Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized fates and patterning

Sophie M. Morgani[1,2], Jakob J. Metzger [3], Jennifer Nichols[2], Eric D. Siggia[3,4] and Anna-Katerina Hadjantonakis[1,4,5]

1. Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
2. Wellcome Trust-Medical Research Council Centre for Stem Cell Research, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK.
3. Center for Studies in Physics and Biology, The Rockefeller University, New York, NY 10065, USA.
4. Corresponding authors: siggiae@mail.rockefeller.edu and hadj@mskcc.org.
5. Lead Contact.

Abstract (149)

During gastrulation pluripotent epiblast cells specify and spatially arrange the three germ layers. Human pluripotent stem cells (PSCs) have been shown to undergo spatially-organized fate specification on micropatterned surfaces. Since *in vivo* validation is not possible for the human, we developed a mouse PSC micropattern system and, by making direct comparisons to mouse embryos, reveal the robust specification of distinct regional identities. BMP, WNT, ACTIVIN and FGF directed mouse epiblast-like cells to undergo an epithelial-to-mesenchymal transition and radially pattern posterior mesoderm fates. Conversely, WNT, ACTIVIN and FGF patterned anterior mesoderm and endoderm identities. By contrast, epiblast stem cells, a more developmentally advanced state, only specified anterior identities but without patterning. The mouse micropattern system offers a robust scalable method to generate regionalized cell types present *in vivo*, resolve how signals promote distinct cellular identities and generate patterns, and compare mechanisms operating *in vivo* and *in vitro* and across species.

Keywords:

Mammalian Embryo, Epiblast, Gastrulation, Pluripotent Stem Cells, Epiblast-like cells, BMP, Activin, Nodal, Wnt, Primitive streak, Mesoderm, Endoderm, Micropatterns

Introduction

Gastrulation is the process of coordinated cell fate specification, spatial patterning and morphogenesis that establishes the blueprint of the adult organism. During gastrulation, the pluripotent epiblast (Epi) differentiates into the three definitive germ layers of the embryo; the ectoderm, mesoderm and endoderm. In the mouse, these events are initiated at approximately embryonic day (E) 6.25 by a convergence of signals emanating from both extraembryonic and embryonic tissues and acting at the proximal, posterior of the embryo. The BMP/Wnt/Nodal/FGF signaling hub drives posterior Epi cells to undergo an epithelial-to-mesenchymal transition (EMT) [1-3], establishing a dynamic territory referred to as the primitive streak (PS). The PS elongates and extends distally as gastrulation proceeds. Distinct cell types are specified depending on the time and position at which they undergo EMT and exit the PS [4, 5]. Emerging mesenchymal cells either move proximally and laterally, forming the extraembryonic mesoderm, or bilaterally in an anterior direction circumnavigating the space between the Epi and outer visceral endoderm (VE) layers, giving rise to the embryonic mesoderm and definitive endoderm (DE). Epi cells that maintain an epithelial state and do not ingress through the PS will form the ectoderm.

Pluripotent stem cells (PSCs) are the *in vitro* counterpart of the pluripotent Epi of the embryo. They can be expanded indefinitely and differentiated into derivatives of all germ layers [6]. However, standard differentiation protocols generate cell fates in a spatially disorganized manner that is incomparable with *in vivo* gastrulation. Recently it was shown that, when human embryonic stem cells (hESCs) are differentiated within geometrically uniform, circular micropatterns, they reproducibly pattern cell fates with radial symmetry [7]. Based on a limited number of markers, human micropatterned colonies were suggested to comprise a central population of ectoderm followed by mesoderm, endoderm, and an outer layer of extraembryonic trophectoderm [7]. This revealed a surprising capacity of the BMP, Wnt and Nodal pathways to collectively organize cell fates. The scalability and reproducibility of this assay, coupled with the ease of genetically modifying PSCs, and the ability to manipulate conditions for cells grown in culture, as well as the simplicity of imaging make this a robust and attractive system to disentangle the cellular behaviors and signaling interactions that pattern mammalian embryos. Even so, this human organotypic system raised many questions,

largely due to the absence of a human *in vivo* standard for direct comparison and assignment of cellular identities.

Here we adapted the micropattern-based system to defined medium conditions to precisely dissect signaling requirements, and to mouse PSCs to for which *in vivo* reference points are accessible to definitively assign cell fates. We started by converting mouse ESCs to epiblast-like cells (EpiLCs), the *in vitro* counterpart of the Epi of the early pre-gastrulation embryo [8]. Mouse EpiLCs seeded onto circular micropatterned surfaces aquired a simple epithelial morphology and flat-disc geometry. By all markers examined, these cells were identical to the Epi of the E5.5-6.0 embryo. When exposed to gastrulation-promoting factors, micropatterned EpiLCs underwent an EMT and recapitulated germ layer differentiation and spatial organization of specific regions of the mouse gastrula. This demonstrated that the cup-shaped geometry of the rodent embryo is not requisite for the spatial patterning of mouse pluripotent cells. Furthermore, the capacity to undergo spatially organized germ layer specification was specific to EpiLCs. Under the same conditions, neither ESCs nor EpiSCs demonstrated robust patterning of cell fates. Hence, the mouse micropattern system represents a defined and quantitative tool to functionally assess the spectrum of mouse pluripotent states that have been described [9], and determine their correlation to hESCs.

In vivo, we observed a proximal-to-distal gradient of BMP signaling activity – cells in the posterior (proximal) PS exhibited high signaling activity, while those in the anterior PS (distal) showed no activity. We hypothesized that by modulating the signals provided to PSCs we could recapitulate the proximal-distal environments operative *in vivo* and generate distinct regional identities *in vitro*. Exposure of micropatterned EpiLCs to BMP, FGF, ACTIVIN (NODAL) and WNT, signals that mirrored the proximal posterior embryonic environment, promoted an EMT and acquisition of posterior Epi, PS, embryonic and extraembryonic mesoderm identities. When BMP was removed, emulating the anterior PS environment (FGF, ACTIVIN and WNT), anterior Epi, DE and axial mesoderm (AxM) cell types were formed instead. Hence, we demonstrated for the first time that *in vitro* micropattern differentiation parallels events occurring during gastrulation *in vivo* in mammalian embryos, and that mouse PSCs residing in a flat-disc geometry can pattern cohorts of neighboring regional identities correlating with those established in the embryo. Utilizing the micropattern system to manipulate the BMP

pathway in isolation allowed us to extend findings made in mouse mutants by addressing the anterior versus posterior requirements for this signaling pathway within the PS. We established a direct requirement for BMP4 in posterior mesoderm formation, and demonstrated that BMP signaling is not required for AxM and DE specification. Further quantitative analysis of the signaling dynamics, the role of secreted inhibitors and cell-cell interactions should reveal how pathways operate in a flat-disc-shaped geometry, resembling the majority of mammalian embryos (including human), that can now be directly correlated to mouse, the best developed mammalian genetic model.

Results

Micropatterned EpiLCs correspond to the pre-gastrulation epiblast

The pluripotent state represents a continuum spanning from Epi cell specification in the pre-implantation blastocyst (at approximately embryonic day (E)3.5) to differentiation at gastrulation which initiates at E6.25 [9] (Figure 1A). Prior to the onset of gastrulation (E5.5-6.0), the Epi is in a formative state of pluripotency whereby naïve pre-implantation markers, present in the blastocyst, have been downregulated but differentiation has not yet commenced (Figure 1B) [9, 10]. To establish an *in vitro* system to model mouse gastrulation, we reasoned that we should start with a PSC population comparable to the *in vivo* Epi at this time. Global transcriptional profiling has identified epiblast-like cells (EpiLCs) as the closest *in vitro* counterpart of the formative Epi [8, 10-12]. We sought to determine whether this correlation held when EpiLCs were grown on micropatterned surfaces. EpiLCs were generated as previously described [8] and plated onto 1000 μm diameter micropatterns (Figure 1C). Fibronectin and Laminin are basement membrane components present at the Epi-VE interface of peri-gastrulation mouse embryos [13, 14]. While EpiLCs are generated on a Fibronectin substrate [8], we noted superior adhesion of cells to the micropatterns when coating them with Laminin.

Like the pre-streak Epi, micropatterned EpiLCs expressed the pluripotency-associated markers OCT4, SOX2, NANOG and OTX2, as well as the post-implantation Epi marker, OCT6 [15] (Figure 1B,D). Neither the Epi of the pre-gastrula embryo nor EpiLCs expressed KLF4, a marker of naïve pluripotency (Figure 1B,D). The pre-streak Epi does not express lineage-associated markers such as GATA6, FOXA2, CDX2 or

BRACHYURY and, these were also absent from EpiLC colonies (Figure 1B,D) except where spontaneous differentiation had occurred. Micropatterned EpiLC colonies formed an epithelial monolayer (marked by E-CADHERIN) (Figure 1E). Cell density within the epithelium increased by nearly a factor of four in 24 hours (Figure S1A). We noted that micropatterned EpiLCs exhibited a slightly lower cell density than the epithelium of the embryonic Epi at pre- and early gastrulation stages (Figure S1A). However, it has been shown that *in vivo* development can also proceed when the cell density of the Epi is experimentally decreased [16]. The uniform size and circular morphology of micropatterned colonies is amenable to the robust quantification of spatial patterning by measuring protein levels as a function of radial position measured from the colony center to the colony edge (Figure S1B) [7, 17]. Such quantification indicated that micropatterned EpiLC colonies were spatially homogeneous (Figure 1F), hence we started with a defined population correlating to the pre-gastrulation Epi at approximately E5.5.

EpiLCs recapitulate differentiation and spatial organization of posterior cell fates

At the onset of gastrulation, the Epi comprises distinct anterior and posterior domains identified by the expression of specific markers. SOX2 is elevated within the anterior Epi, while NANOG is restricted to the posterior Epi (Figure 2A-C, S1C). BRACHYURY expression marks the emerging PS (Figure 2A-C, S1C). When micropatterned EpiLC colonies were cultured with FGF2 and ACTIVIN A for 72 hours, BRACHYURY was induced at the colony periphery (Figure S1E-G), reminiscent of PS formation. The expression of BRACHYURY was abolished, and SOX2 maintained homogeneously, in the presence of the small molecule inhibitor of Wnt signaling, XAV939 [18] (Figure S1E-G). Conversely, BMP inhibition had no obvious effect on BRACHYURY expression (data not shown). Hence, this initial BRACHYURY expression was dependent on endogenous Wnt signaling. Therefore, FGF and ACTIVIN alone support an anterior Epi fate, while FGF and ACTIVIN combined with endogenous WNT stimulate PS gene expression (Figure S1G, Table 1). The later germ layer markers, GATA6, SOX17 and CDX2 (Figure S1E), were not expressed under these conditions indicating that additional factors were required to stimulate the further differentiation of cells.

In vivo, gastrulation is triggered by a combination of signals from both embryonic Epi cells and the extraembryonic tissues that lie adjacent to the proximal posterior Epi.

These extraembryonic signals include WNT3 produced by the VE, and BMP4 produced by the extraembryonic ectoderm (ExE) [2, 19] (Figure 2D). As our pluripotent cell cultures lacked the VE and ExE extraembryonic cell types, we asked whether supplying these signals exogenously could initiate gastrulation-like events *in vitro*. EpiLCs were plated overnight onto micropatterns in defined serum-free medium containing 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A, supplemented the following day with 50 ng/ml BMP4 and 200 ng/ml WNT3A (Figure 2E). Under these conditions, EpiLCs underwent robust and reproducible organized germ layer specification (Figure 2F-H, S2A-C).

