Interplay between Morphogen Signaling and Tissue Geometry in Early Mammalian Development

Citation

Permanent link

Terms of Use

Share Your Story

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Harvard John A. Paulson School of Engineering and Applied Sciences
have examined a dissertation entitled:

"Interplay between Morphogen Signaling and Tissue Geometry in Early Mammalian Development"

presented by: Steven Michael Zwick

Signature _____

    Typed name: Professor S. Ramanathan

Signature _____

    Typed name: Professor D. Needleman

Signature _____

    Typed name: Professor Y.C. Hsu

Signature _____

    Typed name: Professor O. Pourquié

May 5, 2022

Interplay between Morphogen Signaling and Tissue Geometry in Early Mammalian Development

A dissertation presented

by

Steven Michael Zwick

to

The Committee on Higher Degrees in Applied Physics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Applied Physics

Harvard University

Cambridge, Massachusetts

May 2022

Dissertation Advisor: Prof. Sharad Ramanathan   Steven Michael Zwick

Interplay between Morphogen Signaling and Tissue Geometry in Early Mammalian Development

Abstract

Morphogens are long-range signals that induce concentration-dependent responses across developing tissues, directing the patterning and morphogenesis of embryos. Though multiple morphogens have been identified, characterized, and shown to be essential for development, it is often unclear how morphogens move through a complex and dynamic environment within the embryo to establish reliable concentration gradients. Further, little is known about how these gradients are interpreted by progenitor tissues, particularly in mammalian embryos where experimental access is challenging. In this dissertation supervised by Sharad Ramanathan, I, together with Zhechun Zhang, show that formation of a robust BMP signaling gradient in the early mouse embryo is dependent upon the geometric compartmentalization of the embryo and the basolateral localization of BMP receptors, which restricts where BMP morphogens are sensed in vivo. With evidence from mathematical modeling, human embryonic stem cells in vitro, and mouse embryos in vivo, we demonstrate that this mechanism organizes morphogen signaling gradients in epithelial tissues that are buffered against fluctuations. I also describe preliminary work with Tianlei He developing methods to control the spatiotemporal and combinatorial presentation of morphogens to human organoids in vitro, via the precise, rapid deposition of hundreds of morphogen-soaked beads and organoids on a single chip. I conclude by discussing how this approach can be used to interrogate the nature of morphogen interpretation in human tissues at the quantitative scale necessary to understand its complexities.

Table of Contents

List of Figures

Acknowledgments

I would first like to thank my research advisor, Sharad Ramanathan, for his tremendous guidance and support, professional and personal, over the past eight years. I arrived in his lab as a young and naïve physics student interested in doing his own biology experiments but who didn't know how to hold a pipette. Sharad has provided me with not only a rigorous research environment to refine and hone my technical skills within, but also an incredible amount of intellectual freedom to engage my curiosities, "play around" with a system, and learn how to think as a scientist thinks. Sharad has always encouraged me to focus on what "young Steven" would be excited about; in this respect, I have Sharad to thank for helping me realize my dreams.

I would also like to thank my thesis committee members: Olivier Pourquié, Ya-Chieh Hsu, and Daniel Needleman. They have given me fantastic feedback on even my most outlandish research ideas and made me feel welcome within a scientific field that was new to me. Their advice, recommendations, and encouragement were vital to my accomplishments as a Ph.D. student.

The Ramanathan Lab has changed in size and personnel over the past eight years, but throughout it has always been an exciting, loving, and supportive workplace. I want to thank every lab member I have had the fortune of overlapping with: Zhechun (Lance) Zhang, Jeffrey Lee, Sumin Jang, Dann Huh, Ching-Han (Hannah) Shen, Leon Furchtgott, Sandeep Choubey, Abdullah Yonar, Ethan Loew, Jim Valcourt, Sam Melton, Deniz Aksel, Chenguang (Chen-Chen) Li, Tianlei He, Perry Ellis, Brandon Seo, Theresa Weis, Divya Venkatasubramanian, Heitor Megale, Yusef Ilker Yaman, Alexandru Bacanu, Roya Huang, Sean Murphy, Timothy Hallacy, and William Weiter. I especially want to thank Lance: beyond being a valuable lab mate and co-writer of Chapter 2 of

this dissertation, Lance was an invaluable mentor to me early in graduate school – I do not think I would retained what sanity I have left after these years without his calm and cool perspective, and his gentle encouragement to not give up when one experiment fails. Tianlei He has been a wonderful friend, a rockstar engineer, and an incredibly quick learner during our preparation of Chapter 3 in the last year.

Finally, I could not have completed this dissertation without the love and support of my close friends and family. Philip Gosnell and Pearson Miller have been my steadfast Cambridge friends over the past eight years, entertaining my strangest scientific ideas and some strange non-scientific ones too. Tatum Williamson has been my rock throughout this whole process, supporting me with tender care through my lowest times, while helping me to remain present in the moment and appreciate what I have during the best of times. She has been a wonderful life partner, a devoted confidante, a spectacular co-parent of two splendid little cats, and my best friend. My mother and father provided me with unconditional love and encouraged me to follow my curiosities wherever they took me. More than anything in life, I have simply wanted to make them proud; they never fail to remind me they already are.

To my mother, my father, and Tatum

Chapter 1. Introduction

1.1.   A historical perspective on development: the eye

Development is the process by which a single fertilized egg cell gives rise to an

entire multicellular organism, in order that the organism may reproduce to create

another fertilized egg cell. To describe this process as complex would be a vast

understatement. Consider the development of a human being. From a single cell,

trillions of cells are generated, each belonging to any of the hundreds of known cell

types with unique molecular signatures. These trillions of cells geometrically arrange

themselves and structurally connect to form specialized tissues and organs, each with

critical functions to perform within the body. Each organ merits its own scientific field

of study given their physiological complexity; entire research programs and medical

careers are devoted to understanding the most minute details of their functions and

vulnerabilities to disease. And this total assembly of cells, tissues, and organs must

happen repeatedly, for every individual organism across families, generations, and

species, with few enough errors that the number of surviving members does not

dwindle out.

Such is the apparent incomprehensibility of the problem that during the "Age of

Enlightenment" in 17th and 18th century Western Europe, concurrent with advances in

complex analysis and the invention of the condensing steam engine, the predominant

theory of development was preformationism[1]: the adult body is so complex, the newly-

discovered spermatozoa, so simple, that there must be a miniature, preformed version

of the organism within the sperm (Figure 1.1). The alternative theory, genesis of a

living, breathing, thinking human being from a single undifferentiated unit, seemed

comparatively absurd.
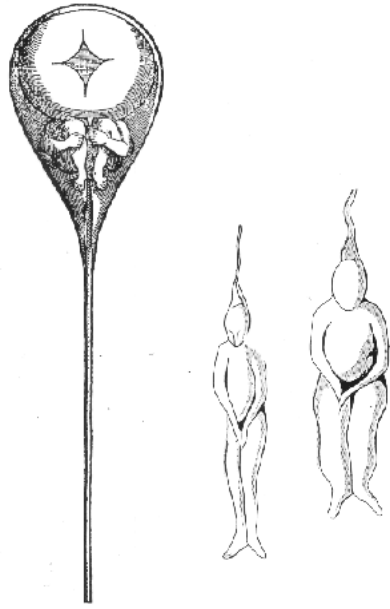
Figure 1.1: Sketch of a tiny human inside a sperm. Drawn by Dutch microscopist Nicolaas Hartsoeker in 1695[1].

In the centuries since, our understanding of how embryonic development occurs has advanced exponentially, although we are far from anything resembling complete understanding. An informative, if ironic, example of this intellectual progress can be found in the study of early eye development. The eye is an organ that functions as an intricate optical measurement system; a state-of-the-art camera capable of adjusting its own aperture and focus, with a sensitivity range exceeding that of any film; the home of an uncountably large number of precise biochemical reactions to convert light waves into electrochemical signals that a brain can process. It is in this light that the eye has often served as the crown jewel of arguments made by advocates of intelligent design and creationism[2]. What but divine intervention could produce such an instrument? British evolutionary biologist Charles Darwin himself wrestled with the problem of explaining how the organ could form through ordinary biological processes. He admitted as such in his seminal *On the Origin of Species*:

To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest possible degree.[3]

And yet, over the past century, studies of eye development have yielded some of our most fruitful insights into how tissues, organs, and bodies form. The first experiments demonstrating *induction*, the process by which one tissue can direct and change the embryonic fate of another, were performed by transplanting an optic cup into the epidermis of an amphibian embryo[4–6]. The result, as found independently by Hans Spemann and Hilde Mangold in 1901 and Warren Lewis in 1904, was the formation of an ectopic eye lens in the animal's skin next to the transplanted optic cup. The lens could not have come from the optic cup itself; the optic cup had induced nearby skin progenitor cells to change their cellular fate altogether and become an eye lens instead. The historical impact of these experiments, to some extent, would later be supplanted by Spemann and Mangold's Nobel Prize winning discovery of the "Spemann-Mangold organizer": a group of cells in the dorsal lip of the amphibian embryo that, when transplanted to the ventral side of the embryo, induced the formation of a second embryonic axis complete with a nervous system and, of course, a pair of eyes[7].

The significance of this finding inspired a decades-long race to understand the mechanism of *neural induction* through a variety of embryological studies. I will highlight here a particular experiment[8,9] performed by developmental biologist Pieter Nieukwoop in 1963. Tissue explants dissected from the ectoderm of early frog embryos and cultured in isolation would give rise to non-neural epidermis. But Nieukwoop found that by dissociating these explants into individual cells and then reaggregating them, something strange happened: the aggregated cells differentiated and self-

organized into anterior neural tissues, including the anterior brain, olfactory placodes, and eyes (Figure 1.2). The presumptive epidermis of the amphibian embryo, the same tissue type into which Lewis, Spemann, and Mangold transplanted optic cups and induced lens differentiation, did not need an inducer for neural induction! All that was needed was to remove the influence of all other tissues, to separate the ectodermal explant from surrounding mesodermal and endodermal tissues and scramble the positional information initially present in the dissected tissue. As it turned out, the ectoderm within normal frog embryos was being induced to differentiate into non-neural epidermis. And the foundational neural induction experiment, the transplant of an optic cup into the future epidermis, was actually a neutralization of this epidermal
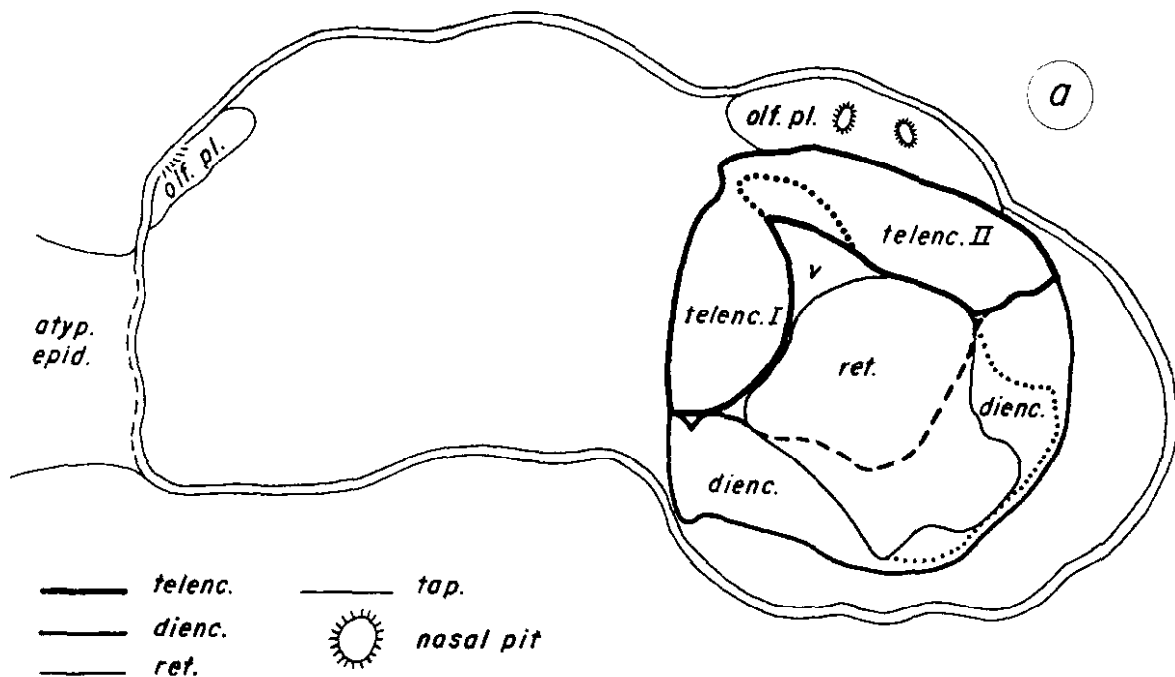
Figure 1.2: Nieukwoop's ectodermal reaggregates. Figure reproduced from reference[8]. Graphical diagram of tissue structures Nieukwoop observed in a reaggregate of ectodermal cells dissected from the *Rana pipiens* embryo and cultured for 7 days. *adh. gl.*, adhesive gland; *atyp. epid.*, atypical epidermis; *dienc.*, diencephalon; *l.*, lens; *olf. pl.*, olfactory placode; *ret.*, retina; *tap.*, tapetum; *telenc.*, telencephalon; *v.*, ventricle.

induction! The optic cup, the original inducing tissue, was repressing some inducing factor.

With the advancement of molecular and genetic techniques later in the 20th century came the identification of many of these inducing factors. Neural induction, the effect of the Spemann-Mangold organizer, is now known to be mediated by proteins such as Noggin, Chordin, Cerberus, Follistatin, and xnr-3[10,11]. Injection of mRNA encoding for any of these proteins into ectodermal explants from the *Xenopus laevis* frog is sufficient to recreate the outcome of Nieukwoop's reaggregation experiment, repressing epidermal differentiation and resulting in the formation of eyes and other anterior neural structures. What these genes share in common is their role as antagonists to the bone morphogenetic protein (BMP) signaling pathway, with BMP protein itself identified as an epidermal inductive factor.

BMP had another significance – it is a member of the transforming growth factor-β (TGF-β) family, the first molecules identified as extracellular *morphogens*[6,12]. The term morphogen was conceived by Alan Turing in 1952 to describe putative diffusive chemicals capable of bringing about changes in organismal form, and how interactions between such chemicals diffusing in space could produce molecular gradients and patterns across tissues[13]. These theoretical ideas were expanded upon by others and synthesized with concepts like induction and positional information[14,15], culminating in Lewis Wolpert's description of the famous "French flag" model[16] in the 1960s. In the model, morphogens are secreted by cells into their extracellular environment and diffuse away from their source, steadily decreasing in concentration as they are absorbed or degraded. This produces a concentration gradient of the morphogen that varies as a function of the distance from the source cells. Cells in surrounding tissues can measure the concentration of the morphogen through some molecular interaction

5

and adopt a particular fate according to that concentration. In this manner, positional information is communicated among cells via the concentrations of diffusive factors, and cell fates can be altered by repositioning the sources from which morphogens are secreted (for instance, during a transplantation experiment). Though the first morphogen factors in embryos would take another two decades to be identified[17,18], multiple molecules have since been shown to act as diffusive morphogens[6], being both necessary for correct patterning in multiple tissue types and sufficient to induce mis-patterning when their concentration gradients are altered.

What then of neural induction and the eye? If Nieukwoop's reaggregation experiment demonstrated that frog ectodermal tissues are fated to become anterior neural tissues in the absence of influence from any surrounding tissues, the French flag model would seem to imply that the frog ectoderm becomes neural in the absence of any secreted morphogen factors from surrounding tissues. In other words, anterior neural fates, including those of the eye, appeared to be something of a default state for early cells of the frog embryo: progenitor cells, with no other information, would steadily differentiate toward anterior neural fates until some other instruction (a morphogen) is received. This so-called *default model*, which has stimulated much discussion within the scientific community[19,20], is simple and tantalizing but, like any model, has serious limitations. Injection of BMP inhibitors into comparable tissues within the chick embryo fail to recreate the neural induction effect seen in frog embryos[21], but neural induction can be reproduced when a chick tissue analogous to the Spemann Mangold organizer, the node, is transplanted instead[21,22]. More importantly, just because Nieukwoop's reaggregated explants were isolated from surrounding tissues does not mean that they were isolated from all morphogen signaling. To the contrary, the strange diversity of cell types and tissues that form de novo in these

aggregates would suggest that cells are in fact receiving several different morphogen signals secreted from other cells within the aggregate, continuously and chaotically differentiating and reshaping cells into increasingly complex structures. This is to say nothing of the biomechanical effect of disaggregating a connected tissue of cells, nor of the crucial roles that cellular junctions[23] and extracellular matrices[24] fulfill in developing tissues. Nieukwoop's reaggregation experiment beautifully demonstrated the sufficiency of ectodermal cells to self-organize into complex neural tissues like the eye, but the "default" model leaves us no closer to understanding the mysterious alchemy contained within a simple ball of frog cells.

It may come as a surprise then to discover that the default model, for all its faults, would have another major success: the self-organization of three-dimensional optic cups from human stem cells in vitro. The advent of pluripotent stem cells[25,26], capable of differentiating into cell types from all germ layers of the embryo, had transformative implications for studies of development and disease in the beginning of the 21st century. There was now an experimental model for human development with human cells, and an established methodology to probe the molecular circuits that organized the formation of human tissues. Nevertheless, it would still be difficult to reliably generate tissues made of stem cells that even distantly resembled embryonic tissues in vivo. The optic cup proved to be one of the early exceptions.

In a landmark pair of papers published respectively in 2011 and 2012, Yoshiki Sasai and collaborators demonstrated that, starting from uniform aggregates of mouse[27] or human[28] stem cells, structures closely resembling optic cups gradually self-organized. They were composed of the correct tissue types too: a neural retinal

epithelium, a retinal pigment epithelium, and, after a long enough time for the human aggregates, rod and cone photoreceptor cells.

The scientific impact of these studies was far beyond the mimicry of an embryonic tissue; they uncovered further insights into how the optic cup forms in vivo, details that were not known from embryological studies. For example, there was no lens epithelium in these optic cups – apparently it was not needed to get the correct structure. Measurement and perturbation of biomechanical protein activity within the cup revealed new details about the physical forces that shaped the progenitor tissues into cups. And perhaps most stunningly, the human optic cups were larger than the mouse optic cups, each with sizes comparable to the optic cups in their respective embryos! This would seem to imply that the size of optic cups is an intrinsic property of the progenitor cells from which they arise; no further instructions from surrounding tissues were needed to start building an eye of the right size. These findings were among the first to illustrate the explanatory power of so-called *organoids*: three-dimensional collections of cells that exploit inherent genetic programs to self-assemble into structures in vitro resembling organs in vivo[29], not unlike Nieukwoop's ectodermal aggregates sixty years ago.

What was the protocol to generate these optic cups in vitro? What sequence of inducing factors did the researchers need to add to their stem cells to kick start this brilliant spontaneous self-organization? The answer: very little. A very basic stem cell culture media with vital amino acids and nutrients; a small amount of fetal bovine serum, or the equivalent knockout serum replacement; a mixture of extracellular matrix proteins to stabilize the three-dimensional structure of the organoid. The human optic cups would require a few extra details – the Wnt morphogen signaling pathway would have to be repressed initially to counteract an effect of the bovine serum, then

8

reactivated near the end of the protocol to get the right proportion of retinal pigment epithelium. Activating the signaling pathway for another morphogen, Hedgehog, further improved the efficiency of the retinal differentiation. But for the most part (and unlike many organoid models developed since), the formation of optic cups in vitro required forming a small aggregate of stem cells, refreshing their media daily, and waiting for magic to happen. The default model of anterior neural induction had not outlived its usefulness yet.

Where are we today in our understanding of development? For many of the tissues that form in the embryo, the expression of every gene has been measured, the responsible morphogen factors have been identified, the relevant knockout mutants have been generated, and the phylogeny has been examined. We are still incredibly far from a quantitative or predictive understanding of how cells assemble into these structures. If we were not, and we truly understood the fundamental algorithm of development, then surely organoids corresponding to all the embryonic tissues would have been generated by now. Further still, these protocols would be increasingly optimized until tissues in vitro and in vivo were indistinguishable. Currently, protocols to generate organoids for new tissues are regularly discovered and published[30], but for most human tissues an analogous organoid system has not yet been developed, let alone an organoid system that recreates the structure or function of a complete organ.

We are still far from understanding how the eye is formed from progenitor cells, though we have learned quite a bit from a century of progress with embryological studies, genetic perturbations, and organoid engineering. This research continues to progress rapidly today. Just this past year, researchers discovered that anterior neural organoids can reproducibly generate a bilaterally symmetric pair of optic vesicles, and

further demonstrated that they are light-sensitive[31]. These optic vesicle-containing brain organoids offer the new potential to study the simultaneous assembly and interaction of multiple organs within the same organoid, opening the door to a more holistic understanding of embryogenesis.

Why have eyes provided such fundamental insights into development over past century, often preceding studies of other organ systems? Maybe it is partly due to their striking, recognizable appearance; it would not take a researcher with the same attention to detail as Pieter Nieukwoop to notice a round, pigmented bulb structure emerging from an otherwise ordinary ball of cells. But I hypothesize there is another reason, by the same logic that underlies the default model of anterior neural induction: by chance or circumstance, the early eye is superb at organizing itself. Experimental induction of other tissues in embryos often required intimate knowledge of the molecular components involved[32], but frog eyes appeared spontaneously from aggregates of ectodermal progenitor cells nearly sixty years ago. The discovery of new organoid systems is now rate-limited by a high-dimensional, Goldilocks-like search for the "just right" combination and timing of signaling factors with which to treat stem cells[33,34]. But organoids mimicking the optic cup emerged almost effortlessly in comparison, and once again new principles of development were discovered.

In 1962, the same year in which Nieukwoop performed his reaggregation experiments, American philosopher of science Thomas Kuhn wrote the following in his book *The Structure of Scientific Revolutions*[35]: "In science… novelty emerges only with difficulty, manifested by resistance, against a background provided by expectation." Perhaps the eye was simply the scientific path of least resistance.

But what about the other tissues and organs of the body? How will we uncover the rules of development for kidneys, muscles, hearts, and lungs? How will we make the organoids that accurately model their structures and functions, or discover new therapeutics to treat the diseases that arise within these tissues? For these questions, I do not think we can rely upon a default model. We will need to learn more about the morphogen signals that pattern and instruct cells to organize into these tissues. We will need to learn more about how these morphogen signals travel through the embryo as it changes shape, and how they are interpreted by progenitor tissues spatially, temporally, and combinatorially. And just as the experimental induction of these tissues often required the advancement of molecular techniques to clone and deliver the right morphogens at the right place and time in the embryo, so too will the discovery of new organoid systems require novel techniques to control morphogen signaling precisely within them.

Addressing these challenges, not the development of the eye, is the central subject of this dissertation. In the remainder of Chapter 1, I introduce several specific concepts and systems that are discussed in subsequent chapters. In Chapter 2, I describe our work investigating how signaling gradients for the morphogen BMP are organized by the geometry of epithelial tissues, in both human pluripotent stem cells and the early mouse embryo. Our findings here demonstrate how even a simple model with a single diffusive morphogen can lead to counterintuitive results that are not easily predicted by developmental genetics alone. Then in Chapter 3, I introduce novel methods we have developed to control the spatial, temporal, and combinatorial presentation of an arbitrary number of morphogens to hundreds of organoids simultaneously on a single chip in vitro. I end by discussing how these tools can be used in the future to discover the morphogen code that gives rise to a tissue phenotype using an iterative

computational inference algorithm; I also describe a few open questions in development that could be addressed through such an approach.

1.2.   Morphogens

Given that the idea of a morphogen has evolved over its history from a useful theoretical concept to a well-characterized class of molecules with strict experimental criteria, it would be useful to reestablish what I mean by the term in this dissertation. Morphogens are signaling molecules, produced by cells, that act over some distance to induce a cellular response in proportion to its concentration[6,36]. I am concerned here primarily with morphogens that move through the extracellular space of tissues. There are some exceptions: in a curious twist of fate, the first factor identified to fulfill the characteristics of a morphogen, the *Drosophila melanogaster* protein Bicoid, is not extracellular. Bicoid has a graded distribution along the anterior-to-posterior axis of the early *Drosophila* embryo that is essential to establish the different domains of gene expression along the axis[17,18,37]; dampening this gradient shifts these transcriptional domains or removes them altogether, resulting in lethal developmental defects. However, Bicoid is an intracellular transcriptional regulator, and the environment through which it travels, the *Drosophila* embryo, is a single multinucleated cell at this stage called a syncytium, generated by repeated nuclear divisions without cellular fission. As a result, Bicoid is unlike most morphogens in that its gradient is established by diffusion through the cytoplasm of a cell. In this respect it is likely not a coincidence that Bicoid was the first discovered morphogen – fixation and visualization of molecules in extracellular space is notoriously difficult[38,39], making the measurement of morphogen concentrations there challenging.

Not to be deterred, embryologists gradually identified various families of extracellular morphogens that were essential for development over the course of the 1990s and 2000s. The first identified extracellular morphogen, the *Drosophila* protein Decapentaplegic, is a member of the TGF-β family responsible for the long range patterning of the dorsal-ventral embryonic axis of flies[12], as well as their wings[40]. TGF-β family signals were found to differentiate *Xenopus* cells into different germ layer fates at different concentrations[41–43], providing circumstantial evidence for a TGF-β family morphogen that organized early fish development[44]. Other morphogen families that were identified include Wnt[45], Hedgehog[46], fibroblast growth factor (FGF)[47], epidermal growth factor (EGF)[48], and retinoic acid (RA)[49]. These morphogens all have strongly conserved roles across a wide range of model organisms. Wnt, for example, is in broad terms a 'posteriorizing' morphogen responsible for inducing many tissues found on the tail-side of vertebrate and invertebrate organisms[50]. Hedgehog signals have crucial roles across the phylogenetic tree in the formation and patterning of the neural tube, as well as the anterior-posterior specification of limbs[51].

The pathways for these morphogens are also organized differently, resulting in some interesting differences in signaling properties among the families. Wnt ligands are lipid-modified, which is required for their secretion and for the ligands to bind their corresponding receptor Frizzled[52]. This hydrophobic modification often limits Wnt proteins to transport along cell membranes, preventing them from diffusing freely over long distances[53]. Hedgehog, on the other hand, functions by repressing the repressor of the signaling pathway to turn it back on, changing how the pathway forms signaling gradients and responds to noise in morphogen production[54].

What is most interesting about the morphogen signaling pathways, in my opinion, is how few have been identified at all. Researchers studying the development

of flies, fish, birds, and mammals alike have repeatedly discovered that the mysterious signal organizing their system of interest is simply another member of a previously identified morphogen family: TGF-β, Wnt, FGF, EGF, Hedgehog, RA.

As always, there are exceptions. Vascular endothelial growth factor (VEGF) is an essential morphogen for the formation and growth of blood vessels within developing tissues[55]. JAK/STAT signaling is indispensable for hematopoiesis in mammals[56], though the evidence demonstrating that its specific ligands behave as morphogens was observed in *Drosophila*[56–58]. VEGF and JAK/STAT signals could be grouped with FGF and EGF within a larger morphogen family of receptor-tyrosine kinases (RTK), given their conserved receptor structure[59]. Delta-Notch signaling mediates cell-cell communication and patterning in a huge variety of embryonic tissues[60], but its ligands are not diffusive and only activate signaling in adjacent cells. There are several signaling pathways that are crucial for development for which a morphogen factor has not been identified. Pathways such as Akt[61] and p53[62] have important roles in developmental processes like proliferation and apoptosis, but it is not clear that any of their wide variety of upstream ligands can perform as morphogens. The more recently discovered Hippo signaling pathway has gained attention as a density-sensing mechanism for developing tissues[63], but again no evidence for a Hippo-specific morphogen has been found, to my knowledge.

As a brief aside, I will also clarify the nature of related morphogens contained within the same family. It is true that most of these families contain many related but distinct proteins that each behave as morphogens. The TGF-β[64,65] and Wnt[66] families contain dozens of different ligands that activate the pathway in a concentration-dependent manner and many are essential for development. Why does it matter that there are only a handful of morphogen pathways if there is a great deal of morphogen

diversity within them? The answer is that these the morphogens are often functionally redundant and primarily differ in where they happen to be expressed in the embryo. There are often bottlenecks in signaling pathways where different morphogens activate the same downstream effectors with minor differences. Almost all morphogens from the TGF-β superfamily, for example, activate one of two canonical signaling cascades by directing the phosphorylation of intracellular effectors SMAD1/SMAD5/SMAD9 (BMP subfamily) or SMAD2/SMAD3 (Activin/Nodal/TGF-β subfamily)[64,65].

Regardless of the exact number of exceptions, it is clear that the number of morphogen signaling families necessary for embryogenesis (let us round up to 10) is orders of magnitude smaller than known number of cell types in the adult human body (over 200, not counting transitory cell types in the embryo) or the number of genes in the human genome (about 30,000). The problem becomes further complicated when considering that morphogen signals are necessary for developing tissues to develop proper structures. Consider the brain, an organ composed of billions of neurons connected by trillions of synapses. The Human Connectome Project, launched in 2009 and envisioned as a five-year project to map this connective network[67], has yet to be declared officially complete as of the date of submission for this dissertation. Recording the position and structural connectivity of every cell, in addition to its cell type and gene expression, would require an immense amount of information storage. How then could ten morphogen signaling pathways possibly encode all that information, information that progenitor cells need during development?

This problem was not unforeseen by theoretical biologists, well before any morphogen had yet to be identified. In his models of morphogen gradients and tissue patterning, Lewis Wolpert clarified that morphogen concentrations provided positional information but not the instructions for molecular differentiation[16]; how cells

interpreted this positional information would depend on the developmental history of the cell and its gene expression. Thus, the same morphogen could be used repeatedly in multiple developmental contexts, activating the same core pathway but inducing different genetic programs depending on what cell is receiving the signal. The dimensionality of a morphogen code can be further expanded by considering the possibility of cross-talk between different families[36,68]. Indeed, morphogen pathways regularly alter between promoting and stifling each other's signaling activity; BMP signals indirectly promote Nodal expression in the mouse embryo where they together induce mesendodermal differentiation, but these two signaling pathways subsequently compete through mutual inhibition to induce the formation of one tissue type over another (posterior mesoderm or anterior mesoderm/endoderm, respectively)[69]. Some combinations of morphogen signals elicit effects that neither morphogen can produce in isolation[68].

A possible solution to the dimensionality problem, then, is that the huge amount of external instruction cells need during embryogenesis is encoded in the spatial, temporal, and combinatorial presentation of morphogens they detect in their extracellular environment. Everything inside the cell – its gene expression, epigenomic state, the number and post-transcriptional states of its proteins – operates as a ledger to track what has happened to it previously and what it should do next, given the morphogen code it receives. Building an embryo, to put it simply, requires getting the right combination of signals at the right place and right time. This hypothesis is not entirely naïve either. Comparative analyses of genomes from different species suggest that evolutionary differences have arisen primarily through the mutation of cis-regulatory regions that govern whether genes are expressed, not through changes in the genes themselves[70]. Morphogen pathway genes typically have strong sequence

conservation throughout the animal kingdom. It is not unreasonable to conjecture that much of the difference between building a frog and building a human is embedded in the morphogen code.

Though the morphogen families have been well-characterized, there remain significant challenges to discovering the morphogen code for developmental processes. It is still difficult today to measure the concentration of extracellular morphogens reliably, occasionally leading to confusion. For example, long-range Wnt signaling has been extensively studied in the wing imaginal disc of *Drosophila*, where was assumed to spread across entire disc to pattern the wing of the fly. However, genetically replacing the main *Drosophila* Wnt gene with a copy that produces a mutant form of the morphogen that is tethered to the cell membrane did not disrupt the patterning of the wing[71]; in fact, these flies were completely viable organisms despite the inability of the mutant Wnt protein to spread away from any cell that secreted it. Even when we know it is essential for a morphogen to be expressed from a particular tissue, we do not always know where it is essential to be sensed.

What options for extracellular morphogen visualization are available? Tissue fixation techniques often do not capture proteins in the extracellular environment with high efficiency[38,39]. Tissue fixation also irreversibly halts the development of the tissue, limiting insight into how morphogen gradients are established and interpreted over time. An alternative genetic approach is to insert a measurable tag, such as a fluorescent protein, into the gene sequence for a morphogen, fusing it to the protein during translation. This method has had success in enabling the live visualization of morphogen gradients, such as the Bicoid gradient in *Drosophila* embryos[72]. Nevertheless, there are still significant issues: genetic modification quickly becomes low-throughput and expensive in model organisms evolutionarily closer to humans. The design of a

17

protein tag must be done carefully, given that interference of a key domain can lead to failure of the morphogen to be secreted, to bind a receptor, or to establish a concentration gradient identical to the untagged morphogen. Visualization of proteins in living embryos is even more difficult in mammalian model organisms where the embryo grows in utero. Imaging of mammalian embryos within the womb requires difficult techniques such as optical coherence microscopy[73] or intravital microscopy[74] that have strict limitations. Researchers have developed methods to culture the mouse embryo ex utero after removal from the mother at various stages of development[75,76], but the effectiveness of this approach is temporally limited by when the embryo is collected and how long it can survive during tissue culture. Finally, there is a general complication for all these approaches: even if ten morphogen signaling pathways appears to a small number in context of the complexity of development, it is a large number of pathways to measure and track dynamic activity within simultaneously. Fluorescence imaging within the visible spectrum of light is generally limited to three or four channels with non-overlapping spectra.

Another strategy to learn about the morphogen code is through its manipulation. By perturbing the presentation or transport of morphogens in embryos and observing the resulting outcome, we can discover where morphogens are sufficient to organize developmental processes. The aforementioned genomic engineering of flies with membrane bound Wnt proteins is powerful example of what can be learned about from such experiments[71]. Nevertheless, it is difficult to perturb the concentrations of morphogens in a high-throughput manner to explore the high dimensional morphogen code space sufficiently quickly. Genomic engineering is a powerful approach, but often requires clever genetic approaches to produce perturbations that are less trivial than complete knockouts. Directly injecting morphogens into tissues is a classic technique

from embryology that today continues to reveal intricate details of how morphogen gradients organize development. Injection of BMP and Nodal mRNA into two different cells in the early zebrafish embryo, creating opposing signaling gradients, is sufficient to induce the organization of an entire second embryonic axis in vivo[77]. Further, by dissection and removal of the tissue encompassing the two ectopic morphogen sources, the authors of this study demonstrate these morphogen gradients are sufficient to organize explants to construct embryonic axes in vitro. Despite the strength of genetic and classic embryology techniques to demonstrate sufficiency in developmental models, they are limited by their throughput and scale poorly to mammalian model organisms where the embryo is harder to modify genetically and difficult to access in utero.

One other classic method of embryologists to administer morphogen gradients to embryonic tissues is by using absorptive beads soaked in the morphogen protein. Such an approach was used to provide the fundamental evidence that the TGF-β signal Activin functions as a morphogen[42,43,78]. Beads soaked in a solution of radioactively labeled Activin protein and transplanted into the early *Xenopus* embryo generate a concentration gradient that can be detected through radiography[43]; these proteins directly induce signaling through Activin receptors to pattern the embryo over long ranges in a concentration-dependent manner, meeting the all the criteria for a morphogen. Beads continue to be used in embryology experiments today to interrogate how morphogen gradients are interpreted by tissues to organize morphogenesis[79–81], sometimes leading to counter-intuitive results that conflict with long-standing models. In a recent study[81] (of which Lewis Wolpert is a co-author), it was shown, using combinations of BMP and Activin beads implanted in the early chicken embryo, that chick progenitor cells measure the local morphogen concentration relative to

19

neighboring cells when deciding their cell fate, rather than measuring the absolute morphogen concentration. However, studies using beads are generally performed in non-mammalian embryos due to their experimental accessibility in ovo. In Chapter 3, I introduce a method to control the positioning of morphogen-soaked beads and human organoids simultaneously on a microfabricated chip in vitro, enabling further investigation of how morphogen gradients organize the patterning of mammalian tissues.

In summary, thirty-five years on from the discovery of the first morphogen, we still know very little about where they are in the embryo and how they induce the developmental programs for which they are required.

1.3.   Embryo geometry and receptor localization

How do morphogen signals travel from cells that secrete them to the cells that sense them in a three-dimensional embryo that is changing shape? It is often not clear how the one- and two-dimensional models of morphogen signaling illustrated in figures extend to three dimensions, where they can make entirely different predictions[16]. These differences can be a consequence of basic mathematics before any biological detail is considered. In one of the simplest models of diffusion, a particle randomly walking on an infinite lattice in one or two dimensions will eventually explore every point in the space, but a particle walking in three dimensions may never reach a particular destination even after an infinite amount of time[82–84].

Further complicating the matter is the complex geometric structure of embryos: they are not simple spherical aggregates of cells but *compartmentalized* structures, filled with pockets and cavities lined by walls of epithelial tissue[85–88]. These epithelial tissues

are made up of cells connected by tight junctions that are largely impermeable to passive fluid flow[89,90], segregating the extracellular space above (apical) and below (basolateral) these junctions. This compartmentalization is essential for proper function of tissues and organs; imagine a circulatory system with blood vessels that leaked, or a gastrointestinal tract that failed to separate waste from nearby organs.

How does this compartmentalization shape the formation of morphogen gradients within the developing embryo? By subdividing the extracellular space within the embryo, epithelial tissues delimit how morphogens are transported and further dictate which tissues can talk to one another[87,91]. Furthermore, consider the case when the tissue detecting a morphogen is itself epithelial. If its tight junctions separate the extracellular space into apical and basolateral compartments, from which compartment should the tissue measure the morphogen concentration?

