boxes was filled with water. On the following day, 100 uL was removed from each well and 100 uL of fresh E7 medium containing Wnt (50 ng/mL final concentration) and the compounds (2 μM final concentration) was added. Compounds were diluted in DMSO and the same volume of DMSO was added to control wells. Aspirating the entire volume of medium from the wells resulted in a drastic loss of cells. Therefore, on the second day the wells were washed by adding and removing 150 μL of E7 +/- Activin (10 ng/mL). 180 uL was then removed and 200 of fresh E7 +/- Activin (10 ng/mL) was added back. RI (10 μM) was maintained throughout the duration of the experiment.

Immunofluorescence

Cells on dishes or coverslips were rinsed once with PBS –/–, fixed with 4% paraformaldehyde (Alfa Aesar, Thermo Fisher Scientific, Tewksbury, MA) for 20 minutes at RT, and then rinsed twice and stored in PBS –/–. Cells were blocked and permeabilized with blocking buffer (2% bovine serum albumin and 0.1% Triton X-100 in PBS –/–) for 30 minutes at RT. Cells were incubated with primary antibodies in blocking buffer overnight at 4 °C and then washed three times with 0.1% Tween-20 in PBS –/– (PBST). Cells were incubated with secondary antibodies (diluted 1:1000): Alexa Fluor 488, 555, or 647-conjugated from (Invitrogen Molecular Probes, Thermo Fisher Scientific) and DAPI nuclear stain (1:10,000 dilution) in blocking buffer for 30 minutes at RT, and then washed twice with PBST and once with PBS –/–. Dishes were stored and imaged in PBS –/–. Coverslips were mounted on slides using Fluoromount-G mounting medium (SouthernBiotech, Birmingham, AL).

Imaging

Wide-field images of fixed samples were acquired on an Olympus IX-70 inverted microscope with a 10x/0.4 numerical aperture objective lens. Tiled image acquisition was used to acquire images of large areas of dishes or coverslips in four channels corresponding to DAPI and Alexa Fluor 488, 555, and 647. Live imaging was performed on a spinning disk confocal microscope equipped with 405-, 488-, and 561-nm lasers and an environmental chamber (CellVoyager

CV1000, Yokogawa). Images were acquired every 10 minutes with a 20x/0.75 numerical aperture objective lens. The cells were maintained at 37 °C and 5% CO_2 during live imaging.

Generation of SMAD1 and SMAD2 reporter cell lines

For the tagRFP-SMAD1 reporter cells, CRISPR/Cas9 mediated genome engineering was used to fuse a cassette containing a blasticidin resistance gene (BsdR), a T2A self-cleaving peptide, and a tagRFP fluorescent protein onto the N-terminus of SMAD1, so that the locus produces both a tagRFP-SMAD1 fusion protein together with BsdR. Similarly, for the mCitrine-SMAD2 reporter line, CRISPR/Cas9 was used to fuse a cassette containing a puromycin resistance gene (PuroR), a T2A self-cleaving peptide, and an mCitrine fluorescent protein onto the N-terminus of SMAD2, so that the locus produces both an mCitrine-SMAD2 fusion protein together with PuroR.

RUES2 hESCs were nucleofected with a pX335 plasmid (Cong et al., 2013) that co-expresses the nickase version of Cas9 and the specific sgRNA of interest (protospacer sequences: 5'-GCAGCACTAGTTATACTCCT-3' for SMAD1 and 5'-GGACGACATGTTCTTACCAA-3' for SMAD2), as well as the appropriate homology donor plasmid. Nucleofection was carried out using the Cell Line Nucelofector Kit L (Lonza, Walkersville, MD) and the B-016 setting of a Nucleofector II instrument. Nucleofected cells were plated into MEF-CM supplemented with 10 μM RI. After 4 days, blasticidin or puromycin was added for 7 days to select for cells that had been targeted. Cells that survived selection were passaged as single cells using Accutase, plated in MEF-CM supplemented with 10 μM RI, and allowed to grow into colonies. Colonies arising from a single cell were handpicked, expanded, and screened for correct targeting by PCR amplification of the genomic region and Sanger sequencing. Correctly targeted clones were subsequently transfected with ePiggyBac plasmids containing either H2B-mCitrine or H2B-mCherry cassettes to enable nuclear labeling for cell tracking (Lacoste, Berenshteyn, & Brivanlou, 2009). Individual clones were again isolated and controlled for normal karyotype (G-banding) and pluripotency maintenance.

Generation of *SMAD3*-/- cell lines

RUES2 hESCs were nucleofected with a pX330 plasmid (Cong et al., 2013) that co-expressed the wild-type version of Cas9 and one of two sgRNAs targeting SMAD3 (protospacer sequences: 5'-CCACCAGATGAACCACAGCA-3' and 5'-TTATTATGTGCTGGGGACAT-3' for sgRNA #1 and #2, respectively). The sgRNAs were designed to target the first or second coding exon shared by all SMAD3 isoforms, a strategy that was successfully used to knockout SMAD3 in human primary cell lines (Voets et al., 2017). Two different sgRNAs were used to control for off-target effects. We modified the pX330 plasmid to also express a puromycin-2A-EGFP cassette to enrich for cells that had been successfully nucleofected. Nucleofection was carried out as described above and cells were plated in MEF-CM supplemented with 10 μM RI. On the following day puromycin was added for 24 hours. Cells that survived selection were allowed to recover for several days. Cells were then passaged as single cells using Accutase, plated in MEF-CM supplemented with μM RI, and allowed to grow into colonies. Colonies arising from a single cell were handpicked, expanded, and screened for correct targeting by PCR amplification of the genomic region and Sanger sequencing. The resulting chromatograms were decomposed using the Tide web-based tool (Brinkman, Chen, Amendola, & van Steensel, 2014)