After 24 hours of differentiation, the EpiLC micropatterned colonies gave rise to two populations – a central population expressing the Epi markers OCT4, SOX2 and NANOG and an outer population expressing the PS marker BRACHYURY. Within the central population, OCT4 and SOX2 levels were reduced relative to the 0 hours starting population (Figure 2F-H, S2D); ~2-fold in the case of SOX2 (Figure 2H), comparable to the difference observed between the anterior and posterior domains of the *in vivo* Epi (Figure 2C). Conversely, there was no change in NANOG expression (Figure 2H, S2D). Hence a NANOG-expressing, SOX2-low state emerged as observed in the posterior Epi of the embryo. Concomitantly, BRACHYURY was induced at the colony edge (Figure 2F-H). BRACHYURY and SOX2 expression was predominantly mutually exclusive but showed a degree of overlap (Figure S2E) that was also observed in PS cells *in vivo* (Figure 2A), and likely corresponded to cells in a transition state between Epi and PS. At this time, no later germ layer-associated markers (GATA6, SOX17, CDX2, FOXA2) were expressed (Figure 2F,H). Hence, 24 hour micropatterned colonies equated to the posterior Epi and emerging PS of the mouse embryo at E6.25-6.5.

In vivo, 24 hours after the initiation of gastrulation (E7.25-7.5), the expression signature of the Epi has changed, and mesoderm and DE populations are being specified as they emerge from the PS. In order to map the *in vitro* micropattern differentiation to the gastrulating embryo, we identified cohorts of markers that would allow us to distinguish between the increasingly complex array of populations – the anterior and posterior Epi, the PS, the embryonic and extraembryonic mesoderm (arising from the posterior PS), and the DE and AxM (arising from the anterior PS) (Figure S8).

In addition to expressing NANOG and SOX2, the posterior Epi also expresses OTX2 (Figure S3A). Like SOX2, OTX2 is expressed highly within the anterior and at lower levels within the posterior Epi (Figure S3A). At E7.0-7.25, the distal posterior Epi begins to express FOXA2 (Figure 3A) and, by E7.5, CDX2 is expressed throughout the posterior Epi (Figure 3B).

The first cells to leave the posterior Epi and migrate through the BRACHYURY-positive PS (Figure 2A, S1C) coexpress BRACHYURY and GATA6 (Figure S3B). Over time, the cells emanating from the PS adopt distinct mesodermal and DE identities. Cells that exit the posterior PS and move proximally into the extraembryonic region generate the extraembryonic mesoderm, which can be subdivided into the allantois core domain (ACD) - expressing BRACHYURY, SOX17, CDX2, FOXF1 (Figure 3C,D, S3C), the allantois outer mesenchyme (AOM) - expressing SOX17, CDX2, GATA6, FOXF1 (Figure 3C,D, S3C) and the yolk sac mesoderm (YSM) - expressing GATA6 and FOXF1 (Figure 3D). Over time, cells of the ACD contribute to the AOM to support allantois elongation [20].

Cells that originate from the posterior PS, but instead move in an anterior direction around the embryo, will form embryonic mesoderm. At E7.5, based on protein expression, two populations of embryonic mesoderm could be distinguished, which we referred to as Mesoderm 1 and Mesoderm 2 (Figure 3C). Mesoderm 1 cells, which exited the PS earlier and were located more anteriorly, expressed GATA6 and OTX2 (Figure 3B-D, S3A). Mesoderm 2 cells, which left the PS later and were more posterior, expressed BRACHYURY and FOXF1 (Figure 3B-D). Cells emerging from the anterior PS give rise to the DE – coexpressing FOXA2 and SOX17 (Figure S3D) and AxM - coexpressing FOXA2 and BRACHYURY (Figure S3E).

Utilizing these classifications (Figure S3F), we assigned cellular identities to the populations arising at 48 and 72 hours of *in vitro* micropattern differentiation, when later germ layer markers started to be expressed. At both 48 and 72 hours, three colony domains (central, mid and outer) were evident (Figure S2A,B) and the populations within these domains were largely conserved between time points. At 48 and 72 hours, the colony center continued to express NANOG and low levels of OCT4 and SOX2 (Fig, 2F-H, S2D), as well as OTX2 (Figure S3G) [21], markers all expressed within the posterior

Epi of the embryo. Furthermore, at 72 hours, FOXA2 began to be heterogeneously expressed within the colony center (Figure 3E), analogous to the onset of FOXA2 expression within the distal posterior Epi at E7.0-7.25 (Figure 3A). While OTX2 and SOX2 are also highly expressed within the *in vivo* anterior Epi (Figure S3A), the expression of the posterior-restricted marker NANOG within the same region (Figure 2A, S2D), as well as FOXA2 (Figure 3E), confirmed that the micropattern center was posterior Epi-like. *In vivo*, CDX2 expression is induced within the posterior Epi at E7.5 (Figure 3B), but was not observed within the center of micropatterned colonies (Figure 2F, 3F). Therefore, based on the expression of FOXA2 but not CDX2, the central population correlated to the posterior Epi after E7.0 but prior to E7.5.

While BRACHYURY initially marked a population at the colony periphery, by 48 and 72 hours it had become localized to an intermediate micropattern domain (Figure 2F-H). BRACHYURY is a marker of the PS but is also expressed within the extraembryonic mesoderm ACD, with CDX2 (Figure 3C, S3C), and in early PS derivatives with GATA6 (Figure S3B). As the majority of BRACHYURY-positive cells did not express CDX2 or GATA6 (Figure S3H,I) they likely represented a PS state. The positional change in BRACHYURY expression over time suggests a wave of expression being propagated throughout the colony or an inwards movement of BRACHYURY-positive PS-like cells.

The outermost micropattern domain comprised several distinct cell populations that corresponded to various subtypes of embryonic and extraembryonic mesoderm. At 48 hours, a large fraction of cells within the outer domain coexpressed BRACHYURY, CDX2 and SOX17 (Figure 2F, 3F,G, S3J), a signature of the extraembryonic mesoderm ACD of the embryo (Figure S3C). However, at 72 hours, this population was no longer observed. Instead a distinct cell state emerged that coexpressed CDX2 and SOX17 (Figure 3F,G), but not BRACHYURY (Figure 2F, S3J). Although CDX2 and SOX17 are both expressed within the hindgut endoderm of the embryo [22], this micropattern population also expressed FOXF1 (Figure 3H), which is expressed within mesoderm but not endoderm tissues (Figure 3D). Hence cells that coexpress CDX2, SOX17 and FOXF1 likely corresponded to the AOM, suggesting a temporal progression of ACD cells to an AOM state, as *in vivo*. We also observed rare cells that coexpressed FOXF1 and GATA6 (Figure 3H, yellow arrowhead) representing YSM.

A distinct population of cells within the outer domain expressed GATA6. A small fraction of these also expressed BRACHYURY, as in the first cells leaving the PS (Figure S3B). Additionally, we observed cells that coexpressed GATA6 and OTX2 (Figure S3G), corresponding to Mesoderm 1 (Figure 3B, S3A). Within the embryo, we could discern a second population of mesoderm, Mesoderm 2, which exited the PS later and expressed FOXF1 and BRACHYURY (Figure 3B,D). However, within the micropatterns, FOXF1 was restricted to the colony periphery, distinct from the intermediate BRACHYURY-expressing domain (Figure 2H, 3H), hence the later Mesoderm 2 population was not formed. While both embryonic and extraembryonic mesoderm populations were present within the outer micropattern domain, they tended to exist within discrete clusters (Figure 3G).

In the presence of FGF, ACTIVIN, BMP and WNT we did not observe DE (FOXA2 and SOX17 positive) or AxM (FOXA2 and BRACHYURY positive) cell types (Figure 3E). Therefore, *in vitro*, a combination of BMP, Wnt, Nodal (Activin), and FGF signaling promoted the specification and spatial organization of posterior Epi (center), PS (mid) and embryonic and extraembryonic mesoderm (outer), recapitulating gastrulation events occurring within the posterior of the mouse gastrula (Figure 3I, Table 1).

Micropattern differentiation involves a TGFβ-regulated EMT

One of the primary hallmarks of gastrulation is an EMT, involving downregulation of the epithelial marker E-CADHERIN and upregulation of the mesenchymal marker N-CADHERIN in cells ingressing through the PS (Figure 4A,B). Epi cells that do not undergo EMT differentiate into NE, while those that undergo EMT emanate from the PS and acquire mesoderm or DE identities [3, 23]. We asked whether micropattern differentiation engaged these same morphogenetic processes. A PS-like population arose after 24 hours of *in vitro* differentiation (Figure. 2F) followed by the formation of a 2-3 cell layer ridge at the colony perimeter at 48 hours (Figure 4C,D). By 72 hours, the ridge was positioned more centrally suggesting an inwards movement and resulting in a volcano-like structure (Figure 4C,D). Initially, the ridge overlapped with the BRACHYURY/CDX2 coexpression domain but, at 72 hours, was positioned at the border between the BRACHYURY PS and CDX2 AOM population (Figure S4A). Cells at the border of the BRACHYURY region lost the epithelial marker E-CADHERIN and upregulated the mesenchymal marker N-CADHERIN (Figure 4D,E, S4B). As *in vivo*, the

outer N-CADHERIN expression domain correlated with the position of the PS (BRACHYURY), embryonic mesoderm (GATA6) and extraembryonic mesoderm (CDX2) populations (Figure 4F). Furthermore, both the intermediate PS-like domain and the outer embryonic and extraembryonic mesoderm domain expressed SNAIL (Figure S4C-E), a transcriptional repressor that regulates the gastrulation EMT [24]. At 48 hours, N-CADHERIN-positive cells emerged at the base of the colony, beneath the E-CADHERIN-positive epithelial layer, and were observed more centrally over time (Figure 4D,E). We also occasionally observed BRACHYURY expressing cells in more central positions at 72 hours (Figure 4F), which could suggest an inwards migration of mesenchymal PS derivatives between the upper epithelium and the surface of the micropattern slide. Conversely, central posterior Epi-like cells, maintained E-CADHERIN (Figure 4D,E).

Various signaling pathways including Wnt, FGF and TGFβ regulate EMT in development and cancer [23, 25, 26]. In particular, the role of TGFβ signaling through SMAD2/3 has been well characterized [27]. Mice with mutations in *Smad2*/*3* or *Nodal* do not gastrulate and lack normal mesoderm structures [28-31]. To determine whether SMAD2/3 signaling regulated EMT in the *in vitro* micropattern system, we cultured EpiLC micropatterned colonies in medium containing FGF2, BMP4, WNT3A but lacking ACTIVIN A and supplemented with a small molecule inhibitor of the ALK5 receptor, SB431542 (referred to as ACTIVINi). In the absence of Activin (Nodal) signaling, cells maintained high levels of E-CADHERIN and accumulated at the edge of colonies (Figure 4G,H, S4C). Furthermore, they failed to downregulate SOX2 and did not differentiate, evidenced by the lack of BRACHYURY, GATA6, CDX2 or SOX17 expression (Figure 4I, S4F). Thus, in these flat-disc-shaped micropatterns, SMAD2/3 signaling regulated the EMT associated with an exit from pluripotency and onset of differentiation confirming that *in vitro* micropattern differentiation and *in vivo* gastrulation are regulated by common pathways and processes even though their geometries (flat-disc versus cup-shaped) are distinct.

Micropattern colonies exhibit position-dependent BMP signaling

While the cell culture medium provided homogeneous signals to the micropatterned colonies, different cell fates emerged within distinct radial domains. To determine whether this patterning correlated to a position-dependent interpretation of signals, we

focused on BMP, a key upstream signal necessary for gastrulation with an effective antibody readout of activity - nuclear localization of phosphorylated SMAD1/5/8 (pSMAD1/5/8). *In vivo*, at the early streak stage (E6.5), BMP4 is expressed by the ExE and later (E7.5) by the allantois and chorion [32] (Figure S5A), and acts on adjacent tissues.

In the micropatterns, at 0 hours, nuclear pSMAD1/5/8 was observed at low levels throughout colonies (Figure 5A). This was in agreement with the low BMP signaling activity present in the proximal Epi of the embryo at E6.5 (Figure S5B). Starting at 24 hours of micropattern differentiation, nuclear pSMAD1/5/8 became elevated at the colony edge within the domains correlating to PS, embryonic and extraembryonic mesoderm cell fates (Figure 5A-C). At 24 and 48 hours, the majority of pSMAD1/5/8-positive cells expressed BRACHYURY (Figure 5A, S5C,D). *In vivo*, nuclear pSMAD1/5/8 was also elevated within the PS, embryonic and extraembryonic mesoderm, correlating with BRACHYURY expression (Figure 5D,E, S5B,E).