Most extracellular morphogens are detected by transmembrane receptor proteins that are expressed by cells and localized to the cell membrane[92]. When the extracellular domain of the receptor binds a ligand in the extracellular space, the interaction induces a conformational change in the intracellular domain of the receptor protein, creating an active site for some downstream protein interaction. For many morphogens, it is not known how receptor proteins are trafficked to the membrane and whether the distribution of receptors is evenly divided around the membrane or specifically localized. It is immediately apparent how such localization could dramatically affect a tissue's response to a morphogen gradient, motivating this as a topic of investigation.

To my knowledge, the first evidence of a localized signaling receptor was found in 1987: Sevenless, a receptor tyrosine kinase expressed in the eye disc epithelium within the developing retina of *Drosophila melanogaster*[93,94]. This receptor specifically localizes to the apical tips of these cells where it is inaccessible to ligands elsewhere in

the extracellular space. Despite belonging to the receptor tyrosine kinase family, Sevenless is not a receptor for a morphogen. Its corresponding ligand, Bride of Sevenless (BOSS), is a transmembrane protein that is expressed and apically localized in the adjacent photoreceptor cell[95], and thus does not induce tissue patterning at range. Other signaling receptors have since been discovered to have an asymmetric membrane localization, though most are receptor tyrosine kinases in *Drosophila*[94,96] and *C. elegans*[97] that are not receptors for a known morphogen. Human airway epithelia express a trio of receptor tyrosine kinases that localize to the basolateral membrane of the tissue, as well as the corresponding ligand that is secreted from the apical membrane[98]. This results in an interesting mechanism for wound healing: the growth factor ligand only contacts the basolateral receptor if a tear form of the epithelium, whereupon the ligand stimulates restoration of the epithelial integrity. In a recent finding, tumor necrosis factor (TNF) ligands and receptors were similarly found to be sequestered on the opposite sides of *Drosophila* imaginal disc epithelia[99]. In this context, the asymmetric localization provides a mechanism to correct oncogenic errors, where cells with polarity defects capable of producing tumors are targeted for apoptosis.

The first published evidence I have found of an asymmetrically localized morphogen receptor in vivo is for the *Drosophila* Wnt receptors, Frizzled-1 and Frizzled-2[100]. Surprisingly, these receptors were found to localize oppositely in the imaginal disc epithelium, with Frizzled-1 localizing at apical junctions and Frizzled-2 evenly distributed along the basolateral membrane. This difference in localization is mediated a difference in the cytoplasmic tail of each receptor. The study's authors further identified that the two receptors differentially stimulate the non-canonical and canonical Wnt signaling pathways, respectively; with such a mechanism, cells could theoretically tune their response to Wnt ligands between the two downstream

pathways, changing their morphogen interpretation by altering the expression ratio for the two receptors. However, perturbation of this differential receptor localization yields only a mild developmental phenotype, altering the organization of hairs on the wing.

TGF-β receptors have been found to localize basolaterally in Madin-Darby canine kidney (MDCK) cells in vitro[101–103]. It was further demonstrated in another study that a short LTAxxVAxxR amino acid motif located near its C-terminus is both necessary and sufficient for the basolateral localization of the type II receptor[104]. Multiple TGF-β family receptors have recently been shown to localize at the basolateral membrane of human embryonic stem cells in vitro[105], suggesting that this localization may be evolutionarily conserved. Nevertheless, it is an open question whether TGF-β family receptors are localized basolaterally in any epithelial tissues in vivo, and whether receptor localization is essential for the proper interpretation of morphogen gradients in any context within mammalian embryos.

In Chapter 2, I report our evidence that the receptor for TGF-β family morphogen BMP is basolaterally localized in human pluripotent stem cells in vitro and in the mouse epiblast in vivo. We further demonstrate that the mis-localization of this receptor via mutation of the LTAxxVAxxR motif is sufficient to induce ectopic BMP signaling around the start of gastrulation, a morphogenetic process for which a BMP gradient is essential. These data also indicate that BMP receptors in the mouse epiblast are shielded from ligands in the apical lumen of the embryo, revealing a previously unappreciated role for the compartmentalization of mouse embryo in organizing a BMP gradient. Concurrent with the original publication of this work, another study was published in 2019 providing evidence for basolateral TGF-β receptor localization in the intestinal epithelium of *C. elegans*[106], with mis-localization leading to defects in developmental patterning. Together, these studies illustrate how embryo geometry and receptor

localization have dramatic implications for understanding how morphogen gradients are formed, detected, and interpreted by developing tissues.

1.4. Mammalian gastrulation and the primitive streak

Here I will give a brief introduction on the process of gastrulation in the mouse and human embryos, which is relevant to the results of Chapters 2 and 3. Mouse and human gastrulation are largely thought to occur in a similar fashion, with the only major difference being the initial geometry of the embryo. Unless stated otherwise, the reader can assume the details I provide are true for gastrulation in both species.

Gastrulation is the process by which the initial undifferentiated embryonic tissue, called the epiblast, differentiates into three germ layers: endoderm, mesoderm, and ectoderm. In mouse, the entire process takes about a single day starting about six days after fertilization[69,107], while human gastrulation occurs throughout the third week of development[108]. This marks the first fate choice made by progenitor cells that determines where in the adult body their progeny will be located. For example, cells in the brain and skin are derived from the ectoderm, while the epithelial lining of many internal organs and the digestive tract arises from the endoderm. The epiblast is not the only tissue in the embryo prior to gastrulation; there are also extra-embryonic tissues, which are also derived from the original fertilized egg cell but do not contribute any cells to the eventual adult body (with some minor exceptions[109]). These extra-embryonic tissues provide essential supportive roles for embryonic development, including during gastrulation.

Prior to gastrulation, the epiblast is organized as an epithelial monolayer of pluripotent cells connected by tight junctions. The mouse epiblast is shaped like a cup,

while the human epiblast is relatively flat. The cup-shape of the mouse epiblast is conserved among rodents but not other mammals like rabbits, which have flatter epiblasts akin to human[110]. In all cases, there is a second epithelial tissue underlying the epiblast, an extra-embryonic tissue called the visceral endoderm (or the hypoblast). In mouse, the epiblast and visceral endoderm form a two-layer cup structure, with another extra-embryonic tissue, the extra-embryonic ectoderm, forming a mirror opposite cup. Between the two cups is a single lumen of extracellular space called the pre-amniotic cavity. This tissue geometry shapes morphogen signaling gradients within the mouse embryo at the start of gastrulation, as discussed in Chapter 2. In human, the epiblast and hypoblast form a bilaminar disc that separates two different lumens, and the extra-embryonic ectoderm equivalent tissue envelops both structures.

In mice, the extra-embryonic ectoderm secretes the TGF-β family morphogen BMP4 to induce signaling in the epiblast, which is essential for gastrulation to initiate[111]. This morphogen induces a cascade of morphogen signaling in the nearby epiblast, inducing the expression of Wnt proteins[112]. Wnt drives the expression of TGF-β family morphogen Nodal, which subsequently promotes own gene expression in a positive feedback loop via an autoregulatory enhancer[113]. Altogether, these three signals drive the differentiation of posterior epiblast cells into mesendoderm, the intermediate progenitor of the mesoderm and endoderm germ layers. Genetic deletion of essential components for any of these three signaling pathways destroys this patterning[111,114–117]. This signaling cascade and its role in initiating gastrulation is thought to be conserved in human embryos[105,118], although direct molecular evidence and perturbations in vivo are lacking.

Around the same time as BMP signals are expressed, inhibitors of TGF-β and Wnt signaling are expressed and secreted from the anterior visceral endoderm[119],

limiting mesendodermal differentiation to the posterior epiblast on the other side. This establishes an anterior-posterior axis in the epiblast that will be maintained throughout the development of the organism. In the mouse embryo, genetic deletion of two TGF-β inhibitors expressed in the anterior visceral endoderm is sufficient to disrupt anterior-posterior axis formation and can result in the formation of multiple embryonic axes[120]. Interestingly, genetic mutations resulting in an overactivation of Wnt signaling can also produce multiple anterior-posterior axes in mice[121–126], suggesting that Wnt morphogens organize the initiation of axis formation while TGF-β inhibitors restrict it from occurring more than once.

The patterning of the epiblast is accompanied by rapid morphogenetic changes. As posterior epiblast cells differentiate into mesendoderm, their changing gene expression induces a sequence of cellular events characterized primarily in mouse[69]. Wnt and Nodal morphogens drive the expression of mesendodermal genes including master regulator Brachyury[115,120,127]. Together with these signals, Brachyury further induces the expression of morphogen FGF8 to initiate the epithelial-mesenchymal-transition (EMT) of differentiating cells[128,129], where cells exit the epithelial layer and ingress beneath the tissue. A variety of events occur during this transition[130]. Nodal signaling in the posterior epiblast directs the expression of matrix metalloproteinases that create necessary perforations in the basement membrane beneath the epiblast[131]. Apical constriction of the surrounding epithelium driven by myosin activity extrude differentiating cells through these basal perforations[132]. Concurrently, FGF signaling drives the expression of transcription factor SNAI1, a repressor of E-cadherin transcription that directly binds E-box sequences in the E-cadherin promoter[133,134]. Loss of E-cadherin subsequently dissolves the adherens and tight junctions connecting mesendodermal cells to the epithelium, releasing them to ingress and migrate beneath

the epiblast where they form the mesoderm and endoderm germ layers. The remaining epiblast cells that do not exit the epithelium during gastrulation give rise to the ectoderm germ layer.

The ingression of mesendodermal cells underneath the epiblast creates an anatomical groove structure in the tissue called the *primitive streak*. In mouse, differentiating cells ingress only through the streak and not the surrounding tissue. Curiously, the primitive streak doesn't spread radially but extends linearly down the middle of the epiblast. In this manner, the primitive streak establishes the midline of the embryo along its anterior-posterior axis, creating a bilateral symmetry. Even though mammalian embryos are shaped differently prior to gastrulation, the linear structure of the primitive streak is strongly evolutionarily conserved among amniotes, including mammals, birds, and some reptiles[110].

It is not clear in any model organism how morphogen signals, which presumably spread radially from their source, organize the formation of a streak that is linear. In chick[79,135] and rabbit[136] embryos, cellular intercalations or processional movements produce the movement of posterior epiblast cells toward the midline and anterior[110]. It has been hypothesized that the Wnt-planar cell polarity pathway organizes these planar movements. However, this interpretation is still a matter of debate since the genetic knockdown of pathway components do not perturb the formation of the primitive streak, only its extension[137]. Live imaging studies of mouse embryos have found no evidence of large-scale cell migrations in the epiblast[138], and cells that enter the primitive streak appear to be uncoordinated in their movements[139,140]. Paradoxically, when *nodal* knockout stem cells are injected into pre-gastrulation mouse embryos and incorporated in the epiblast, they contribute preferentially to anterior epiblast tissues[141], suggesting some role for morphogen signaling in cellular movement toward the streak.

Little is known about the how the primitive streak forms or extends in the human embryo. Existing in vitro human organoid models in which streak fates spontaneously arise fail to produce a structure resembling the primitive streak[142–145]. Authors of a recent review[146] have argued that these results demonstrate that the primitive streak is dispensable for development in vitro. It is not clear to me what these authors mean by "development in vitro", given the understanding that development is a process that results in an organism[147]. Nonetheless, I share their academic view that the demarcation of primitive streak formation as the definitive process establishing an individual human being, with the ethical and legal rights that entails[148], is intellectually arbitrary and misguided.

How does the embryo ensure that only one primitive streak, and thus one anterior-posterior axis, is made? As introduced earlier, genetic perturbations in mice resulting in the misactivation of Wnt signaling[121–126] or the loss of TGF-β inhibitors[120] are the two classes of morphogen signaling pathway mutations known to cause axis duplication through the formation of multiple primitive streaks. This would suggest, at least qualitatively, that the formation of a single primitive streak is a matter of balance between these activators and inhibitors. Still, a quantitative understanding of how these two gradients negotiate the initiation of a streak is missing for mammalian model organisms, and only recently has the importance of such quantitative interactions become appreciated in non-mammalian model organisms.

In the chick embryo, the critical morphogens initiating primitive streak formation in the posterior side and repressing it in the anterior are TGF-β signals VG1/Nodal and BMP4, respectively[149]. Misexpression of VG1 (or Activin) in other regions of the chick embryo is sufficient to induce an ectopic primitive streak[150], while misexpression of BMP antagonizes streak formation[21], providing a simple qualitative model of

competition between the two signals. However, misexpression experiments using combinations of VG1/Activin and BMP sources produce unexpected phenotypes for such a model. For example, while an individual VG1 source transplanted into the anterior epiblast induces an ectopic streak, a group of multiple VG1 sources in the same location does not induce an ectopic streak[81]. Most surprisingly, a weak source of Activin, insufficient to induce an ectopic streak alone, is sufficient to initiate streak formation if it is surrounded by sources of BMP, despite the fact that BMP signaling represses streak initiation in the chick embryo. This striking result indicates that chick epiblast cells effectively interpret the steepness of morphogen gradients rather than their absolute strength when deciding whether to begin forming a primitive streak. How the chick epiblast makes such an interpretation is an open question that merits further study, as is the nature of morphogen gradient interpretation in mammalian model organisms.

In Chapter 3, after describing our novel method to position hundreds of morphogen-soaked beads and human organoids on a single chip in vitro, I discuss how this approach can be used to interrogate the nature of morphogen gradient interpretation in human tissues, particularly in the context of mesendodermal differentiation, EMT, and primitive streak initiation.

1.5.   References

1. Pinto-Correia, C. *The Ovary of Eve: Egg and Sperm and Preformation*. (University of Chicago Press, 2007).

2. Behe, M. J. *Darwin's Black Box: The Biochemical Challenge to Evolution*. (Simon and Schuster, 1996).

3. Darwin, C. *On the Origin of Species by Means of Natural Selection, Or The Preservation of Favoured Races in the Struggle for Life*. (John Murray, Albemarle Street., 1859).

4. H, S. Uber Korrelationen in der Entwicklung des Auges. *Verh Anat Ges* **15**, 61–79 (1901).

5. Lewis, W. H. Experimental studies on the development of the eye in amphibia. I. On the origin of the lens. Rana palustris. *American Journal of Anatomy* **3**, 505–536 (1904).

6. Rogers, K. W. & Schier, A. F. Morphogen gradients: from generation to interpretation. *Annu Rev Cell Dev Biol* **27**, 377–407 (2011).

7. Spemann, H. & Mangold, H. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv f mikr Anat u Entwicklungsmechanik* **100**, 599–638 (1924).

8. Nieuwkoop, P. D. Pattern formation in artificially activated ectoderm (Rana pipiens and Ambystoma punctatum). *Dev Biol* **6**, 255–279 (1963).

9. Chow, R. L. & Lang, R. A. Early eye development in vertebrates. *Annu Rev Cell Dev Biol* **17**, 255–296 (2001).

10. Harland, R. Neural induction. *Curr Opin Genet Dev* **10**, 357–362 (2000).

11. Weinstein, D. C. & Hemmati-Brivanlou, A. Neural induction. *Annu Rev Cell Dev Biol* **15**, 411–433 (1999).

12. Ferguson, E. L. & Anderson, K. V. Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. *Cell* **71**, 451–461 (1992).

13. Turing, A. M. The chemical basis of morphogenesis. 1953. *Bull Math Biol* **52**, 153–197; discussion 119-152 (1990).

14. Stumpf, H. F. Mechanism by which cells estimate their location within the body. *Nature* **212**, 430–431 (1966).

15. Lawrence, P. A. Gradients in the Insect Segment: The Orientation of Hairs in the Milkweed Bug Oncopeltus Fasciatus. *Journal of Experimental Biology* **44**, 607–620 (1966).

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

17. Driever, W. & Nüsslein-Volhard, C. The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell* **54**, 95–104 (1988).

18. Struhl, G., Struhl, K. & Macdonald, P. M. The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57**, 1259–1273 (1989).

19. Muñoz-Sanjuán, I. & Brivanlou, A. H. Neural induction, the default model and embryonic stem cells. *Nat Rev Neurosci* **3**, 271–280 (2002).

20.   Stern, C. D. Neural induction: 10 years on since the 'default model'. *Curr Opin Cell Biol* **18**, 692–697 (2006).

21.   Streit, A. *et al.* Chordin regulates primitive streak development and the stability of induced neural cells, but is not sufficient for neural induction in the chick embryo. *Development* **125**, 507–519 (1998).

22.   Streit, A., Berliner, A. J., Papanayotou, C., Sirulnik, A. & Stern, C. D. Initiation of neural induction by FGF signalling before gastrulation. *Nature* **406**, 74–78 (2000).

23.   Garcia, M. A., Nelson, W. J. & Chavez, N. Cell–Cell Junctions Organize Structural and Signaling Networks. *Cold Spring Harb Perspect Biol* **10**, a029181 (2018).

24.   Walma, D. A. C. & Yamada, K. M. The extracellular matrix in development. *Development* **147**, dev175596 (2020).

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

26.   Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).

27.   Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–56 (2011).

28.   Nakano, T. *et al.* Self-Formation of Optic Cups and Storable Stratified Neural Retina from Human ESCs. *Cell Stem Cell* **10**, 771–785 (2012).

29.   Kim, J., Koo, B.-K. & Knoblich, J. A. Human organoids: model systems for human biology and medicine. *Nat Rev Mol Cell Biol* **21**, 571–584 (2020).

30.   Hayashi, R. *et al.* Generation of 3D lacrimal gland organoids from human pluripotent stem cells. *Nature* (2022) doi:10.1038/s41586-022-04613-4.

31.   Gabriel, E. *et al.* Human brain organoids assemble functionally integrated bilateral optic vesicles. *Cell Stem Cell* **28**, 1740-1757.e8 (2021).

32.   Gurdon, J. B. Embryonic induction — molecular prospects. *Development* **99**, 285–306 (1987).

33.   Hofer, M. & Lutolf, M. P. Engineering organoids. *Nat Rev Mater* **6**, 402–420 (2021).

34.   Tan, T. H., Liu, J. & Grapin-Botton, A. Mapping and exploring the organoid state space using synthetic biology. *Semin Cell Dev Biol* S1084-9521(22)00141–0 (2022) doi:10.1016/j.semcdb.2022.04.015.

35. Kuhn, T. S. *The Structure of Scientific Revolutions*. (University of Chicago Press, 1962).

36. Briscoe, J. & Small, S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009 (2015).

37. Driever, W. & Nüsslein-Volhard, C. A gradient of bicoid protein in Drosophila embryos. *Cell* **54**, 83–93 (1988).

38. Cragg, B. Preservation of extracellular space during fixation of the brain for electron microscopy. *Tissue and Cell* **12**, 63–72 (1980).

39. Chuo, S. T.-Y., Chien, J. C.-Y. & Lai, C. P.-K. Imaging extracellular vesicles: current and emerging methods. *Journal of Biomedical Science* **25**, 91 (2018).

40. Lecuit, T. *et al.* Two distinct mechanisms for long-range patterning by Decapentaplegic in the Drosophila wing. *Nature* **381**, 387–393 (1996).

41. Green, J. B. A. & Smith, J. C. Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature* **347**, 391–394 (1990).

42. Gurdon, J. B., Harger, P., Mitchell, A. & Lemaire, P. Activin signalling and response to a morphogen gradient. *Nature* **371**, 487–492 (1994).

43. McDowell, N., Zorn, A. M., Crease, D. J. & Gurdon, J. B. Activin has direct long-range signalling activity and can form a concentration gradient by diffusion. *Current Biology* **7**, 671–681 (1997).

44. Reilly, K. M. & Melton, D. A. Short-range signaling by candidate morphogens of the TGF beta family and evidence for a relay mechanism of induction. *Cell* **86**, 743–754 (1996).

45. Zecca, M., Basler, K. & Struhl, G. Direct and long-range action of a wingless morphogen gradient. *Cell* **87**, 833–844 (1996).

46. Heemskerk, J. & DiNardo, S. Drosophila hedgehog acts as a morphogen in cellular patterning. *Cell* **76**, 449–460 (1994).

47. Kengaku, M. & Okamoto, H. bFGF as a possible morphogen for the anteroposterior axis of the central nervous system in Xenopus. *Development* **121**, 3121–3130 (1995).

48. Katz, W. S., Hill, R. J., Clandinin, T. R. & Sternberg, P. W. Different Levels of the C. elegans growth factor LIN-3 promote distinct vulval precursor fates. *Cell* **82**, 297–307 (1995).

49. White, R. J., Nie, Q., Lander, A. D. & Schilling, T. F. Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS Biol* **5**, e304 (2007).

50.   Martin, B. L. & Kimelman, D. Wnt-Signaling and the Evolution of Embryonic Posterior Development. *Curr Biol* **19**, R215–R219 (2009).

51.   Huangfu, D. & Anderson, K. V. Signaling from Smo to Ci/Gli: conservation and divergence of Hedgehog pathways from Drosophila to vertebrates. *Development* **133**, 3–14 (2006).

52.   Hosseini, V. *et al.* Wnt lipidation: Roles in trafficking, modulation, and function. *J Cell Physiol* **234**, 8040–8054 (2019).

53.   Serralbo, O. & Marcelle, C. Migrating cells mediate long-range WNT signaling. *Development* **141**, 2057–2063 (2014).

54.   Li, P. *et al.* Morphogen gradient reconstitution reveals Hedgehog pathway design principles. *Science* **360**, 543–548 (2018).

55.   Apte, R. S., Chen, D. S. & Ferrara, N. VEGF in Signaling and Disease: Beyond Discovery and Development. *Cell* **176**, 1248–1264 (2019).

56.   Neubauer, H. *et al.* Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell* **93**, 397–409 (1998).

57.   Ayala-Camargo, A., Ekas, L. A., Flaherty, M. S., Baeg, G.-H. & Bach, E. A. The JAK/STAT pathway regulates proximo-distal patterning in Drosophila. *Developmental Dynamics* **236**, 2721–2730 (2007).

58.   Hayashi, Y. *et al.* Glypicans regulate JAK/STAT signaling and distribution of the Unpaired morphogen. *Development* **139**, 4162–4171 (2012).

59.   Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor-tyrosine kinases. *Cell* **141**, 1117–1134 (2010).

60.   Hori, K., Sen, A. & Artavanis-Tsakonas, S. Notch signaling at a glance. *J Cell Sci* **126**, 2135–2140 (2013).

61.   Manning, B. D. & Toker, A. AKT/PKB Signaling: Navigating the Network. *Cell* **169**, 381–405 (2017).

62.   Bowen, M. E. & Attardi, L. D. The role of p53 in developmental syndromes. *Journal of Molecular Cell Biology* **11**, 200–211 (2019).

63.   Zheng, Y. & Pan, D. The Hippo Signaling Pathway in Development and Disease. *Developmental Cell* **50**, 264–282 (2019).

64.   Massagué, J. TGFβ signalling in context. *Nat Rev Mol Cell Biol* **13**, 616–630 (2012).

65. Mueller, T. D. & Nickel, J. Promiscuity and specificity in BMP receptor activation. *FEBS Lett* **586**, 1846–1859 (2012).

66. Clevers, H. Wnt/β-Catenin Signaling in Development and Disease. *Cell* **127**, 469–480 (2006).

67. Van Essen, D. C. *et al.* The Human Connectome Project: a data acquisition perspective. *Neuroimage* **62**, 2222–2231 (2012).

68. Sagner, A. & Briscoe, J. Morphogen interpretation: concentration, time, competence, and signaling dynamics. *Wiley Interdiscip Rev Dev Biol* **6**, e271 (2017).

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

70. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).

71. Alexandre, C., Baena-Lopez, A. & Vincent, J.-P. Patterning and growth control by membrane-tethered Wingless. *Nature* **505**, 180–185 (2014).

72. Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W. & Tank, D. W. Stability and nuclear dynamics of the Bicoid morphogen gradient. *Cell* **130**, 141–152 (2007).

73. Singh, M. *et al.* Applicability, usability, and limitations of murine embryonic imaging with optical coherence tomography and optical projection tomography. *Biomed. Opt. Express, BOE* **7**, 2295–2310 (2016).

74. Huang, Q. *et al.* Intravital imaging of mouse embryos. *Science* **368**, 181–186 (2020).

75. Piliszek, A., Kwon, G. S. & Hadjantonakis, A.-K. Ex Utero Culture and Live Imaging of Mouse Embryos. *Methods Mol Biol* **770**, 243–257 (2011).

76. Aguilera-Castrejon, A. *et al.* Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature* **593**, 119–124 (2021).

77. Xu, P.-F., Houssin, N., Ferri-Lagneau, K. F., Thisse, B. & Thisse, C. Construction of a vertebrate embryo from two opposing morphogen gradients. *Science* **344**, 87–89 (2014).

78. Gurdon, J. B., Mitchell, A. & Ryan, K. An experimental system for analyzing response to a morphogen gradient. *Proc Natl Acad Sci U S A* **93**, 9334–9338 (1996).

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

80.   Arias, C. F., Herrero, M. A., Stern, C. D. & Bertocchini, F. A molecular mechanism of symmetry breaking in the early chick embryo. *Sci Rep* **7**, 15776 (2017).

81.   Lee, H. C. *et al.* 'Neighbourhood watch' model: embryonic epiblast cells assess positional information in relation to their neighbours. *Development* dev.200295 (2022) doi:10.1242/dev.200295.

82.   Pólya, G. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Mathematische Annalen* **84**, 149–160 (1921).

83.   Doyle, P. G. & Snell, J. L. Random Walks and Electric Networks. *arXiv:math/0001057* (2000).

84.   Ben-Naim, E., Krapivsky, P. L. & Rednor, S. Fundamental Kinetic Processes. http://physics.bu.edu/~redner/542/book/ (2008).

85.   Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).

86.   Bryant, D. M. & Mostov, K. E. From cells to organs: building polarized tissue. *Nat Rev Mol Cell Biol* **9**, 887–901 (2008).

87.   Durdu, S. *et al.* Luminal signalling links cell communication to tissue architecture during organogenesis. *Nature* **515**, 120–124 (2014).

88.   Harmansa, S., Alborelli, I., Bieli, D., Caussinus, E. & Affolter, M. A nanobody-based toolset to investigate the role of protein localization and dispersal in Drosophila. *Elife* **6**, e22549 (2017).

89.   Kovbasnjuk, O., Leader, J. P., Weinstein, A. M. & Spring, K. R. Water does not flow across the tight junctions of MDCK cell epithelium. *Proceedings of the National Academy of Sciences* **95**, 6526–6530 (1998).

90.   Liang, G. H. & Weber, C. R. Molecular aspects of tight junction barrier function. *Curr Opin Pharmacol* **19**, 84–89 (2014).

91.   Vincent, J.-P. & Dubois, L. Morphogen transport along epithelia, an integrated trafficking problem. *Dev Cell* **3**, 615–623 (2002).

92.   Ashe, H. L. & Briscoe, J. The interpretation of morphogen gradients. *Development* **133**, 385–394 (2006).

93.   Banerjee, U., Renfranz, P. J., Pollock, J. A. & Benzer, S. Molecular characterization and expression of sevenless, a gene involved in neuronal pattern formation in the Drosophila eye. *Cell* **49**, 281–291 (1987).

94.   Banerjee, U., Renfranz, P. J., Hinton, D. R., Rabin, B. A. & Benzer, S. The sevenless+ protein is expressed apically in cell membranes of developing Drosophila retina; it is not restricted to cell R7. *Cell* **51**, 151–158 (1987).

95.   Krämer, H., Cagan, R. L. & Zipursky, S. L. Interaction of bride of sevenless membrane-bound ligand and the sevenless tyrosine-kinase receptor. *Nature* **352**, 207–212 (1991).

96.   Zak, N. B. & Shilo, B. Z. Localization of DER and the pattern of cell divisions in wild-type and Ellipse eye imaginal discs. *Dev Biol* **149**, 448–456 (1992).

97.   Simske, J. S., Kaech, S. M., Harp, S. A. & Kim, S. K. LET-23 receptor localization by the cell junction protein LIN-7 during C. elegans vulval induction. *Cell* **85**, 195–204 (1996).

98.   Vermeer, P. D. *et al.* Segregation of receptor and ligand regulates activation of epithelial growth factor receptor. *Nature* **422**, 322–326 (2003).

99.   de Vreede, G., Gerlach, S. U. & Bilder, D. Epithelial monitoring through ligand-receptor segregation ensures malignant cell elimination. *Science* **376**, 297–301 (2022).

100. Wu, J., Klein, T. J. & Mlodzik, M. Subcellular Localization of Frizzled Receptors, Mediated by Their Cytoplasmic Tails, Regulates Signaling Pathway Specificity. *PLoS Biol* **2**, e158 (2004).

101. Murphy, S. J. *et al.* Differential Trafficking of Transforming Growth Factor-β Receptors and Ligand in Polarized Epithelial Cells. *Mol Biol Cell* **15**, 2853–2862 (2004).

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

103. Yin, X. *et al.* Basolateral delivery of the type I transforming growth factor beta receptor is mediated by a dominant-acting cytoplasmic motif. *Mol Biol Cell* **28**, 2701–2711 (2017).

104. Murphy, S. J., Shapira, K. E., Henis, Y. I. & Leof, E. B. A unique element in the cytoplasmic tail of the type II transforming growth factor-beta receptor controls basolateral delivery. *Mol Biol Cell* **18**, 3788–3799 (2007).

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

106. Lin, J., Vora, M., Kane, N. S., Gleason, R. J. & Padgett, R. W. Human Marfan and Marfan-like Syndrome associated mutations lead to altered trafficking of the Type II TGFβ receptor in Caenorhabditis elegans. *PLoS One* **14**, e0216628 (2019).

107. Tam, P. P. L. & Loebel, D. A. F. Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet* **8**, 368–381 (2007).

108. Zhai, J., Xiao, Z., Wang, Y. & Wang, H. Human embryonic development: from peri-implantation to gastrulation. *Trends Cell Biol* **32**, 18–29 (2022).

109. Kwon, G. S., Viotti, M. & Hadjantonakis, A.-K. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev Cell* **15**, 509–520 (2008).

110. Stower, M. J. & Bertocchini, F. The evolution of amniote gastrulation: the blastopore-primitive streak transition: Evolution of amniote gastrulation. *WIREs Dev Biol* **6**, e262 (2017).

111. Mishina, Y., Suzuki, A., Ueno, N. & Behringer, R. R. Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes Dev* **9**, 3027–3037 (1995).

112. Ben-Haim, N. *et al.* The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Dev Cell* **11**, 313–323 (2006).

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

114. Conlon, F. L. *et al.* A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919–1928 (1994).

115. Dunn, N. R., Vincent, S. D., Oxburgh, L., Robertson, E. J. & Bikoff, E. K. Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* **131**, 1717–1728 (2004).

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

117. Kelly, O. G., Pinson, K. I. & Skarnes, W. C. The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development* **131**, 2803–2815 (2004).

118. Martyn, I., Brivanlou, A. H. & Siggia, E. D. A wave of WNT signaling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. *Development* **146**, dev172791 (2019).

119. Perea-Gómez, A., Shawlot, W., Sasaki, H., Behringer, R. R. & Ang, S. HNF3beta and Lim1 interact in the visceral endoderm to regulate primitive streak formation and anterior-posterior polarity in the mouse embryo. *Development* **126**, 4499–4511 (1999).

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

121. Zeng, L. *et al.* The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell* **90**, 181–192 (1997).

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

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

124. Mukhopadhyay, M. *et al.* Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation. *Development* **130**, 495–505 (2003).

125. Escalante-Alcalde, D. *et al.* The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis and axis patterning. *Development* **130**, 4623–4637 (2003).

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

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

128. Sun, X., Meyers, E. N., Lewandoski, M. & Martin, G. R. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes Dev* **13**, 1834–1846 (1999).

129. Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev Cell* **1**, 37–49 (2001).

130. Yang, J. & Weinberg, R. A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* **14**, 818–829 (2008).

131. Kyprianou, C. *et al.* Basement membrane remodelling regulates mouse embryogenesis. *Nature* **582**, 253–258 (2020).

132. Ramkumar, N. *et al.* Crumbs2 promotes cell ingression during the epithelial-to-mesenchymal transition at gastrulation. *Nat Cell Biol* **18**, 1281–1291 (2016).

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

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

135. Voiculescu, O., Bertocchini, F., Wolpert, L., Keller, R. E. & Stern, C. D. The amniote primitive streak is defined by epithelial cell intercalation before gastrulation. *Nature* **449**, 1049–1052 (2007).

136. Halacheva, V. *et al.* Planar cell movements and oriented cell division during early primitive streak formation in the mammalian embryo. *Dev Dyn* **240**, 1905–1916 (2011).

137. Chuai, M. *et al.* Cell movement during chick primitive streak formation. *Dev Biol* **296**, 137–149 (2006).

138. Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M. & Sutherland, A. The mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev Dyn* **241**, 270–283 (2012).

139. Ichikawa, T. *et al.* Live Imaging of Whole Mouse Embryos during Gastrulation: Migration Analyses of Epiblast and Mesodermal Cells. *PLoS One* **8**, e64506 (2013).

140. McDole, K. *et al.* In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell* **175**, 859-876.e33 (2018).

141. Lu, C. C. & Robertson, E. J. Multiple roles for Nodal in the epiblast of the mouse embryo in the establishment of anterior-posterior patterning. *Dev Biol* **273**, 149–159 (2004).

142. Moris, N. *et al.* An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415 (2020).

143. van den Brink, S. C. *et al.* Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409 (2020).

144. Girgin, M. U. *et al.* Bioengineered embryoids mimic post-implantation development in vitro. *Nat Commun* **12**, 5140 (2021).

145. Rossi, G. *et al.* Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell* **28**, 230-240.e6 (2021).

146. Sheng, G., Martinez Arias, A. & Sutherland, A. The primitive streak and cellular principles of building an amniote body through gastrulation. *Science* **374**, abg1727 (2021).

147. Barresi, M. J. F. & Gilbert, S. F. *Developmental Biology*. (Oxford University Press, 2019).

148. Lovell-Badge, R. *et al.* ISSCR Guidelines for Stem Cell Research and Clinical Translation: The 2021 update. *Stem Cell Reports* **16**, 1398–1408 (2021).

149. Raffaelli, A. & Stern, C. D. Signaling events regulating embryonic polarity and formation of the primitive streak in the chick embryo. *Curr Top Dev Biol* **136**, 85–111 (2020).

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

Chapter 2. Mouse embryo geometry drives formation of robust signaling gradients through receptor localization.

Abstract

Morphogen signals are essential for cell fate specification during embryogenesis. Some receptors that sense these morphogens are known to localize to only the apical or basolateral membrane of polarized cell lines in vitro. How such localization affects morphogen sensing and patterning in the developing embryo remains unknown. Here, we show that the formation of a robust BMP signaling gradient in the early mouse embryo depends on the restricted, basolateral localization of BMP receptors. The mis-localization of receptors to the apical membrane results in ectopic BMP signaling in the mouse epiblast in vivo. With evidence from mathematical modeling, human embryonic stem cells in vitro, and mouse embryos in vivo, we find that the geometric compartmentalization of BMP receptors and ligands creates a signaling gradient that is buffered against fluctuations. Our results demonstrate the importance of receptor localization and embryo geometry in shaping morphogen signaling during embryogenesis.

2.1. Introduction

Morphogens are long-range signaling molecules that move in extracellular space to induce concentration-dependent cellular responses in their target tissues[1,2]. Genetic perturbation of morphogens and their cognate receptors often leads to missing cell types and embryonic structures[3–7]. Many mechanisms have been proposed to explain how morphogens induce signaling gradients in target tissues and therefore direct the spatial organization of cell fates[1,2,8–19]. Surprisingly, several morphogen receptors have been found to localize to either the apical or basolateral membrane of epithelial tissues[15,20–24]. Such localization can dramatically affect how the target tissue senses morphogens[15,20,23]. How receptor localization modulates morphogen signaling in developing embryos is not known.