RT-PCR

RUES2 cells were seeded in 6-well plates (100,000 single cells per well) in E7 medium supplemented with 10 μM RI and incubated overnight. On the following day the medium was changed to E7 supplemented with 10 μM RI with or without 10 ng/mL Activin A. Samples (3 pooled-wells) were collected in 1 mL Trizol at 0, 2.5, 4, 6 and 12 hours. The 2-day protocol was carried out similarly: the medium was changed on the day after seeding to E7 supplemented with 10 μM RI with or without 100 ng/mL Wnt3a. On the second day following seeding the wells were washed once with PBS and the media was changed to E7 supplemented with 10 μM RI with or without 10 ng/mL Activin A. Samples were collected in 1 mL Trizol before the addition of Wnt3a, after Wnt3a treatment (referred to as 0 h), and at 2.5, 4, 8 , and 12 hours after the addition of Activin A. Total cellular RNA for each sample was extracted using the RNeasy Mini Kit (QIAGEN, Germantown, MD) and cDNA was synthesized using the Transcriptor First Strand cDNA Synthesis Kit (Roche). RT-PCR for selected genes was performed on 3 technical replicates using the LightCycler 480 SYBR Green I Master mix in a LightCycler 480 instrument (Roche, Basel, Switzerland). Primers were designed using Primer-BLAST (Ye et al., 2012) or obtained from qPrimerDepot (Cui, Taub, & Gardner, 2007) or from previously published sequences (Mendjan et al., 2014). Primer sequences and source are listed in Table 5.2.

Table 5.2. Human RT-PCR primer sequences and source

Gene symbol	Forward primer	Reverse primer	Source
NANOG	TCCAACATCCTGAACCTCAGC	ACCATTGCTATTCTTCGGCCA	Primer-BLAST
OCT4	AAACCCACACTGCAGCAGAT	TGTGCATAGTCGCTGCTTGA	Primer-BLAST
SOX2	TACAGCATGATGCAGGACCA	CCGTTCATGTAGGTCTGCGA	Primer-BLAST
NODAL	AGACATCATCCGCAGCCTAC	CAAAAGCAAACGTCCAGTTCT	Primer-BLAST
GAPDH	AATCCCATCACCATCTTCCA	TGGACTCCACGACGTACTCA	Primer-BLAST
SMAD7	CCAGGCTCCAGAAGAAGTTG	CCAACTGCAGACTGTCCAGA	qPrimerDepot
LEFTY1	CTCCATGCCGAACACCAG	GGAAAGAGGTTCAGCCAGAG	qPrimerDepot
LEFTY2	TCAATGTACATCTCCTGGCG	CTGGACCTCAGGGACTATGG	qPrimerDepot
BRA	CGTTGCTCACAGACCACAG	ATGACAATTGGTCCAGCCTT	qPrimerDepot
GSC	GAGGAGAAAGTGGAGGTCTGGTT	CTCTGATGAGGACCGCTTCTG	Mendjan et al.
EOMES	CACATTGTAGTGGGCAGTGG	CGCCACCAAACTGAGATGAT	Mendjan et al.

Total RNA-sequencing

RUES2 cells were seeded in 6-well plates (200,000 single cells per well) in E7 medium supplemented with 10 μM RI and incubated overnight. On the following day the media was changed to E7 supplemented with 10 μM RI with or without Activin (10 ng/mL). Samples (3 pooled-wells) were collected in 1 mL Trizol at 1, 2.5, 4, 8 and 12 hours for the Activin-treated conditions and after 0, 6 and 12 hours for the untreated conditions (to be used as negative controls). Total cellular RNA for each sample was extracted using RNeasy Mini Kit (QIAGEN) and 2 ug of total RNA was used to prepare each library. Library construction was performed with the TruSeq RNA Library Preparation Kit (Illumina, San Diego, CA) as per the manufacturer's instructions and sequenced on the Illumina HiSeq 2500 platform. Raw reads were mapped to hg19 using STAR aligner, and the gene read counts were normalized using the DESeq2 Bioconductor package (Love, Huber, & Anders, 2014). Library preparation, sequencing, and mapping were performed by the New York Genome Center (New York, NY, USA). All raw data files are available from the GEO database (accession number GSE111717).

ATAC-sequencing

Chromatin accessibility profiling was carried out using the Omni-ATAC-sequencing protocol as an attempt to reduce mitochondrial DNA in our samples (Corces et al., 2017). We obtained single cells using Gentle Cell Dissociation Reagent (STEMCELL Technologies), rather than Accutase or other enzyme-based reagents, which we found disrupted downstream processing steps. Cells (100,000 per condition) were resuspended in 50 µL lysis buffer (0.1% NP-40, 0.1% Tween-20, 0.01% Digitonin) prepared in resuspension buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, and 3 mM $MgCl_2$ in water) for 3 minutes on ice. We optimized the lysis time so that the outer membrane was disrupted as indicated by Trypan blue staining of the nuclei but that the nuclei remained intact. The lysis buffer was removed by washing with 0.1% Tween-20 in resuspension buffer, and the transposition was subsequently carried out in 50 µL reaction volume containing 25 µL 2X TD buffer and 2.5 µL Tn5 transposase (available as individual products upon request from Illuminia). The transposition reaction was incubated at 37 °C for 30 minutes in a thermomixer set to 1,000 r.p.m. The transposed DNA was isolated using the Qiagen MinElute kit and was eluted in 30 µL of water in the final step. Library preparation was carried out using 10 µL of transposed DNA per sample as described previously with 10 - 11 cycles of amplification (Buenrostro, Wu, Chang, & Greenleaf, 2015). Our DNA tended to be under digested, so we performed doubled-sided bead purification using AMPure XP beads (Beckman). Libraries were sequenced as paired-end 75 bp reads, multiplexing all samples per experiment (7) on one lane of the Illumina High-Seq 500 platform at The Rockefeller University Genomics Resource Center. Two biological replicates for each condition were collect, processed, and sequenced from independent experiments.

Image analysis

For images acquired from micropatterned cell culture experiments, stitching and colony detection were carried out as described previously using custom software written in MATLAB (Etoc et al., 2016). For analysis of the SMAD reporter lines on micropatterned colonies a single z-plane through the middle of the colony was analyzed at each time point. Background in each channel was removed by subtracting a minimum intensity image that was generated by taking the minimum value at each pixel over all images in that channel in the same z-plane. Vignetting was corrected by dividing by a flat-field image that was generated by normalizing the intensity of the minimum image to 1. This procedure corrects the intensity drop-off at the border of the images without further altering the average image intensity.