At 72 hours of micropattern differentiation, the fraction of pSMAD1/5/8-positive cells that coexpressed BRACHYURY was significantly reduced (Figure S5C,D). This corresponded to the presence of nuclear pSMAD1/5/8, but not BRACHYURY, within anterior Mesoderm 1 cells of the embryo (Figure S5F-H). There was no evidence of BMP signaling activity within the distal Epi or anterior PS (Figure S5B,F,I), consistent with these cells being positioned furthest from the ExE source of BMP4. Hence, cell types identified within the *in vitro* micropattern system experienced a comparable BMP signaling history to their *in vivo* counterparts. In the embryo and in micropatterned colonies, posterior Epi cells displayed low levels of BMP signaling activity, which was elevated within the posterior PS, embryonic Mesoderm 1 and extraembryonic mesoderm populations. By contrast, the distal Epi and anterior PS of the embryo were devoid of BMP activity (Figure S5B,F,I). However, a comparable distal Epi/anterior PS signaling niche that lacked BMP activity was not present in the micropatterns under these conditions (FGF, ACTIVIN, BMP, WNT).

The spatial organization of hESC-derived cell fates during micropattern differentiation is mediated by a combination of receptor occlusion at the colony center and loss of secreted inhibitors from the colony edge [17, 33]. To test for the involvement of receptors

in the micropattern organization of mouse cell fates, we substituted exogenous WNT3A with a GSK3 inhibitor, CHIR99021 (CHIR), which circumvents the receptor to activate downstream Wnt pathway components (Figure S6A,B). Under these conditions, CDX2, GATA6 and SOX17 were expressed at the outer colony edge indicating that mesoderm differentiation was unaffected (Figure S6C,D). However, the BRACHYURY expression territory was expanded throughout the center of the colony (Figure S6E,F), recapitulating the expansion of BRACHYURY expression in CHIR-cultured embryos [34]. These data suggest that the transmission of signals or activity of inhibitors through receptors is key for setting up distinct cell fate domains within the flat-disc micropatterned colonies.

The absence of BMP allows DE and AxM specification

BMP4, WNT3A, ACTIVIN A and FGF2 directed micropatterned EpiLC differentiation towards posterior embryonic fates (posterior Epi, PS, embryonic and extraembryonic mesoderm), but not cell fates arising from the anterior PS (DE and AxM). Since the anterior PS is devoid of BMP signaling activity, we reasoned that removing BMP would mirror the anterior signaling niche and create an environment permissive to specify anterior DE and AxM fates, but not posterior fates. EpiLCs were plated onto micropatterns and differentiated for 72 hours with FGF2, ACTIVIN A, BMP4 and WNT3A (control), FGF2, ACTIVIN A and WNT3A (no BMP) or FGF2, ACTIVIN A and WNT3A with a small molecule inhibitor of BMP signaling, DMH1 [35] (BMPi) (Figure 6A). In control conditions, nuclear pSMAD1/5/8 was observed in cells at the perimeter of colonies alongside CDX2 and SOX17, followed by a region of BRACHYURY expression and a central region of cells expressing SOX2 and low levels of FOXA2 (Figure 6B-E). In medium conditions lacking BMP, the absence of BMP signaling activity was confirmed by lack of nuclear pSMAD1/5/8 (Figure 6B,E). Under these conditions, the domain of extraembryonic mesoderm, marked by CDX2, was lost (Figure 6C,E). Instead we observed elevated SOX17 in outer cells, which was now robustly coexpressed with FOXA2 (Figure 6D-F) representing DE (Figure S3D). We also observed a separate population of outer cells that coexpressed FOXA2 and BRACHYURY (Figure 6F,G), likely representing cells within the anterior PS (Figure 3A, yellow arrowheads), node or AxM (Figure S3E).

In conditions lacking BMP signaling activity, SOX2 was elevated (Figure 6B-E) suggesting that central cells may represent a more anterior Epi state. To determine

whether this was the case, we utilized a *Sox1*^{GFP} fluorescent reporter embryonic stem cell (ESC) line. *Sox1*^{GFP} marks early neurectoderm specification from cells within the anterior Epi [36]. We differentiated *Sox1*^{GFP} ESCs as described in Figure 6A, either in the presence or absence of BMP. As with other cell lines analyzed, in the presence of BMP cells within the outer domain of *Sox1*^{GFP} EpiLC micropatterned colonies expressed CDX2 and, in the absence of BMP, they expressed FOXA2 (Figure S7A,C). While *Sox1*^{GFP} was largely absent from micropatterned colonies in the presence of BMP, consistent with the colony center representing posterior Epi, in the absence of BMP *Sox1*^{GFP} was expressed at high levels throughout the colony center (Figure S7B,C). Furthermore, in the absence of BMP, OTX2 levels were also elevated, with the highest expression observed at the colony periphery within the domain corresponding to DE and AxM fates (Figure S7B,C). This agrees with the later embryonic expression of OTX2 within the DE (Figure S7D). Hence, removing BMP from the (FGF, ACTIVIN and WNT) growth factor cocktail promoted differentiation towards anterior Epi and anterior PS derivatives (Figure 6H, Table 1).

Epiblast stem cells form definitive endoderm in the presence and absence of BMP

Epiblast stem cells (EpiSCs), maintained under standard FGF and ACTIVIN (F/A) culture conditions [37, 38], correlate to later embryonic stages than EpiLCs. EpiSCs express markers associated with the gastrulating Epi, including those of cell fates emerging from the anterior PS [39]. We therefore asked whether EpiSCs demonstrate a more restricted developmental potential in the context of the micropatterns. EpiSC9 cells [40] were cultured in defined medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A and then differentiated on the micropatterns as described for EpiLCs, for 72 hours in the presence or absence of BMP4 (Figure 7A). In the presence of BMP, EpiLCs formed an outer extraembryonic mesoderm population, marked by CDX2 expression (Figure 3F, 6C, S7A). Under the same conditions, EpiSCs generated few CDX2 positive cells (Figure 7B). Instead, cells within differentiated EpiSC colonies expressed SOX2, representing an Epi state, and GATA6, SOX17 and FOXA2, corresponding to a DE fate or BRACHYURY (Figure 7B,C). BRACHYURY was also expressed at the outer edge of colonies in the same domain as FOXA2 and likely represented AxM (Figure 7B). While BRACHYURY tended to be restricted to the outer colony edge, GATA6, SOX17 and FOXA2 were expressed in patches throughout the colony (Figure 7B,C),. This suggested that DE differentiation occurred without spatial organization. In the absence of BMP, GATA6,

FOXA2 and SOX17 were expressed more uniformly throughout colonies (Figure 7B,C). Therefore, in the presence of BMP, EpiLCs specify posterior PS-derived embryonic and extramebryonic mesoderm fates and in the absence of BMP give rise to anterior PS-derived DE and AxM (Figure 6). Conversely, under either of these conditions EpiSCs predominantly give rise to DE (Figure 7).

We utilizing published microarray data from Hayashi et al [8], comparing the pre-gastrulation E5.75 Epi, EpiLCs and EpiSCs, to ask whether the functional bias of EpiSCs towards anterior PS fates may be due to elevated expression of genes encoding BMP inhibitors, which would render EpiSCs unresponsive to the posteriorization BMP signal. The E5.75 Epi, EpiLCs and EpiSCs all expressed high levels of the pluripotency marker *Oct4*, but EpiSCs also expressed high levels of the anterior markers *Foxa2* and *Sox17* (Figure 7D). Furthermore, we noted a marked increase in the expression of the BMP pathway inhibitor *Chordin*, as well as elevated expression of the Wnt pathway inhibitor *Dkk1*, and the Nodal pathway inhibitor *Lefty2* within EpiSCs (Figure 7D). Together these data suggest that EpiSCs are functionally restricted in their developmental potential, which may result from the elevated expression of signaling pathway inhibitors.

Discussion

We have developed a robust, quantitative and scalable micropattern protocol to differentiate mouse EpiLCs, the *in vitro* counterpart of the pre-gastrulation pluripotent Epi of the embryo [8]. In response to FGF, ACTIVIN (NODAL), BMP and WNT, the critical gastrulation-inducing signals acting in the embryo [3], EpiLCs on circular micropatterns underwent reproducible spatially coordinated cell fate specification comparable to *in vivo* gastrulation. Detailed marker analysis of gastrulating mouse embryos (which allow the mapping not only of expression but also of position) and micropatterns allowed us to map the *in vitro* differentiation in developmental time and space. In the absence of the spatial and temporal information of the embryo, we required a cohort of 15 markers (SOX2, OCT4, NANOG, SOX1, OTX2, BRACHYURY, CDX2, GATA6, SOX17, FOXA2, FOXF1, E-CADHERIN, N-CADHERIN, SNAIL, pSMAD1/5/8) to distinguish between cell fates such as anterior vs. posterior Epi, or extraembryonic mesoderm vs. trophectoderm and DE as these cell types express many common factors. This emphasizes the

480 necessity of expression signatures rather than individual markers to accurately assign
481 cell fates *in vitro*.
482
483 During 72 hours of differentiation, micropatterned colonies advanced from an E5.5
484 pluripotent Epi-like state to comprising an array of populations present in the embryo just
485 prior to E7.5 (Figure 8A). Hence, under these conditions, *in vitro* differentiation was
486 slower than *in vivo* development. Conceivably, further manipulation of the timing, levels
487 and combination of signaling factors provided to EpiLCs, as well as the extracellular
488 matrix and stiffness of the substrate on which cells are maintained, could alter the rate of
489 differentiation and support the specification of cell fates emerging at later gastrulation
490 stages.
491
492 At 72 hours, micropatterned colonies could be divided into 3 spatially distinct domains
493 (central, mid and outer) (Figure 8A). Cells within the colony center showed minimal BMP
494 signaling and expressed posterior Epi markers. PS markers were initially expressed at
495 the periphery, but over time were observed more centrally. This was accompanied by an
496 EMT and the emergence of outer mesenchymal cells, plausibly emanating from the PS-
497 like region. The outer domain displayed elevated BMP activity and contained multiple
498 populations including allantois and yolk sac extraembryonic mesoderm and early
499 embryonic mesoderm.
500
501 In contrast to most gastrulating mammalian embryos, which exhibit a flat-disc geometry,
502 rodents including the mouse are cup-shaped. A conceptual flattening of the cell fate
503 arrangement within the mouse embryo [41], could not fully recapitulate the organization
504 of cell types observed within the flat-disc micropatterns. Therefore the most evident
505 correspondence between embryonic and micropattern cell fates was signaling history.
506 However, while all cells within the outer micropattern domain experienced high levels of
507 BMP signaling, both embryonic and extraembryonic mesoderm fates were specified. It is
508 therefore unclear whether additional morphogens distinguish embryonic and
509 extraembryonic mesoderm, or if factors such as three dimensional growth, migration and
510 extracellular matrix dictate fate. Extension of the micropattern system to different
511 geometries, morphogens and inhibitors should resolve these questions.
512

The spatial organization of cell identities within the micropatterns emerged even though signals were provided uniformly. Thus, epithelial cell cultures can self-organize and the signaling history of a cell depends on its local environment, as well as the external medium. When WNT3 was replaced with CHIR [42] a small molecule that activates the WNT signaling pathway intracellularly, bypassing the receptors and secreted inhibitors acting at the cell surface, the PS region expanded into the colony center. Hence, as with human micropattern differentiation [7], endogenously produced inhibitors likely exclude signals from the colony center to define the inner domains.

In vivo, asymmetries within the Epi, leading to anterior-posterior axis establishment, arise through localized signaling from the adjacent extraembryonic tissues, notably the anterior visceral endoderm (AVE) [43]. As micropatterned cultures do not contain extraembryonic cell types (namely primitive endoderm or trophectoderm derivatives), they instead formed radially symmetric cell fate domains reminiscent of mutant embryos with defects in AVE specification or positioning [44-47]. The cohort of signaling factors and secreted inhibitors expressed by adjacent tissues within the embryo make development robust yet difficult to quantify. For example, the extraembryonic VE is a source of inhibitors including CERBERUS and LEFTY1 on the anterior [43], and WNT3 on the posterior [48] side of the embryo, whereas the Epi and its derivatives express WNT3, LEFTY2 and DKK1 [49, 50]. As our *in vitro* system patterns in the absence of extraembryonic cell types, it allows us to decipher Epi-intrinsic patterning mechanisms.