The early mouse embryo (E6.0–E6.5) adopts an egg-cylinder geometry[5,6,25] (Figure 2.1a). It contains a lumen (the pre-amniotic cavity) encased by two epithelial tissues: the epiblast and extraembryonic ectoderm (ExE). The ExE secretes the morphogen BMP4, which is sensed by receptors in the epiblast[5–7]. The resulting BMP signaling is required for the differentiation of the epiblast into mesoderm[3,4]. Both the epiblast and ExE have stereotyped epithelial tissue geometries[26], with their apical membranes surrounding the lumen and their basolateral membranes facing a narrow interstitial space (between these tissues and the underlying visceral endoderm [VE]). This lumen and interstitial space are separated by impermeable tight junctions present throughout the epithelia except at the border between the ExE and epiblast (Figure 2.1a). Indeed, when small-molecule dye fluorescein was injected into the pre-amniotic cavity of an E6.5 mouse embryo, it did not penetrate the epiblast or ExE but diffused through a channel at the edge of the epiblast (Figure 2.2). Thus, the extracellular space
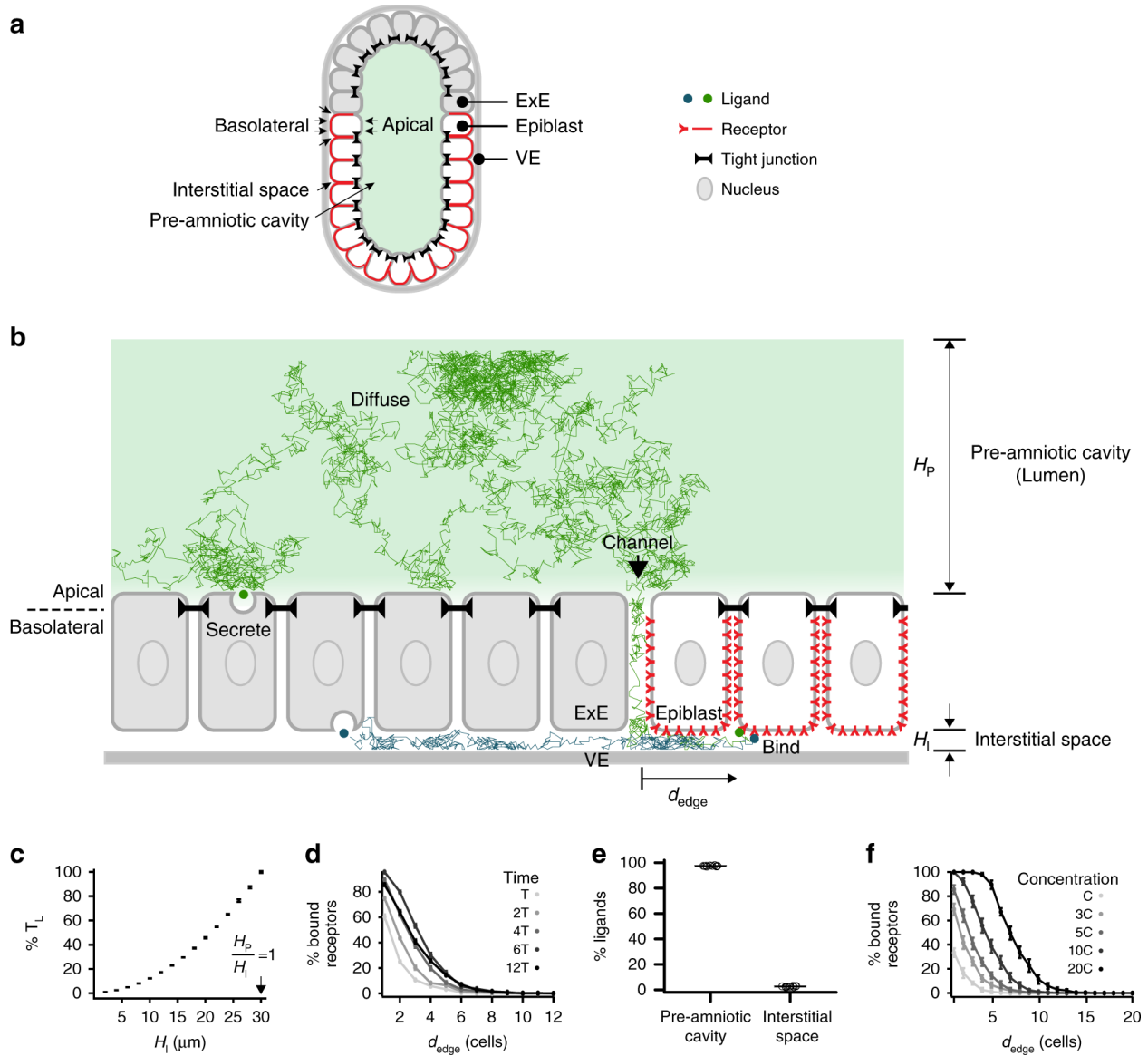
Figure 2.1: Receptor localization facilitates the formation of a robust signaling gradient in early mouse embryo. a, Illustration of pre-gastrulation mouse embryo, with the epiblast (white) and extraembryonic ectoderm (ExE, light gray) together enclosing the pre-amniotic cavity. Apical membranes of epiblast cells face the pre-amniotic cavity whereas basolateral membranes face the interstitial space.
b, Illustration of a simulation with basolateral receptors. ExE cells secrete BMP4 ligands from their apical (green) or basolateral (blue) membranes, while epiblast cells have BMP receptors (red) on their basolateral membranes. Ligands cannot diffuse past tight junctions between cells (black). Simulated ligand trajectories show that ligands diffuse from epiblast edge (black arrow) through interstitial space to approach and bind basolateral receptors. H_P and H_I denote heights of pre-amniotic cavity and interstitial space, respectively. **c,** The time between BMP4 ligands entering interstitial space and being captured by receptors, T_L, increases with the height of the interstitial space.

43

Figure 2.1 (*continued*): d, Percentage of ligand-bound receptors as a function of their distance from epiblast edge, d_{edge}, over time in simulations with apically secreted ligands (T = 7.5 min). **e,** Percentage of ligands in pre-amniotic cavity vs. interstitial space at steady state (90 min) in simulations with apically secreted ligands. **f,** Percentage of ligand-bound receptors as a function of dedge at steady state (90 min) shows signaling gradients at different BMP4 concentrations in simulations with apically secreted ligands (C = 0.16 ng/mL or ligand/receptor ratio of 0.1). Error bars denote SEM

in the embryo through which BMP4 ligands diffuse is compartmentalized into a lumen

and an interstitial space.

Here, by combining mathematical modeling, quantitative imaging,

embryological perturbation, and microfluidics, we demonstrate that restricted receptor

localization in conjunction with the compartmentalized embryo geometry constrains

the diffusion of and therefore response to BMP4 ligands. We show that the BMP

signaling gradient arises from the edge of the epiblast even under conditions of uniform

BMP4 stimulation. Further, the interplay between restricted receptor localization and

the compartmentalized embryo geometry buffers BMP4 ligands in the pre-amniotic

cavity through an entropic effect. This entropic buffering renders the formation of BMP

signaling gradient robust to fluctuations in BMP4 level. Consistently, mis-localizing

BMP receptors in the mouse embryo leads to ectopic BMP signaling. Thus, receptor

localization and embryo geometry together play an essential role in regulating

morphogen signaling during early development.

Figure 2.2 (*following page*): An extracellular channel is present between the epiblast and the extraembryonic ectoderm. a, Illustration of pre-gastrulation mouse embryo, with the epiblast (white) and extraembryonic ectoderm (ExE, light gray) together enclosing the pre-amniotic cavity. Apical membranes of epiblast cells face the pre-amniotic cavity whereas basolateral membranes face the visceral endoderm (VE, gray). **b**, Phase, fluorescence, and color-combined images of an E6.5 mouse embryo after microinjection of fluorescein into pre-amniotic cavity. Epiblast is impermeable, but fluorescein diffuses through the gap at the edge of epiblast (border between epiblast and ExE, pink arrow). A and P denote anterior and posterior, respectively. These images are representative of three images from three embryos. **c**, Top: Four sagittal sections of an E6.5 embryo stained for DNA, epiblast marker OCT4, and tight junction marker ZO-1. Dotted boxes denote different areas of epiblast, in which ZO-1 intensity was quantified. Bottom: Average ZO-1 intensity at the edge (pink) and the rest of epiblast (gray). These images are representative of two sets of images from 2 embryos. Scale bar 20 μm. Error bars denote SEM

Figure 2.2 (*continued*):
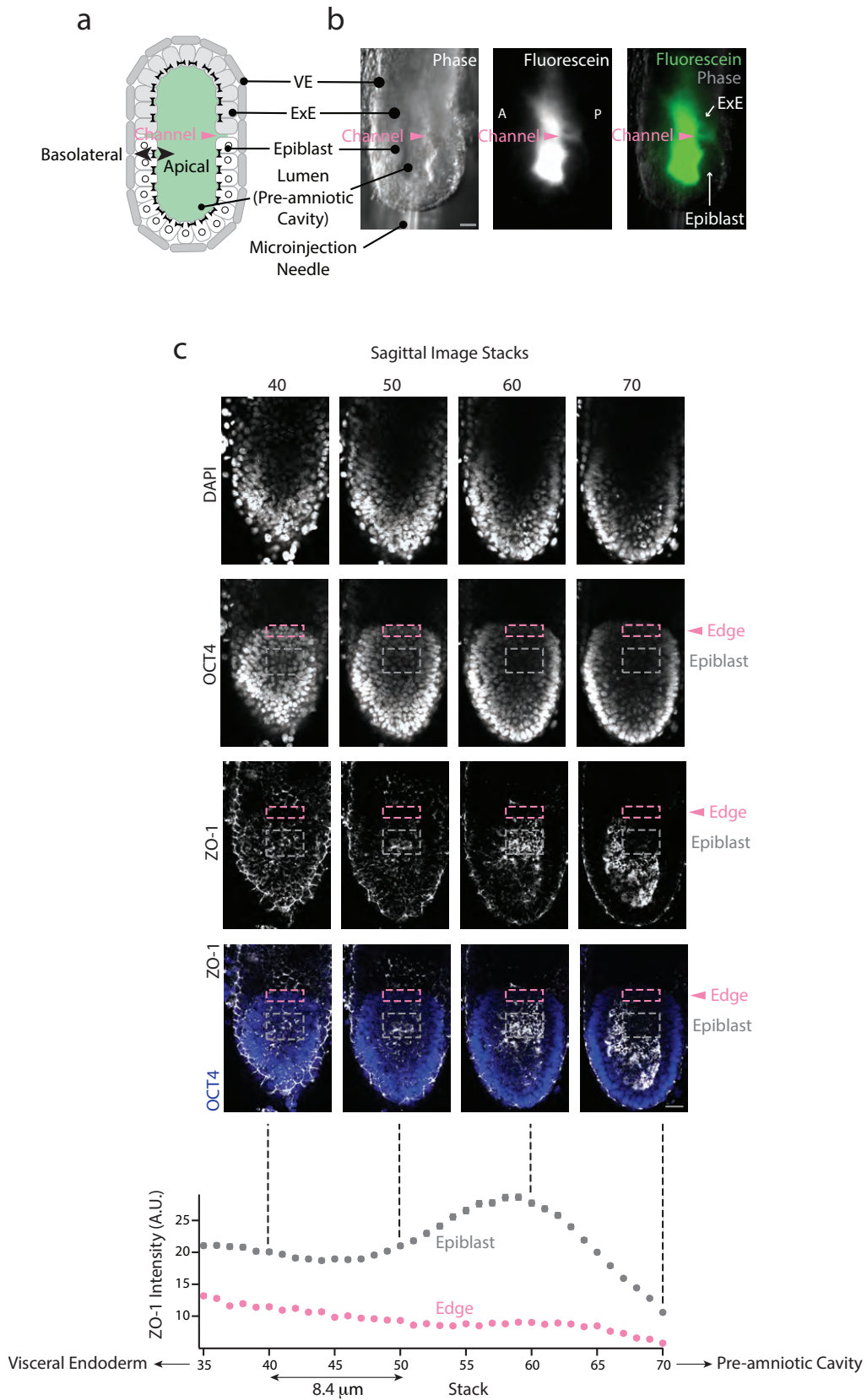
2.2. Results

2.2.1. Receptor localization facilitates the formation of signaling gradient

To understand how receptor localization impacts BMP signaling between the ExE and epiblast, we simulated the movement of individual BMP4 ligands in the early mouse embryo (E6.0–E6.5) from secretion to receptor binding, using Brownian dynamics[27]. Given the evidence of polarized ligand secretion by epithelial cells in vitro[20,21] (Figure 2.3), we modeled different instances in which BMP4 ligands were secreted apically (into the pre-amniotic cavity) or basolaterally (into the interstitial space) by the ExE (Figure 2.1b). After secretion, ligands diffused through extracellular space in the embryo. Due to tight junctions in the simulation, ligands could move between the pre-amniotic cavity and the interstitial space only by diffusing through the channel between the ExE and epiblast. Some morphogen receptors are known to localize to only the apical or basolateral membranes of epithelial cells[15,20,22,24,28,29]; such localization could determine the compartment from which ligands are sensed by receptors in the epiblast. Therefore, we also performed simulations with BMP receptors localized exclusively on the apical membrane (facing the pre-amniotic cavity) or basolateral membrane (facing the interstitial space) of epiblast cells. Finally, our model assumed that once BMP4 ligands bound their receptors, signaling activity was induced and the ligands were cleared.

Our simulations show that if the BMP receptors are basolaterally localized in the epiblast, the compartmentalized geometry of the embryo naturally results in the formation of a robust BMP signaling gradient. This occurs despite the absence of other regulatory mechanisms such as signaling inhibitors[2,10–12,15]. The basolateral localization of BMP receptors requires that ligands diffuse through the interstitial space between the

47
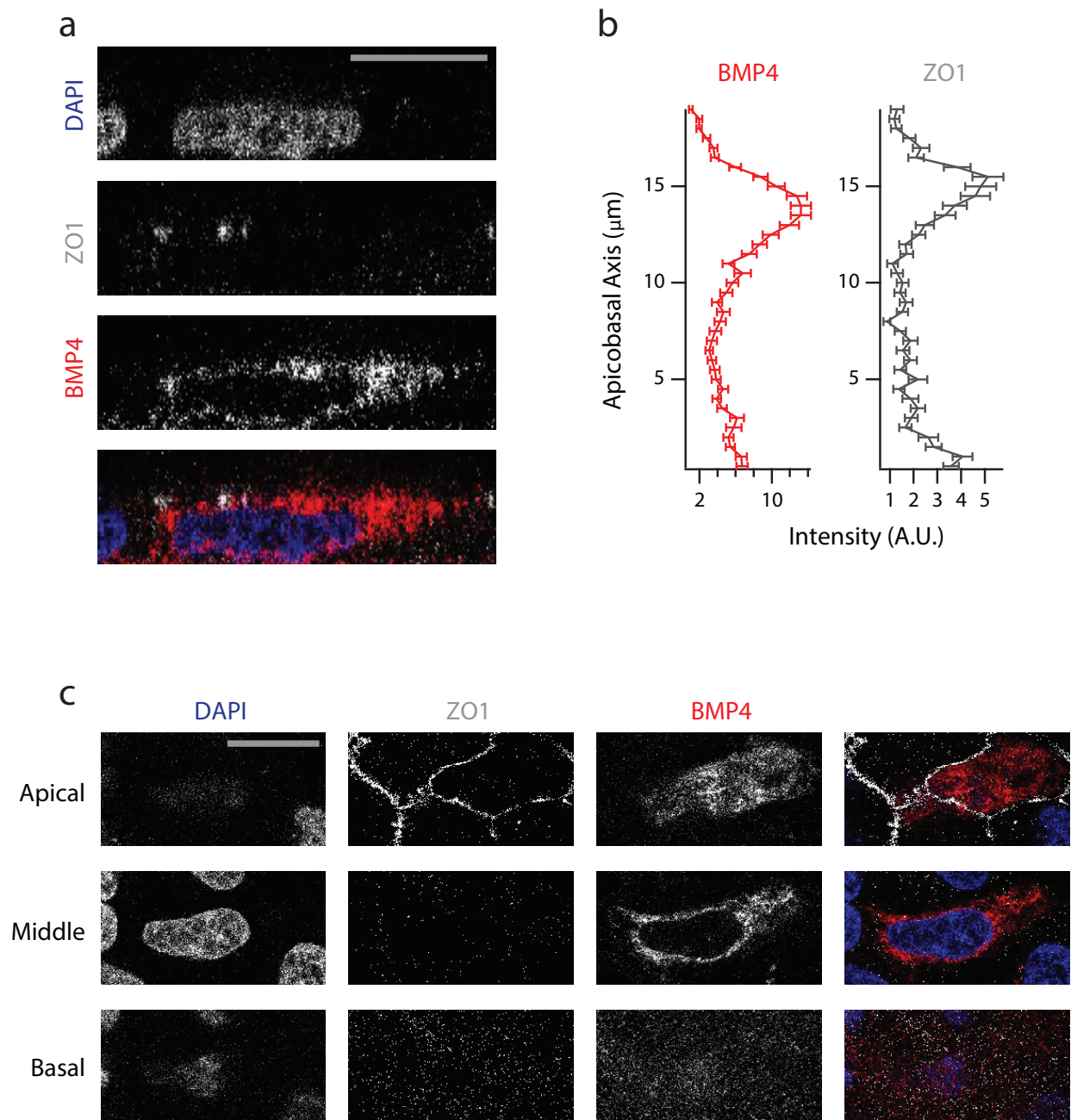
Figure 2.3: BMP4 localization in hESCs. a, Confocal images in lateral view of a hESC transfected with plasmid expressing GFP-BMP4 (red), immunostained for DNA (blue) and ZO-1 (white). **b,** Plots of GFP-BMP4 (left) and ZO-1 (right) intensity along the apicobasal axis show presence of GFP-BMP4 near apical membrane (n=6 cells from 3 experiments). **c,** Apical, middle, and basal confocal stacks of a transfected hESC immunostained as in **(a).** Error bars denote SEM. Scale bar 10 μm

epiblast and VE to access them (Figure 2.1b). The height of this interstitial space, H_I, regulates the time, T_L, and hence the distance a ligand can diffuse before being captured by a receptor (Figure 2.1c). As a consequence, BMP4 ligands are more likely to bind receptors that are closer to the epithelial edge, giving rise to a BMP signaling gradient from the edge of the epiblast inward (Figure 2.1d). The signaling gradient forms regardless of whether BMP4 ligands are secreted from the apical or basolateral membrane of the ExE and arises even if ligands are imposed to be uniformly distributed in the pre-amniotic cavity (Figure 2.4a).

The basolateral localization of BMP receptors, in conjunction with the asymmetric compartmentalization of the embryo, also makes formation of this BMP signaling gradient robust to fluctuations in the BMP4 source strength. Due to the large volume difference between the pre-amniotic cavity and interstitial space and the channel (between the ExE and epiblast) that connects them, the majority of BMP4 ligands accumulate in the cavity on the apical side of the epiblast (Figure 2.1e, Figure 2.4b). This is an entropic effect: the entropy of BMP4 ligands is maximized when the ligands are uniformly distributed between the pre-amniotic cavity and the interstitial space. In other words, the accumulation of BMP4 ligands in the cavity is driven by the same physical forces that allows ink to diffuse through water and ultimately reach a uniform distribution independent of where ink is dropped initially. Consistently, BMP4 ligands accumulate in the pre-amniotic cavity, regardless of whether the ligands are secreted apically or basolaterally from the ExE in the simulation (Figure 2.1e, Figure 2.4b).

This accumulation results in an entropic buffering effect: the pre-amniotic cavity serves as a ligand reservoir that buffers the signaling gradient against fluctuations in the BMP source strength. Indeed, if the total ligand concentration is increased by

Figure 2.4 (*following page*): Formation of robust BMP signaling gradient is insensitive to BMP4 secretion mechanism. a, Percentage of ligand-bound receptors as function of distance from epiblast edge, d_{edge}, over time in simulations where ligands are secreted apically (left), secreted basolaterally (middle), or presented uniformly in lumen (right). T=5 min. **b**, Percentage of unbound ligands in pre-amniotic cavity (P) and interstitial space (I) at steady state (30 min) in simulations where ligands are secreted apically or basally from ExE. Here, C=0.32 ng/mL. **c**, Percentage of ligand-bound receptors as function of d_{edge} at steady state in simulations where ligands are secreted apically (left), secreted basolaterally (middle), or presented uniformly in lumen (right). Rows correspond to simulations where receptors are localized basolaterally (top) or apically (center), or where tight junctions are absent (bottom). Error bars denote SEM

Figure 2.4 (*continued*):
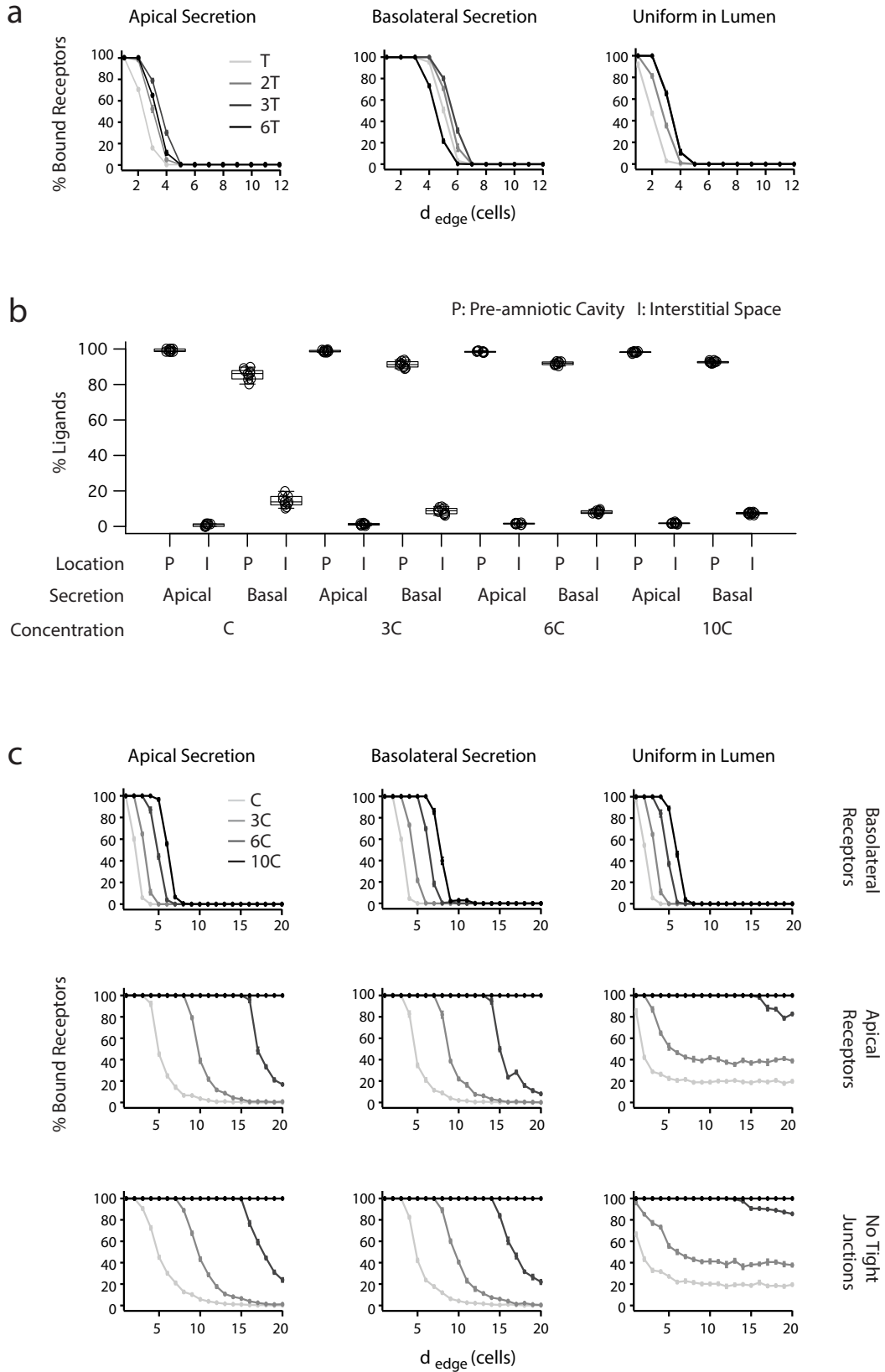
Figure 2.5 (*following page*): BMP signaling gradient as a function of pre-amniotic cavity size. a, Variation of steady state (90 min) BMP signaling depth as a function of pre-amniotic cavity height (H_p) in simulations where ligands are secreted apically (left) or basolaterally (right) from ExE. Darkness of points and curves corresponds to increased ratio of secreted ligands to receptors. **b**, Percentage of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, in simulations at steady state where ligands are secreted apically (left) or basolaterally (right). Rows correspond to different pre-amniotic cavity heights. Error bars denote SEM

Figure 2.5 (*continued*):

tenfold in a simulation with basolateral receptors, the signaling gradient shifts inward by only a few cell widths (Figure 2.1f, Figure 2.4c). However, as the size of the pre-amniotic cavity is reduced in the simulation, increases in ligand concentration shift the signaling gradient significantly further into the epiblast (Figure 2.5), demonstrating the buffering effect. Strikingly, if BMP receptors are apically localized in the epiblast or if tight junctions are absent, this tenfold increase is sufficient to saturate all receptors in the simulation and destroy the signaling gradient (Figure 2.4c). Thus, the entropic buffering of the BMP signaling gradient relies upon both the basolateral receptor localization and embryonic geometry in our simulation. Variations in other simulation parameters, such as the ligand diffusion coefficient D, the probability of binding between ligand and unbound receptors upon contact $P_{binding}$, and the turnover rate of ligand–receptor pairs T_t, do not similarly disrupt the formation of this signaling gradient (Figures 2.6-2.8). Likewise, the signaling gradient forms regardless of whether the embryo is rotationally symmetric or if the channel between the ExE and the epiblast is present only at the posterior side in the simulation (Figure 2.9), even though in the latter case the gradient is more prominent at the posterior side. In summary, our simulations demonstrate that the formation of the signaling gradient is robust in that it can form under wide variety of condition; and further, while the scale of the gradient increase with source strength, this increase is limited by basolateral receptor localization and the asymmetric compartmentalization of the embryo.

Assuming that the BMP receptors are basolaterally localized, our model provides three experimentally testable predictions. First, a BMP signaling gradient will form inward from the epiblast edge even if ligands are present at high concentration throughout the lumen (Figure 2.4a). Second, formation of this signaling gradient will be robust to fluctuations in BMP concentration (Figure 2.1f). Third, the mis-localization of

54

Figure 2.6 (*following page*): BMP signaling gradient as a function of simulation parameters. a, Percentage of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, in simulations at steady state (90 min) with different BMP4 diffusion coefficient, D. Here, D is varied from 10 (lightest) to 80 μm^2/s (darkest). **b,** Percentage of ligand-bound receptors as a function of d_{edge} in simulations at steady state with different turnover times of ligand-receptor pairs, T_t. Here, T_t is varied from 15 min (lightest) to 60 min (darkest). **c,** Percentage of ligand-bound receptors as a function of d_{edge} in simulations at steady state with different probability of ligand binding to an unbound receptor upon contact ($P_{binding}$). Here, $P_{binding}$ is varied from a diffusion-limited regime (dark) to a regime that is not diffusion-limited (light). **d,** BMP signaling depth as a function of D and T_t in simulations at steady state in which $P_{binding}$=0.2. **e,** Same as **(d)** but for $P_{binding}$=0.02. **f,** Same as **(d)** but for $P_{binding}$ =0.002. Here, ligand-receptor ratio is 0.5. Error bars denote SEM

Figure 2.6 (*continued*):
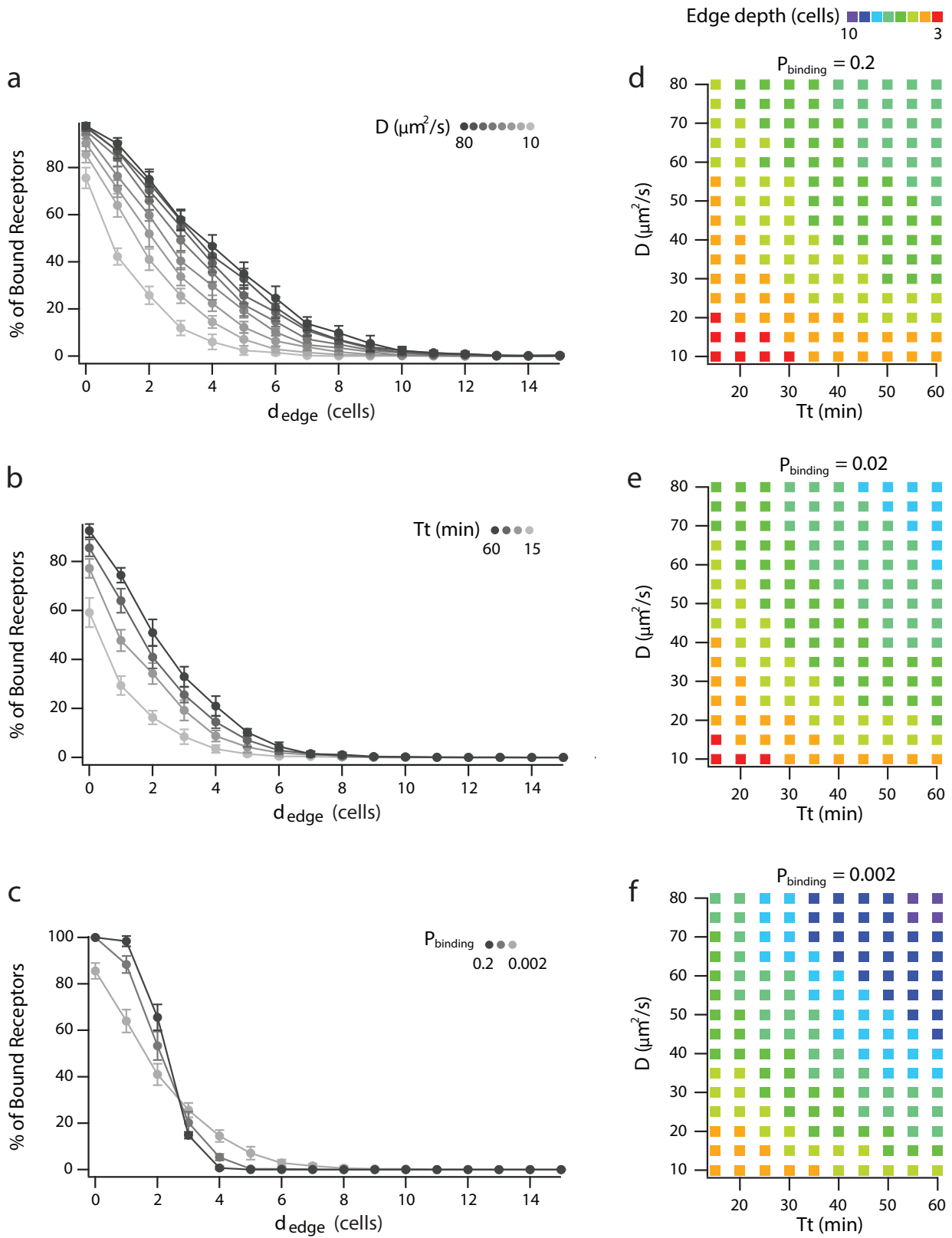
Figure 2.7 (*following page*): BMP signaling gradient in equilibrium and non-equilibrium simulations. a, Percentage of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, in simulations at steady state (90 min). Ligand secretion and ligand-receptor turnover rates are coupled so that the total number of ligands remains constant. Darkness of points and curves corresponds to increased ratio of secreted ligands to receptors. **b,** Same as in **(a)** but for non- equilibrium simulation at T=90 min in which total ligand number steadily decreases. Here, no new ligands are secreted after an initial burst at T=0 min. **c,** Same as in **(a)** but for non-equilibrium simulation at T=90 min in which total ligand number steady increases. Here, ligand secretion and ligand-receptor turnover rates are uncoupled, and the secretion rate is significantly higher than the turnover rate. Error bars denote SEM

Figure 2.7 (*continued*):

a — $K_{ligand-release} = K_{receptor-turnover}$
Total concentration remains constant

b — $K_{ligand-release} < K_{receptor-turnover}$
Total concentration decreases by 18% over 90 min

c — $K_{ligand-release} > K_{receptor-turnover}$
Total concentration increases by 30% over 90 min

Ligand/Receptor: 2 1 0.5 0.1

% of Bound Receptors

d_{edge} (cells)

58
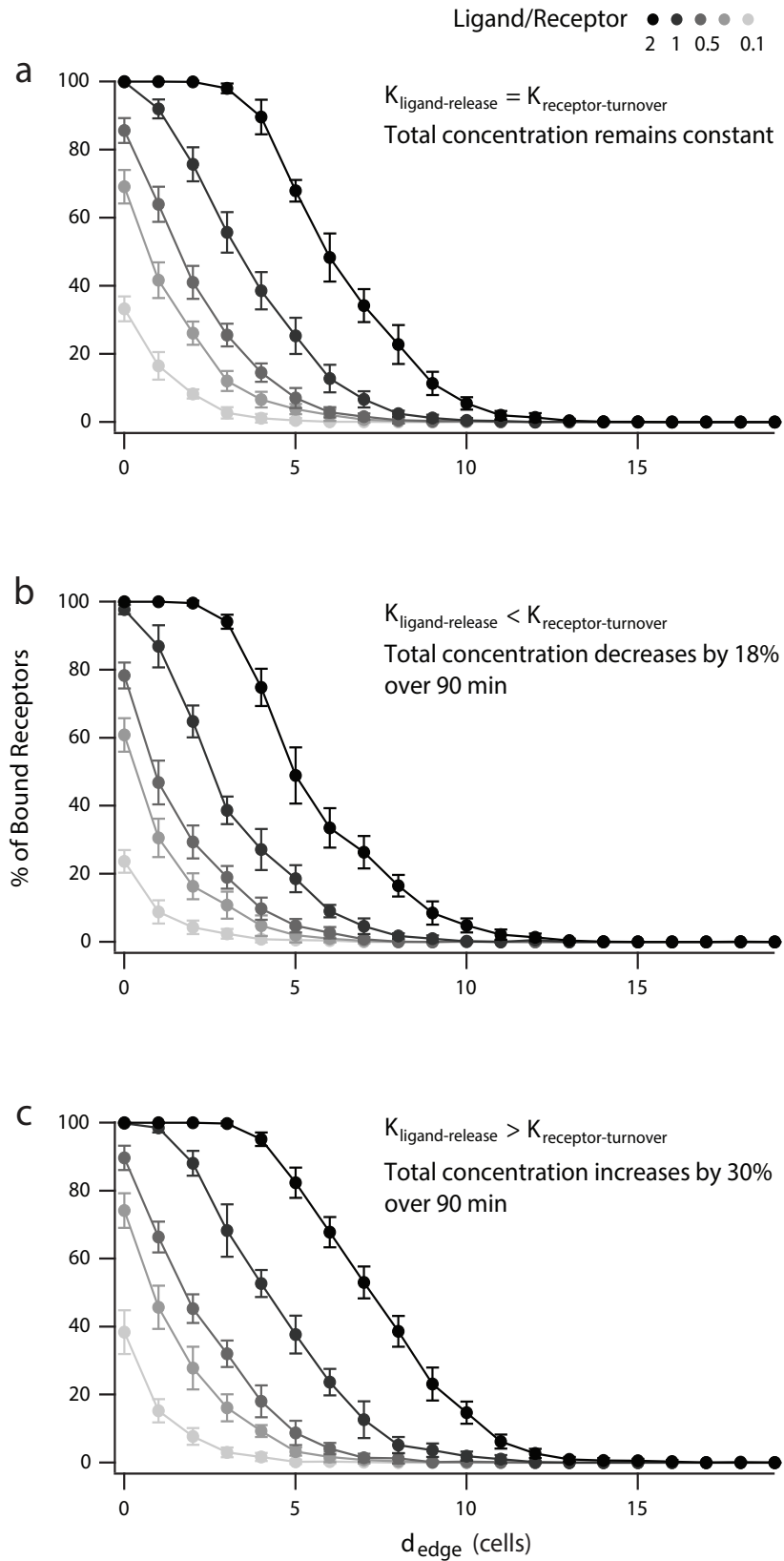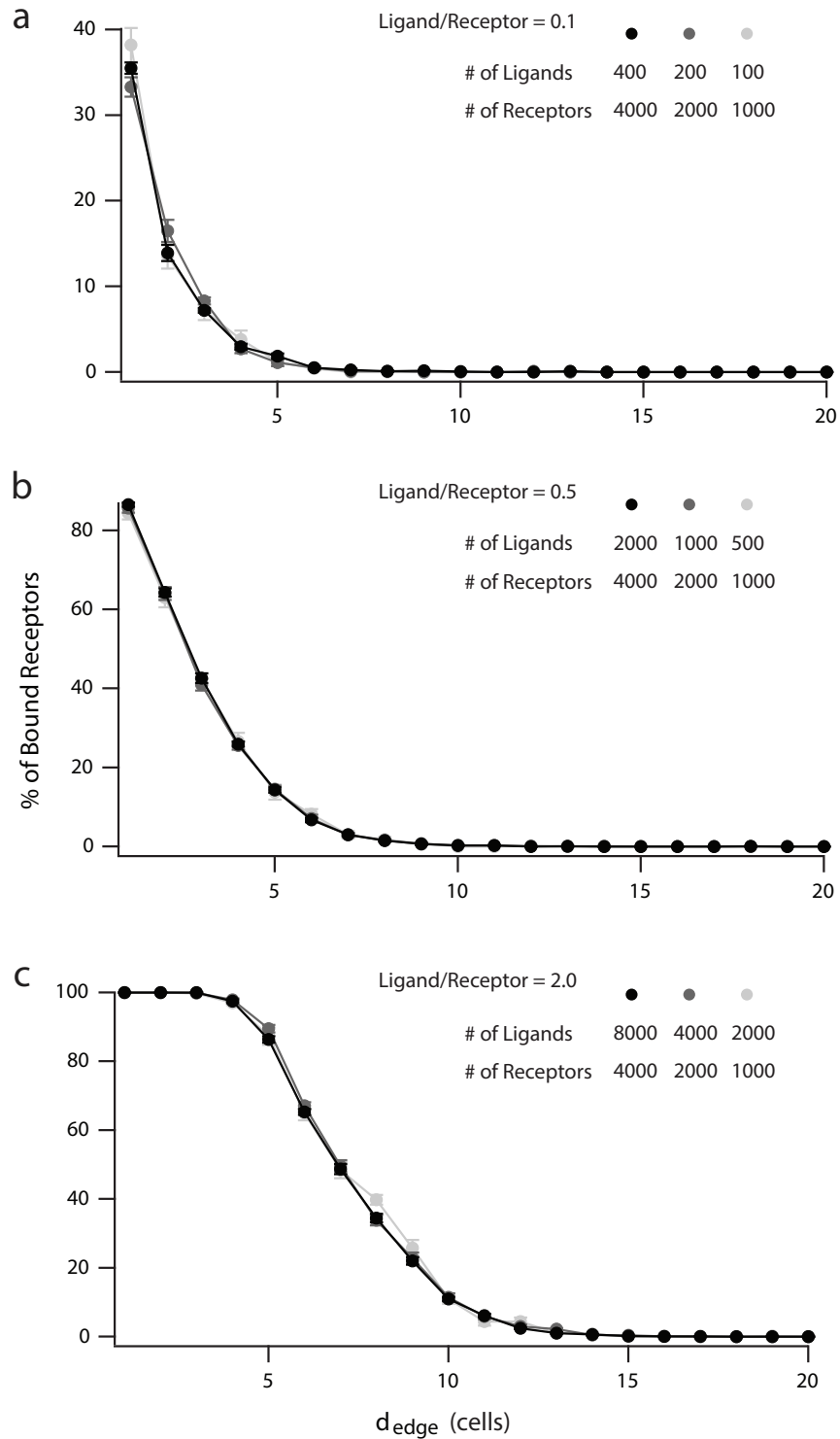
Figure 2.8: BMP signaling gradient in simulation as a function of total ligand and receptor numbers. a, Percentage of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, in simulations at steady state (90 min) with a ligand-receptor ratio of 0.1, for different total numbers of ligands and receptors. **b,** Same as **(a)** but for a ligand-receptor ratio of 0.5. **c,** Same as **(a)** but for a ligand-receptor ratio of 2. Error bars denote SEM
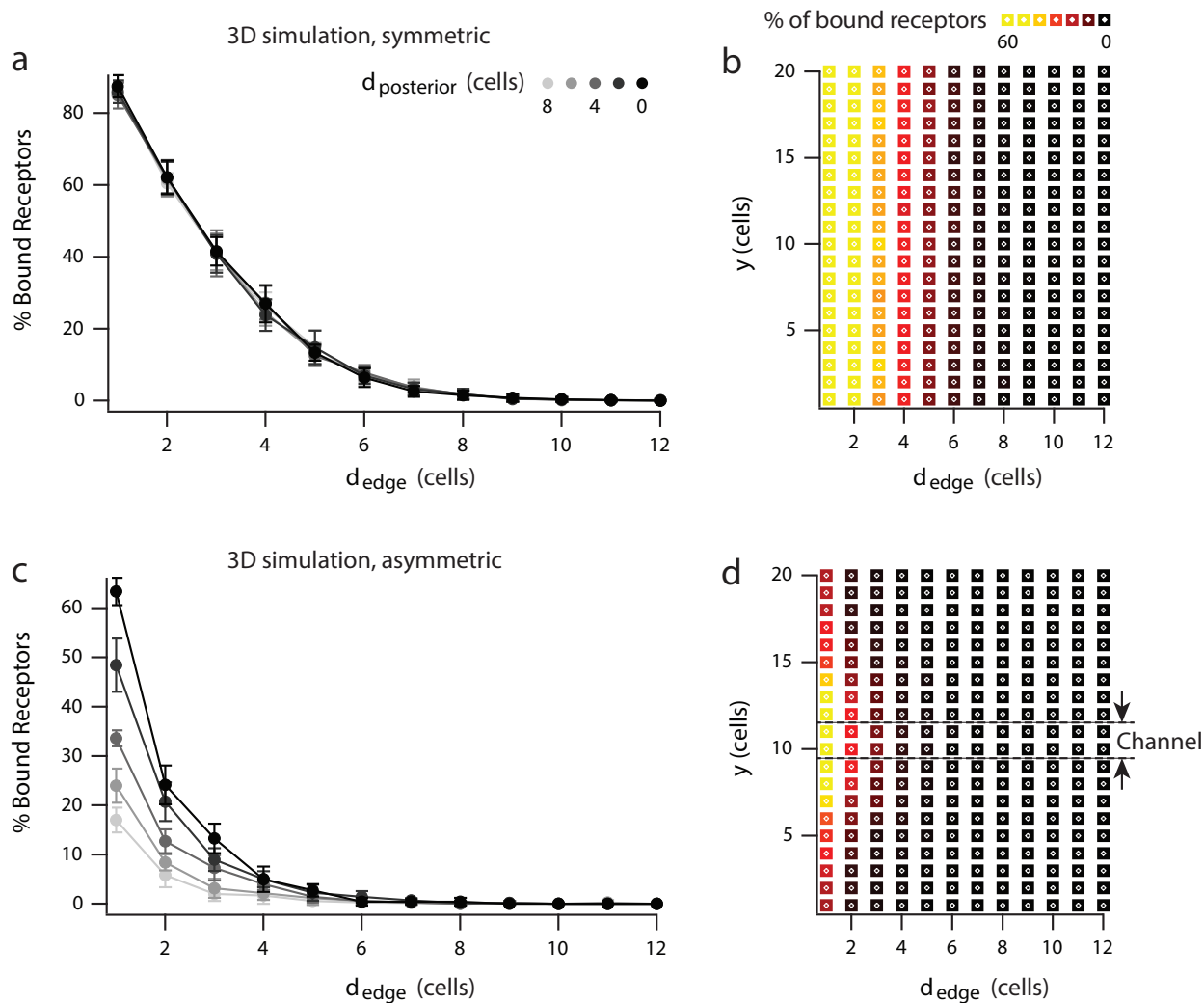
Figure 2.9: BMP signaling gradient forms from epiblast edge in symmetric and asymmetric 3D simulations. a, Percentage of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, at steady state (90 min) for radially symmetric three-dimensional (3D) simulations, in which channel between epiblast and ExE is present around the embryo. Here ligand/receptor ratio is 0.5. Darkness of curve corresponds to shorter distance of cells from posterior end of embryo ($d_{posterior}$). **b,** Percentage of ligand-bound receptors as a function of dedge and the distance along the anteroposterior axis (y) for individual epiblast cells in symmetric 3D simulations at steady state. **c, d,** Same as in **(a, b)** but for asymmetric 3D simulations in which channel between epiblast and ExE is restricted to posterior end, as seen in fluorescein injection experiments (Figure 2.2b). Dashed lines mark where channel is open. Error bars denote SEM