Nuclei segmentation and signal quantification were performed on the corrected images as follows. The H2B image was thresholded to generate a binary image separating the foreground (nuclei) from the background. The original, corrected H2B image was then filtered with a median and sphere filter with parameters matching the expected size of individual nuclei. Local maxima corresponding to individual nuclei were detected using the MATLAB extended-maxima transform function. Maxima were dilated to increase the likelihood of obtaining a single maximum per nuclei. Maxima falling within the foreground were used as seeds for watershed segmentation, which was also restricted to the foreground and was used to obtain a labeled object corresponding to each nucleus within the image. The segmented objects were further processed to eliminate objects much larger or smaller than the expected size of individual nuclei. The results of the segmentation were used as a mask to obtain the median per cell nuclear intensity in each channel. Images acquired from fixed samples that were stained by immunofluorescence

were similarly segmented and analyzed. The DAPI channel was used to perform nuclei segmentation, and the segmented objects were subsequently used to obtain the median per cell signal intensity in each channel.

For the mCitrine-SMAD2 reporter line an enrichment of the mCitrine signal in the cytoplasm relative to the nucleus prior to ligand presentation could be detected, which decreased concomitantly with an increase in the nuclear signal following Activin presentation. Therefore, the mCitrine-SMAD2 response was quantified as the nuclear-to-cytoplasmic ratio, which was used previously as a readout for TGFβ pathway activity (Warmflash et al., 2012). In order to estimate the cytoplasmic signal for each cell, a narrow donut surrounding the nuclear mask was formed by dilating the nuclear mask once by an inner radius and a second time by an outer radius and subtracting the first dilated object from the second. The donut, which formed the cytoplasmic mask, was not restricted to the foreground pixels (H2B signal), but it was prevented from overlapping with the masks of neighboring nuclei. The median mCitrine intensity within the nuclear mask was divided by the median intensity within the cytoplasmic mask on a per cell basis. The qualitative behavior of the SMAD2 nuclear-to-cytoplasmic response was not sensitive to the exact size of the inner and outer radius of the donut. Therefore, the values were chosen manually and kept fixed throughout all analyses. For the RFP-SMAD1 reporter line a faint nuclear signal could be detected prior to ligand presentation. However, no cytoplasmic signal could be detected above the background. Therefore, the RFP-SMAD1 response was quantified as the median nuclear signal normalized to the median H2B signal to normalize for cells moving in and out of the z-plane. In order to analyze the SMAD response as a function of radial position within the micropatterned colonies, the cells within a single colony were binned based on the radial position of their center and the average response per cell within each bin was calculated. The radial profile for individual colonies was then averaged over several colonies. Analysis of the SMAD reporter lines in single-cell culture was carried out in a similar manner. The maximum intensity projection image, rather than a single z-plane was analyzed. Since cells remained flat under these conditions it was not necessary to normalize the SMAD1 signal by H2B.

Toatl RNA-sequencing analysis and clustering

For each gene, a baseline expression profile, which was calculated using a linear interpolation between the 0, 6 and 12 h control samples, was subtracted from the expression values of the Activin-treated samples. The gene list was filtered to contain only those genes that: 1) had at least one time point with an absolute fold-change larger than 2 (up- or down-regulated) compared to 0 h and 2) had at least one time point with a normalized read count higher than 100. That generated a list of 3,529 genes of interest, which was then hierarchically clustered by their z-scored expression values, using Cluster 3.0 (de Hoon, Imoto, Nolan, & Miyano, 2004) with the following options: centered correlation as the similarity metric and average linkage as clustering method. The resulting hierarchical tree was visualized using Java TreeView (Saldanha, 2004) to identify the minimal clusters of interest.

Motif analysis

To identify motifs enriched within gene clusters, the 2000 bp upstream sequences for all genes were extracted using the PWMEnrich Bioconductor package. Motif enrichment of each cluster's sequence set was performed using AME (Bailey, Johnson, Grant, & Noble, 2015) with the

HOCOMOCOv10 database (Kulakovskiy et al., 2016) against the background (upstream sequences of all genes).

Gene enrichment analysis

We assessed the enrichment of the genes in each of the groups identified in Figure 2.9A for previously defined marker gene sets from isolated endoderm, mesoderm, and ectoderm/epiblast tissue from E7.5 mouse embryos (Lu et al., 2018) using the GOseq Bioconductor package (Young, Wakefield, Smyth, & Oshlack, 2010). Mouse genes were mapped to human orthologues using data downloaded from the Mouse Genome Informatics database: http://www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt.

Frog embryo collection and manipulation

Xenopus laevis embryos were obtained by *in vitro* fertilization according to standard protocols (Sive, Grainger, & Harland, 1994). At the 2-cell stage embryos were transferred to 0.1X MMR containing BET inhibitors or the equivalent of volume of DMSO and cultured at 18 °C until late tailbud stages (Nieuwkoop and Faber stages 33-36). Every 24 hours embryos were moved to fresh 0.1X MMR medium with BET inhibitors or DMSO. Embryos were fixed with 3.7% PFA in 1X MEM salts for 1 hour at RT and then stored at -20 °C in 100% ethanol. For animal cap explant experiments, embryos were treated with BET inhibitors and 0.1 M LiCl for 1 hour beginning at the 2-cell cell stage. Embryos were washed and transferred to 0.1X MMR + BET inhibitors and cultured at RT (20-22 °C). At stage 9, the vitelline membrane was removed with forceps and animal cap explants were cut in 0.5X MMR with an eyebrow knife (8 caps per condition). Caps were transferred to 0.5X MMR + 0.05% BSA +/– Activin A (0.5 ng/mL) and cultured for 3 hours corresponding to stage 10.5 - 11. Caps were frozen at -20 °C for subsequent RNA extraction.

Animal cap RNA extraction and RT-PCR

Solution D was prepared fresh with 2-mercaptoethanol (35 µL in 5 mL solution D, Table 5.3). Animal caps for each condition were lysed by pipetting in 50 µL solution D. Each sample was mixed with 850 uL Trizol and incubated on ice for 5 minutes. Each sample was then mixed with 100 uL chloroform and incubated on ice for an additional 5 minutes followed by centrifugation for 5 minutes. The aqueous phase (~400 µL) was removed, transferred to a fresh tube, and mixed with an equal volume of isopropanol. Samples were incubated on ice for 10-15 minutes and then centrifuged for 15 minutes at 4 °C. The supernatant was removed and the pellet was washed with 75% ethanol. The pellet was dried and then resuspended in 25 µL of RNase/DNase free water. cDNA generation and RT-PCR was carried out as described above for hESCs using the primer sequences listed in Table 5.4.