The micropattern system can be used to extend findings in animal models to a defined, serum-free environment where signaling modulation can be unambiguously interpreted to reveal how timing and levels of signaling influence cell fate. As a first step in this direction, we analyzed the effect of manipulating the BMP pathway. Embryos with disrupted BMP signaling do not form a morphological PS and predominantly arrest at early gastrulation [51-54], obscuring the assessment of a role for BMP in later mesoderm and endoderm specification. When we applied FGF, ACTIVIN and WNT alone (in the absence of BMP) to micropatterned colonies, DE and AxM were specified rather than extraembryonic mesoderm (Figure 7B). These data reveal that BMP is not significantly induced by WNT and its absence does not perturb anterior cell fate specification. In the future, the micropattern assay could be used as a robust, efficient and scalable way to

546 survey signaling conditions and systematically screen interactions between individual
547 genes and pathways.
548
549 Due to a paucity of data on gastrulating human embryos, cell fates arising during hESC
550 micropattern differentiation could only been predicted [7, 17]. The mouse micropattern
551 differentiation provides the essential missing link between *in vitro* gastrulation models in
552 mouse and human, and *in vivo* mouse development. The identification of an
553 extraembryonic mesoderm population within mouse, but not human micropatterns
554 prompts an analysis of the human system with equivalent marker combinations under
555 comparable serum-free medium conditions containing both BMP4 and WNT3A to
556 determine whether populations such as extraembryonic mesoderm can be generated, or
557 whether in human, extraembryonic mesoderm, as has been shown for the amnion [55],
558 does not arise from the Epi at gastrulation. Human and mouse embryos are of a similar
559 size (Figure 8C,D), undergo micropattern differentiation within equivalent diameter
560 colonies and specify cell fates as a function of distance from the colony edge,
561 suggesting that these species use common mechanisms to regulate specification and
562 patterning. While hESCs are more similar to mouse EpiSCs than mouse ESCs [37, 38],
563 it is not clear where they lie on the pluripotency continuum [9] in relation to other
564 pluripotent states. As we observed spatially organized germ layer specification when
565 starting with EpiLCs but not EpiSCs, in this context hESCs show a stronger functional
566 correlation to mouse EpiLCs than EpiSCs. The further correlation of mouse and human
567 *in vitro* micropattern data, in the context of different pluripotent starting states, and with
568 *in vivo* mouse embryos should yield insights into the conserved and divergent
569 mechanisms regulating fundamental aspects of early mammalian development.
570
571

Materials and Methods

Gene and gene product nomenclature

Genes and gene products are referred to using guidelines set by the International Committee on Standardized Genetic Nomenclature for Mice - gene symbols are italicized with only the first letter upper case while proteins are all upper case and no italics (http://www.informatics.jax.org/mgihome/nomen/gene.shtml). Cytokines are referred to as proteins (all upper case) while the corresponding signaling pathways are referred to in lower case, non-italic.

Cell culture

ESC lines used for this study include E14 (129/Ola background) [56], R1 (129/Sv background) [57], *Sox17*$^{GFP/+}$ (R1, 129/Sv background) [58], *T*$^{GFP/+}$ (E14.1, 129/Ola background, also known as GFP-Bry) [59] and *Sox1*^{GFP} (E14Tg2a background, also known as 46C) [36]. ES cells were routinely cultured on 0.1% gelatin coated tissue culture grade plates (Falcon) in serum and LIF medium as previously described [60]. Serum and LIF medium was comprised of Dulbecco's modified Eagle's medium (DMEM) (Gibco) containing 0.1 mM non-essential amino-acids (NEAA), 2mM glutamine and 1mM sodium pyruvate, 100 U/ml Penicillin, 100 µg/ml Streptomycin (all from Life Technologies), 0.1 mM 2-mercaptoethanol (Sigma), and 10% Fetal Calf Serum (FCS, F2442, Sigma) together with 1000U/ml LIF. They were passaged every 2 days upon reaching approximately 80% confluency by washing with phosphate buffered saline (PBS) before adding 0.05% Trypsin (Life Technologies) for 3 minutes at 37°C and dissociating into a single cell suspension by pipetting. Trypsin activity was then neutralized with serum-containing medium. Cells were collected at 1300 rpm for 3 minutes and 1/5 of cells transferred to a new plate.

For this study, the EpiSC9 epiblast stem cell line was used (129SvEv x ICR background) [40]. EpiSCs were cultured under standard conditions as previously described [38], in defined, serum-free N2B27 medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A. EpiSCs were passaged upon reaching approximately 80% confluency by washing with PBS then replacing with Accutase (Sigma) and scraping cells from the plate. Cells were pipetted gently to avoid single cell dissociation. Cells were collected at

1300 rpm for 3 minutes and 1/5 of cells transferred to a new plate. ESCs and EpiSCs were maintained at 37°C at 5% CO_2 and 90% humidity.

EpiLC conversion

Prior to plating on micropatterns, ES cells were converted to a transient EpiLC state as previously described [8]. First, 10 cm plates were coated with 16 µg/ml of Fibronectin for 1 hour at room temperature followed by two washes with PBS. ES cells were collected by trypsinization (see above), counted and 1.6 x 10^6 cells plated onto the Fibronectin-coated plates for 48 hours in EpiLC medium, N2B27 medium containing 20 ng/ml ACTIVIN A and 12 ng/ml FGF2. Medium was changed daily.

Micropattern differentiation

To coat micropatterned surfaces, a solution was prepared of 20 µg/ml Laminin (L20202, Sigma) in PBS without calcium and magnesium (PBS-/-). A 15 cm tissue culture plate was lined with Parafilm (Pechiney Plastic Packaging) and 700 µl drops were made onto the Parafilm surface. Micropatterned chips (Arena A, CYTOO) were washed once with PBS-/- and then inverted on top of the drops followed by incubation at 37°C for 2 hours. Micropatterns were then washed 5 times with 5 ml of PBS-/-. EpiLCs were collected by trypsinization and a single cell suspension generated. Cells were counted and 2 x 10^6 EpiLCs were evenly plated onto micropatterns within 6-well plates (Falcon) in EpiLC medium. Medium was supplemented with a small molecule inhibitor of Rho-associated kinase (ROCK, 10 µM Y-27632) for the first 2 hours after plating, to reduce apoptosis [61, 62]. Plates were maintained in the tissue culture hood for 30 minutes after plating to allow time for cells to evenly adhere to the micropatterns before moving to the incubator. After 2 hours, medium containing ROCKi was exchanged for N2B27 medium containing 12 ng/ml FGF2, 20 ng/ml ACTIVIN A, 50 ng/ml BMP4 (Peprotech) and 200 ng/ml WNT3A (R&D). Cells were maintained for up to 72 hours in this state. To determine the effect of BMP signaling on the differentiation, cells were differentiated as described above for 72 hours with FGF, ACTIVIN, BMP and WNT (control condition) or with FGF2, ACTIVIN A and WNT3A without BMP4 (no BMP) or FGF, ACTIVIN and WNT with the addition of 2 µM DMH1 (Sigma) (BMPi).

Immmunostaining, imaging and quantification of cells

Prior to immunostaining, cells were either grown on micropatterns or in 1 µ-slide 8 well IbiTreat plates (Ibidi). Cells were washed twice with PBS before being fixed with 4% paraformaldehyde (PFA) (Electron Microscopy Sciences) at room temperature for 15 minutes. Cells were then washed a further two times with PBS followed by permeabilization with PBS containing 0.1% Triton-X (Sigma) (PBS-T) for 10 mins at room temperature. Cells were then blocked in PBS-T with 1% bovine serum albumin (BSA, Sigma) and 3% donkey serum (Sigma) for 30 minutes at room temperature. Primary antibodies were added overnight at 4°C, diluted to the appropriate concentration in PBS-T with 1% BSA. Details of primary antibodies are supplied in Supplementary Table 1. The following day, cells were washed three times for 15 minutes with PBS followed by incubation with secondary antibodies (1:500, Alexa Fluors, Life Technologies, Dylight, Jackson ImmunoResearch) in PBS-T with 1% BSA for 2 hours at room temperature. Finally, cells were washed three times for 15 minutes with PBS with the final wash containing 5 µg/ml Hoechst (Life Technologies). Cells grown on micropatterns were then mounted onto glass slides (Fisher Scientific) with Fluoromount-G (Southern Biotech). Cells were imaged using a LSM880 confocal (Zeiss). Brightfield-only images were acquired using a Zeiss Axio Vert.A1.

Quantitative analysis of micropattern differentiation

For micropattern image analysis and quantification, tiled Z-stack images of individual colonies were collected using a LSM880 confocal microscope (Zeiss) at 512 x 512 format using a 20x objective. The background signal was subtracted using ImageJ software and each channel saved as a separate tiff file. Tiff files containing the Hoechst nuclear staining of each colony were classified into regions containing nuclei and those that did not using Ilastik [63], an interactive image classification software. Using this information, a probability mask was generated and analysis carried out using custom software written in Python. All analysis was carried out on entire Z-stacks of multiple colonies and an average of results across colonies displayed.

Segmentation of individual cells within images of colonies proved problematic due to the large number and high density of cells. For these reasons, manual correction of segmentation, as routinely used in smaller systems [64], was not feasible. Therefore quantification of immunostaining fluorescence intensity across the radii of colonies as well as coexpression analysis was completed on a voxel basis to eliminate segmentation

672 artifacts. To generate plots of radial immunostaining fluorescence intensity, each voxel
673 within a colony was assigned a distance from the colony center. The fluorescence
674 intensity for each marker was measured per voxel and then the average fluorescence
675 intensity of voxels at a particular radial position (binned into discrete radial bands) was
676 calculated for each colony. The average radial fluorescence intensity across multiple
677 colonies was then calculated. To display the expression of multiple different markers
678 across the radii of colonies on the same scale, the relative level of each marker was
679 quantified by normalizing to the highest level of expression (shown as 100) either across
680 a time-course or within an individual time-point. Spatial patterning across multiple
681 colonies was also demonstrated by generating average colony images for individual
682 markers where each segmented cell was represented as a dot whose color indicates its
683 fluorescence in the specified channel.
684
685 Coexpression analysis was carried out on a voxel level, i.e. the fluorescence level of
686 each marker within a single voxel was calculated and plotted. For genes that were not
687 expressed, or only expressed at low levels, at the start of the differentiation, gates could
688 be drawn based on the fluorescence at 0 hours and used to quantify the percentage of
689 total voxels expressing a particular marker at later time points.
690
691 **Nuclear density measurements in micropatterns and embryos**
692 The number of nuclei per 100 μm was quantified for 0 hours and 24 hours of
693 micropattern differentiation utilizing the colony side view (z-axis) from confocal images
694 acquired using a 40x objective at 0.5 μm interval steps. The number of nuclei was
695 quantified across the entire width of the colony at 10 distinct positions and the average
696 number of nuclei per 100 μm distance were calculated. For E5.5 embryos, the number of
697 nuclei per 100 μm was quantified on sagittal confocal optical sections based on the
698 number of nuclei within a sagittal optical section of the epiblast and the distance around
699 the epiblast within the same section, manually measured using ImageJ software. For
700 E6.5 the same was done using confocal images of transverse cryosections. Only cells
701 within the epiblast were counted. Five embryos at E5.5 and five at E6.5 were analyzed in
702 this manner.
703
704 Inter-nuclear distance was manually measured using ImageJ software. A line was drawn
705 from the center of one nuclei to the center of the adjacent nuclei. For micropattern

differentiation, 150 measurements were made per time point (0 hours and 24 hours). For *in vivo* data, 5 different embryos were measured at each time point (E5.5 and E6.5). At E5.5, 125 measurements were made and at E6.5, 189 measurements were made.