BMP receptors to the apical membrane should lead to ectopic BMP signaling in the epiblast (Figure 2.10), since apically localized receptors will be able to detect BMP4 ligands that are buffered in the lumen (Figure 2.1e).

a
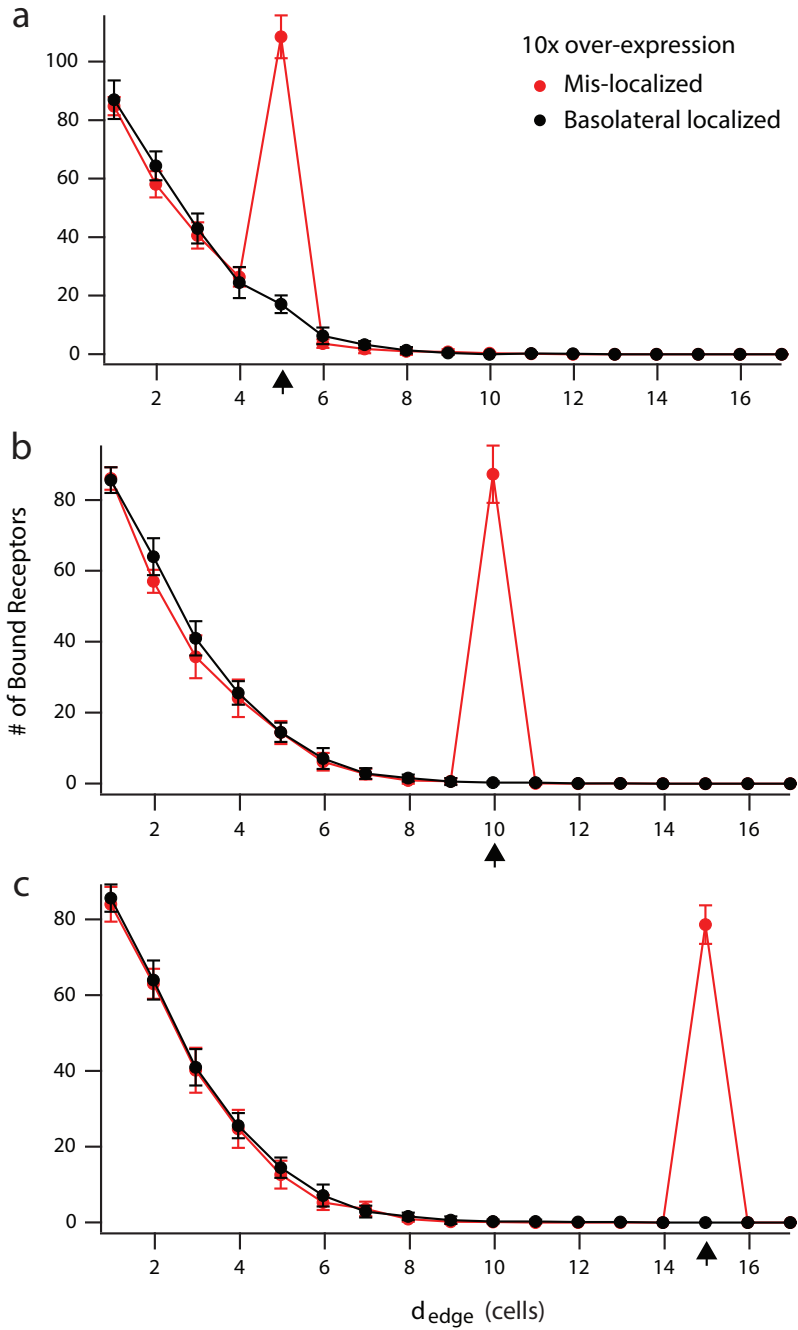
Figure 2.10: Effect of receptor overexpression and mis-localization on BMP signaling gradient in simulation. a, Number of ligand-bound receptors as a function of distance from epiblast edge, d_{edge}, in simulations at steady state (90 min). In these

Figure 2.10 (*continued*): simulations, a single cell (black arrow) at 5 cell widths from the epiblast edge overexpresses tenfold more wild type receptors that localize basolaterally (black) or mutant receptors that are mis-localize apically and basolaterally (red). For consistency with other figures, number of ligand-bound receptors is rescaled so that "100" corresponds to the default number of receptors in epiblast cells. **b,** Same as **(a)** but receptors are overexpressed in a cell at 10 cell widths from epiblast edge. **c,** Same as **(a)** but receptors are overexpressed in a cell at 15 cell widths from epiblast edge. Error bars denote SEM

2.2.2. *Basolateral localization of BMP receptors in hESCs and mouse epiblast*

We asked whether BMP receptors are indeed basolaterally localized in mammalian cells. We measured the localization of these receptors through surface immunostaining[22,24] as well as by imaging GFP- and epitope-tagged receptors (see Methods). The BMP co-receptors BMPR1A (Figure 2.11a, b, g-j) and BMPR2 (Figure 2.11k, l) are basolaterally localized in human embryonic stem cells (hESCs)[15]. We moreover found that the majority of TGF-β family receptors (including BMP receptors) in sequenced vertebrates contain a conserved LTA amino acid motif near their C-terminus (Figure 2.12). This motif has been shown to be necessary and sufficient for the basolateral localization of TGFBR2 in MDCK cells, and the mutation of this motif to an LTG sequence leads to the receptor's mis-localization to the apical membrane[22]. Consistently, we found that TGFBR2 and its co-receptor TGFBR1 are localized at the basolateral membrane of epithelial human hESCs (Figure 2.11c-f). Furthermore, the ACTIVIN/NODAL receptors ACVR1B and ACVR2B have also been found to be basolaterally localized in studies using human gastruloids[15], consistent with the fact that these receptors have LTA motifs (Figure 2.12). Thus, an evolutionarily conserved LTA motif is present in all of these receptors that are exclusively localized along the basolateral membrane in hESCs.
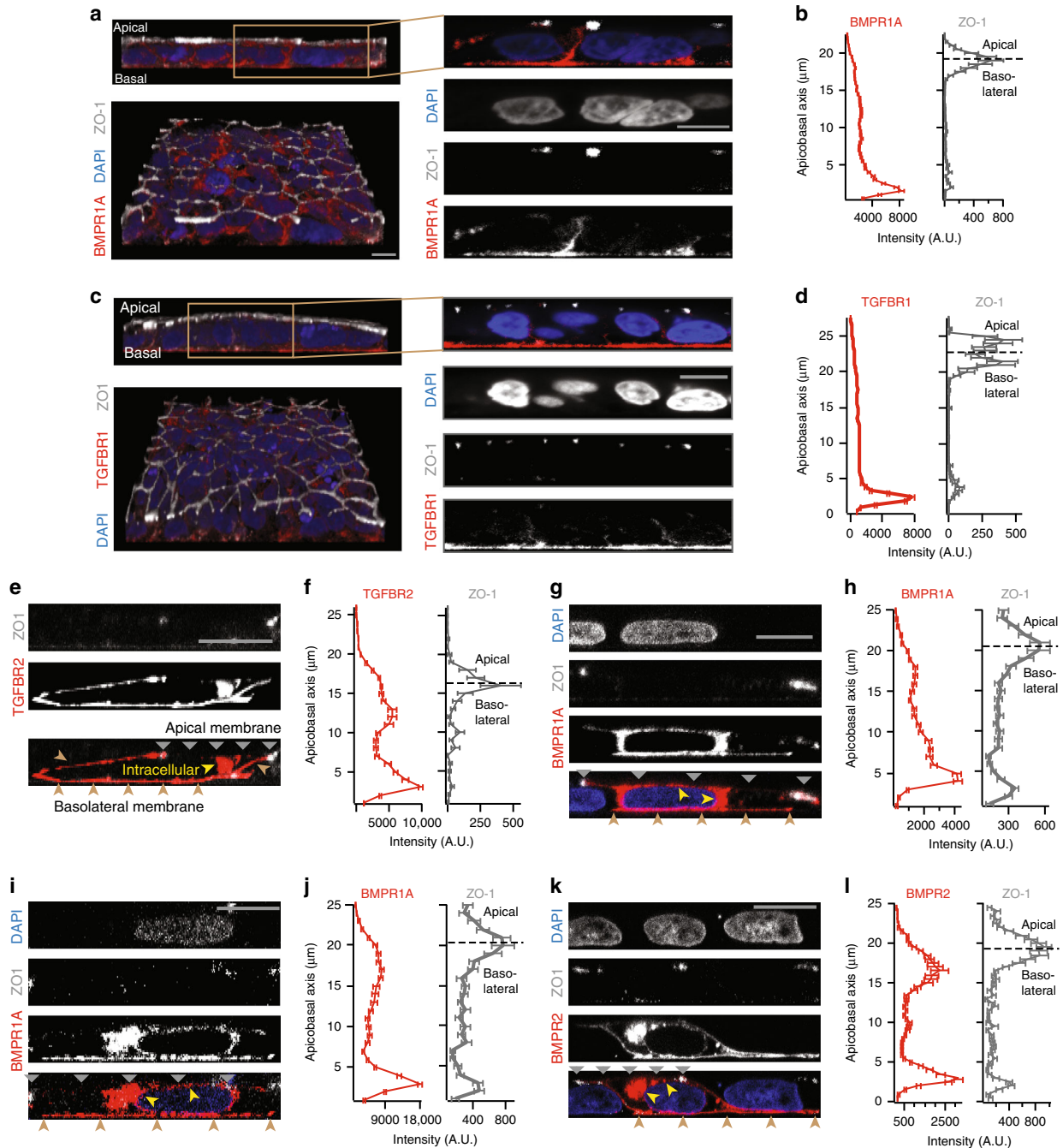
Figure 2.11: BMP and TGF-β receptors localize at basolateral membrane of hESCs in vitro. a, Left column: 3D confocal images of hESC colony stained for BMPR1A (red), tight junction marker ZO-1 (white), and DNA (blue), in lateral (top) and tilted view (bottom). Right column: Zoomed-in lateral images. Scale bar 10 μm. **b,** Plots of BMPR1A (left) and ZO-1 (right) staining intensity along apicobasal axis show BMPR1A localized beneath tight junctions (n = 38 cells from two experiments). **c,** Left column: 3D confocal image of hESC colony stained for DNA (blue), TGFBR1 (red), and ZO-1 (white) in lateral (top) and tilted view (bottom). Right column: Zoomed-in lateral section.

Figure 2.11 (*continued*): d, Plots of TGFBR1 (left) and ZO-1 levels (right) against apicobasal axis show that TGFBR1 is localized below tight junctions (n = 51 cells from two experiments). **e**, Confocal image of a hESC expressing TGFBR2-Clover (red), stained for ZO-1 (white). **f**, TGFBR2-Clover (left) and ZO-1 levels (right) against apicobasal axis (n = 4 cells from two experiments). **g, h,** Same as **(e, f)** but for BMPR1A-Clover (n = 2 cells from two experiments). **i, j,** Same as **(e, f)** but for BMPR1A-HA (n = 3 cells from two experiments). **k, l,** Same as **(e, f)** but for BMPR2-Clover (n = 3 cells from two experiments). Yellow arrows in **(e, g, i, k)** denote intracellular receptors in secretory pathway. Gray and brown arrows indicate apical and basolateral membranes, respectively. Scale bar 10 μm. Error bars denote SEM

Figure 2.12 (*following page*): LTA motif in TGF-β superfamily receptors. Protein sequence alignment of TGF-β superfamily receptors shows conservation of LTA motif in nine receptors and a LSA motif with a conservative substitution of serine for threonine in two receptors

Figure 2.12 (*continued*):

TGFBR2 / BMPR1A / BMPR2

	TGFBR2	BMPR1A	BMPR2	
C. elegans	EEMWDPEACAR TAGCAFARVWNH	EDSWHSIPHFR ALKLKKEMAEL	EEMWDPEACAR TAGCAFARVWNH	Roundworm
H. vulgaris	EECWDRDPEAR ANCIHERLLAI	SECWYNDSAARL ALRIKKTLLTL	EECWEMEADARL AWCVYERLVEL	Hydra
B. glabrata	KECWYSNAAARLTALRIRKTLACL	SECWNPSPAARLT LRIKKTLGKM	EDCWDTDAEARLTALCVQERLSDM	Snail
S. purpuratus	VDCWDEDPEARLTASCVEARFQEL	SECWNQNPAAR T LRVKKTLGKL	EDCWDHDAEARLTALCVEERIIEL	Sea urchin
D. melanogaster	EECWDHDAEARL SCVMERFAQL	QECWHPNPTVRLTALRVKKTLGRL	EDCWDHDADARLT LCAEERMQEM	Fruit fly
O. dioica	TECWTEEPTWRL ALKMKKDLSHE	EECWMESPEARLT MRIRMTAQKL	SDCWCDDPSARLT VRIKKTVARL	Tunicate
D. rerio	NECWDHDPEARLTAQCVAERFNEM	SECWAHNPASRLT LRVKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Zebrafish
X. laevis	TECWDQDPEARLTAQCVDERFNEM	AECWAQNPASRLTALRIKKTLAKM	DDCWDQDAEARLTAQCAEERMAEL	Frog
G. gallus	IECWDHDPEARLTAQCVAERFSEF	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Chick
C. lupus	TECWDHDPEARLTAQCVAERFSEL	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Dog
S. scrofa	AECWDHDPEARLTAQCVAERFSEL	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Pig
M. musculus	TECWDHDPEARLTAQCVAERFSEL	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Mouse
P. troglodytes	TECWDHDPEARLTAQCVAERFSEL	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Chimpanzee
H. sapiens	TECWDHDPEARLTAQCVAERFSEL	SECWAHNPASRLTALRIKKTLAKM	EDCWDQDAEARLTAQCAEERMAEL	Human

ACVR1 / ACVR1B / ACVR1C

	ACVR1	ACVR1B	ACVR1C	
C. elegans	EDSWHSIPHFR ALKLKKEMAEL	KTCWNGNPSAR T YICRKRMDER	KTCWNGNPSAR T YICRKRMDER	Roundworm
H. vulgaris	RECWREEPTARL ALRVKKSLSKL	SECWYSDSAARL ALRIKKTLLTL	SECWYSDSAARL ALRIKKTLLTL	Hydra
B. glabrata	KECWYSNAAARLTALRIRKTLACL	KECWYSNAAARLTALRIRKTLACL	KECWYSNAAARLTALRIRKTLACL	Snail
S. purpuratus	KECWCPTASAR T LRVKKTLMKI	KECWYQNGAARLTALRIKKTLAGL	KECWYQNGAARLTALRIKKTLAGL	Sea urchin
D. melanogaster	KECWHQNPDVRL ALRIKKTIHKL	KECWYPNPVARLTALRIKKTLASI	KECWYPNPVARLTALRIKKTLASI	Fruit fly
O. dioica	EECWMESPEARLT MRIRMTAQKL	EECWMESPEARLT MRIRMTAQKL	EECWMESPEARLT MRIRMTAQKL	Tunicate
D. rerio	KECWYQNPSARLTALRIKKTLDKI	RECWYANGAARLTALRIKKTLSQL	RECWHANPAARLTALRVKKTISQV	Zebrafish
X. laevis	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGKARLT LRIKKTISQL	Frog
G. gallus	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Chick
C. lupus	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Dog
S. scrofa	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Pig
M. musculus	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Mouse
P. troglodytes	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Chimpanzee
H. sapiens	KECWYQNPSARLTALRIKKTLTKI	RECWYANGAARLTALRIKKTLSQL	RECWYANGAARLTALRIKKTISQL	Human

ACVRL1 / BMPR1B / TGFBR1

	ACVRL1	BMPR1B	TGFBR1	
C. elegans	EDSWHSIPHFR ALKLKKEMAEL	EDSWHSIPHFR ALKLKKEMAEL	KTCWNGNPSAR T YICRKRMDER	Roundworm
H. vulgaris	RECWREEPTARL ALRVKKSLSKL	SECWYSDSAARL ALRIKKTLLTL	SECWYNDSAARL ALRIKKTLLTL	Hydra
B. glabrata	KECWYSNAAARLTALRIRKTLACL	SECWNPSPAARLT LRIKKTLGKM	KECWYSNAAARLTALRIKKTLACL	Snail
S. purpuratus	KECWCPTASAR T LRVKKTLMKI	SECWNQNPAAR T LRVKKTLGKL	KECWYQNGAARLTALRIKKTLAGL	Sea urchin
D. melanogaster	KECWHQNPDVRL ALRIKKTIHKL	QECWHPNPTVRLTALRVKKTLGRL	KECWYPNPVARLTALRIKKTLASI	Fruit fly
O. dioica	EECWMESPEARLT MRIRMTAQKL	EECWMESPEARLT MRIRMTAQKL	EECWMESPEARLT MRIRMTAQKL	Tunicate
D. rerio	KECWYQNPSARLTALRIKKTLDKI	TECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Zebrafish
X. laevis	RECWYQSPSARLTALRIKKALNKL	TECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Frog
G. gallus	KECWYQSPAARLTALRIKKSLQKL	MECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Chick
C. lupus	RECWYPNPSARLTALRIKKTLQKL	MECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Dog
S. scrofa	RECWYPNPSARLTALRIKKTLQKL	TECWAQNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Pig
M. musculus	RECWYPNPSARLTALRIKKTLQKL	TECWAQNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Mouse
P. troglodytes	RECWYPNPSARLTALRIKKTLQKI	TECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Chimpanzee
H. sapiens	RECWYPNPSARLTALRIKKTLQKI	TECWAHNPASRLTALRVKKTLAKM	RECWYANGAARLTALRIKKTLSQL	Human

ACVR2A / ACVR2B

	ACVR2A	ACVR2B	
C. elegans	EEMWDPEACAR AGCAFARVWNH	EEMWDPEACAR AGCAFARVWNH	Roundworm
H. vulgaris	EECWEMEADARLSA CVYERLVEQ	EECWEMEADARLSAWCVYERLVEQ	Hydra
B. glabrata	SECWNPSPAARL LRIKKTLGKM	KECWYSNAAARL ALRIRKTLACL	Snail
S. purpuratus	EECWDQDAEARLSAGCVEERIAQF	EECWDQDAEARLSAGCVEERIAQF	Sea urchin
D. melanogaster	EECWDHDAEARLS SCVMERFAQL	EECWDHDAEARLS SCVMERFAQL	Fruit fly
O. dioica	KDCWDCDAEARL AGCVLKRVQEI	KDCWDCDAEARL AGCVLKRVQEI	Tunicate
D. rerio	EECWDHEAEARLSAGCVEERVVQM	EECWDHDAEARLSAGCVQERISQI	Zebrafish
X. laevis	EECWDHDAEARLSAGCVEERIIQM	EECWDHDAEARLSAGCVEERISQI	Frog
G. gallus	EECWDHDAEARLSAGCVEERIIQM	EECWDHDAEARLSAGCVEERIAQI	Chick
C. lupus	EECWDHDAEARLSAGCVGERITQM	EECWDHDAEARLSAGCVEERVALI	Dog
S. scrofa	EECWDHDAEARLSAGCVGERITQM	EECWDHDAEARLSAGCVEERVSLI	Pig
M. musculus	EECWDHDAEARLSAGCVGERITQM	EECWDHDAEARLSAGCVEERVSLI	Mouse
P. troglodytes	EECWDHDAEARLSAGCVGERITQM	EECWDHDAEARLSAGCVEERVSLI	Chimpanzee
H. sapiens	EECWDHDAEARLSAGCVGERITQM	EECWDHDAEARLSAGCVEERVSLI	Human

We next explored whether BMP receptors are similarly localized in the basolateral membrane of mouse epiblast cells in vivo. To visualize receptors specifically on the cell membrane, we developed a protocol for surface immunostaining the mouse epiblast around the start of gastrulation (see Methods). After collection of E6.5 mouse embryos, we surgically removed the ExE from each embryo and exposed the epiblast to BMPR1A antibodies. We subsequently fixed and permeabilized the embryos and immunostained them for tight junction protein ZO-1 and epiblast marker OCT4. Light-sheet microscopy of the immunostained embryos shows that BMPR1A receptors in epiblast cells are localized on the basolateral membrane facing the underlying VE (Figure 2.13, Figure 2.14).
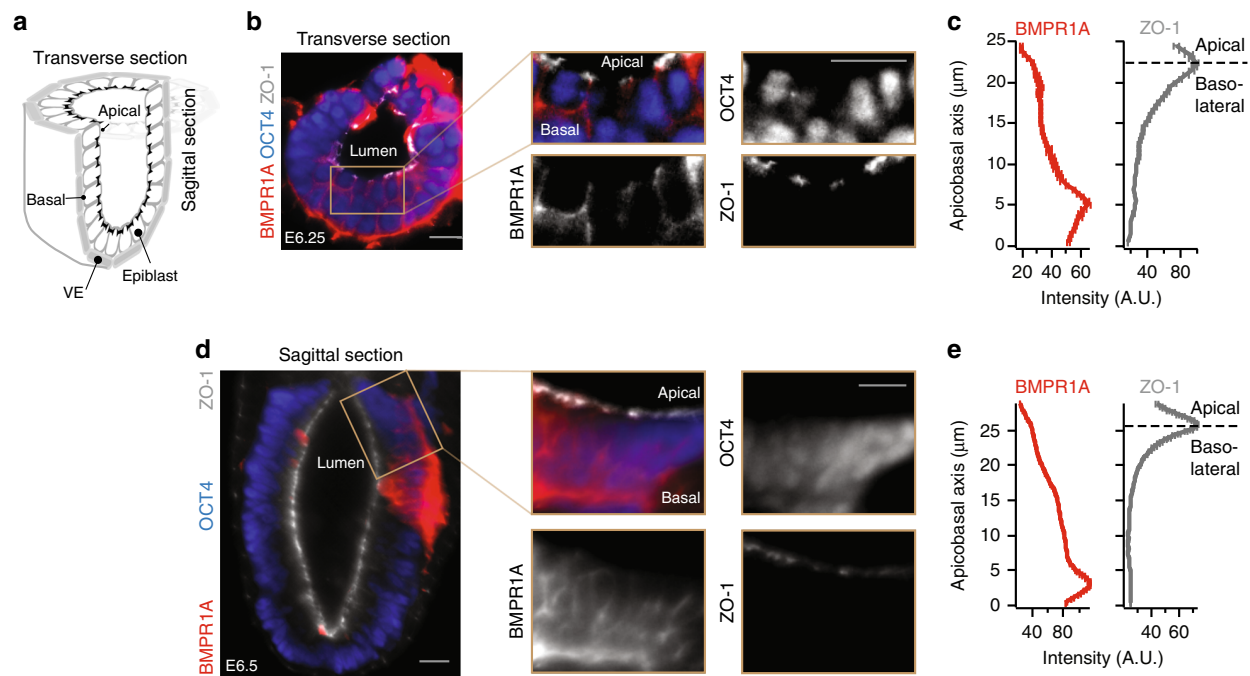
Figure 2.13: BMP receptors localize at basolateral membrane of mouse epiblast in vivo. a, Illustration of pre-gastrulation mouse embryo, showing transverse and sagittal sections. **b,** Transverse section of an E6.25 mouse embryo stained for epiblast marker OCT4, BMPR1A, and ZO-1. **c,** Plots of BMPR1A (left) and ZO-1 (right) staining intensity along apicobasal axis for transverse section from b show BMPR1A localized beneath tight junctions. **d, e,** Same as in (**b, c**) but for a sagittal section of an E6.5 mouse embryo. These images are representative of two sets of images in two embryos/experiments. In mouse data, scale bar 20 μm. Error bars denote SEM
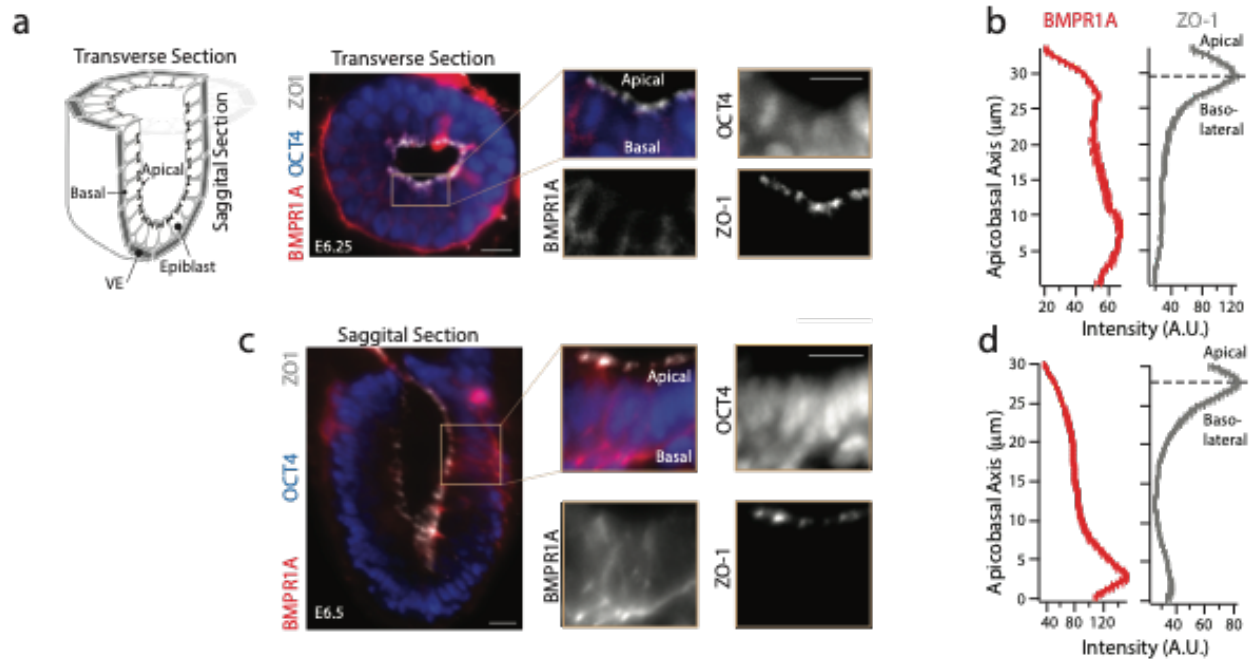
Figure 2.14: Additional evidence that BMP receptors localize at basolateral membrane of mouse epiblast in vivo. a, Transverse section of an E6.25 mouse embryo stained for OCT4 (blue), BMPR1A (red), and ZO-1 (white) shows receptors localized at basolateral membrane of epiblast. **b**, BMPR1A (left) and ZO-1 (right) levels along apicobasal axis. **c, d**, same as **(a, b)** but for a sagittal section of an E6.5 mouse embryo. These images are representative of two sets of images from two embryos. Scale bar 20 μm. Error bars denote SEM

2.2.3. *A robust BMP signaling gradient forms from the epiblast edge*

We asked whether the predicted formation of a robust BMP signaling gradient would occur in the epiblast. We first measured the distribution of phosphorylated SMAD1/5 (pSMAD1/5, the downstream effectors of the BMP signaling pathway) in epithelial hESC colonies exposed to BMP4 ligands. These epithelial colonies have impermeable tight junctions and a narrow, permeable basement membrane matrix underneath mimicking an interstitial space. The tissue geometry therefore is comparable to the geometry of the epiblast in mammalian embryos[25,30]. Akin to the simulation, we observed pSMAD1/5 gradients organized from the edges of epithelial hESC colonies exposed to spatially uniform concentrations of BMP4 (Figure 2.15a, b, Figure 2.16a, b). The formation of these BMP signaling gradients was robust to changes in ligand concentration: colonies exposed to BMP4 concentrations across a 1000-fold range displayed stable pSMAD1/5 gradients inward from colony edges, with the depth of the gradient varying only between 2 and 10 cell widths (Figure 2.15c, d, Figure 2.16e). This limited change in depth is consistent with our simulation results (Figure 2.17). The formation of these robust gradients was dependent on the segregation of apical and basolateral extracellular compartments by tight junctions. When tight junctions were disturbed by a brief treatment of passaging reagent ReLeSR or calcium chelator EGTA[31] before BMP4 induction, signal response occurred throughout hESC colonies (Figure 2.16c). Further, if hESCs were exposed to uniform BMP4 shortly after single-cell passaging, cells that had not yet formed tight junctions with adjacent cells showed significantly higher pSMAD1/5 activity than those surrounded by tight junctions (Figure 2.16d).
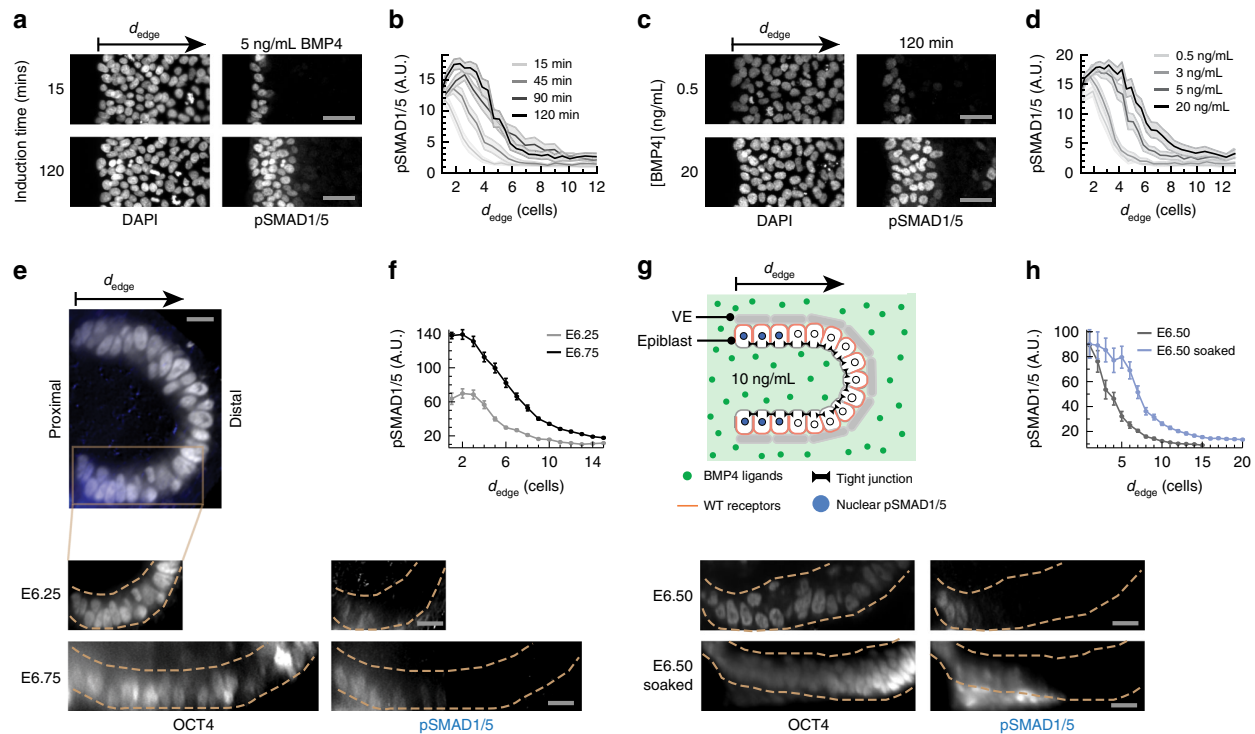
Figure 2.15: A robust BMP signaling gradient forms from the epiblast edge.
a, hESC colony stained for DNA and pSMAD1/5 after 15 or 120 min BMP4 induction. These images are representative of five sets of images from two experiments. **b,** pSMAD1/5 level of single hESCs as a function of their distance from the nearest colony edge, or d_{edge}, after 15–120 min of BMP4 induction. **c,** hESC colony stained for DNA and pSMAD1/5 after 120 min exposure to 0.5–20 ng/mL BMP4. These images are representative of five sets of images from two experiments. **d**, pSMAD1/5 level of single hESCs as a function of d_{edge} after 120 min exposure to 0.5–20 ng/mL BMP4. Shaded error bars denote 95% confidence intervals (paired t-test). In hESC data, scale bar 50 μm. **e,** E6.25–E6.75 mouse embryos stained for OCT4 (white) and pSMAD1/5 (blue). Dotted yellow lines indicate epiblast boundary. **f,** pSMAD1/5 level of mouse epiblast cells as a function of their distance from the posterior proximal edge of the epiblast, or d_{edge}, for E6.25 (two embryos per experiment) and E6.75 embryos (two embryos per experiment). **g,** Top: Illustration of BMP4 exposure experiment. ExE is surgically removed, and remaining epiblast-VE cup is soaked in media containing 10 ng/mL BMP4 for 30 min. Bottom: Intact E6.5 mouse embryo and BMP4-soaked E6.5 mouse embryo, both stained for OCT4 and pSMAD1/5. Dotted yellow lines indicate epiblast boundary. **h,** pSMAD1/5 intensity of epiblast cells as a function of d_{edge} for intact (three embryos per experiment) and BMP4-exposed E6.5 embryos (two embryos per experiment). In mouse data, error bars denote SEM and scale bar 20 μm