Table 5.3. Animal cap lysis solution D

Component	Amount
Guanidinium thiocyanate	12.5 g
10% Sodium N-lauroylsarcosinate	1.32 mL
Sodium citrate (1 M, pH 7)	880 µL
Water	14.65 mL

Table 5.4. *Xenopus laevis* RT-PCR primer sequences and source

Gene symbol	Forward primer	Reverse primer
Xodc	GCCATTGTGAAGACTCTCTCCATTC	TTCGGGTGATTCCTTGCCAC
Xgsc	GCTGATTCCACCAGTGCCTCACCAG	GGTCCTGTGCCTCCTCTTCCTCCTG
Xeomes	GCCTACGAAACAGACTACTCCT	TAATGGAGGGAGGGGTTTCTAC
Xbra	TTCTGAAGGTGAGCATGTCG	GTTTGACTTTGCTAAAAGAGACAGG
Xnr3	ATGGGATCATTTAAGACCAAGCTTTCT	GCAACAAGATGGTGAGATTGAAGAC

REFERENCES

Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., Craigie, A., Larroux, C., & Degnan, B. M. (2007). Wnt and TGF-β Expression in the Sponge Amphimedon queenslandica and the Origin of Metazoan Embryonic Patterning. *PLoS One*, *2*(10), e1031. http://doi.org/10.1371/journal.pone.0001031

Alexandre, C., Baena-Lopez, A., & Vincent, J. P. (2013). Patterning and growth control by membrane-tethered Wingless. *Nature*, *505*(7482), 180–185. http://doi.org/10.1038/nature12879

Arnold, S. J., & Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nature Reviews Molecular Cell Biology*, *10*(2), 91–103. http://doi.org/10.1038/nrm2618

Bachiller, D., Klingensmith, J., Kemp, C., Belo, J. A., Anderson, R. M., May, S. R., et al. (2000). The organizer factors Chordin and Noggin are required for mouse forebrain development. *Nature*, *403*(6770), 658–661. http://doi.org/10.1038/35001072

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, *43*(W1), W39–W49. http://doi.org/10.1093/nar/gkv416

Beers, J., Gulbranson, D. R., George, N., Siniscalchi, L. I., Jones, J., Thomson, J. A., & Chen, G. (2012). Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nature Protocols*, *7*(11), 2029–2040. http://doi.org/10.1038/nprot.2012.130

Belkina, A. C., & Denis, G. V. (2012). BET domain co-regulators in obesity, inflammation and cancer. *Nature Reviews Cancer*, *12*(7), 465–477. http://doi.org/10.1038/nrc3256

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., et al. (2006). The Nodal Precursor Acting via Activin Receptors Induces Mesoderm by Maintaining a Source of Its Convertases and BMP4. *Developmental Cell*, *11*(3), 313–323. http://doi.org/10.1016/j.devcel.2006.07.005

Blythe, S. A., Cha, S. W., Tadjuidje, E., Heasman, J., & Klein, P. S. (2010). β-Catenin Primes Organizer Gene Expression by Recruiting a Histone H3 Arginine 8 Methyltransferase, Prmt2. *Developmental Cell*, *19*(2), 220–231. http://doi.org/10.1016/j.devcel.2010.07.007

Bolce, M. E., Hemmati-Brivanlou, A., Kushner, P. D., & Harland, R. M. (1992). Ventral ectoderm of Xenopus forms neural tissue, including hindbrain, in response to activin. *Development*, *115*(3), 681–688.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell*, *122*(6), 947–956. http://doi.org/10.1016/j.cell.2005.08.020

Brinkman, E. K., Chen, T., Amendola, M., & van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research*, *42*(22), e168–e168. http://doi.org/10.1093/nar/gku936

Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C. H. H., Lim, B., et al. (2011). Activin/Nodal Signaling Controls Divergent Transcriptional Networks in Human Embryonic Stem Cells

and in Endoderm Progenitors. *Stem Cells*, *29*(8), 1176–1185. http://doi.org/10.1002/stem.666

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology*, *109*(1), 21.29.1–21.29.9. http://doi.org/10.1002/0471142727.mb2129s109

Butler, M. T., & Wallingford, J. B. (2017). Planar cell polarity in development and disease. *Nature Reviews Molecular Cell Biology*, *18*(6), 375–388. http://doi.org/10.1038/nrm.2017.11

Chen, B., Dodge, M. E., Tang, W., Lu, J., Ma, Z., Fan, C. W., et al. (2009). Small molecule–mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nature Chemical Biology*, *5*(2), 100–107. http://doi.org/10.1038/nchembio.137

Chen, C., & Shen, M. M. (2004). Two Modes by which Lefty Proteins Inhibit Nodal Signaling. *Current Biology*, *14*(7), 618–624. http://doi.org/10.1016/j.cub.2004.02.042

Chen, G., Gulbranson, D. R., Hou, Z., Bolin, J. M., Ruotti, V., Probasco, M. D., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nature Methods*, *8*(5), 424–429. http://doi.org/10.1038/nmeth.1593

Chhabra, S., Liu, L., Goh, R., Kong, X., & Warmflash, A. (2019). Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids. *PLoS Biology*, *17*(10), e3000498. http://doi.org/10.1371/journal.pbio.3000498

Clevers, H., Loh, K. M., & Nusse, R. (2014). An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science*, *346*(6205), 1248012. http://doi.org/10.1126/science.1248012

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. http://doi.org/10.1126/science.1231143

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., & Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development*, *120*(7), 1919–1928.