Mice

All mice used in this study were of a wild-type CD1 background. Mice were maintained in accordance with the guidelines of the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee (IACUC). Mice were housed under a 12-hour light/dark cycle in a specific pathogen free room in the designated facilities of MSKCC. Natural matings of CD1 males and 4-6 weeks old virgin CD1 females were set up in the evening and mice checked for copulation plugs the next morning. The date of vaginal plug was considered as E0.5.

Immunostaining and imaging of embryos

To analyze the expression of markers within post-implantation embryos, the uterus of pregnant mice was dissected and deciduae removed. Embryos were dissected from the deciduae and the parietal endoderm removed. Embryos were washed twice in PBS and fixed in 4% PFA for 30 minutes at room temperature. Embryos were permeabilized in PBS with 0.5% Triton-X for 30 minutes followed by blocking overnight in PBS-T with 5% horse serum (Sigma). Primary antibodies were added the following day, diluted in blocking buffer at the appropriate concentration (details can be found in Supplementary Table 1) and incubated overnight at 4°C. The next day, embryos were washed 3 times for 15 minutes in PBS-T and then blocked for a minimum of 2 hours. Embryos were then incubated with the secondary antibodies diluted in blocking buffer overnight at 4°C. Alexa Fluor® (Thermo Fisher Scientific) secondary antibodies were diluted 1:500. The following day, embryos were washed 3 times for 15 minutes in PBS-T with the last wash containing 5 µg/ml Hoechst. Embryos were imaged in PBS-T in glass bottom dishes (MatTek) using an LSM880 confocal (Zeiss).

Cryosectioning and quantitative embryo measurements

For cryosectioning, embryos were incubated in a 30% sucrose solution until they sank to the bottom of the vial. Embryos were then transferred to optimal cutting temperature compound (OCT, Tissue-Tek) overnight. The following day, embryos were transferred to mounting molds (Fisher Scientific) containing OCT and appropriately oriented to give

sagittal or transverse sections. Embryo-containing molds were carefully transferred to dry ice until frozen and then temporarily to -80°C until cryosectioning. Cryosections of 10 µm were cut using a Leica CM3050S and imaged using a confocal microscope as described above.

To quantify immunostaining within gastrulating mouse embryos, transverse cryosections were imaged by confocal microscopy. For quantification of the relative levels of SOX2 and NANOG within different regions of the Epi, the anterior and posterior regions were manually selected using ImageJ software and immunostaining fluorescence levels in arbitrary units. Five cryosections were quantified per embryo and the levels normalized to the fluorescence levels of the Hoechst nuclear stain. At E6.5, 3 embryos were quantified, while as E7.5, 2 embryos were quantified. For quantification of the levels of pSMAD1/5/8 within different cell types within the gastrulating mouse embryo, transverse cryosections through the PS of E6.5 embryos were selected. Individual cells within the Epi, PS and mesodermal wings were manually selected using ImageJ software and fluorescence levels in arbitrary units. Data was normalized to the fluorescence level of the Hoechst nuclear stain. The PS was defined as BRACHYURY expressing cells within the posterior Epi while the mesodermal wings were identified as cells that had left the Epi epithelial layer and were migrating between the Epi and VE. Three cryosections were quantified per embryo and 3 embryos were analyzed.

The diameter of embryos at different developmental stages was measured on acquired images using ImageJ software. Measurements were made along the anterior-posterior axis of transverse cryosections of embryos. Multiple cryosections at the widest region of the embryo were utilized and multiple embryos per developmental stage.

Declarations

Acknowledgements

We thank Paul Tesar for EpiSC lines; Kathryn Anderson, Ali Brivanlou and members of the Hadjantonakis and Brivanlou-Siggia labs for critical discussions and comments on the manuscript. SMM is supported by a Wellcome Trust Sir Henry Wellcome postdoctoral fellowship. Work in the Hadjantonakis lab was supported by grants from NYSTEM (C029568) and the NIH (R01DK084391 and P30CA008748). Work in the Siggia lab was supported by the NSF (PHY1502151) and NIH (R01HD080699 – jointly held with Ali H. Brivanlou).

Competing interests

The authors have no competing interests to declare.

Author contributions

SMM, EDS, AKH conceived of and designed experiments. SMM carried out all experiments. JM developed image analysis code. SMM and AKH wrote the manuscript with input from all co-authors. All authors read and approved the final manuscript.

References

1. Hashimoto, K. and N. Nakatsuji, *Formation of the Primitive Streak and Mesoderm Cells in Mouse Embryos - Detailed Scanning Electron Microscopical Study.* Development Growth & Differentiation, 1989. **31**(3): p. 209-218.
2. Tam, P.P. and D.A. Loebel, *Gene function in mouse embryogenesis: get set for gastrulation.* Nat Rev Genet, 2007. **8**(5): p. 368-81.
3. Arnold, S.J. and E.J. Robertson, *Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo.* Nature Reviews Molecular Cell Biology, 2009. **10**(2): p. 91-103.
4. Kinder, S.J., et al., *The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo.* Development, 1999. **126**(21): p. 4691-701.
5. Lawson, K.A., *Fate mapping the mouse embryo.* Int J Dev Biol, 1999. **43**(7): p. 773-5.
6. Keller, G., *Embryonic stem cell differentiation: emergence of a new era in biology and medicine.* Genes Dev, 2005. **19**(10): p. 1129-55.
7. Warmflash, A., et al., *A method to recapitulate early embryonic spatial patterning in human embryonic stem cells.* Nat Methods, 2014. **11**(8): p. 847-54.
8. Hayashi, K., et al., *Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells.* Cell, 2011. **146**(4): p. 519-32.
9. Morgani, S., J. Nichols, and A.K. Hadjantonakis, *The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states.* BMC Dev Biol, 2017. **17**(1): p. 7.
10. Smith, A., *Formative pluripotency: the executive phase in a developmental continuum.* Development, 2017. **144**(3): p. 365-373.
11. Kalkan, T., et al., *Tracking the embryonic stem cell transition from ground state pluripotency.* Development, 2017.
12. Kalkan, T. and A. Smith, *Mapping the route from naive pluripotency to lineage specification.* Philos Trans R Soc Lond B Biol Sci, 2014. **369**(1657).
13. Viotti, M., S. Nowotschin, and A.K. Hadjantonakis, *SOX17 links gut endoderm morphogenesis and germ layer segregation.* Nature Cell Biology, 2014. **16**(12): p. 1146-U54.
14. Bedzhov, I. and M. Zernicka-Goetz, *Self-Organizing Properties of Mouse Pluripotent Cells Initiate Morphogenesis upon Implantation.* Cell, 2014. **156**(5): p. 1032-1044.
15. Zhu, Q., et al., *The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways.* Elife, 2014. **3**.
16. Kojima, Y., O.H. Tam, and P.P. Tam, *Timing of developmental events in the early mouse embryo.* Semin Cell Dev Biol, 2014. **34**: p. 65-75.
17. Etoc, F., et al., *A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization.* Developmental Cell, 2016. **39**(3): p. 302-315.
18. Huang, S.M., et al., *Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling.* Nature, 2009. **461**(7264): p. 614-20.
19. Robertson, E.J., *Dose-dependent Nodal/Smad signals pattern the early mouse embryo.* Semin Cell Dev Biol, 2014. **32**: p. 73-9.

20. Downs, K.M., et al., *The Allantoic Core Domain: New Insights Into Development of the Murine Allantois and Its Relation to the Primitive Streak.* Developmental Dynamics, 2009. **238**(3): p. 532-553.
21. Auman, H.J., et al., *Transcription factor AP-2 gamma is essential in the extra-embryonic lineages for early postimplantation development.* Development, 2002. **129**(11): p. 2733-2747.
22. Lewis, S.L. and P.P.L. Tam, *Definitive endoderm of the mouse embryo: Formation, cell fates, and morphogenetic function.* Developmental Dynamics, 2006. **235**(9): p. 2315-2329.
23. Ferrer-Vaquer, A., M. Viotti, and A.K. Hadjantonakis, *Transitions between epithelial and mesenchymal states and the morphogenesis of the early mouse embryo.* Cell Adhesion & Migration, 2010. **4**(3): p. 447-457.
24. Carver, E.A., et al., *The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition.* Mol Cell Biol, 2001. **21**(23): p. 8184-8.
25. Kang, Y.B. and J. Massague, *Epithelial-mesenchymal transitions: Twist in development and metastasis.* Cell, 2004. **118**(3): p. 277-279.
26. Ciruna, B. and J. Rossant, *FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak.* Dev Cell, 2001. **1**(1): p. 37-49.
27. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-72.
28. Dunn, N.R., et al., *Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo.* Development, 2004. **131**(8): p. 1717-28.
29. Nomura, M. and E. Li, *Smad2 role in mesoderm formation, left-right patterning and craniofacial development.* Nature, 1998. **393**(6687): p. 786-90.
30. Weinstein, M., et al., *Failure of egg cylinder elongation and mesoderm induction in mouse embryos lacking the tumor suppressor smad2.* Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9378-83.
31. Conlon, F.L., et al., *A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse.* Development, 1994. **120**(7): p. 1919-28.
32. Lawson, K.A., et al., *Bmp4 is required for the generation of primordial germ cells in the mouse embryo.* Genes & Development, 1999. **13**(4): p. 424-436.
33. Tewary, M., et al., *A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning.* Development, 2017.
34. Sumi, T., et al., *Epiblast ground state is controlled by canonical Wnt/beta-catenin signaling in the postimplantation mouse embryo and epiblast stem cells.* PLoS One, 2013. **8**(5): p. e63378.
35. Ao, A., et al., *DMH1, a novel BMP small molecule inhibitor, increases cardiomyocyte progenitors and promotes cardiac differentiation in mouse embryonic stem cells.* PLoS One, 2012. **7**(7): p. e41627.
36. Ying, Q.L., et al., *Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture.* Nature Biotechnology, 2003. **21**(2): p. 183-186.
37. Tesar, P.J., et al., *New cell lines from mouse epiblast share defining features with human embryonic stem cells.* Nature, 2007. **448**(7150): p. 196-9.
38. Brons, I.G., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-5.

39. Kojima, Y., et al., *The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak.* Cell Stem Cell, 2014. **14**(1): p. 107-20.
40. Najm, F.J., et al., *Rapid and robust generation of functional oligodendrocyte progenitor cells from epiblast stem cells.* Nat Methods, 2011. **8**(11): p. 957-62.
41. Behringer, R.R., et al., *A flattened mouse embryo: leveling the playing field.* Genesis, 2000. **28**(1): p. 23-30.
42. Wray, J., et al., *Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation.* Nat Cell Biol, 2011. **13**(7): p. 838-45.
43. Stower, M.J. and S. Srinivas, *Heading forwards: anterior visceral endoderm migration in patterning the mouse embryo.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2014. **369**(1657).
44. Nowotschin, S., et al., *The T-box transcription factor Eomesodermin is essential for AVE induction in the mouse embryo.* Genes & Development, 2013. **27**(9): p. 997-1002.
45. Migeotte, I., et al., *Rac1-Dependent Collective Cell Migration Is Required for Specification of the Anterior-Posterior Body Axis of the Mouse.* Plos Biology, 2010. **8**(8).
46. Kimura-Yoshida, C., et al., *Canonical Wnt signaling and its antagonist regulate anterior-posterior axis polarization by guiding cell migration in mouse visceral endoderm.* Developmental Cell, 2005. **9**(5): p. 639-650.
47. Ding, J.X., et al., *Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo.* Nature, 1998. **395**(6703): p. 702-707.
48. Rivera-Perez, J.A. and T. Magnuson, *Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3.* Developmental Biology, 2005. **288**(2): p. 363-371.
49. Meno, C., et al., *Mouse lefty2 and zebrafish antivin are feedback inhibitors of nodal signaling during vertebrate gastrulation.* Molecular Cell, 1999. **4**(3): p. 287-298.
50. Peng, G.D., et al., *Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo.* Developmental Cell, 2016. **36**(6): p. 681-697.
51. Gu, Z.Y., et al., *The type I serine threonine kinase receptor ActRIA (ALK2) is required for gastrulation of the mouse embryo.* Development, 1999. **126**(11): p. 2551-2561.
52. Mishina, Y., et al., *Multiple roles for activin-like kinase-2 signaling during mouse embryogenesis.* Developmental Biology, 1999. **213**(2): p. 314-326.
53. Mishina, Y., et al., *Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis.* Genes Dev, 1995. **9**(24): p. 3027-37.
54. Winnier, G., et al., *Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse.* Genes Dev, 1995. **9**(17): p. 2105-16.
55. Dobreva, M.P., et al., *On the origin of amniotic stem cells: of mice and men.* Int J Dev Biol, 2010. **54**(5): p. 761-77.
56. Hooper, M., et al., *HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells.* Nature, 1987. **326**(6110): p. 292-5.
57. Nagy, A., et al., *Derivation of completely cell culture-derived mice from early-passage embryonic stem cells.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8424-8.