70
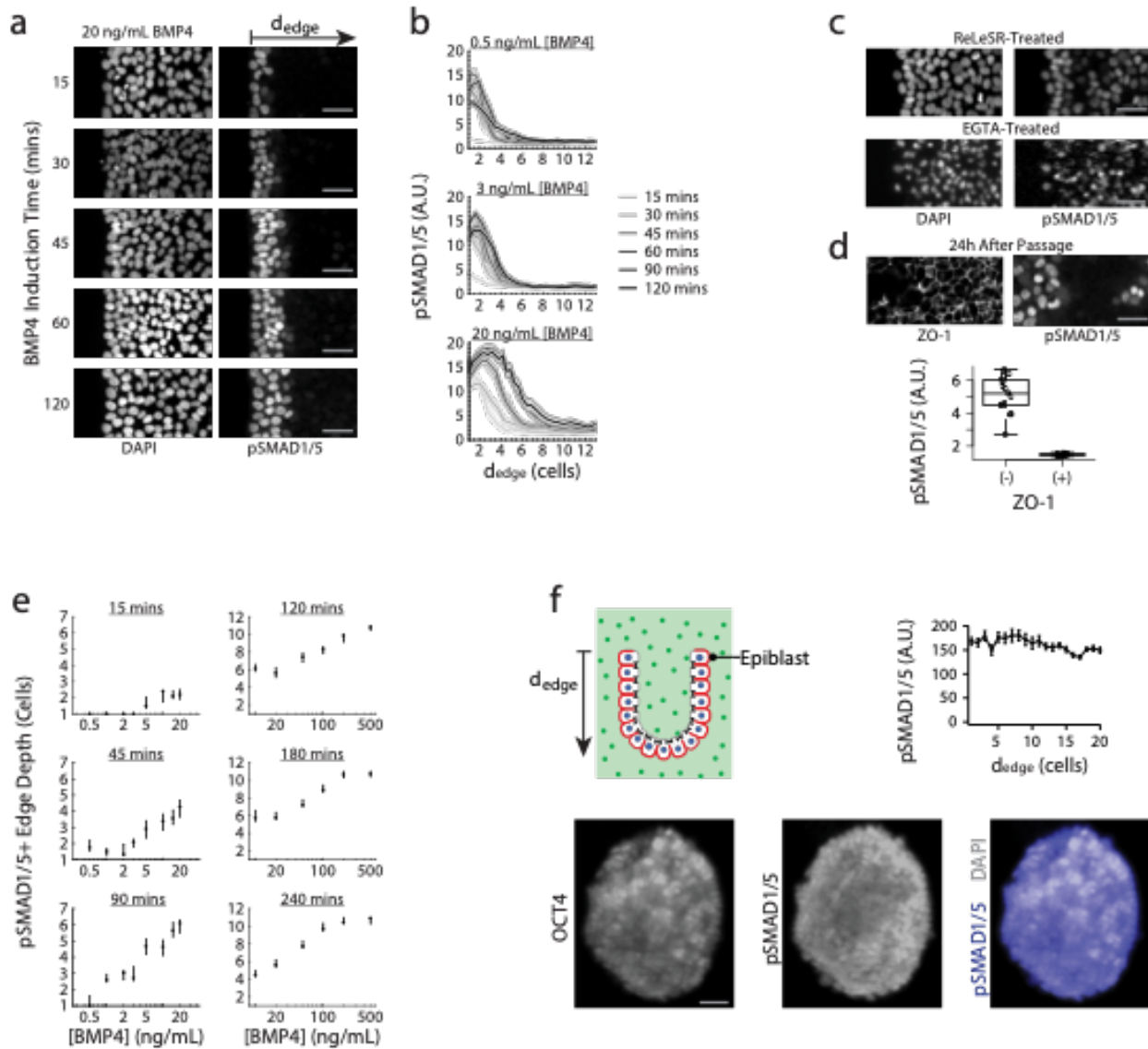
Figure 2.16: Dependence of BMP signaling gradient on epithelial integrity and embryo geometry. a, hESC colonies treated with 20 ng/mL BMP4 for 15-120 min and stained for DNA and pSMAD1/5. **b**, pSMAD1/5 levels of hESCs as a function of their distance from the nearest colony edge, dedge, after 15-120 min of BMP4 treatment. **c**, hESCs stained for DNA and pSMAD1/5 after treatment with ReLeSR (top) or EGTA (bottom) (see Methods) followed by 20 min BMP4 induction at 5 ng/mL. **d**, Top: hESCs exposed to 5 ng/mL BMP4 for 20 min and stained for ZO-1 and pSMAD1/5 at 24 hours after single cell passage. Bottom: Average pSMAD1/5 levels of hESCs surrounded by tight junctions (+) and hESCs not surrounded by tight junctions (-), as indicated by ZO-1 immunostain (20 cells of each condition, from 2 experiments). **e**, BMP signaling depth (as indicated by pSMAD1/5 levels of hESCs) as a function of BMP4 concentration after 15- 240 min of treatment. In hESC data, error bars denote 95% confidence intervals and scale bar 50 μm. **f**, Top: Illustration of mouse epiblast after removal of ExE and VE (see Methods), soaking in media containing 10 ng/mL BMP4. Bottom: E6.25 mouse embryo stained for OCT4 and

Figure 2.16 (*continued*): pSMAD1/5 shows BMP4 concentration is sufficient to induce signaling activity in all epiblast cells if ExE and VE are removed. Top right: Average pSMAD1/5 level of epiblast cells as function of d_{edge} in soaked embryos with ExE and VE removed. Here, d_{edge} represents cell's distance from proximal edge of epiblast (10 cells each condition, from 1 experiment). In mouse data, error bars denote SEM and scale bar 20 μm
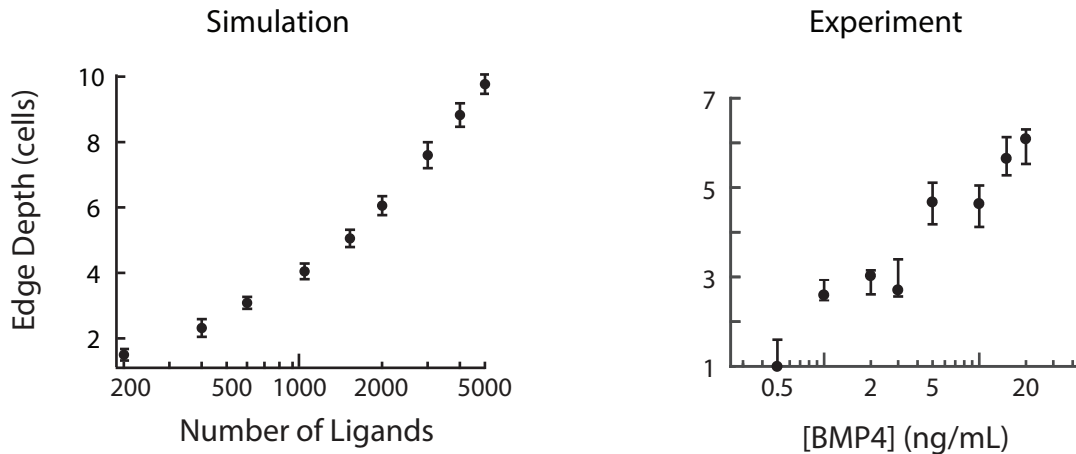
Figure 2.17: BMP signaling depth in simulation and in hESCs. BMP signaling depth from epithelial edge as a function of BMP4 concentration in mouse embryo simulation (left) and in hESC colonies (right) after 90 min of exposure. The BMP signaling depth is defined by 50% of maximum BMP4 signaling. Simulation error bars denote SEM, hESC data error bars denote 95% confidence intervals

We observed similar BMP signaling gradients in early mouse embryos as well. In harvested mouse embryos stained for pSMAD1/5, we observed a gradient of pSMAD1/5 activity inward from the proximal edges of the epiblast at the pre-streak (~E6.25) through the early streak (~E6.75) stages of development (Figure 2.15e, f, Figure 2.18). To test whether this signaling gradient is maintained even in uniformly high concentrations of BMP4, we surgically removed the ExE from E6.5 mouse embryos, exposing the remaining epiblast-VE cup. We then soaked the cup in media containing 10 ng/mL BMP4 for 30 min before fixing and immunostaining for pSMAD1/5 (Figure 2.15g). In these BMP-soaked embryos, the pSMAD1/5 gradient reached only a few cell
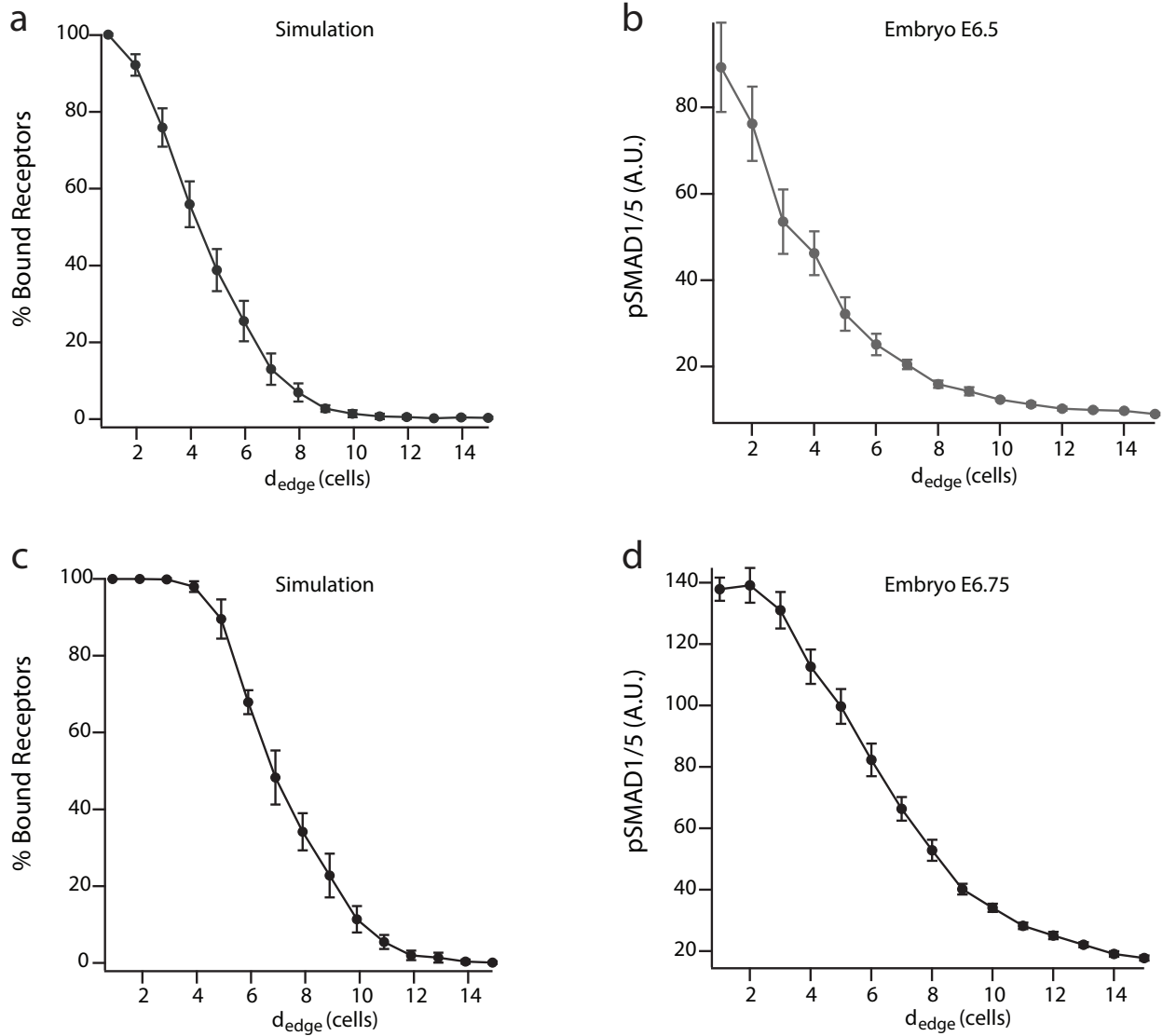
Figure 2.18: Simulated BMP signaling gradient resembles BMP signaling gradient in mouse embryos. a, Percentage of ligand-bound receptors at steady state (90 min) as a function of distance from epiblast edge, d_{edge}, in simulation with apically secreted ligands, as in Fig. 1f. Here ligand-receptor ratio is 1.0. **b,** pSMAD1/5 intensity of epiblast cells as a function of d_{edge} in E6.5 mouse embryo, as in Fig. 3h. Here, d_{edge} represents distance of cell from posterior proximal edge of the epiblast. **c,** Same as **(a)** but for ligand-receptor ratio of 2.0. **d,** Same as **(b)** but for E6.75 mouse embryo, as in Fig. 3f. Error bars denote SEM

widths further from the proximal epiblast edge as compared to wild-type embryos

(Figure 2.15g, h). This restriction of BMP signaling was maintained despite the fact that

the BMP4 concentration was sufficiently high to induce pSMAD1/5 activity uniformly

throughout the epiblast if its basolateral surface was exposed to ligands (Figure 2.16f). In summary, our results in vitro and in vivo show that gradients of BMP signaling activity robustly form inward from the edges of epithelial tissues with basolateral receptor localization.

2.2.4. Mis-localization of receptors leads to ectopic BMP signaling

Having verified the first two predictions of the model, we next tested whether the mis-localization of BMP receptors to the apical membrane results in ectopic BMP signaling. To do so, we designed a plasmid expressing epitope-tagged mutant copies of both *BMPR1A* and *BMPR2*, in which their LTA motifs were mutated into an LTG sequence (see Methods). Unlike the wild-type receptors, these mutant receptors localized at both the apical and basolateral membranes of hESCs transfected with these plasmids (Figure 2.19a, b, Figure 2.20a). The transfected hESCs, in the absence of exogenous BMP4 ligands, did not show any significant BMP signaling activity (Figure 2.19c, Figure 2.20b). After BMP4 exposure, however, cells expressing the mis-localizing receptors had significantly higher levels of nuclear pSMAD1/5 than their neighboring non-transfected cells (Figure 2.19b, c, Figure 2.20a). The pSMAD1/5 levels of these transfected cells were comparable to that of non-transfected cells at colony edges (Figure 2.19c). In contrast, overexpression of wild-type receptors did not lead to a comparable increase in pSMAD1/5 levels of transfected cells in vitro, as predicted by our simulation (Figure 2.10, Figure 2.21). Thus, while basolaterally localized wild-type BMP receptors in the interior of hESC colonies were insulated from apical ligands by tight junctions, cells with mis-localized BMP receptors could sense and respond to these ligands.
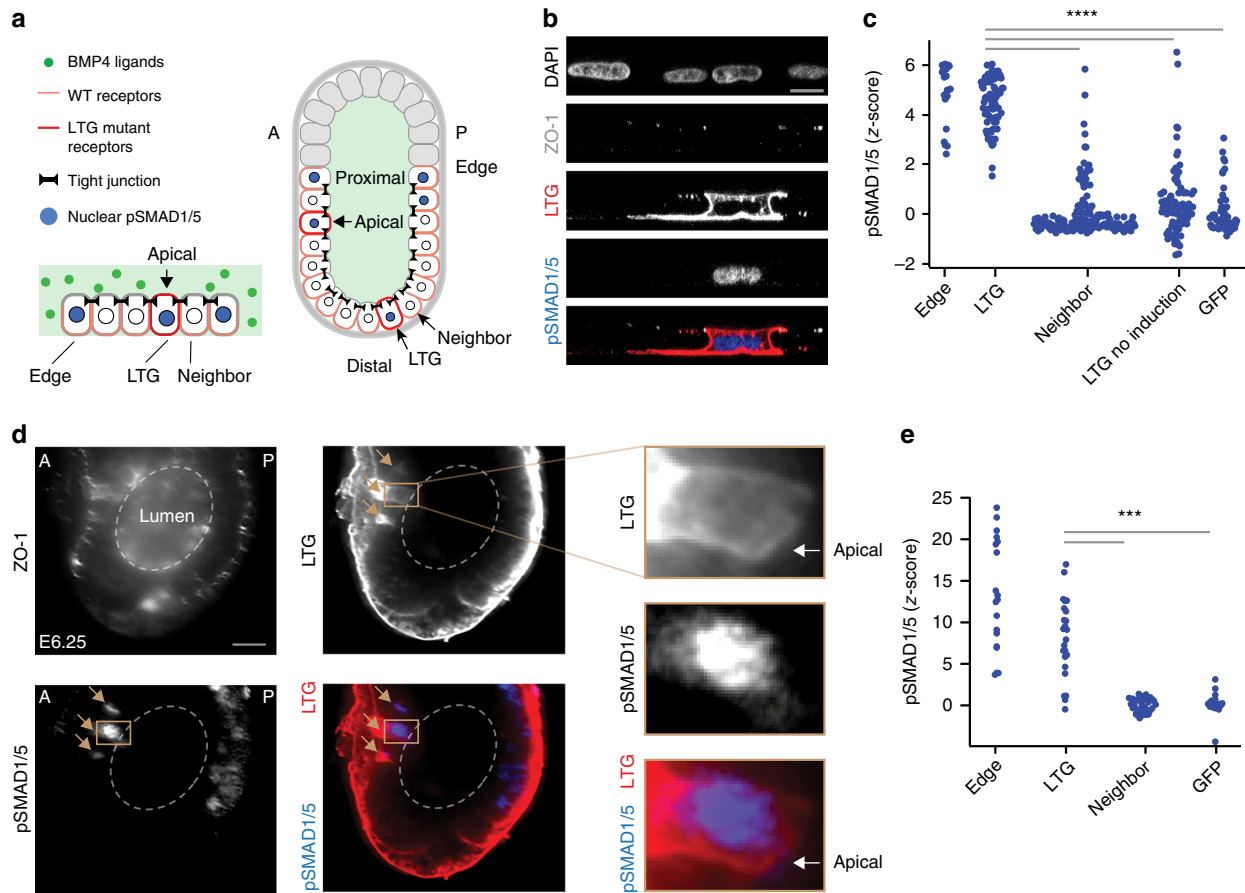
Figure 2.19: Mis-localization of receptors leads to ectopic BMP signaling. a, Left: Illustration of a hESC colony containing a single cell with mis-localized LTG mutant BMP receptors, exposed to BMP4 ligands. Right: Illustration of mouse embryo with two cells expressing mis-localized LTG mutant BMP receptors, leading to ectopic pSMAD1/5 activity. A and P denote anterior and posterior, respectively. **b,** Confocal image of hESC colony transfected with mutant receptor plasmid (BMPR1A^{A514G}-Clover-IRES-BMPR2^{A494}), immunostained for pSMAD1/5 after a 30 min BMP4 induction. From top: DNA, ZO-1, Clover (LTG mutant receptors), pSMAD1/5, and color-combined image. Scale bar 10 μm. **c,** pSMAD1/5 intensities of hESCs after 30 min BMP4 induction: cells at edge of colony (Edge, n = 23 from two experiments); non-edge cells expressing mutant BMPR1A^{A514G} and BMPR2^{A494G} receptors (LTG, n = 73 from two experiments); non-transfected neighbors of transfected cells (Neighbor, n = 166 from two experiments); cells expressing mutant receptors but without BMP4 induction (LTG no induction, n = 96 from two experiments); and cells transfected with GFP plasmid (n = 51 from two experiments). Z-score denotes number of standard deviations beyond background mean (of neighboring non-transfected cells), ****p < 0.0001 (paired t-test). **d,** E6.25 mouse embryo transfected with mutant receptor plasmid (BMPR1A^{A514G}-Clover-IRES-BMPR2^{A494G}), immunostained for ZO-1, Clover (LTG mutant receptors), and pSMAD1/5. Images show localization of mutant receptors at both apical and basolateral membrane and pSMAD1/5 activity in a transfected cell. Brown arrows indicate transfected cells. Scale bar 20 μm

75

Figure 2.19 (*continued*): e, pSMAD1/5 intensity epiblast cells: cells at edge of epiblast (Edge, n = 42 from 13 embryos); non-edge cells transfected with mutant receptor plasmid (LTG, n = 27 from 13 embryos); their neighboring non-transfected cells (Neighbor, n = 52 from 13 embryos); and cells transfected with a GFP plasmid (GFP, n = 20 from 3 embryos). Edge and LTG cells with z-score greater than 25 are not shown (n = 26), [***]$p < 0.001$ (paired t-test)

Figure 2.20 (*following page*): Receptor mis-localization results in ectopic BMP signaling in vitro and in vivo. a, Apical (top row), middle (middle row), and basal (bottom row) confocal image stacks of hESC colony transfected with mutant receptor plasmid (BMPR1A[A514G]-Clover-IRES- BMPR2[A494G]) and exposed to 10 ng/mL BMP4 for 30 min. Two cells in center of colony expressed mis-localized BMP receptors. (Left to Right): DNA, Clover (LTG mutant receptors; red), pSMAD1/5 (blue), and color-combined channels. These images show that receptor mis- localization leads to ectopic pSMAD1/5 activation in the center of hESC colonies. Scale bar 10 μm. These images are representative of 8 images from 2 experiments. **b**, Same as **(a)** but without BMP4 induction. These images show that the expression of mutant receptors in the absence of exogenous BMP4 ligands does not result in ectopic pSMAD1/5 activation. Scale bar 40 μm. These images are representative of 4 images from 2 experiments. **c**, Custom-made device for microinjection (left) and electroporation (right) of mouse embryos (see Methods). **d**, Left: Illustration of mouse embryo with a single cell expressing mis-localized LTG mutant BMP receptors, leading to ectopic pSMAD1/5 activity. A and P denote anterior and posterior, respectively. Right: Two E6.5 mouse embryos transfected with mutant receptor plasmid and immunostained for OCT4, Clover (LTG mutant receptors, red), and pSMAD1/5 (blue). Insets show transfected cells at increased magnification and contrast. Cells expressing mutant receptors have ectopic pSMAD1/5 activity in both the anterior (top row) and distal (bottom row) epiblast. These images are representative of 13 images from 13 experiments. Scale bar 20 μm

Figure 2.20 (*continued*):
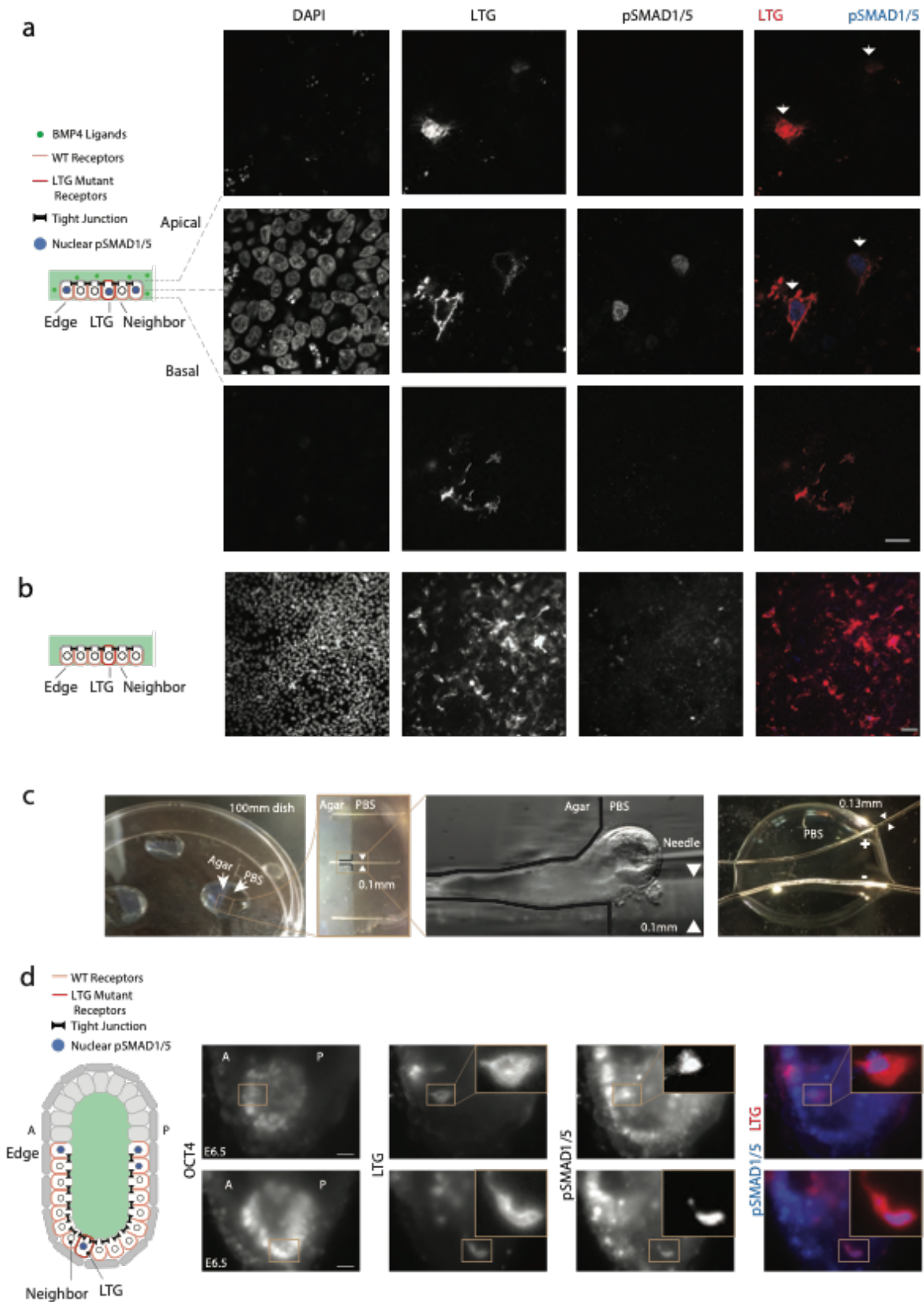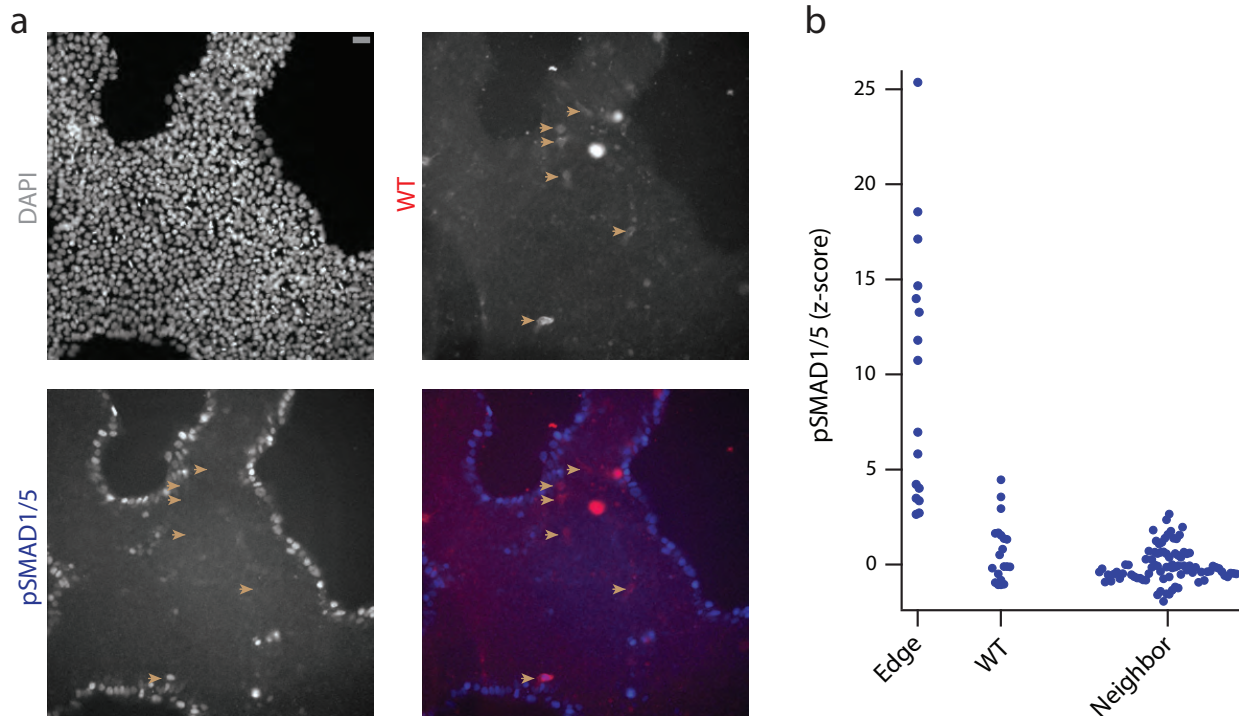
Figure 2.21: Effect of wild type receptor overexpression on BMP signaling. a, hESC colony transfected with wild type receptor plasmid (BMPR1A-Clover-IRES-BMPR2), exposed to 10 ng/ml BMP4 for 30 min, and stained for DNA and pSMAD1/5 (blue). Brown arrows indicate transfected cells expressing wild type receptors (WT, red). **b,** pSMAD1/5 intensities of hESCs after 30 min BMP4 induction: cells at colony edge (Edge, n=21 in 2 experiments); non- edge cells overexpressing wild type BMPR1A and BMPR2 receptors (WT, n=21 in 2 experiments); and non-transfected neighbors of transfected cells (Neighbor, n=86 in 2 experiments). Z-score denotes number of standard deviations beyond background mean (of neighboring non-transfected cells). Edge (n=6) and WT cells (n=1) with z-score greater than 25 are not shown. Scale bar 20 μm

To test the effect of receptor mis-localization in vivo, we developed a method to deliver our mutant BMP receptor plasmid to anterior and distal regions of the epiblast that do not normally show BMP signaling activity, while leaving the rest of the mouse embryo unperturbed (see Methods, Figure 2.20c). Consistent with our results in hESCs, mutant BMP receptors were localized at both the apical and basolateral membranes of transfected epiblast cells in vivo (Figure 2.19d). This mis-localization led to ectopic BMP signaling in cells in the anterior and distal regions of the epiblast, where neighboring

non-transfected cells showed no signal response (Figure 2.19d, e, Figure 2.20d). pSMAD1/5 levels in electroporated cells resembled that of cells at the epiblast edge (Figure 2.19e). These data support our simulation results, in which BMP4 ligands can be present throughout the pre-amniotic cavity while basolateral BMP receptors in the epiblast are insulated from these signals.

2.2.5. *Distance from tissue edge governs patterning of epithelial tissues*

In summary, our results in silico, in vitro, and in vivo demonstrate how basolateral receptor localization and embryo geometry together, through an entropic buffering mechanism, result in the formation of robust BMP signaling gradients at tissue edges. Consistently, our mathematical model argues that an epithelial cell's distance from the tissue edge (d_{edge}) predicts the cell's signaling response better than its distance from the source of the signal (d_{source}, Figure 2.22a, b). Here, the predictive power is quantified by the proficiency (the mutual information shared between the coordinate of a cell and its pSMAD1/5 levels, given as a percentage out of the total information entropy of pSMAD1/5 levels[32]). While studies in multiple model organisms have shown that d_{source} is a critical determinant of patterning[1,8,9,13,33], our results argue that d_{edge} could also be an important developmental coordinate for the patterning of epithelial tissues. To test how epithelial cell fate decisions are organized along d_{source} and d_{edge}, we developed microfluidic devices capable of producing precise morphogen gradients (Figure 2.22c, Figure 2.23a-c). The environment within the device mimics that of a morphogen gradient produced by a signal source at the left end of the device. We exposed hESC colonies to a BMP4 gradient from 10 to 0 ng/mL for 30 min. Consistent with our previous results (Figure 2.15, Figure 2.16a, b, Figure 2.24), signaling activity
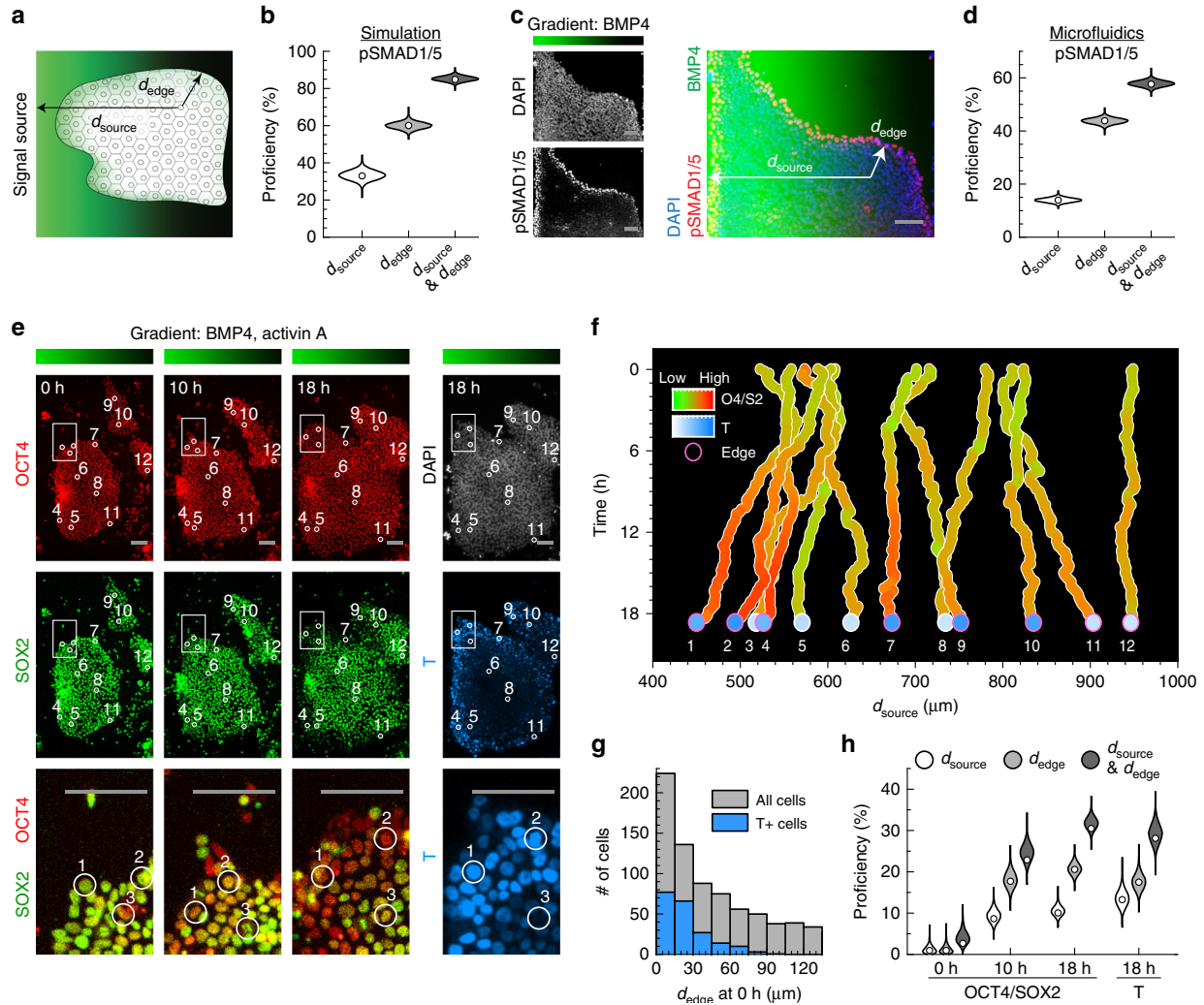
Figure 2.22: Distance from tissue edge and distance from signal source govern patterning of epithelial tissue. a, Illustration of epithelial tissue within a morphogen gradient emanating from a source to the left. The coordinates d_{edge} and d_{source} denote a cell's distance from the nearest tissue edge and from the signal source, respectively. **b,** Proficiency of d_{source}, d_{edge}, or both coordinates to predict pSMAD1/5 level of epiblast cells in simulation (Figure 2.1). **c,** Epithelial hESC colony exposed to BMP4 gradient in microfluidic device for 30 min and stained for DNA and pSMAD1/5. BMP4 gradient ranges from 10 ng/mL (green) to 0 ng/mL (black). Coordinates d_{source} and d_{edge} are depicted for a single cell as in (**a**). These images are representative of five sets of images from two experiments. **d,** Proficiency of d_{source}, d_{edge}, or both coordinates to predict pSMAD1/5 levels in hESCs exposed to microfluidic BMP4 gradient (n = 13,828 cells). **e,** OCT4-RFP (red) SOX2-YFP (green) double reporter hESCs within microfluidic gradient after 0, 10, and 18 h of differentiation. DNA, white; BRACHYURY/T, teal. Inset highlights differentiation at colony edges. Position of 12 sample cells labeled by circles. Bar above shows BMP4 and ACTIVIN A gradient within microfluidic device, ranging from 10 ng/mL (green) to 0 ng/mL (black) of each. These images are representative of two sets of images from two experiments. **f,** d_{source} of 12 tracked cells from (**e**) throughout time-lapse, colored by

Figure 2.22 (*continued*): OCT4/SOX2 ratios (red/green) and BRACHYURY/T level (teal) at end of time-lapse. Pink circles mark cells with d_{edge} less than 52 μm at 10 h of differentiation. **g,** Distribution of d_{edge} at start of differentiation, with teal marking cells that were BRACHYURY/T+ after 18 h of differentiation. **h,** Proficiencies of d_{source}, d_{edge}, or both coordinates to predict OCT4/SOX2 ratios and BRACHYURY/T levels (n = 1275 cells). Violin plots denote Gaussian kernel density estimates. Scale bar 100 μm