Conlon, P., Gelin-Licht, R., Ganesan, A., Zhang, J., & Levchenko, A. (2016). Single-cell dynamics and variability of MAPK activity in a yeast differentiation pathway. *Proceedings of the National Academy of Sciences*, *113*(40), E5896–E5905. http://doi.org/10.1073/pnas.1610081113

Corces, M. R., Trevino, A. E., Hamilton, E. G., Greenside, P. G., Sinnott-Armstrong, N. A., Vesuna, S., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature Methods*, *14*(10), 959–962. http://doi.org/10.1038/nmeth.4396

Cui, W., Taub, D. D., & Gardner, K. (2007). qPrimerDepot: a primer database for quantitative real time PCR. *Nucleic Acids Research*, *35*(Database), D805–D809. http://doi.org/10.1093/nar/gkl767

D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E., & Baetge, E. E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature*

Biotechnology, *23*(12), 1534–1541. http://doi.org/10.1038/nbt1163

Datto, M. B., Frederick, J. P., Pan, L., Borton, A. J., Zhuang, Y., & Wang, X.F. (1999). Targeted Disruption of Smad3 Reveals an Essential Role in Transforming Growth Factor β-Mediated Signal Transduction. *Molecular and Cellular Biology*, *19*(4), 2495–2504. http://doi.org/10.1128/MCB.19.4.2495

de Hoon, M. J. L., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. *Bioinformatics*, *20*(9), 1453–1454. http://doi.org/10.1093/bioinformatics/bth078

Deglincerti, A., Etoc, F., Guerra, M. C., Martyn, I., Metzger, J., Ruzo, A., et al. (2016). Self-organization of human embryonic stem cells on micropatterns. *Nature Protocols*, *11*(11), 2223–2232. http://doi.org/10.1038/nprot.2016.131

Demagny, H., Araki, T., & De Robertis, E. M. (2014). The Tumor Suppressor Smad4/DPC4 Is Regulated by Phosphorylations that Integrate FGF, Wnt, and TGF-β Signaling. *Cell Reports*, *9*(2), 688–700. http://doi.org/10.1016/j.celrep.2014.09.020

Di Micco, R., Fontanals-Cirera, B., Low, V., Ntziachristos, P., Yuen, S. K., Lovell, C. D., et al. (2014). Control of Embryonic Stem Cell Identity by BRD4-Dependent Transcriptional Elongation of Super-Enhancer-Associated Pluripotency Genes. *Cell Reports*, *9*(1), 234–247. http://doi.org/10.1016/j.celrep.2014.08.055

Diaz-Cuadros, M., Wagner, D. E., Budjan, C., Hubaud, A., Tarazona, O. A., Donelly, S., et al. (2020). In vitro characterization of the human segmentation clock. *Nature*, *580*(7801), 113–118. http://doi.org/10.1038/s41586-019-1885-9

Dunn, N. R., Vincent, S. D., Oxburgh, L., Robertson, E. J., & Bikoff, E. K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development*, *131*(8), 1717–1728. http://doi.org/10.1242/dev.01072

Estarás, C., Benner, C., & Jones, K. A. (2015). SMADs and YAP Compete to Control Elongation of β-Catenin:LEF-1-Recruited RNAPII during hESC Differentiation. *Molecular Cell*, *58*(5), 780–793. http://doi.org/10.1016/j.molcel.2015.04.001

Etoc, F., Metzger, J., Ruzo, A., Kirst, C., Yoney, A., Ozair, M. Z., et al. (2016). A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Developmental Cell*, *39*(3), 302–315. http://doi.org/10.1016/j.devcel.2016.09.016

Evans, M. J., & Kaufman, M. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154–156.

Finley, L. W. S., Vardhana, S. A., Carey, B. W., Alonso-Curbelo, D., Koche, R., Chen, Y., et al. (2018). Pluripotency transcription factors and Tet1/2 maintain Brd4-independent stem cell identity. *Nature Cell Biology*, *20*(5), 565–574. http://doi.org/10.1038/s41556-018-0086-3

Fong, C. Y., Gilan, O., Lam, E. Y. N., Rubin, A. F., Ftouni, S., Tyler, D., et al. (2015). BET inhibitor resistance emerges from leukaemia stem cells. *Nature*, *525*(7570), 538–542. http://doi.org/10.1038/nature14888

Francois, P., & Siggia, E. D. (2008). A case study of evolutionary computation of biochemical adaptation. *Physical Biology*, *5*(2), 026009. http://doi.org/10.1088/1478-3975/5/2/026009

Fuerer, C., Nostro, M. C., & Constam, D. B. (2014). Nodal·Gdf1 Heterodimers with Bound Pro

Domains Enable Serum-Independent Nodal Signaling and Endoderm Differentiation. *Journal of Biological Chemistry*, *289*(25), jbc.M114.550301–17871. http://doi.org/10.1074/jbc.M114.550301

Funa, N. S., Schachter, K. A., Lerdrup, M., Ekberg, J., Hess, K., Dietrich, N., et al. (2015). β-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell*, *16*(6), 639–652. http://doi.org/10.1016/j.stem.2015.03.008

Gilbert, S. F. (2000). Developmental Biology (6 ed.). Sunderland: Sinauer Associates.

Gonzales-Cope, M., Sidoli, S., Bhanu, N. V., Won, K. J., & Garcia, B. A. (2016). Histone H4 acetylation and the epigenetic reader Brd4 are critical regulators of pluripotency in embryonic stem cells. *BMC Genomics*, *17*(1), 154. http://doi.org/10.1186/s12864-016-2414-y

Green, J. B. A., & Smith, J. C. (1990). Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature*, *347*(6291), 391–394. http://doi.org/10.1038/347391a0

Gurdon, J. B., & Bourillot, P. Y. (2001). Morphogen gradient interpretation. *Nature*, *413*(6858), 797–803. http://doi.org/10.1038/35101500

Harland, R., & Gerhart, J. (1997). FORMATION AND FUNCTION OF SPEMANN'S ORGANIZER. *Annual Review of Cell and Developmental Biology*, *13*(1), 611–667. http://doi.org/10.1146/annurev.cellbio.13.1.611

Heasman, J. (2006). Patterning the early Xenopus embryo. *Development*, *133*(7), 1205–1217. http://doi.org/10.1242/dev.02304

Heasman, J., Wylie, C. C., Hausen, P., & Smith, J. C. (1984). Fates and states of determination of single vegetal pole blastomeres of X. laevis. *Cell*, *37*(1), 185–194.