58. Kim, I., T.L. Saunders, and S.J. Morrison, *Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells.* Cell, 2007. **130**(3): p. 470-83.
59. Fehling, H.J., et al., *Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation.* Development, 2003. **130**(17): p. 4217-27.
60. Morgani, S.M., et al., *Totipotent embryonic stem cells arise in ground-state culture conditions.* Cell Rep, 2013. **3**(6): p. 1945-57.
61. Ohgushi, M., et al., *Molecular Pathway and Cell State Responsible for Dissociation-Induced Apoptosis in Human Pluripotent Stem Cells.* Cell Stem Cell, 2010. **7**(2): p. 225-239.
62. Watanabe, K., et al., *A ROCK inhibitor permits survival of dissociated human embryonic stem cells.* Nat Biotechnol, 2007. **25**(6): p. 681-6.
63. Sommer, C., et al., *Ilastik: Interactive Learning and Segmentation Toolkit.* 2011 8th Ieee International Symposium on Biomedical Imaging: From Nano to Macro, 2011: p. 230-233.
64. Saiz, N., et al., *Quantitative Analysis of Protein Expression to Study Lineage Specification in Mouse Preimplantation Embryos.* J Vis Exp, 2016(108): p. 53654.

Figure Legends

Figure 1. EpiLCs represent a pluripotent state correlating to the pre-streak epiblast of the embryo. A. Development of the mouse pluripotent epiblast (Epi) from embryonic day (E) 4.5 to 7.5 and correlating *in vitro* pluripotent states. ESCs, embryonic stem cells; EpiLCs, epiblast-like cells; EpiSCs, epiblast stem cells; TE/ExE, trophectoderm/extraembryonic ectoderm; PrE/VE, primitive endoderm/visceral endoderm; DE, definitive endoderm; A, anterior; P, posterior; Pr, proximal; D, distal. **B.** Sagittal sections of immunostained E5.5-6.0 embryos. Yellow dashed line demarcates Epi. Scale bars, 25 μm. Non-nuclear anti-BRACHYURY/CDX2/OCT6 VE fluorescence represents non-specific binding. **C.** ESCs were converted to EpiLCs on Fibronectin in N2B27 with FGF2 and ACTIVIN A (F/A) and knockout serum replacement (KOSR) for 48 hours. EpiLCs were plated onto Laminin-coated micropatterns overnight and analyzed the following day (0 hours). **D.** Maximum intensity projections of immunostained 1000 μm diameter EpiLC micropatterned colonies. Scale bars, 100 μm. **E.** Confocal image showing a z-axis (side profile) region of an immunostained EpiLC micropatterned colony. **F.** Quantification of immunostaining voxel fluorescence intensity from center (0) to edge (500). Data represents average voxel intensity across multiple colonies. Dashed line represents average fluorescence of Hoechst nuclear stain. n = 6 NANOG/KLF4/SOX2/nuclei; n = 14 GATA6/OTX2/OCT6; n = 14 BRACHYURY/FOXA2. BRA, BRACHYURY; E-CAD, E-CADHERIN.

Figure 2. Micropatterned EpiLCs undergo spatially organized differentiation. A. Maximum intensity projection (MIP), sagittal and transverse sections of an embryonic day (E) 6.5 mouse embryo. Dashed line marks transverse plane. Non-nuclear anti-BRACHYURY/CDX2 VE fluorescence represents non-specific binding. ExM, extraembryonic mesoderm; PS, primitive streak; A, anterior; P, posterior; Pr, proximal; D, distal. Scale bars, 50 μm. **B.** Lookup table (LUT) of SOX2 marking anterior Epi (A-Epi) and NANOG marking posterior Epi (P-Epi). **C.** Quantification (5 sections /embryo /stage) of SOX2 and NANOG in manually selected (panel B) anterior (A) and posterior (P) Epi of E6.75 and E7.5 embryos, normalized to Hoechst fluorescence. Data depicts mean fluorescence intensity +/- S.D. N, number of embryos. No NANOG was observed in the A-Epi hence ~0.5 a.u. equates to background signal. **D.** BMP, Wnt, Nodal, FGF signaling initiates gastrulation at the P-Epi - extraembryonic ectoderm (ExE) boundary.

BMP4 produced by ExE stimulates *Wnt3* expression within proximal Epi. WNT3 produced by Epi and visceral endoderm (VE) triggers *Nodal* and *Fgf8* expression. NODAL promotes *Bmp4* expression in the ExE. The anterior VE (AVE) expresses Wnt and Nodal pathway antagonists, restricting signaling activity to P-Epi. **E.** EpiLCs were plated onto Laminin-coated micropatterns overnight (-24 hours) in N2B27 with F/A. The following day medium was changed to F/A, BMP4, WNT3A for 72 hours. Colonies were analyzed at 24-hour intervals. **F.** MIPs of immunostained 1000 μm diameter colonies. All subsequent data represents 1000 μm diameter colonies. Second panel shows high magnification of colony edge. Scale bars, 100 μm. BRA, BRACHYURY. **G.** Depiction of average positional marker expression across multiple colonies. Each dot represents a single cell. **H.** Quantification of voxel fluorescence intensity from colony center (0) to edge (500). Data represents average voxel intensity relative to maximum voxel intensity across time course/marker. For 0,24,48,72 hrs respectively, OCT4/NANOG n = 5,3,3,3, SOX2 n = 15,7,21,20, BRACHYURY n = 11,9,10,12, GATA6/SOX17/CDX2 n = 3,5,6,5. Markers grouped by spatial distribution within colonies. OTX2 and FOXF1 only analyzed at 72 hours.

Figure 3. Assignment of cell identities to micropattern-differentiated EpiLC populations. A,B,D. Confocal maximum intensity projections (MIP), sagittal optical sections and transverse cryosections of immunostained gastrulating embryos. Dashed lines mark transverse plane. Epi, Epiblast; PS, primitive streak; M1, Mesoderm1; M2, Mesoderm2; ACD, allantois core domain; AOM, allantois outer mesenchyme; ExE, extraembryonic ectoderm; Al, allantois; ExM, extraembryonic mesoderm; A, anterior; P, posterior; Pr, proximal; D, distal; R, right; L, left. Scale bars, 50 μm. **A.** Yellow arrowheads mark rare BRACHYURY and FOXA2 coexpressing cells within the anterior PS. **C.** LUT of immunostaining of BRACHYURY marking extraembryonic mesoderm allantois core domain (ACD) and CDX2 expressed highly in allantois outer mesenchyme (AOM) (upper panels) as well as GATA6 marking anteriorly migrated embryonic mesoderm (Mesoderm 1) and BRACHYURY marking embryonic mesoderm close to the PS (Mesoderm 2) (lower panels). **E,F,H.** MIPs of immunostained micropatterns. High magnification shows region at the colony edge. Yellow arrowhead in I marks GATA6/FOFX1 yolk sac mesoderm (YSM) cell. Scale bars, 100 μm. **G.** High magnification of colony edge. Outer domain represents a mixture of populations outlined with dashed lines. At 48 hours the ACD population coexpressed BRACHYURY and

CDX2, M1 expressed GATA6. By 72 hours, outer cells expressed CDX2 (AOM) or GATA6 (M1). BRACHYURY marked PS cells. **I.** Cell identities at 48 and 72 hours of *in vitro* differentiation, under conditions described in B and corresponding *in vivo* fates. Dashed lines mark transverse plane. ExM, extraembryonic mesoderm.

Figure 4. EMT is associated with micropatterned EpiLC differentiation. Data from colonies differentiated as in Figure 2E. **A,B.** Sagittal (A) and transverse sections (B) of late streak embryo. Dashed box marks high magnification region in lower panel. Dashed lines mark transverse planes in B. Non-nuclear anti-BRACHYURY VE fluorescence represents non-specific binding. A, anterior; P, posterior; Pr, proximal; D, distal; L, left; R, right; VE/DE, visceral endoderm/definitive endoderm; ExE, extraembryonic ectoderm; ExM, extraembryonic mesoderm; Epi, epiblast; Meso, mesoderm. Scale bars, 50 μm. **C.** Quantification of colony height from colony center (0) to edge (500) across multiple colonies, 3 independent experiments, 0 hours: n=11, 24 hours: n=15, 48 hours: n=17, 72 hours: n=18. **D.** Time-course showing brightfield images (upper panels) and MIPs of comparable immunostained colonies (lower panels). Scale bars, 100 μm. **E-G.** Images of z-axis profile from colony center (left) to edge (right). **G-I.** EpiLCs were plated onto micropatterns overnight with F/A. The following day medium was changed to F/A, BMP4, WNT3A (**E,F**) or medium blocking Activin/Nodal signaling - FGF2, BMP4, WNT3A, SB431542 (ACTIVINi, **G-I**). **H.** brightfield image of ACTIVINi colony. **I.** MIPs of immunostained ACTIVINi colonies at 72 hours differentiation. Scale bars, 100 μm. BRA, BRACHYURY; E-CAD, E-CADHERIN; N-CAD, N-CADHERIN.

Figure 5. BMP signaling in micropatterns and embryos correlates with embryonic and extraembryonic mesoderm fates. A. MIPs of immunostained colonies differentiated as in Figure 2E. Second panel depicts high magnification of colony edge. Scale bars, 100 μm. BRA, BRACHYURY; pS1/5/8, phosphorylated SMAD1/5/8. **B.** Depiction of spatial patterning across multiple colonies. Each dot represents a single cell. **C.** Quantification of voxel fluorescence intensity of pSMAD1/5/8 from colony center (0 μm) to edge (500 μm). Data represents average voxel intensity across multiple colonies. pSMAD1/5/8 colony numbers (n) in upper right corner. Data relative to maximum voxel intensity across the time course for each marker. **D.** Transverse cryosection of immunostained embryo in Figure S5B. Scale bar, 50 μm. **E.** Quantification of pSMAD1/5/8 and BRACHYURY fluorescence intensity in E6.5 embryos. Cells within

the epiblast, primitive streak (PS) and mesoderm were manually selected on confocal images of transverse cryosections in ImageJ as shown in right-hand panel. PS = BRACHYURY positive cells at embryo posterior. Mesoderm = cells positioned between VE and epiblast. Quantification was carried out on 3 cryosections per embryo. N, number of cells. Data represents mean fluorescence intensity +/- S.D. normalized to Hoechst fluorescence.

Figure 6. Anterior primitive streak fates are specified in the absence of BMP. A. EpiLCs generated as in Figure 1C were plated overnight onto Laminin-coated micropatterns (-24 hours) in N2B27 medium with F/A. Various conditions were used for further differentiation - F/A, BMP4, WNT3A (+BMP), F/A, WNT3A (-BMP) or F/A, WNT3A with DMH1 BMP signaling inhibitor (BMPi). Colonies were analyzed after 72 hours differentiation. **B-D, G.** MIPs of immunostained 72 hour colonies. Scale bars, 100 μm. **E.** Quantification of immunostaining. Voxel fluorescence intensity was measured from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies relative to maximum voxel intensity for each marker. **F.** Quantification of marker coexpression by voxel. Each dot indicates fluorescence intensity of a single voxel. Color represents voxel density within the plot. Numbers within quadrants show % of voxels within the gate. N, number of colonies. **H.** Cell fates after 72 hours *in vitro* differentiation under conditions described in A and corresponding *in vivo* cell types. The outer domain of the micropattern colony comprises cells that coexpress SOX17 and FOXA2, representing definitive endoderm and cells that coexpress BRACHYURY and FOXA2, representing either anterior primitive streak cells (left panel) or axial mesoderm cells (middle panel).