Figure 2.23 (*following page*): Distance from tissue edge and distance from signal source govern signaling and patterning of epithelial tissue. a, Diagram of microfluidic device. (Inset) Visualization of fluorescein gradient at top and middle of cell chamber in microfluidic device. **b**, Visualization of microfluidic gradient after 0, 8, and 15 hours of flow. **c**, Level of fluorescein gradient over 15 hours (left) and its maximum range of fluctuation (right) across cell chamber as function of distance from left side of the chamber (d_{source}). Fluctuation range is given as percentage of total gradient range. **d**, (Left) pSMAD1/5 levels of hESCs exposed to BMP4 gradient for 30 min as a function of d_{source} and d_{edge}. (Right) pSMAD1/5 levels of edge and interior hESCs as a function of d_{source}. Edge cells are defined as cells within 2 cell widths of a colony edge, while interior cells are those further than 6 cell widths from the nearest colony edge. **e**, OCT4/SOX2 ratios of 12 tracked cells from Figure 2.22e over time-lapse experiment. **f**, BRACHYURY/T levels of 12 tracked cells from Figure 2.22e at end of time-lapse experiment. **g**, OCT4/SOX2 ratios at 0, 10, and 18 hours and BRACHYURY/T levels at 18 hours during time-lapse differentiation of edge and interior hESCs as a function of d_{source} (n=1,275 cells). **h**, OCT4-RFP SOX2-YFP hESCs after 48 hours of differentiation in BMP4 and ACTIVIN A gradient (above) as in Figure 2.22e, stained for DNA and BRACHYURY/T. **i**, BRACHYURY/T levels of edge and interior hESCs after 48-hour gradient differentiation as a function of dsource. Here, edge cells are defined as cells within 126 μm of colony edge, while interior cells are those further than 283 μm from nearest colony edge. **j**, ReLeSR-treated hESCs exposed to uniform BMP4 and ACTIVIN A (10 ng/mL of each) for 18 hours and stained for DNA and BRACHYURY/T. hESCs were treated with ReLeSR for 1 min after 0, 4, and 8 hours of differentiation. The images are representative of 4 images from 2 experiments. Error bars denote 95% confidence intervals and scale bar 100 μm

Figure 2.23 (*continued*):
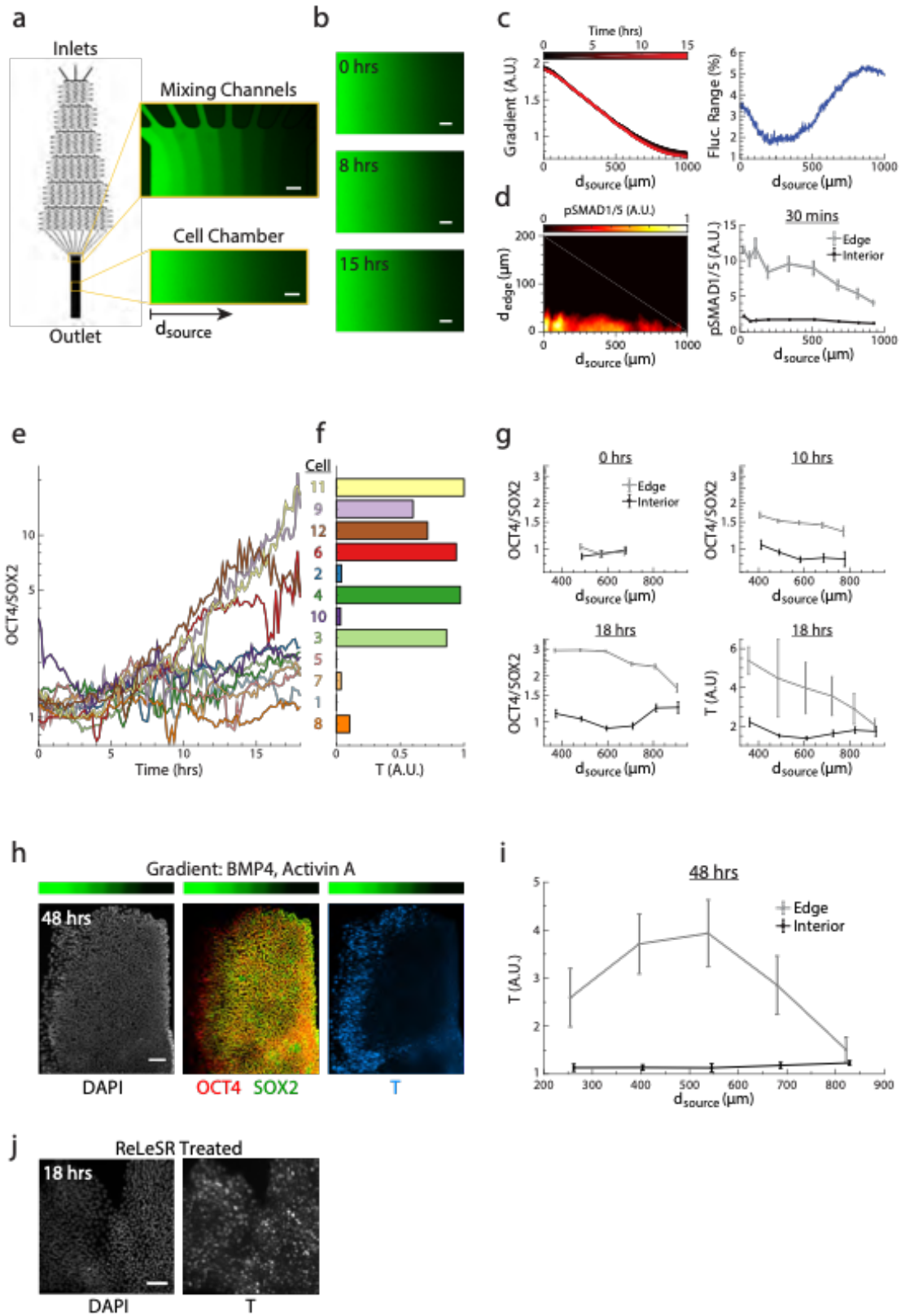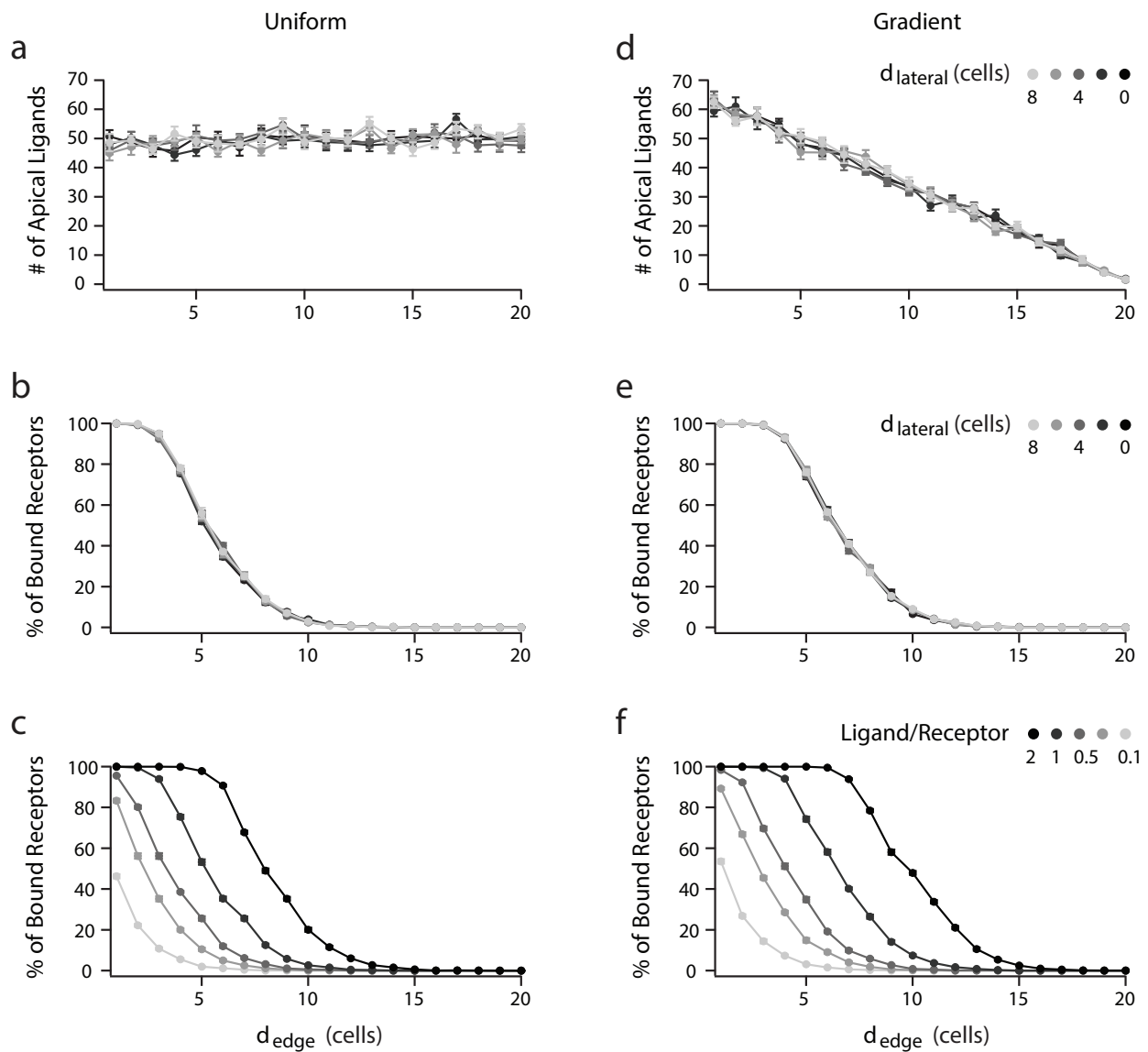
Figure 2.24: BMP signaling gradient in simulations based on hESC colony geometry in vitro. a, Number of apical ligands as a function of distance from nearest colony edge, d_{edge}, in simulations of epithelial hESC colonies exposed to uniform concentration of BMP4 ligands in culture media. Colors of points and curves correspond to position along lateral axis of colony ($d_{lateral}$) perpendicular to d_{edge} axis. **b,** Percentage of ligand-bound receptors as a function of d_{edge}, in simulations from **(a)**. **c,** Percentage of ligand-bound receptors as a function of d_{edge} in simulations from **(a)**, where darkness of points and curves corresponds to higher ligand-receptor ratio. Here ligand-receptor ratio of 0.1 correspond to approximately BMP4 concentration of 0.16 ng/ml. **d-f,** Same as **(a-c)** but for simulations of hESC colonies exposed to gradient of BMP4 ligands along d_{edge} axis. Error bars denote SEM

depended strongly on d_{edge} (Figure 2.22c, d, Figure 2.23d). In fact, a cell's d_{edge} had a significantly higher proficiency than d_{source} in predicting its signaling response to the BMP gradient (Figure 2.22d, Figure 2.25).

To determine how d_{source} and d_{edge} correlate with cell fate decision dynamics, we built a dual-color OCT4-RFP SOX2-YFP hESC line, in which OCT4 and SOX2 are tagged with fluorescent proteins at their endogenous loci (Figure 2.26). OCT4 and SOX2 are co-expressed in the pluripotent state (OCT4+ and SOX2+) but are differentially regulated during mesodermal differentiation (OCT4+ and SOX2−); this differential regulation is essential for the cell's germ layer fate choice[34]. We then cultured epithelial colonies of this hESC line in the microfluidic device, exposing them to gradients of BMP4 and NODAL-analog ACTIVIN A (from 10 to 0 ng/mL, of each). We measured the OCT4 and SOX2 levels of individual cells in these gradients as well as their d_{source} and d_{edge} for 18 h using time-lapse microscopy (Figure 2.22e, f, Figure 2.23e). At the end of the time-lapse, we immunostained the cells in situ for mesodermal progenitor marker BRACHYURY/T to determine their fate choice (Figure 2.22e, f, Figure 2.23f).

We found that cells with comparable d_{source} but different d_{edge} often adopted distinct cell fates (Figure 2.22e, f, Figure 2.23g). In many cases, cells near colony edges had higher BRACHURY/T and lower SOX2 levels than cells in colony interiors that had a smaller d_{source} throughout the time-lapse. Furthermore, 95% of cells that expressed BRACHYURY/T at the end of time-lapse were initially located near colony edges ($d_{edge} < 66.5$ μm or approximately 5.1 cell widths, Figure 2.22g), where signaling is most active at the start of differentiation (Figure 2.22c). After 48 h of exposure to BMP4 and ACTIVIN A gradients, hESCs with high BRACHYURY/T and low SOX2 levels continued to be located predominantly at the colony edges, while cells in colony interiors remained undifferentiated (Figure 2.23h, i). Like pSMAD1/5, the dependence
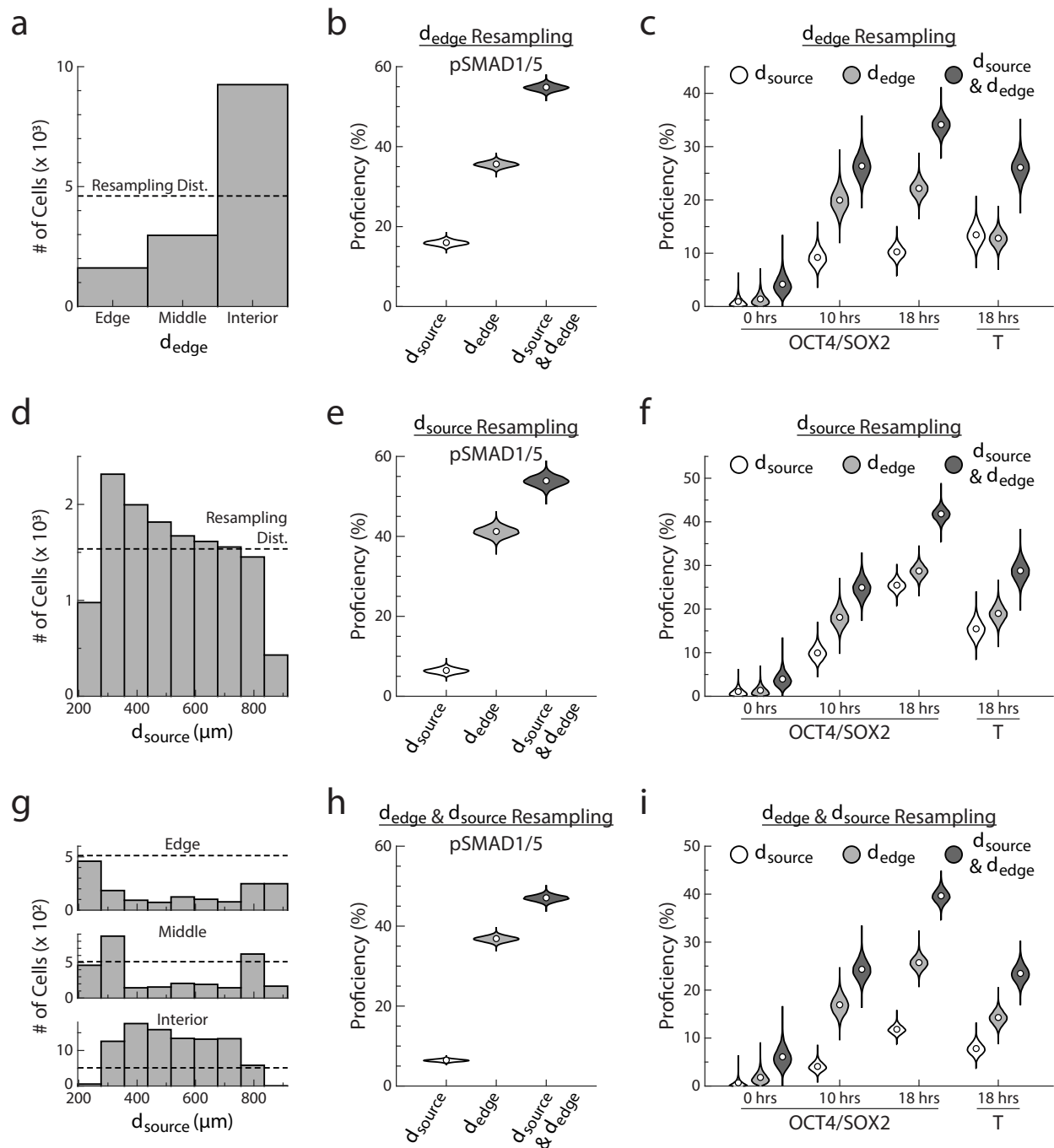
Figure 2.25: Proficiencies do not change significantly after resampling.
a, Distribution of edge cells (d_{edge} < 2 cell widths), middle cells (2 cell widths < d_{edge} < 6 cell widths), and interior cells (d_{edge} > 6 cell widths) in BMP4 gradient hESC microfluidic experiment (Figure 2.22c, d). Dashed line indicates distribution after resampling a uniform number of edge, middle, and interior cells (n = 13,828 cells). **b,** Distribution of cells along d_{source} in BMP4 gradient hESC microfluidic experiment. Dashed line indicates distribution after resampling cells uniformly along d_{source}. **c,** From top to bottom, distribution of edge, middle, and interior cells along d_{source} in

Figure 2.25 (*continued*): BMP4 gradient hESC microfluidic experiment. Dashed lines indicate distribution after resampling cells uniformly along both d_{edge} and d_{source}. **d,** As in Figure 2.22d, proficiency of d_{source}, d_{edge}, or both coordinates to predict pSMAD1/5 levels, calculated after resampling uniformly from edge, middle, and interior cells. **e,** Same as in **(d)** but calculated after resampling cells uniformly along d_{source}. **f,** Same as in **(d)** but calculated after resampling cells uniformly along both d_{edge} and d_{source}. **g, h, i,** Same as in **(d-f)** but for OCT4/SOX2 and T proficiencies in microfluidic time-lapse experiment from Figure 2.22e-h (n = 1,275 cells). Violin plots denote Gaussian kernel density estimates
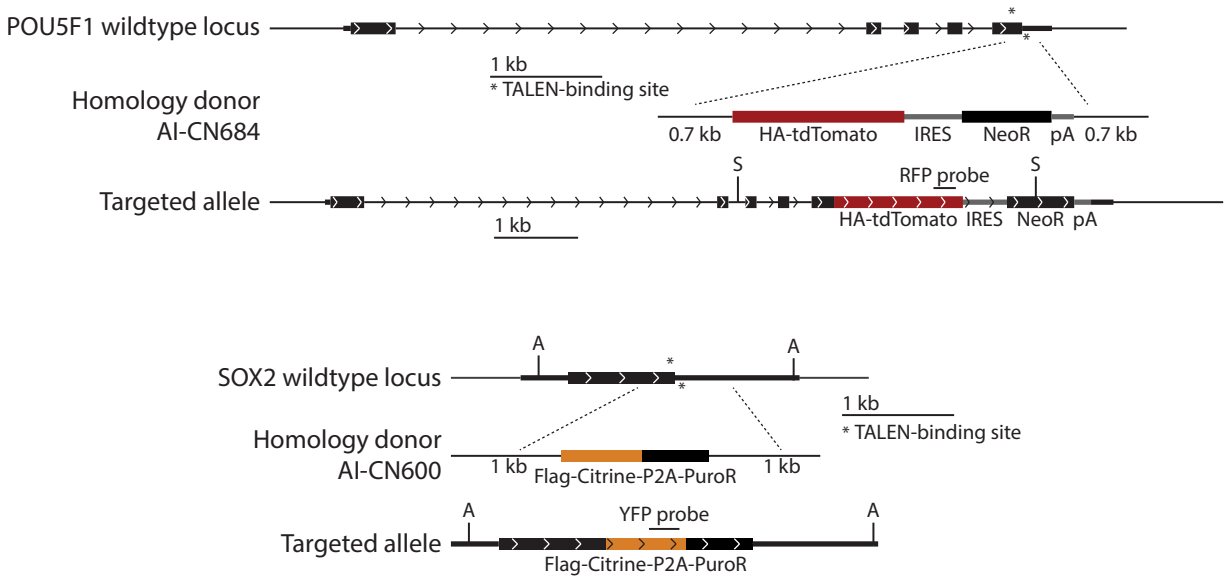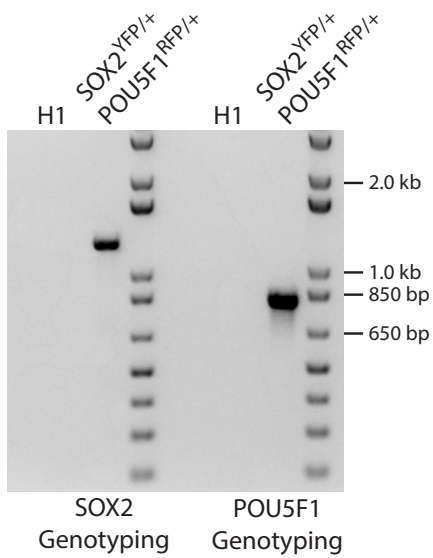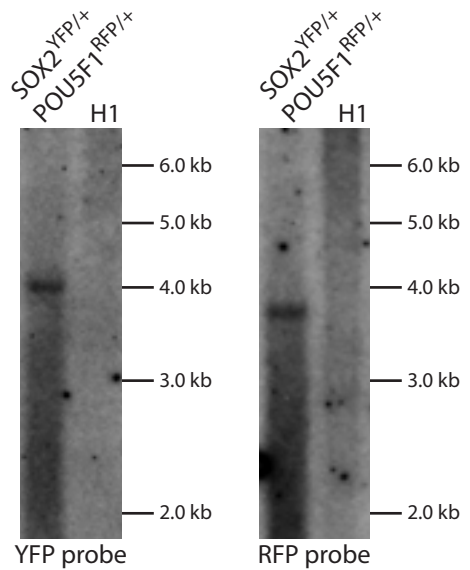
Figure 2.26: Transgenic $SOX2^{YFP/+}$ $POU5F1^{RFP/+}$ **H1 human embryonic stem cell reporter line. a,** Targeting strategy to generate $SOX2^{YFP/+}$ $POU5F1^{RFP/+}$ reporter cell line. **b,** 5′ junction genotyping PCR of the resulting line. **c,** Southern blots of the resulting line. Restriction enzyme sites (A, AflII; S, SacI) are indicated. POU5F1 is also known as OCT4.

of BRACHYURY/T levels on d_{edge} also required epithelial integrity. If hESC colonies were treated with ReLeSR during the first 8 h of differentiation, cells in colony interiors also had high BRACHYURY/T levels (Figure 2.23j).

These data argue that the organization of BMP signaling inward from epithelia tissue edges has significant implications for cell fate decisions. Indeed, we found that d_{source} and d_{edge} each carried independent information about cells' fate choices in the microfluidic device (Figure 2.22h, Figure 2.25g-i). Furthermore, the d_{edge} of hESCs had a significantly higher proficiency of predicting their OCT4/SOX2 and BRACHYURY/T levels than their d_{source}, demonstrating the importance of a cell's distance from epithelial edges as a developmental coordinate.

2.3. Discussion

Our results identify that the interplay between receptor localization and embryo geometry leads to the formation of a robust BMP signaling gradient. Specifically, the compartmentalized geometry of the early mammalian embryo requires BMP4 ligands to diffuse through a narrow interstitial space to approach basolateral receptors. This constraint limits the time and distance ligands can travel before being captured by receptors, which are spatially restricted. As a result, a signaling gradient naturally arises, even when ligands are present uniformly in the lumen on the apical side of the epiblast. Furthermore, through a geometry-related entropic effect, BMP4 ligands accumulate in the apical lumen. Consequently, this lumen serves as a reservoir that buffers the signaling gradient against fluctuations in BMP4 concentration. Due to this entropic buffering mechanism, the channel between the ExE and epiblast that connects the apical lumen and basal interstitial space acts like a stable BMP4 source for the

epiblast. Therefore, a robust BMP signaling gradient forms spontaneously due to the compartmentalized embryo geometry and basolateral receptor localization. While receptor localization has been shown to regulate morphogen signaling in cells in vitro[15,21,23,28] and adult tissues[20], this study demonstrates the effects of receptor localization on morphogen signaling in the developing embryo.

Our current model neglects the possible effects of other regulators of BMP signaling, such as BMP activators and inhibitors. In particular, TGF-β family inhibitors LEFTY1 and CER1 are expressed in the anterior VE of the mouse embryo at E5.75 where they are required for proper patterning during gastrulation[35]. We anticipate that inclusion of such regulators to the model would restrict BMP signaling more to the posterior edge of the epiblast and could contribute further robustness to the BMP signaling gradient against fluctuations in ligand concentration[5,6,10,12,14,15,36]. Nevertheless, our results show that embryonic geometry and receptor localization are sufficient to produce robust gradients of BMP signaling and to explain how mis-localization of BMP receptors leads to ectopic signaling in anterior and distal epiblast cells (Figure 2.19). It would be particularly interesting to incorporate BMP regulators in future versions of the model given that they too can be constrained by the compartmentalization of the embryo[15,37].

This study leads to the question of how access to BMP4 ligands changes as the geometry of the embryo rapidly transforms during gastrulation. As a consequence of BMP signaling, epiblast cells undergo epithelial–mesenchymal transitions (EMT) and ingress within the primitive streak[3–6]. Our results suggest that the breaking of tight junctions during EMT may create additional channels between the pre-amniotic cavity and interstitial space, allowing BMP4 ligands to reach previously inaccessible receptors near sites of ingression. Therefore, our results indicate a possible feedback loop between

embryo geometry and signaling, in which the epithelial integrity of the epiblast restricts BMP signaling while BMP signaling promotes breakdown of the epithelium. How this geometric feedback, in conjunction with the mechanical forces present during EMT[38–40], regulates the distal extension of the primitive streak and the patterning of the streak fates along the anterior–posterior axis is an important subject of future investigation.

We expect that our mechanism of signaling gradient formation will be broadly applicable in many developmental contexts. Epithelial tissues naturally compartmentalize the extracellular space of developing embryos[25,26,41], as found elsewhere in the migrating zebrafish lateral line primordium[41] and the imaginal wing disc of flies[42]. In particular, while the shape of the epiblast differs among mammalian species, its role in compartmentalizing the pre-gastrulation embryo into an apical lumen and basolateral interstitial space is conserved[6,25,30,43]. The observation that TGF-β family receptors in species ranging from flies to mammals contain the LTA amino acid motif suggests that their basolateral localization may also be evolutionarily conserved. It has been shown in a concurrent study that mutation of the LTA motif results in apical mis-localization of the TGF-β receptor in the polarized intestinal cells of *C. elegans*[44], supporting this hypothesis. Moreover, the LTA motif of human TGF-β receptors overlaps with several germline missense mutations associated with Marfan-like syndromes[44], indicating the potential importance of receptor mis-localization for developmental diseases.

Our results demonstrate how the combination of compartmentalization and receptor localization restricts the sensing of morphogens in developing tissues, which can dramatically modulate signaling and downstream tissue patterning. Therefore, future studies should take these factors into account when considering how morphogen signals pattern the embryo during development.

2.4. Methods

2.4.1 Simulation of BMP4 dynamics

In 2D simulation, BMP4 ligands are secreted by six ExE cells and received by 20 epiblast cells arranged linearly along the proximal–distal axis. The 3D simulation contains 20 such linearly arranged arrays of cells in parallel along the anterior–posterior axis. As a result, the 3D simulation contains 120 ExE cells and 400 epiblast cells. Periodic boundary condition was applied along the anterior–posterior axis. Each cell is 8 μm wide and 18 μm tall. The pre-amniotic cavity above the cells is 260 μm wide and 30 μm tall. The interstitial space is 260 μm wide and 2 μm tall. The lateral separation between cells is 2 μm. The simulation setup is therefore comparable to the geometry of the pre-gastrulation mouse embryo (300 μm in length and 100 μm in width). The height of the interstitial space and lateral separation between cells were estimated from fluorescein injection experiments and images of embryos stained for BMPR1A. For 3D simulations of BMP signaling in hESC colonies, we removed ExE cells from the simulation and used the same parameters.

Each epiblast cell has 100 receptors. In 2D simulation, by default 1000 ligands are initially secreted uniformly from the ExE at either the apical membrane or the basal membrane. Although the true number of ligands and receptors may likely be different in the mouse embryo, our simulation results hold for a wide range of scenarios, from the regime where ligands (4000) heavily outnumber receptors (2000) to the regime where receptors (2000) heavily outnumber ligands (200). 3D simulation, in comparison, contains 40,000 receptors and 4000–80,000 ligands.

After secretion, BMP4 ligand diffusion is simulated as a random walk following Brownian dynamics. Ligand positions are updated after each time step h according to

the equation $\vec{r}_i(t+h)=\vec{r}_i(t) + \vec{\Gamma}_i(t)$, where $\vec{r}_i(t)$ is the position of ligand i at time t. $\vec{\Gamma}_i(t)$

is a random Brownian force acting on ligand i that satisfies constraints

$\vec{\Gamma}_i(t)=0$ and $\vec{\Gamma}_i(t)\vec{\Gamma}_j(t') = cDh\delta_{ij}\delta_{tt}$, where D is the diffusion coefficient and $c = 4$ for 2D

simulations or $c = 6$ for 3D simulations[27]. We estimate $D = 20\,\mu m^2/s$ by default based on

diffusion measurements of BMP homolog Dpp in the larval wing disc of *Drosophila*

melanogaster[18]. We use the "local" diffusion coefficient measured by fluorescence

correlation spectroscopy rather than the "global" diffusion coefficient measured by

FRAP since our simulation explicitly models ligand–receptor binding, which has been

shown to slow ligand diffusion at larger length scales[2,18]. Each simulation integration

step occurs after $h = 3$ ms.

The diffusing ligands are not allowed to diffuse through tight junctions between

cells, cell membranes, or the outer boundaries of the pre-amniotic cavity and interstitial

space. Incoming ligands are instead reflected at these surfaces. Given that tight

junctions are absent between the ExE and the epiblast, ligands in the pre-amniotic

cavity are allowed to reach the interstitial space, and vice versa, through the gap at the

edge of epiblast.

If a ligand contacts an epiblast cell membrane that has any unbound receptors,

the ligand binds the receptor with probability $P = P_{binding}R_{unbound}$, where $P_{binding}$ is the

probability a ligand binds a nearby unbound receptor and $R_{unbound}$ is the fraction of

receptors on the membrane that are not bound by ligand. By default, $P_{binding} = 0.002$. For

both 2D and 3D simulations, each epiblast cell has 80 receptors on lateral membrane

and 20 receptors on its basal membrane.

After $T_t = 45$ min, a timescale related to the endocytosis and recycling of ligand-

bound receptors[17,45,46], each receptor–ligand pair is replaced by an unbound receptor on

the same epiblast cell membrane and an unbound ligand secreted by the same ExE cell. This coupling between releasing of unbound receptor and unbound ligand was to maintain the total ligand concentration as a constant. As a control, we also performed simulations in which (a) no unbound ligand was secreted by ExE cell upon endocytosis of ligand–receptor pair and total ligand concentration is slowly decreasing over time (b) unbound ligands were constantly secreted by ExE cells and total ligand concentration is slowly increasing over time.

For any given set of parameters, simulations were repeated 10 times. The simulation had no other parameters and was coded in C. The code was commented and available at 10.6084/m9.figshare.9684992. To run the code: (1) compile the C-code in terminal (on Mac OSX or cluster) by: g++-o test.exe comment2-loop-simulp3d7tov.cpp. (2) Run the exe by: ./test.exe.

A particle diffusion simulation was utilized rather than a reaction-diffusion model to study the effects of (i) embryonic geometry (Figure 2.5), (ii) polarized ligand secretion (Figure 2.3), (iii) receptor mis-localization (Figure 2.10) on BMP signaling with an intuitively understandable approach.

Although direct quantitative comparison between the model and experiment is not possible without precise knowledge of biochemical parameters, we expect our model to agree with experiment qualitatively in the following five criteria: (i) pSMAD1/5 as a function of time (Figure 2.1d, Figure 2.15b, f); (ii) pSMAD1/5 as a function of concentration (Figure 2.1f, Figure 2.15d, h, Figure 2.20), (iii) when tight junctions are broken (Figure 2.4c, Figure 2.16c,d), (iv) when receptors are mis-localized (Figure 2.19, Figure 2.4c, and Figure 2.10), (v) mutation information between pSMAD1/5 and distance from epithelial edge (Figure 2.22b, d).

2.4.2. Cell lines used in the study

All hESC experiments were performed with WA01 (H1) cells or SOX2-YFP, OCT4-RFP double reporter cells (see below) in an H1 background.

2.4.3. Cell culture and passage

hESCs were maintained in the feeder-free cell culture medium mTeSR1 (STEMCELL Technologies) with daily media changes. For passaging, cells were dissociated en bloc with ReLeSR (STEMCELL Technologies) following the manufacturer's protocol, and detached ES cell clumps were broken into smaller pieces (10–20 cells) by tapping the plate or gently pipetting several times with a wide-bore P1000 micropipette (Corning). Cells were passaged at a 1:12 split ratio onto Matrigel-coated (Corning) plates. Immediately following passage, cells were maintained in mTeSR1 supplemented with $10\,\mu$M ROCK inhibitor Y-27632 (STEMCELL Technologies) for 24 h before returning to culture in mTeSR1 alone.

2.4.4. Surface immunostaining of hESCs

Before surface receptor staining[21], cells were rinsed once in 1× PBS (Lonza). Cells were incubated with primary antibodies diluted in mTeSR1 with 1% BSA and 5% normal donkey serum at 37 °C for 45 min. Afterward, cells were rinsed two times in PBS and subsequently fixed in 4% formaldehyde for 20 min at room temperature. Secondary stains were then performed (see 2.4.4. Intracellular immunostaining of hESCs).

2.4.5. Intracellular immunostaining of hESCs

Cells were fixed for 20 min at room temperature in 4% formaldehyde and rinsed three times with PBS. Permeabilization and blocking were performed simultaneously by incubating cells in blocking buffer (PBS with 5% normal donkey serum and 0.3% Triton X-100) for 1 h at room temperature. Primary antibody incubation was performed overnight at 4 °C in antibody dilution buffer (PBS plus 1% BSA, and 0.3% Triton X-100). The next day, cells were washed with PBS three times and then incubated with DAPI and secondary antibodies in antibody dilution buffer (as above) for 1 h at room temperature. After secondary stain, cells were washed with PBS three times before imaging.

2.4.6. Antibodies

BMPR1A (1:20, sc20736; Santa Cruz); BRACHYURY/T (1:400, AF2085; R&D); Clover (1:600, EMU101; Kerafast); OCT4 (1:800, sc8628; Santa Cruz); pSMAD1/5 (1:800, 13820s; Cell Signaling); TGFBR1 (1:20, sc9048; Santa Cruz); ZO-1 (1:100, 33-9100; Thermo Fisher); ZO-1-FITC (1:100, 33-9111; Thermo Fisher).

2.4.7. Plasmid construction and transient expression of receptors

Receptor genes (BMPR1A and BMPR2) were cloned into the plasmid pCAGIP-TGFBR2-Clover (a gift from the Jeff Wrana lab at Lunenfeld Tanenbaum Research Institute) between restriction sites *Xho*I and *Age*I. To visualize receptors using small epitope tags, Clover was replaced by an Myc or HA tag between restriction sites *Age*I and *Not*I. To minimize side effects caused by plasmid expression of tagged protein, we excluded cells with excessive levels of expression, aggregates of fluorescent proteins, and membrane blebbing from downstream analysis.

2.4.8. Plasmid construction and receptor mis-localization

To mis-localize receptors, LTA motifs in both *BMPR1A* and *BMPR2* were mutated into LTG sequences[22] in our plasmids by site-directed mutagenesis (NEB). The puromycin in the pCAGIP-BMPR1A-Clover plasmid was replaced by BMPR2-Myc between restriction sites *Bmg*BI and *Sac*I. To minimize side effects caused by plasmid expression of tagged protein, we excluded cells with excessive levels of expression, aggregates of fluorescent proteins, or membrane blebbing from downstream analysis.

2.4.9. hESC transfection

Transfection of hESCs was performed using jetPrime (Polyplus-transfection) or the Amaxa Nucleofector II (Lonza). For jetPrime transfection, hESCs were transfected within 2 days of passage following the manufacturer's protocol. For nucleofection, hESC cell colonies were dissociated into single cells (see Single-cell passaging) and split into aliquots of 800,000 cells. Aliquots were spun for 3 min at $200 \times g$ before resuspension in $82\,\mu L$ human stem cell Nucleofector Solution 2 (Lonza) and $18\,\mu L$ Supplement 1 (Lonza) with 1–5 μg of DNA. The cell suspension was added to a nucleofection cuvette, and transfection was carried out using nucleofection program B016. Immediately following transfection, $500\,\mu L$ of mTeSR1 culture medium (STEMCELL Technologies) supplemented with $10\,\mu M$ ROCK inhibitor (STEMCELL Technologies) was added to the cuvette, and cells were seeded into a 15 mm well (Corning) coated with Matrigel (Corning).

2.4.10.   Breaking tight junctions

hESC colonies were washed once with PBS and then treated with ReLeSR (STEMCELL Technologies) for 1–2 min at 37 °C. Alternatively, cells were washed once with PBS and then treated with 2 mM EGTA (SIGMA) for 20 min at 37 °C[47].

2.4.11.   Single-cell passaging

hESC colonies were dissociated into single cells by adding 1 mL of 0.05% Trypsin-EDTA (Life Technologies) or 1 mL Accutase (Innovative Cell Technologies) to cells in a 9.6 cm² well, incubating cells for 5–7 min at 37 °C, and quenching with 1 mL of ES-qualified FBS (Millipore). Cell clumps were broken up by gently flushing cells 5–10 times with a P1000 micropipette. Afterward, cells were collected, centrifuged at $200 \times g$ for 3 min, and re-suspended in mTeSR1 supplemented with 10 μM ROCK inhibitor. In total, 200,000 to 1,200,000 cells were seeded into a 15 mm well coated with Matrigel.