Heemskerk, I., Burt, K., Miller, M., Chhabra, S., Guerra, M. C., Liu, L., & Warmflash, A. (2019). Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells. *eLife*, *8*, 91. http://doi.org/10.7554/eLife.40526

Hill, M. A. (2020). Embryology. Retrieved March 10, 2020, from https://embryology.med.unsw.edu.au/embryology/index.php/Embryonic_Development

Holland, A. J., Fachinetti, D., Han, J. S., & Cleveland, D. W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proceedings of the National Academy of Sciences*, *109*(49), E3350–E3357. http://doi.org/10.1073/pnas.1216880109

Huang, S. M. A., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature*, *461*(7264), 614–620. http://doi.org/10.1038/nature08356

James, D., Levine, A. J., Besser, D., & Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, *132*(6), 1273–1282. http://doi.org/10.1242/dev.01706

Jukam, D., Shariati, S. A. M., & Skotheim, J. M. (2017). Zygotic Genome Activation in

Vertebrates. *Developmental Cell*, *42*(4), 316–332. http://doi.org/10.1016/j.devcel.2017.07.026

Kubo, A., Shinozaki, K., Shannon, J. M., Kouskoff, V., Kennedy, M., Woo, S., et al. (2004). Development of definitive endoderm from embryonic stem cells in culture. *Development*, *131*(7), 1651–1662. http://doi.org/10.1242/dev.01044

Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Soboleva, A. V., Kasianov, A. S., Ashoor, H., et al. (2016). HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models. *Nucleic Acids Research*, *44*(D1), D116–25. http://doi.org/10.1093/nar/gkv1249

la Cova, de, C., Townley, R., Regot, S., & Greenwald, I. (2017). A Real-Time Biosensor for ERK Activity Reveals Signaling Dynamics during C. elegans Cell Fate Specification. *Developmental Cell*, *42*(5), 542–553.e4. http://doi.org/10.1016/j.devcel.2017.07.014

Lacoste, A., Berenshteyn, F., & Brivanlou, A. H. (2009). An Efficient and Reversible Transposable System for Gene Delivery and Lineage-Specific Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell*, *5*(3), 332–342. http://doi.org/10.1016/j.stem.2009.07.011

Larabell, C. A., Torres, M., Rowning, B. A., Yost, C., Miller, J. R., Wu, M., et al. (1997). Establishment of the Dorso-ventral Axis in Xenopus Embryos Is Presaged by Early Asymmetries in β-Catenin That Are Modulated by the Wnt Signaling Pathway. *The Journal of Cell Biology*, *136*(5), 1123–1136. http://doi.org/10.1083/jcb.136.5.1123

Lawrence, P. A., & Levine, M. (2006). Mosaic and regulative development: two faces of one coin. *Current Biology*, *16*(7), R236–R239. http://doi.org/10.1016/j.chom.2020.01.025

Liu, W., Stein, P., Cheng, X., Yang, W., Shao, N. Y., Morrisey, E. E., et al. (2014). BRD4 regulates Nanog expression in mouse embryonic stem cells and preimplantation embryos. *Cell Death & Differentiation*, *21*(12), 1950–1960. http://doi.org/10.1038/cdd.2014.124

Loh, K. M., Ang, L. T., Zhang, J., Kumar, V., Ang, J., Auyeong, J. Q., et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell*, *14*(2), 237–252. http://doi.org/10.1016/j.stem.2013.12.007

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *15*(12), 31. http://doi.org/10.1186/s13059-014-0550-8

Lu, X., Zhao, Z. A., Wang, X., Zhang, X., Zhai, Y., Deng, W., et al. (2018). Whole-transcriptome splicing profiling of E7.5 mouse primary germ layers reveals frequent alternative promoter usage during mouse early embryogenesis. *Biology Open*, *7*(3), bio032508. http://doi.org/10.1242/bio.032508

MacDonald, B. T., Tamai, K., & He, X. (2009). Wnt/β-Catenin Signaling: Components, Mechanisms, and Diseases. *Developmental Cell*, *17*(1), 9–26. http://doi.org/10.1016/j.devcel.2009.06.016

Marshall, C. J. (1995). Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell*, *80*(2), 179–185.

http://doi.org/10.1016/0092-8674(95)90401-8

Martyn, I., Brivanlou, A. H., & Siggia, E. D. (2019a). A wave of WNT signaling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. *Development*, *146*(6), dev172791. http://doi.org/10.1242/dev.172791

Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D., & Brivanlou, A. H. (2018). Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature*, *558*(7708), 132–135. http://doi.org/10.1038/s41586-018-0150-y

Martyn, I., Siggia, E. D., & Brivanlou, A. H. (2019b). Mapping cell migrations and fates in a gastruloid model to the human primitive streak. *Development*, *146*(17), dev179564. http://doi.org/10.1242/dev.179564

Massague, J. (2005). Smad transcription factors. *Genes & Development*, *19*(23), 2783–2810. http://doi.org/10.1101/gad.1350705

Massagué, J. (2012). TGFβ signalling in context. *Nature Reviews Molecular Cell Biology*, *13*(10), 616–630. http://doi.org/10.1038/nrm3434

Massey, J., Liu, Y., Alvarenga, O., Saez, T., Schmerer, M., & Warmflash, A. (2019). Synergy with TGFβ ligands switches WNT pathway dynamics from transient to sustained during human pluripotent cell differentiation. *Proceedings of the National Academy of Sciences*, *116*(11), 4989–4998. http://doi.org/10.1073/pnas.1815363116

McLean, A. B., D'Amour, K. A., Jones, K. L., Krishnamoorthy, M., Kulik, M. J., Reynolds, D. M., et al. (2007). Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed., *25*(1), 29–38. http://doi.org/10.1634/stemcells.2006-0219

Mendjan, S., Mascetti, V. L., Ortmann, D., Ortiz, M., Karjosukarso, D. W., Ng, Y., et al. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. *Cell Stem Cell*, *15*(3), 310–325. http://doi.org/10.1016/j.stem.2014.06.006

Muñoz-Sanjuán, I., & Brivanlou, A. H. (2002). Neural induction, the default model and embryonic stem cells. *Nature Reviews Neuroscience*, *3*(4), 271–280. http://doi.org/10.1038/nrn786

Müller, P., Rogers, K. W., Ben M Jordan, Lee, J. S., Robson, D., Ramanathan, S., & Schier, A. F. (2012). Differential Diffusivity of Nodal and Lefty Underlies a Reaction-Diffusion Patterning System. *Science*, *336*(6082), 721–724. http://doi.org/10.1126/science.1221920

Nakanishi, N., Sogabe, S., & Degnan, B. M. (2014). Evolutionary origin of gastrulation: insights from sponge development. *BMC Biology*, *12*(1), 1–9. http://doi.org/10.1186/1741-7007-12-26

Nemashkalo, A., Ruzo, A., Heemskerk, I., & Warmflash, A. (2017). Morphogen and community effects determine cell fates in response to BMP4 signaling in human embryonic stem cells. *Development*, *144*(17), 3042–3053. http://doi.org/10.1242/dev.153239

Nichols, S. A., Dirks, W., Pearse, J. S., & King, N. (2006). Early evolution of animal cell signaling and adhesion genes. *Proceedings of the National Academy of Sciences*, *103*(33), 12451–12456. http://doi.org/10.1073/pnas.0604065103

Nieuwkoop, P. D. (1969). The Formation of the Mesoderm in Urodelean Amphibians. *Wilhelm Roux Archiv Für Entwicklungsmechanik Der Organismen*, *163*(4), 298–315. http://doi.org/10.1007/BF00577017

Norris, D. P., Brennan, J., Bikoff, E. K., & Robertson, E. J. (2002). The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. *Development*, *129*(14), 3455–3468.