Figure 7. Epiblast stem cells undergo definitive endoderm differentiation in the presence or absence of BMP. A. Epiblast stem cells (EpiSCs) of the EpiSC9 line [40] were cultured in the presence of 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A on fibronectin. EpiSCs were then plated overnight onto Laminin-coated micropatterns (-24 hours) in N2B27 medium with F/A. Various conditions were used for further differentiation - F/A, BMP4, WNT3A (+BMP) or F/A, WNT3A (-BMP) or F/A. Colonies were analyzed after 72 hours differentiation. **B.** MIPs of immunostained 72 hour colonies. Scale bars, 100 μm. **C.** Quantification of immunostaining. Voxel fluorescence intensity was measured from colony center (0) to edge (500). Data represents average voxel intensity across multiple

colonies (n = 10 per condition) and is shown relative to maximum voxel intensity for each marker across both conditions. **D.** Graphs showing the expression level of a number of genes from the published microarray dataset of Hayashi et al. from E5.75 *in vivo* epiblast, EpiLCs and EpiSCs [8]. Data shown is from amplified RNA samples and represents the mean +/- S.D for two independent replicates.

Figure 8. Micropattern differentiation of mouse pluripotent stem cells recapitulates cell fate specification in the posterior or anterior primitive streak. A. Summary of embryo gastrulation (upper) and correlation with *in vitro* micropattern differentiation (lower). With FGF2, ACTIVIN A, BMP4 and WNT3A, mouse PSC differentiation recapitulated differentiation in the proximal posterior of the gastrulating embryo. Epi-like cells (EpiLCs) correlated to the embryonic day (E) 5.5-6.0 pre-streak epiblast (Epi). After 24 hours, cells in the colony center adopted a posterior Epi (P-Epi) identity and a primitive streak (PS)-like population arose at the colony edge, as E6.25-6.5. After 48 hours, clusters of cell populations emerged at the outer colony edge correlating to embryonic Mesoderm 1 (Meso), and extraembryonic mesoderm (ExM) allantois core domain (ACD) arising at E6.75-E7.0. After 72 hours, cells in the colony center represented the distal P-Epi of E7.0-7.25 embryos. Meso, ExM and PS populations were maintained. However, ACD cells were replaced by cells with an allantois outer mesenchyme (AOM) identity. Cells were highly confluent and could not be maintained under these conditions after 72 hours. LO, low expression. Dashed lines mark transverse plane shown below. **B.** Summary of correlation between *in vitro* micropattern differentiation with FGF, ACTIVIN and WNT and *in vivo* gastrulating embryos. Under these conditions, mouse PSCs recapitulated differentiation of distal posterior (left panels) or distal posterior and anterior of embryo. After 72 hours, the central population expressed elevated levels of SOX2 compared to BMP4 conditions, likely representing anterior Epi (A-Epi). Cells coexpressed FOXA2 and SOX17, representing definitive endoderm (DE), and FOXA2 and BRACHYURY, either representing anterior PS (A-PS) or axial mesoderm (AxM) cells. HI, high expression. A, anterior; P, posterior; Pr, proximal; D, distal. Color-coded legends highlight key markers of different cell states at each time point. **C.** Box plots showing Epi length along the anterior-posterior (A-P) axis at pre- (E5.5-6.0) early (E6.5-6.75), mid- (E7.0-7.25) and late (E7.5-7.75) streak stages of mouse embryonic development. The A-P length was measured on sagittal confocal optical sections through the middle of the embryo with ImageJ software, as depicted in

1129 the schematic diagram. N, number of embryos. **D.** Box plots showing human embryonic
1130 disc measurements compiled from human embryo data collections. Abnormal embryo
1131 data were excluded. D = embryonic day. Carnegie stage 6a, pre-streak; 6b, early streak;
1132 7, early-mid gastrulation.
1133
1134
1135

Table 1. Summary of cell fates arising in the presence of various signaling factors.

Cell fates generated after 72 hours of mouse micropattern differentiation with described cytokine combinations. It should be noted that cells do not detect these signals homogeneously. WNT inhibition (WNTi) refers to XAV treatment; ACTIVIN inhibition (ACTIVINi) refers to the absence of ACTIVIN and SB431542.

Signaling pathways	Outcome
FGF, ACTIVIN, (WNTi)	Epiblast
FGF, ACTIVIN, endogenous WNT	Epiblast PS
FGF, ACTIVIN, BMP, WNT	Posterior epiblast PS Extraembryonic mesoderm Embryonic mesoderm
FGF, ACTIVIN, WNT	Anterior epiblast Definitive endoderm Axial mesoderm or anterior PS
FGF, BMP, WNT (ACTIVINi)	Epiblast

Figure S1, Morgani et al.

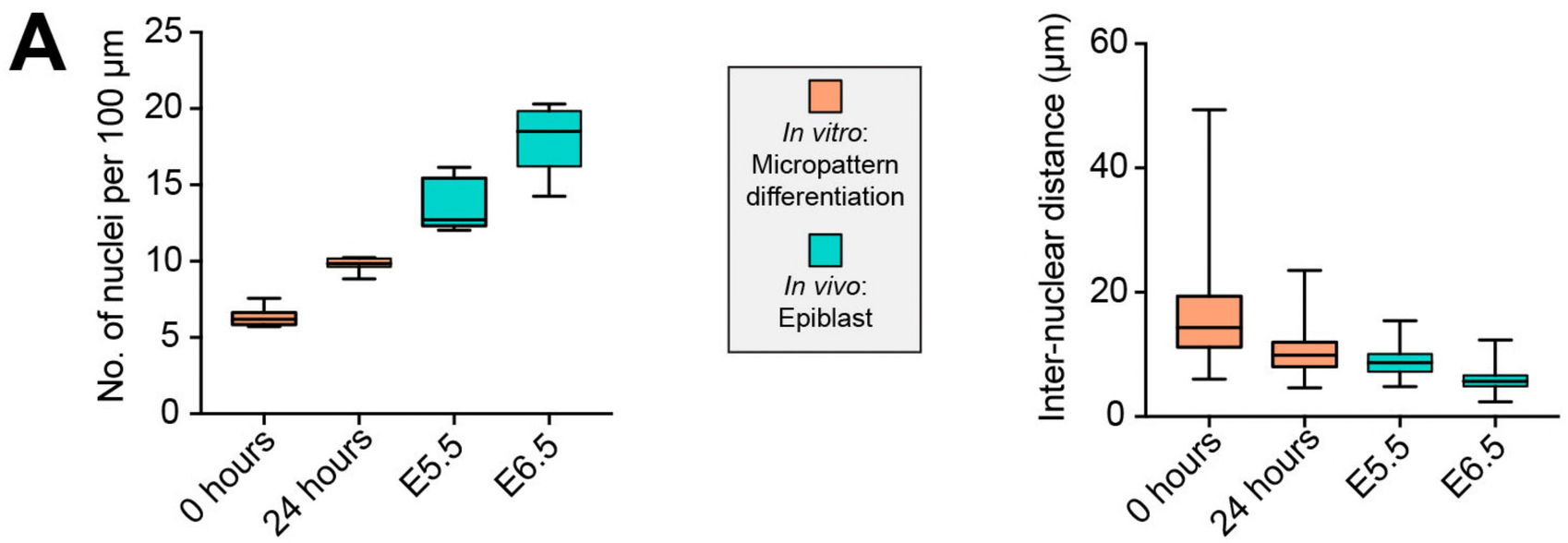
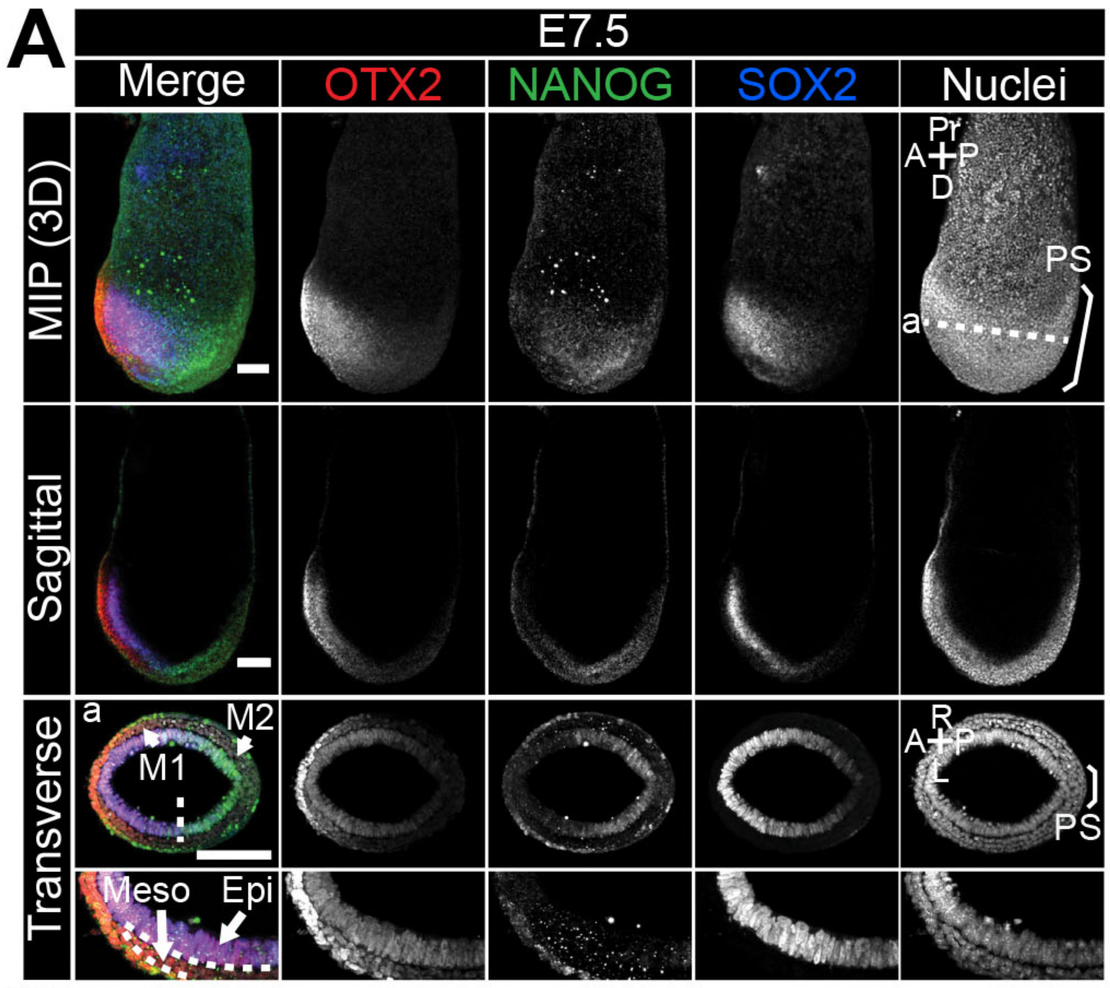
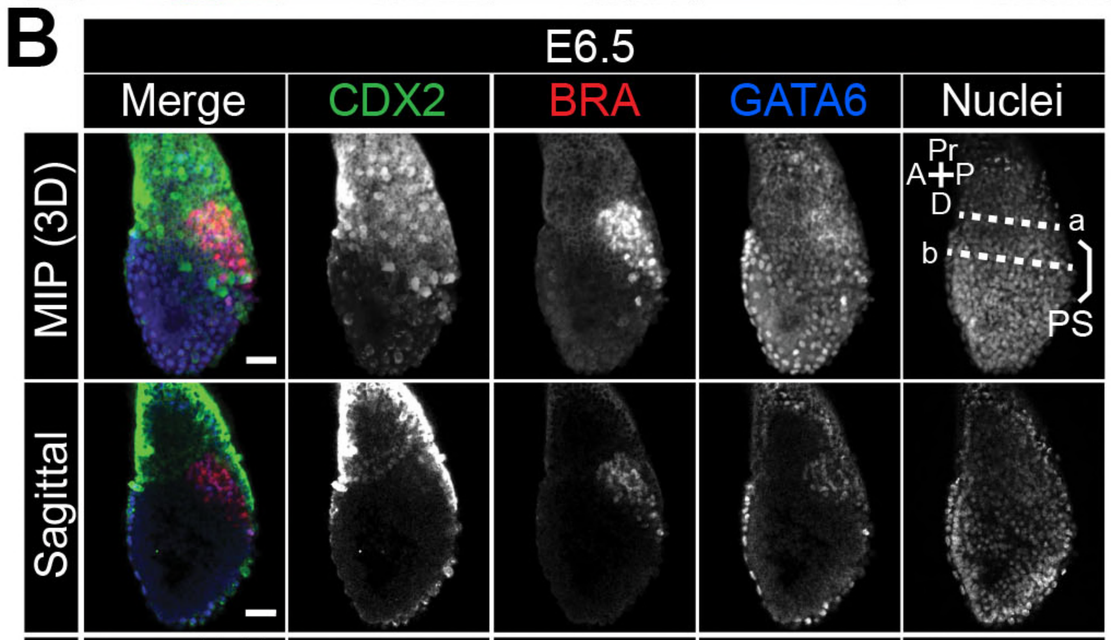
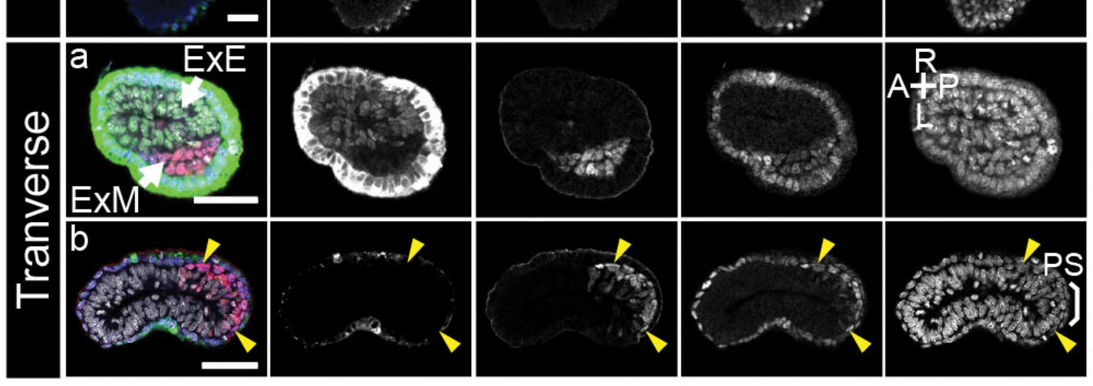