2.4.12.   Epifluorescence imaging of hESCs

hESCs were imaged on a Zeiss Axiovision inverted microscope with Zeiss ×10 and ×20 plan apo objectives (NA 1.3) using the appropriate filter sets and an Orca-Flash 4.0 camera (Hamamatsu). The 38 HE GFP/43 HE DsRed/46 HE YFP/47 HE CFP/49 DAPI/50 Cy5 filter sets from Zeiss were used.

2.4.13.   Confocal imaging of hESCs

Cells were imaged on a Zeiss LSM 700 confocal microscope with Zeiss ×40 and ×63 oil objectives (NA 1.3) with the appropriate filter sets and a back-thinned Hamamatsu EMCCD camera.

2.4.14.   Mouse embryo recovery

Eight-week-old adult C57BL/6J female mice were naturally mated and sacrificed at 6 a.m. (E6.25), 12 p.m. (E6.5), or 6 p.m. (E6.75) on the sixth day post coitum. In each case, the uterus was recovered, and embryos were dissected from the deciduae[48,49] in embryo culture buffer (see 2.4.16. Mouse embyro culture).

2.4.15.   Mouse embryo microinjection

Embryos were transferred to a microinjection chamber immersed in PBS. These microinjection chambers were made with 0.4% agarose and had multiple channels for holding embryos (Figure 2.20c). They were specifically designed to minimize the movement and deformation of embryos during microinjection. Microinjection needles were made by pulling glass capillaries (Kwik-Fil, 1B100F-4, World precision instruments) in a micropipette puller (Model P-97, Sutter instrument) using a custom program (Heat 516, Pull 99, Vel 33, and Time 225). The needle was back-filled with 1.5–2.0 $\mu g/\mu L$ plasmid purified using an endotoxin-free maxiprep kit (NucleoBond Xtra Maxi Plus EF, 740426.10, Macherey-Nagel). To reduce jamming during microinjection, the plasmid solution was centrifuged at $5000 \times g$ for 10 min, and the supernatant was loaded into the needle. The microinjection needle was inserted into the pre-amniotic cavity, and the plasmid solution was injected using air pressure (XenoWorks digital microinjector, Sutter instrument) so that the cavity expanded slightly.

2.4.16.   Mouse embryo electroporation

Microinjected embryos were transferred to the electroporation chamber immersed in PBS (Figure 2.20c). Electrodes in the chamber were made of 0.127 mm

platinum wires (00263, Alfa Aesar). Embryos were placed at the center of the chamber, either parallel or perpendicular to platinum wires. Three electric pulses[50] (30 V, 1 ms duration, 1 s apart) were delivered using a square wave electroporator (ECM 830, BTX).

2.4.17. *Mouse embryo culture*

Electroporated embryos were transferred to a 12-well cell culture dish containing embryo culture media at 37 °C and 5% CO_2. This media contains 50% rat serum (AS3061; Valley Biomedical) and 50% Ham's F12 (31765035; Thermo Fisher) supplemented with N-2 (17502048; Thermo Fisher)[51]. The media was equilibrated in the incubator for 1 h prior to embryo addition. E7.5 embryos cultured in this media developed heartbeats after 24–36 h. Electroporated embryos were cultured for 4 h before immunostaining. Only embryos without visible defects were subjected to downstream analysis.

2.4.18. *Surface immunostaining of embryos*

The EXE and underlying VE were removed from embryos using fine forceps (1125200; Dumont). The remaining epiblast and VE were incubated in primary antibodies diluted in embryo culture media with 1% BSA and 5% normal donkey serum for 45 min at 37 °C and 5% CO_2. The embryos were subsequently washed three times with PBS before being fixed for 30 min at room temperature with 4% formaldehyde. Due to this fixation step, occasionally aggregates of unbound antibodies were retained inside the pre-amniotic cavity. These large aggregates, having no DAPI or OCT4 stain, were excluded from downstream analysis.

2.4.19.   Intracellular immunostaining of embryos

Embryos were fixed for 30 min at room temperature in 4% formaldehyde and rinsed three times with PBS. Permeabilization and blocking were performed simultaneously by incubating cells in 5% normal donkey serum, 1% BSA, and 0.3% Triton X-100 in PBS for 1 h at room temperature. Primary antibody incubation was performed overnight with 1% BSA, and 0.3% Triton X-100 in PBS at room temperature. In the morning following primary incubation, embryos were washed three times with PBS and then incubated with secondary antibodies in staining buffer (as above) for 1 h at room temperature. After secondary stain, embryos were washed three times with PBS before imaging.

2.4.20.   Light-sheet imaging of embryos

Stained embryos were embedded into low-melting agarose (BP165-25; Thermo Fisher) containing 0.1 µm fluorescent beads (F8801; Thermo Fisher). The embedded embryos were then imaged in a Zeiss Light-sheet Z1 microscope under ×20 water objective from four angles. The resulting multi-view images were registered using ImageJ plugin multi-view reconstruction.

2.4.21.   Fabrication of microfluidic devices

Microfluidic devices were fabricated in poly(dimethylsiloxane) PDMS using rapid prototyping and soft lithography following published procedures[52]. A photomask was designed to create microfluidic devices that generate linear concentration gradients. A 100-μm-thick "negative" master mold was fabricated from the photomask by patterning SU-8 3050 photoresist on an Si wafer through photolithography.

"Positive" replicas were generated by molding PDMS against the master. After devices were cured, three inlets and one outlet with 0.5 mm diameters were punched. The mold-side surfaces of devices were rendered hydrophilic by plasma oxidation through a 5 min plasma treatment in room air with a plasma cleaner (Harrick Plasma) at high RF power. Immediately after plasma treatment, devices were submerged in deionized water and autoclaved at 121 °C and 100 kPa for 20 min in liquid cycle to simultaneously sterilize the devices and remove toxic non-cross-linked monomers. Bubbles were removed from the autoclaved devices by vacuum desiccation for 30 min. Afterward, autoclaved Tygon tubing (Saint Gobain) was inserted into inlets and outlets, and the entire device was sterilized again with 30 min of UV light in a Class II Biological Safety Cabinet. For all experiments using the microfluidic devices, the amount of time the microfluidic devices spent not submerged in water or cell culture media after plasma treatment was minimized to maintain the hydrophilicity of the molded surface.

2.4.22.   *Culture of hESCs in microfluidic devices*

hESCs to be cultured in microfluidic devices were passaged and maintained in dish culture as described earlier in Methods. At 1 h prior to application of microfluidic devices, cell culture media was changed to mTeSR supplemented with penicillin–streptomycin solution (×100; Sigma Aldrich). Immediately prior to application of microfluidic devices, the tubing of microfluidic devices was filled with mTeSR + penicillin–streptomycin and clamped closed at ends. Devices were then directly attached to the hESC dish using an aluminum clamp custom-designed to fit the dish. Microfluidic devices were positioned with their molded surface over the hESCs and gently clamped downward onto the dish such that cells were located in the cell

chamber. Afterward, inlet tubing was connected to media reservoirs containing mTeSR + penicillin–streptomycin, and outlet tubing was connected to a 3 mL syringe loaded on a syringe pump (Harvard Apparatus). The syringe pump was set to withdraw fluid at a flow rate of 20 μL/min or less. The clamped dish was then placed back into an incubator or loaded onto a Zeiss Axiovision inverted microscope for time-lapse imaging, followed by unclamping all attached tubing and starting the syringe pump. After an hour of flow through the microfluidic device to prime the gradient over the cells, the media in reservoirs was changed to the appropriate differentiation conditions either by adding chemicals directly or by progressive dilution. At the end of microfluidic experiments, 1 mg/mL fluorescein isothiocyanate-dextran (Sigma Aldrich) was added to inlet reservoirs to measure the gradient profile within the device. Once a stable gradient was detected and imaged, the microfluidic device was unclamped from the plate, and cells were fixed and immunostained in situ following standard procedures (see 2.4.4. Intracellular immunostaining of hESCs).

2.4.23.   *Construction of dual-color hESCs*

TALEN genes targeting *POU5F1* (AI-CN330 targeting TCTGGGCTCTCCCAT; AI-CN331 targeting TCCCCCATTCCTAGAAGG) were prepared using the REAL method[53] to match reported target sites[54]. The TALEN genes targeting *SOX2* (AI-CN298 targeting TTAACGGCACACTGCCC; AI-CN299 targeting TCCAGTTCGCTGTCCGGC) were made by the Joung lab (Massachusetts General Hospital) using the FLASH method (PMID: 22484455). *POU5F1* homology-directed repair (HDR) donors AI-CN623 and AI-CN684 were used for constructing the *POU5F1^{RFP/+}* and *SOX2^{YFP/+}POU5F1^{RFP/+}* lines, respectively. The *SOX2* HDR donor was AI-CN600.

H1 cells at p38-39 were treated with 1 μM thiazovivin (StemRD) one day prior to electroporation (Neon; Invitrogen; resuspension buffer R; 100 μL electroporation tip; 1050 V, 30 ms pulse width, 2 pulses; 1.5 or 2 $\times$ 10^6 cells) as single cells (StemPro Accutase, Life Technologies) with 1.5 or 3 μg of each TALEN plasmid and 6 or 12 μg of the HDR donor plasmid. The cells were treated with 2 μM thiazovivin for 24 h following electroporation. After recovery, cells were treated with 1 μg/mL puromycin (Life Technologies) for 3 days. Following 3 days of recovery, dual *SOX2* and *POU5F1*-targeted cells were treated with 75 μg/mL G418 sulfate (Life Technologies) for 3 days. Fluorescent colonies were validated by PCR (*SOX2* 5′ junction primers: CCTGATTCCAGTTTGCCTCTCTCTTTTTTTC, CTTATCGTCGTCATCCTTGTAATCAGATCTCC; *POU5F1* 5′ junction primers: ATGCTGTTACTCAGCAAGTCCAAAGCTTG, GCGTAGTCTGGGACGTCGTATGGGTAAG), had normal karyotypes (Cell Line Genetics), and Southern blots (Lofstrand) confirmed insertion of fluorescent protein transgenes at only the targeted loci in *SOX2*$^{YFP/+}$*POU5F1*$^{RFP/+}$ (AI01e-SOX2OCT4) and *POU5F1*$^{RFP/+}$ (AI05e-OCT4RFP). Silencing of SOX2-YFP was occasionally observed in a small fraction of *SOX2*$^{YFP/+}$*POU5F1*$^{RFP/+}$ cells. This silenced population was regularly removed by fluorescence-activated cell sorting (FACS).

2.4.24.   *Time-lapse microscopy*

For live-cell microscopy, a Zeiss Axiovision microscope was enclosed with an environmental chamber in which CO_2 and temperature were regulated at 5% and 37 °C, respectively. Time-lapse images were acquired every 10 min for 18–48 h. Image

acquisition was controlled by Zen (Zeiss); all cell tracking was manually performed using the TrackMate package in ImageJ (NIH).

2.4.25. *Image analysis of hESCs*

Cell segmentation and fluorescence measurements were done using CellProfiler[55]. All other hESC image data analysis was performed using custom code written in Matlab (MathWorks). Distance from edge (d_{edge}) was calculated as the raw distance of a cell from the colony edge normalized by the average cell diameter (13 μm). P values and confidence intervals were determined by paired t-test.

2.4.26. *Proficiency calculation*

Segmented cells from microfluidic experiments were binned according to their d_{source}, d_{edge}, pSMAD1/5 level, OCT4/SOX2 ratio and/or BRACHYURY/T level into 6, 3, 2, 4, and 2 bins, respectively. The bins for d_{edge} were $d_{edge} < 2$, $2 < d_{edge} < 6$, and $d_{edge} > 6$, where d_{edge} is in units of cell widths. Bins for pSMAD1/5 and BRACHYURY/T levels were calculated by fitting the null distribution to a Gaussian distribution and binning cells as less than or greater than 10 standard deviations from the null distribution mean. Bins for d_{source} and OCT4/SOX2 ratios were determined as evenly distributed percentiles of the total data. Our results did not qualitatively vary with the number of bins or binning algorithm. For each binned variable X and each pair of binned variables X and Y, the discrete marginal probability distribution $P_x(x)$ and joint probability distribution $P_{(X,Y)}(x,y)$ were calculated from the corresponding bin frequencies. The mutual information between variables X and Y was calculated as $I(X;Y)=\sum_y\sum_x P_{(X,Y)}(x,y)\log(P_{(X,Y)}(x,y)/P_X(x)P_Y(y))$, and the entropy of a variable Y was

calculated as $H(Y)=-\sum_y P_Y(y)\log P_Y(y)$. The proficiency[32] for X to predict Y (also called the uncertainty coefficient or entropy coefficient) was then calculated as $U(Y|X)=I(X;Y)/H(Y)$. The proficiency can be intuitively understood as the mutual information shared between variables X and Y normalized by the information entropy of Y, describing the fraction of bits of information about Y that can be predicted by knowing the value of X. Distributions for proficiencies were determined via bootstrapping by resampling cells 10,000 times with replacement.

2.4.27.   Compliance to ethical and other regulations

We have complied with all relevant ethical regulations for animal testing and research. Our use of animal is approved by Harvard University IACUC (protocol #15-01-229). Our use of human embryonic stem cells is approved by Harvard University IRB (protocol #IRB18-0665) and Harvard University ESCRO (protocol E00065).

2.4.28.   Code availability

The C-code for simulating BMP4 dynamics in mouse embryo is available at https://doi.org/10.6084/m9.figshare.9684992. The Matlab code for image data analysis is available on FigShare: at https://doi.org/10.6084/m9.figshare.9805298

2.5.   Acknowledgements

2.6.   References

1. Rogers, K. W. & Schier, A. F. Morphogen gradients: from generation to interpretation. *Annu Rev Cell Dev Biol* **27**, 377–407 (2011).

2. Muller, P., Rogers, K. W., Yu, S. R., Brand, M. & Schier, A. F. Morphogen transport. *Development* **140**, (2013).

3. Mishina, Y., Suzuki, A., Ueno, N. & Behringer, R. R. Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes Dev* **9**, 3027–3037 (1995).

4. Beppu, H. BMP type II receptor is required for gastrulation and early development of mouse embryos. *Dev. Biol.* **221**, (2000).

5. Tam, P. P. L. & Loebel, D. A. F. Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet* **8**, 368–381 (2007).

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

7. Rivera-Perez, J. A. & Hadjantonakis, A. K. The dynamics of morphogenesis in the early mouse embryo. Cold Spring Harb. Perspect. Biol. 7, https://doi.org/10.1101/cshperspect.a015867 (2014).

8. Entchev, E. V., Schwabedissen, A. & Gonzalez-Gaitan, M. Gradient formation of the TGF-beta homolog Dpp. *Cell* **103**, (2000).

9. Teleman, A. A. & Cohen, S. M. Dpp gradient formation in the Drosophila wing imaginal disc. *Cell* **103**, (2000).

10.   Ben-Zvi, D., Shilo, B. Z., Fainsod, A. & Barkai, N. Scaling of the BMP activation gradient in Xenopus embryos. *Nature* **453**, (2008).

11.   Wang, X., Harris, R. E., Bayston, L. J. & Ashe, H. L. Type IV collagens regulate BMP signalling in Drosophila. *Nature* **455**, (2008).

12.   Zakin, L. & Robertis, E. M. Extracellular regulation of BMP signaling. *Curr. Biol.* **20**, (2010).

13. Kicheva, A., Cohen, M. & Briscoe, J. Developmental pattern formation: insights from physics and biology. *Science* **338**, (2012).

14. Muller, P. Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning system. *Science* **336**, (2012).

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

16. Shilo, B. Z. & Barkai, N. Buffering global variability of morphogen gradients. *Dev. Cell* **40**, (2017).

17. Kicheva, A. Kinetics of morphogen gradient formation. *Science* **315**, (2007).

18. Zhou, S. Free extracellular diffusion creates the Dpp morphogen gradient of the Drosophila wing disc. *Curr. Biol.* **22**, (2012).

19. Shahbazi, M. N., Siggia, E. D. & Zernicka-Goetz, M. Self-organization of stem cells into embryos: a window on early mammalian development. *Science* **364**, (2019).

20. Vermeer, P. D. *et al.* Segregation of receptor and ligand regulates activation of epithelial growth factor receptor. *Nature* **422**, 322–326 (2003).

21. Murphy, S. J. *et al.* Differential Trafficking of Transforming Growth Factor-β Receptors and Ligand in Polarized Epithelial Cells. *Mol Biol Cell* **15**, 2853–2862 (2004).

22. Murphy, S. J., Shapira, K. E., Henis, Y. I. & Leof, E. B. A unique element in the cytoplasmic tail of the type II transforming growth factor-beta receptor controls basolateral delivery. *Mol Biol Cell* **18**, 3788–3799 (2007).

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

24. Yin, X. *et al.* Basolateral delivery of the type I transforming growth factor beta receptor is mediated by a dominant-acting cytoplasmic motif. *Mol Biol Cell* **28**, 2701–2711 (2017).

25. Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).

26. Bryant, D. M. & Mostov, K. E. From cells to organs: building polarized tissue. *Nat Rev Mol Cell Biol* **9**, 887–901 (2008).

27. Zhang, Z., Goldtzvik, Y. & Thirumalai, D. Parsing the roles of neck-linker docking and tethered head diffusion in the stepping dynamics of kinesin. *Proc. Natl. Acad. Sci. USA* **114**, (2017).

28. Stull, M. A. Mammary gland homeostasis employs serotonergic regulation of epithelial tight junctions. *Proc. Natl. Acad. Sci. USA* **104**, (2007).

29. Banerjee, U., Renfranz, P. J., Hinton, D. R., Rabin, B. A. & Benzer, S. The sevenless+ protein is expressed apically in cell membranes of developing Drosophila retina; it is not restricted to cell R7. *Cell* **51**, 151–158 (1987).

30. O'Rahilly, R., Müller, F. & Streeter, G. L. Developmental Stages in Human Embryos: Including a Revision of Streeter's 'Horizons' and a Survey of the Carnegie Collection (Carnegie Institution of Washington, 1987).

31. Stelwagen, K., Farr, V. C., Davis, S. R. & Prosser, C. G. EGTA-induced disruption of epithelial cell tight junctions in the lactating caprine mammary gland. *Am. J. Physiol.* **269**, (1995).

32. Cover, T. M. & Thomas, J. A. Elements of Information Theory. 2nd edn (Wiley-Interscience, 2006).

33. Zagorski, M. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* **356**, (2017).

34. Thomson, M. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**, (2011).

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

36. Ben-Zvi, D., Shilo, B. Z. & Barkai, N. Scaling of morphogen gradients. *Curr. Opin. Genet Dev.* **21**, (2011).

37. Mesnard, D., Donnison, M., Fuerer, C., Pfeffer, P. L. & Constam, D. B. The microenvironment patterns the pluripotent mouse epiblast through paracrine Furin and Pace4 proteolytic activities. *Genes Dev.* **25**, (2011).

38. Pinheiro, D. & Bellaiche, Y. Mechanical force-driven adherens junction remodeling and epithelial dynamics. *Dev. Cell* **47**, (2018).

39. Vianello, S. & Lutolf, M. P. Understanding the mechanobiology of early mammalian development through bioengineered models. *Dev. Cell* **48**, (2019).

40. Hannezo, E. & Heisenberg, C. P. Mechanochemical feedback loops in development and disease. *Cell* **178**, (2019).

41. Durdu, S. *et al.* Luminal signalling links cell communication to tissue architecture during organogenesis. *Nature* **515**, 120–124 (2014).

42.   Harmansa, S., Alborelli, I., Bieli, D., Caussinus, E. & Affolter, M. A nanobody-based toolset to investigate the role of protein localization and dispersal in Drosophila. Elife6, https://doi.org/10.7554/eLife.22549 (2017).

43.   Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. & Zernicka-Goetz, M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. Science356, https://doi.org/10.1126/science.aal1810 (2017).

44.   Lin, J., Vora, M., Kane, N. S., Gleason, R. J. & Padgett, R. W. Human Marfan and Marfan-like Syndrome associated mutations lead to altered trafficking of the Type II TGFβ receptor in Caenorhabditis elegans. *PLoS One* **14**, e0216628 (2019).

45.   Schwartz, A. L., Fridovich, S. E. & Lodish, H. F. Kinetics of internalization and recycling of the asialoglycoprotein receptor in a hepatoma cell line. *J. Biol. Chem.* **257**, (1982).

46.   Ciechanover, A., Schwartz, A. L., Dautry-Varsat, A. & Lodish, H. F. Kinetics of internalization and recycling of transferrin and the transferrin receptor in a human hepatoma cell line. *Eff. lysosomotropic agents J. Biol. Chem.* **258**, (1983).

47.   Rothen-Rutishauser, B., Riesen, F. K., Braun, A., Gunthert, M. & Wunderli-Allenspach, H. Dynamics of tight and adherens junctions under EGTA treatment. *J. Membr. Biol.* **188**, (2002).

48.   Bedzhov, I., Leung, C. Y., Bialecka, M. & Zernicka-Goetz, M. In vitro culture of mouse blastocysts beyond the implantation stages. *Nat. Protoc.* **9**, (2014).

49.   Behringer, R. Manipulating the Mouse Embryo : A Laboratory Manual. 4th edn (Cold Spring Harbor Laboratory Press, 2014).

50.   Soares, M. L., Torres-Padilla, M. E. & Zernicka-Goetz, M. Bone morphogenetic protein 4 signaling regulates development of the anterior visceral endoderm in the mouse embryo. *Dev. Growth Differ.* **50**, (2008).

51.   Glanville-Jones, H. C., Woo, N. & Arkell, R. M. Successful whole embryo culture with commercially available reagents. *Int J. Dev. Biol.* **57**, (2013).

52.   Dertinger, S. K. W., Chiu, D. T., Jeon, N. L. & Whitesides, G. M. Generation of gradients having complex shapes using microfluidic networks. *Anal. Chem.* **73**, (2001).

53.   Sander, J. D. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* **29**, (2011).

54.   Hockemeyer, D. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, (2011).

55.   Kamentsky, L. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27**, (2011).

Chapter 3. High throughput control of morphogen gradients to direct patterning and morphogenesis of human organoids in vitro.

Abstract

Morphogens are long range signaling molecules that induce cellular responses in proportion to their concentration, directing the patterning and morphogenesis of the embryo. Surprisingly, only a handful of morphogen pathway families to date have been identified, suggesting this is a possible bottleneck in the process of embryogenesis that can be exploited to control tissue development. Nevertheless, even after considering this reduced dimensionality, there do not yet exist methods precise enough to control morphogen signaling in multiple pathways simultaneously, while rapid enough to explore this spatiotemporal and combinatorial parameter space at a depth sufficient to understand the complexities of development. Here, we introduce a novel method to control the precise presentation of multiple morphogens simultaneously using absorptive beads as morphogen sources. By fabricating chip with microwells to capture beads, we can rapidly position hundreds of human organoids and beads on a single chip in predefined locations to direct the signaling, patterning, and morphogenesis of organoids. We further discuss an iterative Bayesian framework that could be

implemented with this method to infer the morphogen conditions sufficient to elicit an organoid phenotype in a minimal number of experiments.

3.1. Introduction

A fundamental goal of developmental biology is to explain how undifferentiated progenitor cells give rise to complete tissues and organs, each with distinct cell types arranged in complex physical architectures. These processes are directed by morphogens, signaling molecules that are secreted by cells and travel long ranges to pattern nearby tissues in a concentration-dependent manner[1–4]. Morphogens provide crucial positional information that cells need to determine their correct fate, and genetic perturbation of morphogen concentrations or signaling pathway components leads to serious defects in embryonic development. Interestingly, only about six families of morphogen signaling pathways have been shown in this way to be essential for nearly all of embryogenesis[4], with few exceptions across a wide variety of model organisms (see Chapter 1.2). It is therefore not simply the identity of morphogens but their spatial, temporal, and combinatorial profiles that encodes the information by which cells determine their cell type and location within the developing embryo. What is this spatiotemporal, combinatorial 'morphogen code,' and how is it interpreted by tissues precisely and reliably for robust development?

From an engineering perspective, the constrained number of morphogen signaling families represents a "bottleneck" in the process of development, which involves hundreds of cell types, thousands of genes, and a level of structural complexity that is difficult to parameterize. We hypothesize that this bottleneck can be exploited to control embryonic development through the precise spatiotemporal presentation of

114

morphogens to progenitor cells. This approach has already had success in embryological experiments: for example, it has been shown that an artificial pair of opposing gradients of morphogens BMP and NODAL is sufficient to induce the genetic programs that organize the complete zebrafish embryo in vivo and in vitro[5]. Nevertheless, the specific combination of morphogen gradients sufficient to organize development is not known for most embryonic tissues and organs. Further, even in well-studied systems where the essential morphogens have been identified, such as the gastrulation-stage chick embryo, a predictive understanding of how morphogen gradients are translated into morphogenesis is still missing, leading to counterintuitive experimental results[6].

Human organoids, three-dimensional tissues that self-organize from human stem cells in vitro and mimic the anatomy of human organs[7–10], represent a new class of promising model systems in which fundamental questions about development and disease can be explored without the use of model organisms. Research in this field has advanced exponentially in the past decade, with the establishment of protocols to produce human organoids resembling the gastrointestinal tract[11–14], the kidney[15–17], and even the brain[18–21]. There is still a limited understanding of the morphogen gradients that organize these structures, as evidenced by the great amount of heterogeneity seen among organoids formed under the same protocol[20–22]. Moreover, it is challenging to control this morphogen signaling within organoids precisely and rapidly with existing methods; as a result, most organoid protocols are discovered through a slow search for the correct progenitor cells that, when aggregated, spontaneously form pattern and structure[10].

To understand how morphogen profiles are translated into organogenesis, and to leverage the power of organoid systems to reproduce human tissues in a dish, new

techniques are needed to control the presentation of morphogens to organoids in a rapid, quantitative, and combinatorial manner. Here, we introduce a novel method to control the patterning and morphogenesis of human organoids using a "morphogen-beads-on-a-chip" approach. By trapping morphogen-soaked beads in microwells and micropatterning human embryonic stem cells (hESCs) through the humified microcontact printing[23] of Matrigel, we can precisely localize hundreds of organoids and morphogen sources on a single chip. We demonstrate that this technique can used to control a gradient of BMP signaling and mesodermal gene expression across hESC organoids, further inducing cells to undergo epithelial-mesenchymal-transitions (EMT) akin to the primitive streak in vivo. We conclude by outlining an iterative strategy to infer the combinatorial morphogen code that initiates the developmental program responsible for a given organoid phenotype, and by discussing possible applications of this strategy to generate novel organoid types.

3.2. Results

To control the patterning and morphogenesis of human pluripotent stem cell (hPSC) organoids reliably and reproducibly, we sought to develop a method that achieved the following three criteria: (1) The method should enable precise spatial and temporal control over the presentation of a morphogen to the organoid. (2) The method should enable control of multiple morphogen profiles simultaneously and independently of one another. (3) The method be sufficiently high throughput to explore the parameter space of spatiotemporal, combinatorial morphogen profiles in a reasonable number of experiments.
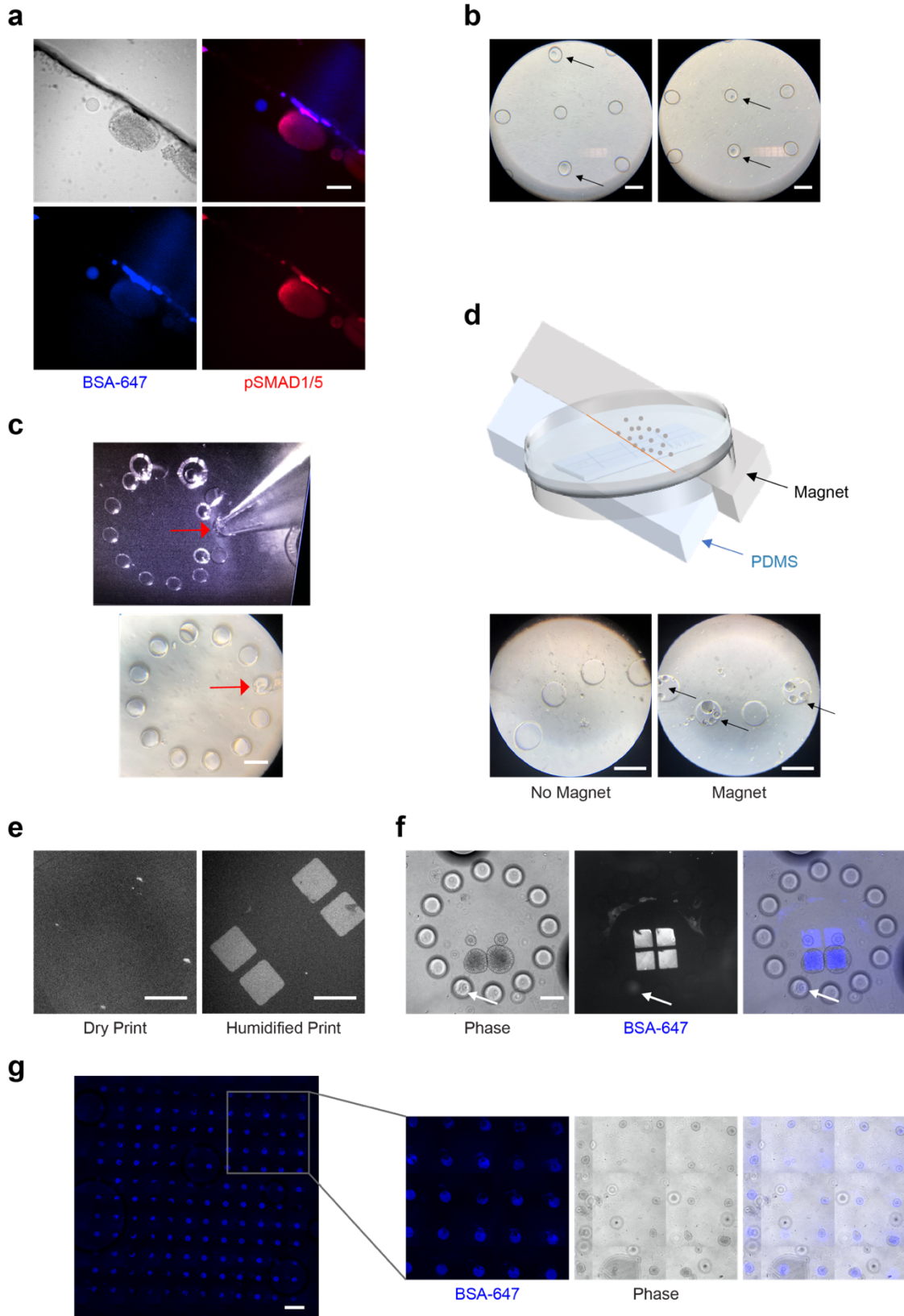
A classic experimental technique to deliver morphogens into embryonic tissues is through the use of absorptive hydrogel beads soaked in a solution of the morphogen[6,24–30]. When beads are implanted into the target tissue, the morphogen will diffuse from the bead and produce a concentration gradient in the surrounding tissues[28]. This method, despite its simplicity, has yielded insightful results: for example, embryological experiments using beads provided the foundational evidence that the protein Activin can act as a morphogen in *Xenopus* embryos[24,28].

We investigated whether a similar approach could be used to control morphogen presentation to organoids in vitro. Given the slow speed of manually implanting beads directly into tissues as well as the small initial size of organoids, we instead sought to control the position of the beads in the surrounding environment of organoids. Beads soaked in recombinant human BMP4 protein induce a gradient of intracellular effector pSMAD1/5 in nearby but non-adjacent organoids (Figure 3.1a), demonstrating that morphogen proteins can passively diffuse from beads through media to induce signaling in nearby organoids in vitro. For these experiments, we used heparin agarose beads (Sigma), which have been used previously in embryological experiments[26] to enhance protein recruitment[31–34]. It is therefore just a matter of positioning beads relative to organoids precisely.

To localize beads in pre-determined locations, we designed PDMS chips with microwells to capture beads through gravity and basic fluid dynamics (Figure 3.1b). We found that circular microwells with diameter and depth both slightly larger than the diameter of a bead performed optimally in capturing a single bead and preventing its release even after multiple washes in PBS. Chip design is cheap, flexible, and parallelizable, with microwells able to be positioned in any configuration above the length scale of the wells themselves (150 μm).

Figure 3.1 (*following page*): Precise, high-throughput deposition of morphogen-soaked beads and organoids on microfabricated PDMS chips. a, Phase contrast (top left) and immunofluorescence images of BMP-soaked bead inducing a pSMAD1/5 (red) gradient in nearby but non-adjacent human embryonic stem cell (hESC) organoid, six hours after bead addition. Heparin-agarose beads were soaked overnight in a solution of 5 μg/mL BMP4 and Alexa Fluor 647-conjugated bovine serum albumin (BSA-647, blue) for visualization**.** Organoid and bead were loosely trapped using edge of a glass slide. **b,** Beads (arrow) trapped in PDMS microwells with diameter and depth 150 μm, after three washes in PBS. **c,** Microinjection of two beads (arrow) directly into a well, after three washes. **d,** Above: illustration of bar magnet/PDMS positioned beneath PDMS microwell chip before deposition of magnetic heparin agarose beads. Below: beads (arrows) trapped in wells above magnet (right) but not wells above PDMS (left), after three washes. **e,** Alexa Fluor 647-conjugated BSA printed in squares onto PDMS using conventional micro-contact printing ("Dry Print", left) and humidified printing (right) protocols. **f,** Localization of four spherical organoids in middle of "clock" arrangement of PDMS microwells, as aligned by eye. BSA-647 marks stamped Matrigel and bead. White arrow: bead trapped in seven o'clock well 24 hours after bead deposition. Beads were added four days after initial cell deposition. **g,** 12 x 12 array of microcontact printed Matrigel circles using humified printing. Inset: spherical organoids are attached where Matrigel is printed. In **(a-f)**, scale bar 200 μm. In **(g)**, scale bar 1 mm

Figure 3.1 (*continued*):

a

BSA-647   pSMAD1/5

b

c

d

Magnet

PDMS

No Magnet   Magnet

e

Dry Print   Humidified Print

f

Phase   BSA-647

g

BSA-647   Phase

119

We deposited beads in microwells by three different approaches with varying degrees of precision and efficiency. First, beads can be directly pipetted into the solution above the chip to achieve random deposition for a fraction of microwells within a minute (Figure 3.1b). This fraction of wells with captured beads is easily controlled by adjusting the concentration of beads added. Nearly all uncaptured beads can be washed out of the media without dislodging trapped beads, ensuring the resulting morphogen profile arises predominantly from captured beads. Alternatively, we directly pipetted into microwells using a microinjection set up (Figure 3.1c). This injection setup can also be used to remove beads or to dislodge any undesired material (e.g. cell aggregates) trapped within wells.

Spatial and temporal control of the system can be further enhanced using commercially available magnetic heparin agarose beads (BeaverBio). After placement of a magnet beneath the chip, beads localized to microwells within domain of the chip above the magnet (Fig. 1d) and were further stabilized from the chance of becoming dislodged from the microwell. Even more, placement of a magnet above the chip quickly removes beads from microwells without disturbing the substrate, allowing for a 'reset' of the morphogen sources along the chip. In this manner, morphogen-soaked beads can be added and removed repeatedly to change the morphogen presentation to organoids continuously over time, limited only by the time-scale of passive diffusion to reestablish a concentration gradient after bead deposition.

The last challenge to control morphogen presentation is to localize the receptive organoids in specific locations on the chip relative to the microwells. Cell micropatterning using PDMS stamps to deposit matrix proteins on glass slides has long been an effective method to position cells[35,36]. However, conventional microcontact printing techniques fail to transfer protein from PDMS stamps to PDMS surfaces

120

(Figure 3.1e), due to the low difference of free energies between the two hydrophobic surfaces[23]. Humified microcontact printing enables the transfer of proteins onto hydrophobic surfaces with significantly improved efficiency[23], allowing us to transfer extracellular matrix proteins (Matrigel, Figure 3.1e) and seed cells (Fig 3.1f, g) consistently onto PDMS chips. Using this method, we can position hundreds of organoids on a single PDMS chip (Figure 3.1g). Organoids on these chips can later be fixed and stained *in situ* for visualization and subsequent analyses.

An illustration of a typical experimental workflow is shown in Figure 3.2. Following this procedure, beads soaked in BMP4 protein were trapped in wells and produced a pSMAD1/5 gradient in nearby organoids oriented toward the beads, as measured by immunofluorescence (Figure 3.3). This strength of this pSMAD1/5 gradient decays as a function of distance from the bead, and can be experimentally varied by changing the BMP4 concentration in which beads are soaked.