Pagliuca, F. W., Millman, J. R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J. H., et al. (2014). Generation of Functional Human Pancreatic β Cells In Vitro. *Cell*, *159*(2), 428–439. http://doi.org/10.1016/j.cell.2014.09.040

Pang, K., Ryan, J. F., Baxevanis, A. D., & Martindale, M. Q. (2011). Evolution of the TGF-β Signaling Pathway and Its Potential Role in the Ctenophore, Mnemiopsis leidyi. *PLoS One*, *6*(9), e24152. http://doi.org/10.1371/journal.pone.0024152

Rowning, B. A., Wells, J., Wu, M., Gerhart, J. C., Moon, R. T., & Larabell, C. A. (1997). Microtubule-mediated transport of organelles and localization of β-catenin to the future dorsal side of Xenopus eggs. *Proceedings of the National Academy of Sciences*, *94*(4), 1224–1229. http://doi.org/10.1073/pnas.94.4.1224

Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N., & Derynck, R. (2013). Smad2 is essential for maintenance of the human and mouse primed pluripotent stem cell state. *The Journal of Biological Chemistry*, *288*(25), 18546–18560. http://doi.org/10.1074/jbc.M112.446591

Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data. *Bioinformatics*, *20*(17), 3246–3248. http://doi.org/10.1093/bioinformatics/bth349

Sato, N., Sanjuan, I. M., Heke, M., Uchida, M., Naef, F., & Brivanlou, A. H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. *Developmental Biology*, *260*(2), 404–413.

Schmierer, B., & Hill, C. S. (2005). Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor beta-dependent nuclear accumulation of Smads. *Molecular and Cellular Biology*, *25*(22), 9845–9858. http://doi.org/10.1128/MCB.25.22.9845-9858.2005

Schneider, S., Steinbeisser, H., Warga, R. M., & Hausen, P. (1996). β-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mechanisms of Development*, *57*(2), 191–198. http://doi.org/10.1016/0925-4773(96)00546-1

Sharpe, J. (2019). Wolpert"s French Flag: what"s the problem? *Development*, *146*(24), dev185967. http://doi.org/10.1242/dev.185967

Shen, M. M. (2007). Nodal signaling: developmental roles and regulation. *Development*, *134*(6), 1023–1034. http://doi.org/10.1242/dev.000166

Shi, Y., Wang, Y. F., Jayaraman, L., Yang, H., Massagué, J., & Pavletich, N. P. (1998). Crystal Structure of a Smad MH1 Domain Bound to DNA: Insights on DNA Binding in TGF-β Signaling. *Cell*, *94*(5), 585–594. http://doi.org/10.1016/S0092-8674(00)81600-1

Shimizu, K., & Gurdon, J. B. (1999). A quantitative analysis of signal transduction from activin receptor to nucleus and its relevance to morphogen gradient interpretation. *Proceedings of*

the National Academy of Sciences, *96*(12), 6791–6796. http://doi.org/10.1073/pnas.96.12.6791

Shultz, M. D., Cheung, A. K., Kirby, C. A., Firestone, B., Fan, J., Chen, C. H. T., et al. (2013). Identification of NVP-TNKS656: The Use of Structure–Efficiency Relationships To Generate a Highly Potent, Selective, and Orally Active Tankyrase Inhibitor. *Journal of Medicinal Chemistry*, *56*(16), 6495–6511. http://doi.org/10.1021/jm400807n

Sick, S., Reinker, S., Timmer, J., & Schlake, T. (2006). WNT and DKK Determine Hair Follicle Spacing Through a Reaction-Diffusion Mechanism. *Science*, *314*(5804), 1447–1450. http://doi.org/10.1126/science.1130088

Singh, A. M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A. L., Sun, Y., et al. (2012). Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell*, *10*(3), 312–326. http://doi.org/10.1016/j.stem.2012.01.014

Sive, H. L., Grainger, R. M., & Harland, R. M. (1994). Early Development of *Xenopus laevis*: Course Manual Cold Spring Harbor Laboratory.

Slack, J. M. W. (1991). From Egg to Embryo (2nd ed.). Cambridge University Press.

Smith, J. C. (2009). Forming and Interpreting Gradients in the Early Xenopus Embryo. *Cold Spring Harbor Perspectives in Biology*, *1*(1), a002477–a002477. http://doi.org/10.1101/cshperspect.a002477

Smith, J. C., Price, B. M. J., Van Nimmen, K., & Huylebroeck, D. (1990). Identification of a potent Xenopus mesoderm-inducing factor as a homologue of activin A. *Nature*, *345*(6277), 729–731. http://doi.org/10.1038/345729a0

Sokol, S., & Melton, D. A. (1991). Pre-existent pattern in Xenopus animal pole cells revealed by induction with activin. *Nature*, *351*(6325), 409–411. http://doi.org/10.1038/351409a0

Spemann, H., & Mangold, H. (1924). über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv Für Mikroskopische Anatomie Und Entwicklungsmechanik*, *100*(3), 599–638. http://doi.org/10.1007/BF02108133

Sumi, T., Tsuneyoshi, N., Nakatsuji, N., & Suemori, H. (2008). Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/beta-catenin, Activin/Nodal and BMP signaling. *Development*, *135*(17), 2969–2979. http://doi.org/10.1242/dev.021121

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. http://doi.org/10.1016/j.cell.2006.07.024

Tam, P. P. L., & Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mechanisms of Development*, *68*(1-2), 3–25. http://doi.org/10.1016/S0925-4773(97)00123-8

Tam, P. P. L., & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nature Reviews Genetics*, *8*(5), 368–381. http://doi.org/10.1038/nrg2084

Teo, A. K. K., Arnold, S. J., Trotter, M. W. B., Brown, S., Ang, L. T., Chng, Z., et al. (2011).

Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes & Development*, *25*(3), 238–250. http://doi.org/10.1101/gad.607311

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science*, *282*(5391), 1145–1147. http://doi.org/10.1126/science.282.5391.1145

Tsankov, A. M., Gu, H., Akopian, V., Ziller, M. J., Donaghey, J., Amit, I., et al. (2015). Transcription factor binding dynamics during human ES cell differentiation. *Nature*, *518*(7539), 344–349. http://doi.org/10.1038/nature14233

Turing, A. M. (1952). The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *237*(641).

Valenta, T., Hausmann, G., & Basler, K. (2012). The many faces and functions of β-catenin. *The EMBO Journal*, *31*(12), 2714–2736. http://doi.org/10.1038/emboj.2012.150

Vallier, L., Alexander, M., & Pedersen, R. A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *Journal of Cell Science*, *118*(19), 4495–4509. http://doi.org/10.1242/jcs.02553

van den Eijnden-Van Raaij, A. J. M., van Zoelent, E. J. J., Van Nimmen, K., Koster, C. H., Snoek, G. T., Durston, A. J., & Huylebroeck, D. (1990). Activin-like factor from a Xenopus laevis cell line responsible for mesoderm induction. *Nature*, *345*(6277), 732–734. http://doi.org/10.1038/345732a0

Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P., & Robertson, E. J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes & Development*, *17*(13), 1646–1662. http://doi.org/10.1101/gad.1100503

Vizán, P., Miller, D. S. J., Gori, I., Das, D., Schmierer, B., & Hill, C. S. (2013). Controlling Long-Term Signaling: Receptor Dynamics Determine Attenuation and Refractory Behavior of the TGF-β Pathway. *Science Signaling*, *6*(305), ra106–ra106. http://doi.org/10.1126/scisignal.2004416

Voets, O., Tielen, F., Elstak, E., Benschop, J., Grimbergen, M., Stallen, J., et al. (2017). Highly efficient gene inactivation by adenoviral CRISPR/Cas9 in human primary cells. *PLoS One*, *12*(8), e0182974. http://doi.org/10.1371/journal.pone.0182974

Wang, Q., Zou, Y., Nowotschin, S., Kim, S. Y., Li, Q. V., Soh, C. L., et al. (2017). The p53 Family Coordinates Wnt and Nodal Inputs in Mesendodermal Differentiation of Embryonic Stem Cells. *Cell Stem Cell*, *20*(1), 70–86. http://doi.org/10.1016/j.stem.2016.10.002

Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D., & Brivanlou, A. H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature Methods*, *11*(8), 847–854. http://doi.org/10.1038/nmeth.3016

Warmflash, A., Zhang, Q., Sorre, B., Vonica, A., Siggia, E. D., & Brivanlou, A. H. (2012). Dynamics of TGF-β signaling reveal adaptive and pulsatile behaviors reflected in the nuclear localization of transcription factor Smad4. *Proceedings of the National Academy of Sciences*, *109*(28), E1947–E1956. http://doi.org/10.1073/pnas.1207607109

Wilson, P. A., & Melton, D. A. (1994). Mesodermal patterning by an inducer gradient depends on secondary cell–cell communication. *Current Biology*, *4*(8), 676–686.

http://doi.org/10.1016/S0960-9822(00)00152-4

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, *25*, 1–47.

Wu, T., Pinto, H. B., Kamikawa, Y. F., & Donohoe, M. E. (2015). The BET Family Member BRD4 Interacts with OCT4 and Regulates Pluripotency Gene Expression. *Stem Cell Reports*, *4*(3), 390–403. http://doi.org/10.1016/j.stemcr.2015.01.012

Xu, R. H., Sampsell-Barron, T. L., Gu, F., Root, S., Peck, R. M., Pan, G., et al. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell*, *3*(2), 196–206. http://doi.org/10.1016/j.stem.2008.07.001

Yan, M., Li, G., & An, J. (2017). Discovery of small molecule inhibitors of the Wnt/β-catenin signaling pathway by targeting β-catenin/Tcf4 interactions:. *Experimental Biology and Medicine*, *242*(11), 1185–1197. http://doi.org/10.1177/1535370217708198

Yang, J., Tan, C., Darken, R. S., Wilson, P. A., & Klein, P. S. (2002). β-Catenin/Tcf-regulated transcription prior to the midblastula transition. *Development*, *129*(24), 5743–5752. http://doi.org/10.1242/dev.00150

Yang, X., Letterio, J. J., Lechleider, R. J., Chen, L., Hayman, R., Gu, H., et al. (1999). Targeted disruption of SMAD3 results in impaired mucosal immunity and diminished T cell responsiveness to TGF-β. *The EMBO Journal*, *18*(5), 1280–1291. http://doi.org/10.1093/emboj/18.5.1280

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*(1), 134. http://doi.org/10.1186/1471-2105-13-134

Yoney, A., Etoc, F., Ruzo, A., Carroll, T., Metzger, J. J., Martyn, I., et al. (2018). WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids. *eLife*, *7*. http://doi.org/10.7554/eLife.38279

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias, *11*(2), R14. http://doi.org/10.1186/gb-2010-11-2-r14

Zhang, Z., Zwick, S., Loew, E., Grimley, J. S., & Ramanathan, S. (2019). Mouse embryo geometry drives formation of robust signaling gradients through receptor localization. *Nature Communications*, *10*(1), 1–14. http://doi.org/10.1038/s41467-019-12533-7

Zhu, H., Kavsak, P., Abdollah, S., Wrana, J. L., & Thomsen, G. H. (1999). A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. *Nature*, *400*(6745), 687–693. http://doi.org/10.1038/23293

Zhu, Y., Richardson, J. A., Parada, L. F., & Graff, J. M. (1998). Smad3 Mutant Mice Develop Metastatic Colorectal Cancer. *Cell*, *94*(6), 703–714. http://doi.org/10.1016/S0092-8674(00)81730-4