To track the mesodermal differentiation of organoids in response to this BMP signaling gradient, we used a novel human embryonic stem cell line with mesodermal marker BRACHYURY fused to a small domain of fluorescent protein YFP and the corresponding large fluorescent protein domain inserted in the AAVS locus under a constitutive promoter (unpublished). Splitting the two fluorescent protein domains in this manner enables rapid fluorescent tagging of endogenous proteins during cell line construction, since the small domain can be much more easily inserted into the endogenous gene locus than the full-length fluorescent protein. Using the BRACYURY-YFP hESC line with our bead-microwell chip, we found that BMP4-soaked beads induce mesendoderm differentiation of organoids on their bead-facing side, as marked by BRACHYURY expression and the observation of differentiating cells undergoing EMT (Figure 3.4a).
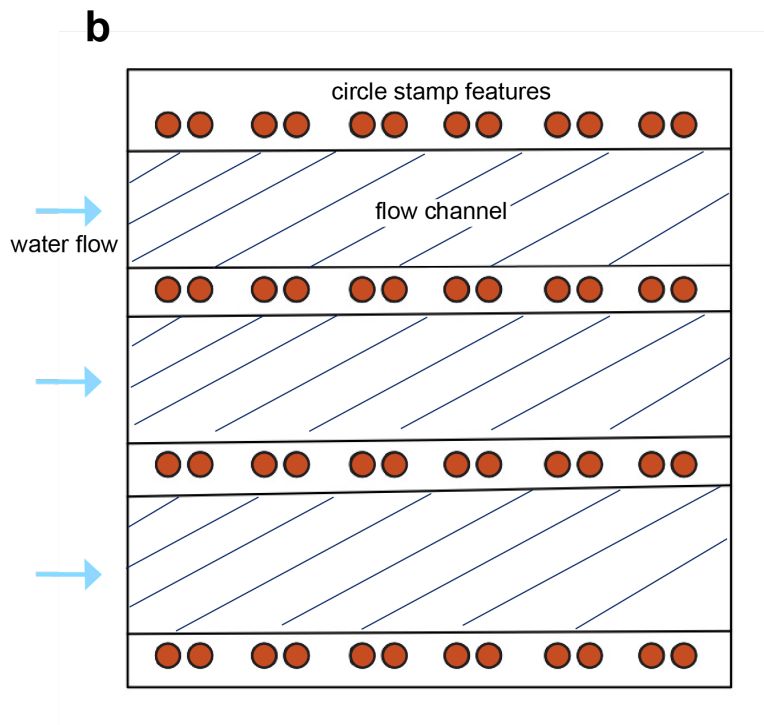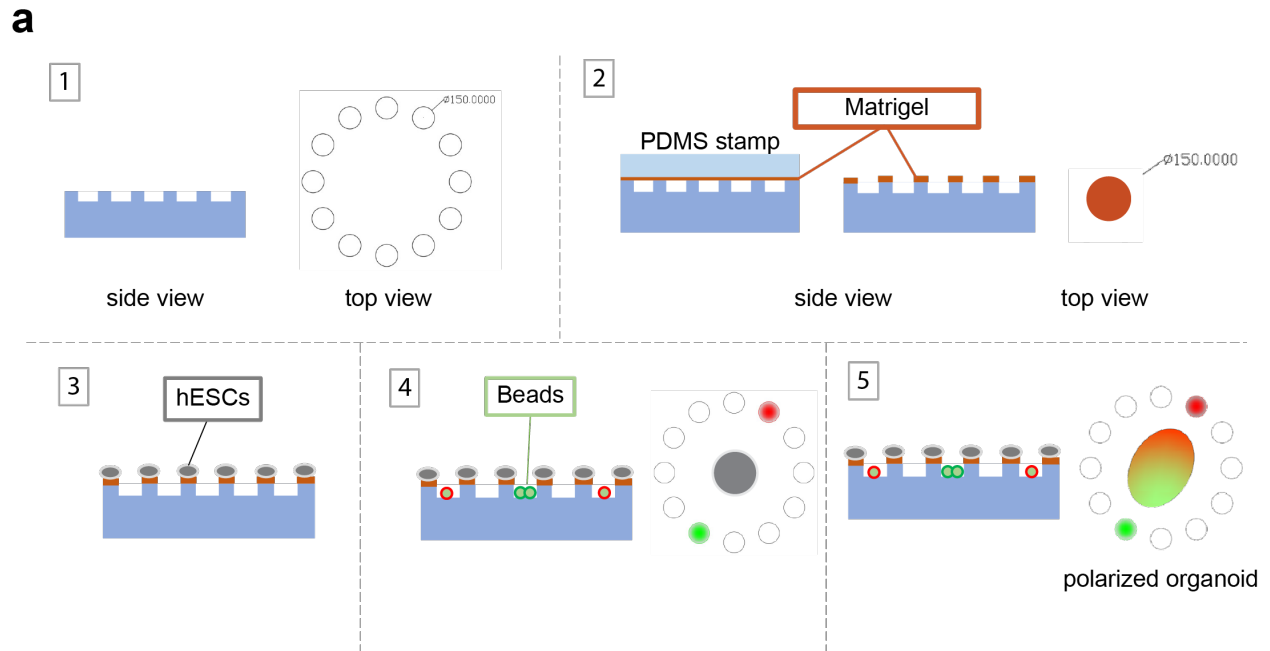
Figure 3.2: Experimental outline using morphogen-soaked beads to pattern organoids on PDMS microwell chip. (a) Outline of a sample experiment. A PDMS chip (blue) with microwells is fabricated using a pre-designed photomask. Matrigel (dark orange) is deposited in defined regions on the PDMS chip by humified microcontact printing using a stamp (light blue), and hESCs (gray) are passaged onto the chip using conventional cell culture methods. Later, hESCs are assembled into organoids and morphogen-soaked beads (red, green) are deposited into microwells

Figure 3.2 (*continued*): on the chip. These beads establish a morphogen gradient across the organoids, directing their patterning accordingly. **(b)** Diagram of two-layer stamp for humified microcontact printing of Matrigel onto PDMS chip. The most positive features (dark orange) mark the regions of contact between stamp and chip where Matrigel is transferred. The most negative features (hash marks) define channels for water flow during microcontact printing. In brief, the stamp is aligned and allowed to contact the PDMS chip, and then water is flowed through channels, diffusing through the stamp to displace proteins chemically from the stamp surface onto the chip.

Figure 3.3 (*following page*): Beads enable the controlled formation of BMP signaling gradients across hESC organoids grown on PDMS chips. (a-c) Beads (arrows) soaked in 10 μg/mL BMP4 were trapped in wells by microinjection near organoids grown on PDMS chips. Organoids were fixed six hours after injection and stained for DAPI (grey) and pSMAD1/5 (red). Top left: phase-contrast image of organoid, top middle: phase-contrast image of bead in focal plane of microwells. BSA-647 (blue) marks where Matrigel is deposited by humified microcontact printing and beads. Scale bar 200 μm

Figure 3.3 (*continued*):

a

| DAPI | BSA-647 | pSMAD1/5 |

b

| DAPI | BSA-647 | pSMAD1/5 |

c

| DAPI | BSA-647 | pSMAD1/5 |

125
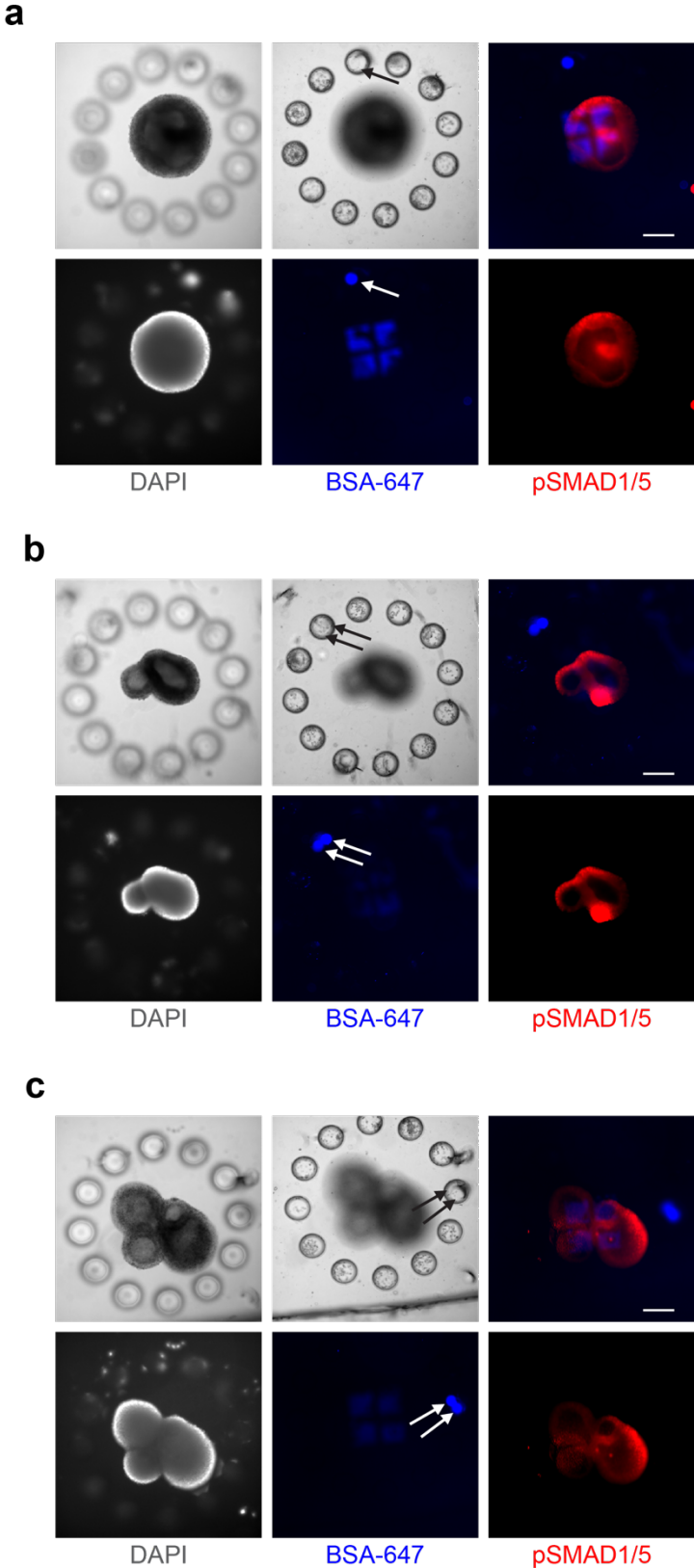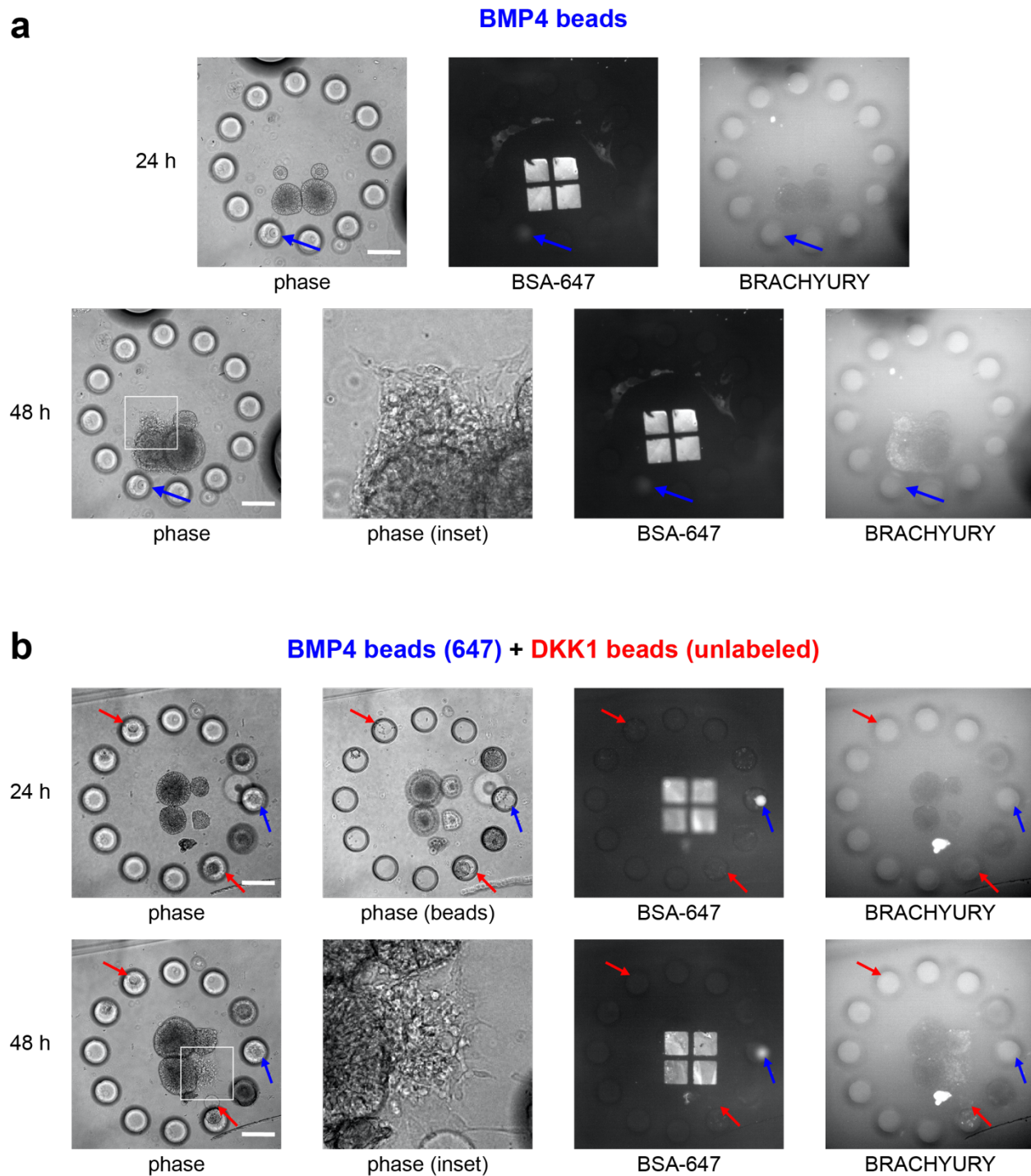
Figure 3.4: Beads organize asymmetric patterning and morphogenesis of organoids. (a) BMP4-soaked bead (blue arrows) induces mesendoderm differentiation of nearby organoids on bead-facing side between 24 and 48 hours after bead deposition, as marked by BRACHYURY expression and EMT (inset). **(b)** Deposition of both BMP-soaked beads and beads soaked in WNT inhibitor DKK1 (red arrows) shapes mesendoderm differentiation of organoids. BMP4-beads, but not DKK1-beads, were soaked in BSA-647 to distinguish them. BRACHYURY expression

Figure 3.4 (*continued*): and EMT (inset) roughly peaked on side of organoid facing BMP4 bead and away from DKK1 beads. Note: dividing cells have displaced the five o'clock DKK1 bead from the well, but the bead is still embedded in a layer of Matrigel, trapping it.

An additional level of control over the system can be obtained by fluorescently labeling beads to distinguish multiple bead types. Beads can be soaked in a solution containing fluorescently-conjugated bovine serum albumin (BSA) to easily visualize the transparent beads during microscopy (Figure 3.1a, f, Figure 3.3, Figure 3.4). Furthermore, we could use this fluorescent label to distinguish beads soaked in BMP4 from beads soaked in DKK1, a protein inhibitor of WNT co-receptor LRP5/6 that represses WNT signaling and mesendoderm differentiation (Figure 3.4b). In this manner, we could direct the patterning of organoids along two morphogen axes simultaneously, easily extending the power of our methodology. Given the wide range of fluorescently-conjugated proteins that are commercially available (Thermofisher) as well as the possibility of fluorescently labeling beads with combinations of colors, it is trivial in practice to extend this method to distinguishing sixteen bead types simultaneously (i.e. the sum of every combination of four different fluorescent colors), likely exceeding the number of morphogens necessary to understand any developmental process in isolation.

3.3. Discussion

To summarize, we have described a novel method we have developed to control systematically the presentation of multiple morphogens to organoids in vitro. Hundreds of organoids and bead wells can be allocated on a single chip, enabling the rapid collection of large amounts of data sharing similar systematic biases (e.g. media).

Given the simplicity of soaking beads in commercially available recombinant morphogens and unique fluorescent labels, there are a huge variety of experimental combinations that can be performed. The question remains: if several experiments, which one should we do first?

I will describe two possible approaches using this method to control the patterning and morphogenesis of human organoids in vitro. The first is to extend the method to existing organoid protocols and to make intelligent guesses about the morphogens that control diversity within them, given what is known about the developmental biology of comparable tissues in model organisms. Even the control of a single morphogen pathway across an organoid has the potential to uncover new biology. Consider, for example, the development of gut tube organoids over the past decade. Like many embryonic tissues, the gut tube arises from common progenitors that become increasingly specialized along the anterior-posterior axis of the embryo – this is essential for the formation of distinct tissues along this axis from the esophagus[13,37] and the lungs[38,39] to the stomach[40], intestines[41,42], and colon[12]. While many of the tissues arising along the gut tube have been modeled using organoid systems in the past decade, these organoids are often spherical and fail to reproduce the anterior-posterior asymmetry that characterizes the gut tube in vivo. This is not simply an engineering exercise: it has been shown that by physically merging foregut and midgut organoids to create this anterior-posterior asymmetry, new cell and tissue types arise at the boundary between the tissues corresponding to the emergence of the liver, pancreas, and biliary ducts from the analogous region in the gut tube[14]. In this manner, asymmetric organoids have the potential to uncover molecular interactions between adjacent organs that are essential for development but impossible to study in humans. Moreover, the difference in protocols between the foregut and midgut organoids is

largely a difference in WNT agonists alone, with increased WNT signaling driving gut progenitors more toward midgut or hindgut fates.

Could the effect of such asymmetric organoids be reproduced using morphogen beads to control a gradient of WNT signaling, for example using beads soaked in DKK1? It is not immediately apparent if this is true, but nonetheless the relative simplicity of this experiment using beads rather than slowly collecting and merging organoids one at a time illustrates the power of this approach to test simple hypotheses. Given what we know about the development of tissues in vivo, we should have more control over the morphogenesis of organoids than we have by waiting for them to self-assemble. Can we direct the formation of optic vesicles from specific domains of brain organoids through the presentation of BMP signals[19,43]? Or can we stimulate the specific folding of intestinal crypts by mimicking the WNT signals secreted by Paneth cells that organize them[44]? We might not need to control five morphogen pathways simultaneously to make significant advancement in long-standing questions about organogenesis, given the difficulty of controlling morphogen signals spatiotemporally in organoids with existing methods.

Suppose however we do not have the luxury of knowing which morphogen signals we should modulate. This is true in many cases – human brain organoids can grow for over a hundred days before displaying a phenotype of interest and have cell types and tissues that are not similarly present in commonly studied model organisms like mice[19,21]. Progress has been made in standardizing the initial conditions that generate these organoids as reproducibly as possible, but they largely still function as black boxes. Transcriptional and epigenetic profiling may reveal clues about the cell types and morphogen signals within them, but these high-dimensional data sets often hide significantly more red herrings[45,46]. How then are we to proceed?

129

Whether the eventual answer involves this method or another, my educated guess is that techniques harnessing the power of organoid systems to obtain large amounts of quantitative data and perform large numbers of perturbations systematically will ultimately be necessary before much significant progress will be made. The promise of organoid systems cannot be simply limited to the same class of low-throughput techniques used for in vivo systems – in the case of embryos, there is at least a ground truth, that this structure will give rise to a complete organism. The French biologist and philosopher Jean Rostand has been quoted to have said, "Theories pass. The frog remains." The power of organoid systems then, in my opinion, must be rapid hypothesis generation. We must find new ways to harness the power of these systems to do more experiments than previously possible, to explore the parameter space of developmental components more fully, and to find the specific trajectories that life happened to take.

Putting aside my fear of egotism momentarily, I would like to put forward a possible class of experiments that could be performed to interrogate this vast parameter space as intelligently and quickly as possible, to get to back to the few ground truths as quickly as possible with a pair of alternate hypotheses in tow. I summarize the approach in Figure 3.5, which does not involve any revolutionary ideas or techniques that have not been shown already in this dissertation. The power of the idea is solely that it is an iterative closed-loop strategy: every piece of data informs our next experiment, so that we might, with sweat and luck, converge upon something interesting or useful. In brief, by controlling the spatial and temporal allocation of beads on our microwell chip, we can rapidly present organoids with a variety of morphogen combinations and sequences and observe the resulting outcomes. Bead microwells could be position randomly at first with a uniform prior expectation on where they

should be, or they could be intelligently placed to test specific hypotheses. Regardless, after some amount of time and some combination of morphogen gradients, we observe and score the resulting organoid phenotype however we choose to measure it. Perhaps it is by some distinct morphological structure that appears, or by the expression of a key factor that can be quickly detected. In either case, if we can reproduce the experiment across enough chips with enough organoids with minimal systematic errors between samples, we now have a framework for statistical modeling of developmental phenotypes, using the power of organoid systems to generate them quickly and cheaply.

The data will be large and noisy, to be sure, but with enough signal perhaps we can begin to cut swaths in this morphogen parameter space and identify broad alternate models that explain our data[6]. The power of a statistical approach then is, without any inherent assumptions about the underlying mechanics of the system, the inverse models we develop suggest the best possible forward models to rule them out. In more technical language, through a Bayesian approach new data will produce new posterior distributions, given our possible prior assumptions. And those posteriors can be used to design the next set of experiments that best distinguish our possible priors. Perhaps the key experiment is to determine whether it is the activator or the inhibitor that is diffusing faster than the other, or perhaps it is a matter of whether this organoid measures the absolute value or the gradient of a particular morphogen when initiating some process. Perhaps the experiment is to determine whether the receptor for a morphogen is localized. The power of such an approach is that we do not need to decide for now, until we have collected the sufficient data to necessitate the eventual in vivo experiment, to demonstrate the hypothesis means something in a living, developing embryo.

Figure 3.5 (*following page*): Iterative strategy for rapid inference of morphogen conditions producing an organoid phenotype. (a) Through use of magnetic heparin agarose beads, morphogen gradients can be repeatedly administered to organoids and later removed using a magnet, enabling complex temporal control of morphogen presentation. Such temporal control will be necessary for more advanced organoid models that require specific sequences or combinations of morphogen signals to produce the required cell fates. **(b)** Example of an iterative strategy to infer the morphogen conditions that optimize the probability of an organoid displaying a phenotype of interest. Differentiating hundreds of organoids with hundreds of beads on a single PDMS chip, across dozens of chips, quickly provides a large amount of data within the spatial, temporal, and combinatorial parameter space of morphogen signals. An iterative Bayesian strategy updating a prior expectation through repeated experiments would provide a possible framework to explore this large parameter in an unbiased manner and converge toward the optimal organoid conditions for a phenotype. While perhaps naïve and ambitious, this approach would identify new hypotheses faster than traditional approaches changing organoid reagents one at a time based on educated guesses informed by the stem cell and developmental biology literature. This approach could be strengthened further by application of that same biological knowledge to constrain the parameter space to be explored.

Figure 3.5 (*continued*):

a

Microwell chip

Print **organoids**

Add **morphogen beads**

Observe **phenotype**

Remove **beads**

Add **morphogen #2 beads**

Observe **phenotype #2**

Remove **beads**

Can repeat **iteratively**

b

Record **bead positions** and **morphogen concentrations**

Design new microwell chip

Score organoid **phenotype** (e.g. morphology, immunostain)

Design maximally informative **next experiment** to refine model parameters

Train **statistical inference** model to optimize organoid phenotype score given bead positions and morphogen concentrations

To conclude this dissertation, I admit that I am not sure whether this exact method will work. There are always pitfalls and limitations, and like any good graduate student I have seen too many of my ideas fail to rely too much on any one to succeed. Still, I feel a strange confidence in the fact that engineering approaches have something to offer the developmental biologist, today more than ever. The molecular and genetic revolution has brought with it huge amounts of data that, while important, have not always revealed their secrets. But the advent of organoids and in vitro studies of development feels like, in my eyes, a return to the time of classic embryologists when scientists had the freedom to "play around" in a system. Engineering gives us as scientists the power to move things around and see what happens with a youthful curiosity. In that sense, I've come over time to think of the original embryologists like Mangold and Spemann and Nieuwkoop as true engineers themselves. Their tools were simpler: the scalpel, the forceps, the hair loop. But they built amazing things with them; they learned how to build eyes, limbs, and brains, all from a little ball of frog cells. I do not know the exact tools we will find most useful in the future, and I do not know what questions we will be able to answer about the mysteries of development. But I do know that I am excited to get back into the lab and try and find out.

3.4. Methods

3.4.1. Design and fabrication of 2-layer stamp with water channels

All microfluidic stamps used were generated by soft lithography using polydimethylsiloxane (PDMS). PDMS base was mixed with curing agent at a 10:1 ratio for all devices. This mix was desiccated in a vacuum desiccator for 30 minutes to

remove air bubbles, poured in a premade mode with silicon wafer and then cured in an oven at 80 °C for 2 hours.

The 2-layer stamp was designed with three layers. The most negative layer defined vascularized channels 300 μm in width, while the most positive layer, located between the vascularized water channels, contained 150 μm diameter circles or 150 μm x 150 μm squares. To generate the silicon wafer made for the PDMS device, a 3'' silicon wafer was coated with a 50 μm SU-8-3050 photoresist film by spin coating, which was subsequently baked until dry. The SU-8 layer was polymerized via photolithography by UV light exposure to polymerize the SU-8 layer through an initial photomask. After a post-exposure baking step, another layer of 50 μm SU8-3050 photoresist was coated onto the once-etched silicon wafer by spin coating. The wafer features were aligned with a second photomask using a microscope, then exposed to UV light again to polymerize another new layer of SU-8 onto the first polymerized layer. Propylene glycol methyl ether acetate (PGMEA) was used as a development agent to remove all unexposed SU-8 on both layers from the wafer altogether. The silicon wafer was then placed in a silane-only desiccator with 2 drops of silane on a glass slide and left in vacuum overnight. After the silicon master is fabricated and silanized, two-layer PDMS stamps were created by pouring PDMS into the mode, desiccating, curing, and gently peeling PDMS from the master.

3.4.2. Preparation of substrate with clock arrangement of wells

The "clock" chips for cell culture and bead deposition contained groups of 12 circular features with 150 μm diameter arranged in clock structures. It was designed together with the 2-layer stamp so that circle/square positive features from the stamps

were aligned in the center of the clock structures on the substrate. This surface was designed as a single layer PDMS device, generated by a conventional photolithography technique. In brief, silicon wafers were coated in SU-8, baked, exposed to UV light with a photomask, post-exposure baked, developed in PGMEA, and silanized overnight before being used as a positive master to cast the PDMS device.

To make the PDMS layer thinner on the chip, the wafer was placed in a plastic dish, submerged in mixed PDMS, and desiccated; glass slide was lightly placed onto the PDMS before the substrate is cured in an oven. The weight of the glass slide gently spread the PDMS layer underneath, creating a thinner layer of PDMS on the master from gravity alone. The thin PDMS layer was then peeled off the master and the glass slide for subsequent use.

3.4.3. Preparation of substrate surface with walls

To keep the chip flat in the well, the PDMS substrate was bonded to a plasma activated 22 mm x 22 mm glass coverslip. Two walls were created around the chip by plasma bonding two 3 mm x 22 mm x 1 mm pieces of flat PDMS on the edge of the coverslip. These two edges create a small chamber when a coverslip is placed on them; this step was crucial to prevent drying of the hydrophobic PDMS chip after aspiration during media changes on the chip since the coverslip on top of the walls provides a hydrophilic surface. Any drying of the PDMS substrate during later cell culture instantly killed any stem cells adhered to the substrate.

3.4.4. Humified microcontact printing

First, 100 μL of Matrigel (Corning) was prepared at 5x the concentration used for 2D cell culture and kept at 4 °C , mixed with 5 μg/ml Alexa Fluor conjugated Bovine Serum (Thermofisher). The Matrigel solution was pipetted onto the surface of flat PDMS stamps. A plasma activated coverslip was placed on top of the Matrigel solution, and the stamp was left at room temperature overnight. Afterward, the stamp was washed in MilliQ water and left to air dry for 30 s. The plasma activated 2-layer stamp was gently pressed onto the flat stamps for 30 s for dry inking of Matrigel. The Matrigel-patterned 2-layer stamps were then quickly aligned with the clock well substrate and pressed down with a 20 g weight. Immediately afterward, the vascularized water channels were be filled with 3 mL MilliQ water prewarmed to 37 °C, and the chip was further incubated at 37 °C for 15 min. After incubation, the remaining water was aspirated and the 2-layer stamp and weight were carefully removed to prevent sliding between the stamp and substrate. For cell deposition, 3 mL of warm mTeSR(+) media (STEMCELL) supplemented with 1x penicillin-streptomycin (Pen-Strep, Sigma-Aldritch) and 25 μL Rock Inhibitor (Stemgent) was added to cover and submerge the chip completely. H1 human embryonic stem cells were passaged onto the chip using Accutase (Innovative Stem Cell Technologies) following the methods described in Chapter 2.4. The media was switched to 3 mL mTeSR(+) + 1x Pen-Strep one day after culturing and changed daily.

3.4.5. Organoid folding

After two days of cell culture on the substrate, a gentle wash with mTeSR(+) media + Pen-Strep was performed to wash away any unattached dead cells. During all media changes and washes, a coverslip was placed onto the substrate (resting on the

outside walls) to avoid any drying of the substrate surface. The dish was placed on ice for one minute. Organoid folding medium was prepared cold by mixing 2.5 mL mTeSR(+) media with 1x Pen-Strep and 75 μL/mL Matrigel (about 20x concentration used for 2D cell culture). Media was then added to the substrate well on ice, and coverslip was removed, and the dish was immediately moved back to the incubator. Monolayer organoids folded into spheres with a single lumen were observed 24-48 hours after addition of the organoid folding medium.

3.4.6. *Preparation of beads coated with morphogen*

Heparin-agarose beads (Sigma) or magnetic heparin-agarose beads (BeaverBio) were washed three times with PBS. Between each wash, the beads were centrifuged in a microcentrifuge tube and the supernatant was aspirated. Beads were then transferred to a tube of PBS with 10 μg/mL BMP4 and 125 ng/mL fluorescent conjugated BSA and rotated at 4 °C overnight.

3.4.7. *Beads deposition*

Before beads were deposited into wells, beads were washed 3x in PBS (with brief centrifugation and supernatant aspiration between each wash). Afterward beads were resuspended into a small volume of PBS (20 uL beads in 200 μL PBS for six chips). Beads deposition is done by one of two different processes. For random bead deposition, 5 μL of the bead solution was added per well. The dish was shaken gently by hand for about 10 s to ensure an even distribution of beads, and then washed three times to removed untrapped beads. The first two washes were done with cold DPBS, and the third was done with E6 media pre-chilled to 4 °C. Finally, 2.5 mL of cold E6

with 1x Pen-Strep and 75 μL/mL Matrigel was added to each well and the plate was returned to the incubator to start differentiation.

For manual bead injection, we pulled a glass capillary tube and polished the tip with sandpaper into a 100 μm diameter. With a pressure-induced injection machine, we precisely injected beads into specific wells near organoids under microscopy. As before, the chip was washed one time with E6 media prechilled to 4 °C, submerged in 2.5 mL E6 media with 1x Pen-Strep and 75 μL/mL Matrigel, and returned to the incubator to start differentiation.

3.4.8. Remaining methods

All remaining methods were performed as described in Chapter 2.4.

3.5.   References

1. Rogers, K. W. & Schier, A. F. Morphogen gradients: from generation to interpretation. *Annu Rev Cell Dev Biol* **27**, 377–407 (2011).

2. Muller, P., Rogers, K. W., Yu, S. R., Brand, M. & Schier, A. F. Morphogen transport. *Development* **140**, (2013).

3. Briscoe, J. & Small, S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009 (2015).

4. Sagner, A. & Briscoe, J. Morphogen interpretation: concentration, time, competence, and signaling dynamics. *Wiley Interdiscip Rev Dev Biol* **6**, e271 (2017).

5. Xu, P.-F., Houssin, N., Ferri-Lagneau, K. F., Thisse, B. & Thisse, C. Construction of a vertebrate embryo from two opposing morphogen gradients. *Science* **344**, 87–89 (2014).

6. Lee, H. C. *et al.* 'Neighbourhood watch' model: embryonic epiblast cells assess positional information in relation to their neighbours. *Development* dev.200295 (2022) doi:10.1242/dev.200295.

7. Fowler, J. L., Ang, L. T. & Loh, K. M. A critical look: Challenges in differentiating human pluripotent stem cells into desired cell types and organoids. *Wiley Interdiscip Rev Dev Biol* **9**, e368 (2020).

8. Hofer, M. & Lutolf, M. P. Engineering organoids. *Nat Rev Mater* **6**, 402–420 (2021).

9. Kim, J., Koo, B.-K. & Knoblich, J. A. Human organoids: model systems for human biology and medicine. *Nat Rev Mol Cell Biol* **21**, 571–584 (2020).

10. Tan, T. H., Liu, J. & Grapin-Botton, A. Mapping and exploring the organoid state space using synthetic biology. *Semin Cell Dev Biol* S1084-9521(22)00141–0 (2022) doi:10.1016/j.semcdb.2022.04.015.

11. McCracken, K. W. *et al.* Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* **516**, 400–404 (2014).

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

13. Trisno, S. L. *et al.* Esophageal Organoids from Human Pluripotent Stem Cells Delineate Sox2 Functions during Esophageal Specification. *Cell Stem Cell* **23**, 501-515.e7 (2018).

14. Koike, H. *et al.* Modelling human hepato-biliary-pancreatic organogenesis from the foregut–midgut boundary. *Nature* **574**, 112–116 (2019).

15. Taguchi, A. *et al.* Redefining the In Vivo Origin of Metanephric Nephron Progenitors Enables Generation of Complex Kidney Structures from Pluripotent Stem Cells. *Cell Stem Cell* **14**, 53–67 (2014).

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

17. Howden, S. E. & Little, M. H. Generating Kidney Organoids from Human Pluripotent Stem Cells Using Defined Conditions. in *Stem Cells and Tissue Repair : Methods and Protocols* (ed. Kioussi, C.) 183–192 (Springer US, 2020). doi:10.1007/978-1-0716-0655-1_15.

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

19. Gabriel, E. *et al.* Human brain organoids assemble functionally integrated bilateral optic vesicles. *Cell Stem Cell* **28**, 1740-1757.e8 (2021).

20. Quadrato, G. *et al.* Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48–53 (2017).

21. Velasco, S. *et al.* Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570**, 523–527 (2019).

22. Quadrato, G., Brown, J. & Arlotta, P. The promises and challenges of human brain organoids as models of neuropsychiatric disease. *Nat Med* **22**, 1220–1228 (2016).

23. Ricoult, S. G., Sanati Nezhad, A., Knapp-Mohammady, M., Kennedy, T. E. & Juncker, D. Humidified Microcontact Printing of Proteins: Universal Patterning of Proteins on Both Low and High Energy Surfaces. *Langmuir* **30**, 12002–12010 (2014).

24. Gurdon, J. B., Harger, P., Mitchell, A. & Lemaire, P. Activin signalling and response to a morphogen gradient. *Nature* **371**, 487–492 (1994).

25. Gurdon, J. B., Mitchell, A. & Mahony, D. Direct and continuous assessment by cells of their position in a morphogen gradient. *Nature* **376**, 520–521 (1995).

26. López-Martínez, A. *et al.* Limb-patterning activity and restricted posterior localization of the amino-terminal product of Sonic hedgehog cleavage. *Current Biology* **5**, 791–796 (1995).

27. Gurdon, J. B., Mitchell, A. & Ryan, K. An experimental system for analyzing response to a morphogen gradient. *Proc Natl Acad Sci U S A* **93**, 9334–9338 (1996).

28. McDowell, N., Zorn, A. M., Crease, D. J. & Gurdon, J. B. Activin has direct long-range signalling activity and can form a concentration gradient by diffusion. *Current Biology* **7**, 671–681 (1997).

29. Pagán-Westphal, S. M. & Tabin, C. J. The transfer of left-right positional information during chick embryogenesis. *Cell* **93**, 25–35 (1998).

30. Arias, C. F., Herrero, M. A., Stern, C. D. & Bertocchini, F. A molecular mechanism of symmetry breaking in the early chick embryo. *Sci Rep* **7**, 15776 (2017).

31. Yan, D. & Lin, X. Shaping morphogen gradients by proteoglycans. *Cold Spring Harb Perspect Biol* **1**, a002493 (2009).

32. Li, Y. *et al.* Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification. *Open Biol* **6**, 150275 (2016).

33. Hettiaratchi, M. H. *et al.* Heparin-mediated delivery of bone morphogenetic protein-2 improves spatial localization of bone regeneration. *Sci Adv* **6**, eaay1240 (2020).

34. Aykul, S., Maust, J. & Martinez-Hackert, E. BMP-4 Extraction from Extracellular Matrix and Analysis of Heparin-Binding Properties. *Mol Biotechnol* **64**, 156–170 (2022).

35. Wilbur, J. L., Kumar, A., Kim, E. & Whitesides, G. M. Microfabrication by microcontact printing of self-assembled monolayers. *Advanced Materials* **6**, 600–604 (1994).

36. Ruiz, S. A. & Chen, C. S. Microcontact printing: A tool to pattern. *Soft Matter* **3**, 168–177 (2007).

37. Zhang, Y. *et al.* 3D Modeling of Esophageal Development using Human PSC-Derived Basal Progenitors Reveals a Critical Role for Notch Signaling. *Cell Stem Cell* **23**, 516-529.e5 (2018).

38. Dye, B. R. *et al.* In vitro generation of human pluripotent stem cell derived lung organoids. *eLife* **4**, e05098 (2015).

39. Chen, Y.-W. *et al.* A three-dimensional model of human lung development and disease from pluripotent stem cells. *Nat Cell Biol* **19**, 542–549 (2017).

40. McCracken, K. W. *et al.* Wnt/β-catenin promotes gastric fundus specification in mice and humans. *Nature* **541**, 182–187 (2017).

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

42. Workman, M. J. *et al.* Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. *Nat Med* **23**, 49–59 (2017).

43. Pandit, T., Jidigam, V. K., Patthey, C. & Gunhaga, L. Neural retina identity is specified by lens-derived BMP signals. *Development* **142**, 1850–1859 (2015).

44. Altay, G. *et al.* Self-organized intestinal epithelial monolayers in crypt and villus-like domains show effective barrier function. *Sci Rep* **9**, 10140 (2019).

45. Furchtgott, L. A., Melton, S., Menon, V. & Ramanathan, S. Discovering sparse transcription factor codes for cell states and state transitions during development. *Elife* **6**, e20488 (2017).

46. Melton, S. & Ramanathan, S. Discovering a sparse set of pairwise discriminating features in high-dimensional data. *Bioinformatics* **37**, 202–212 (2021).