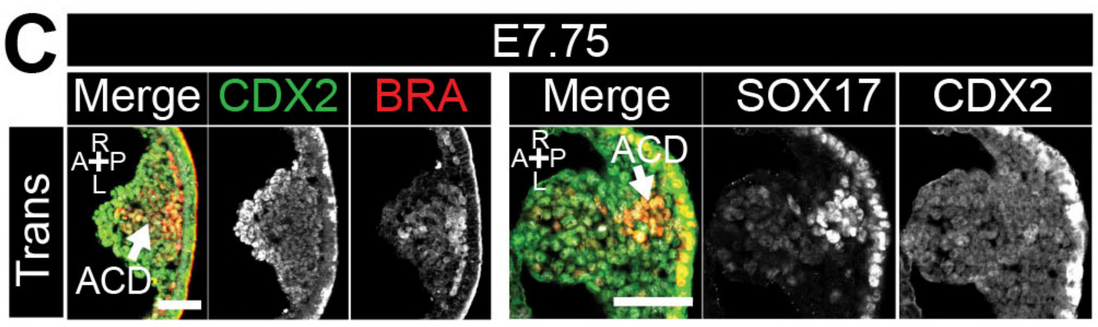
Figure S2, Morgani et al.

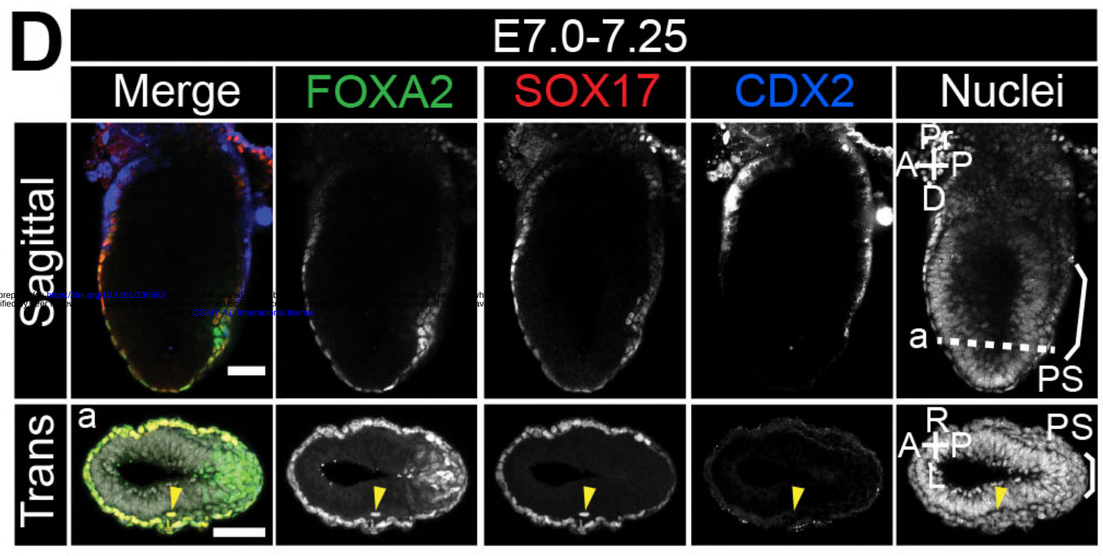
Figure S3, Morgani et al.

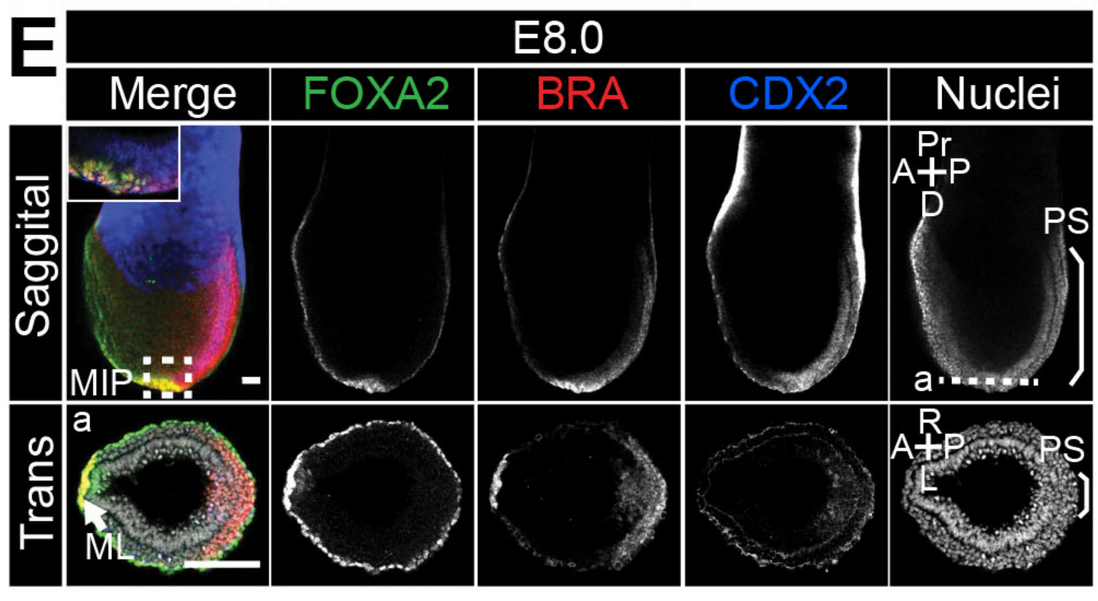


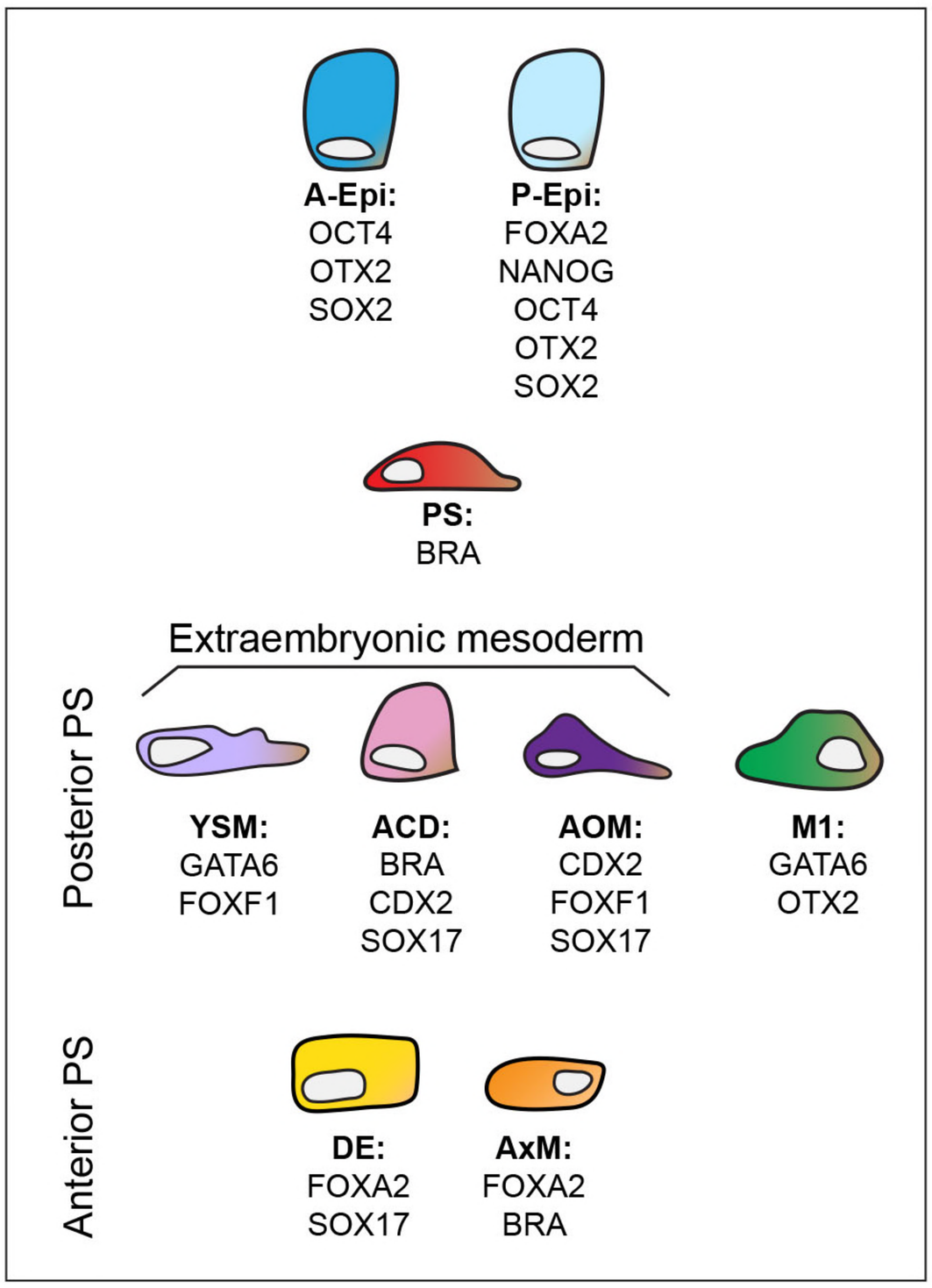


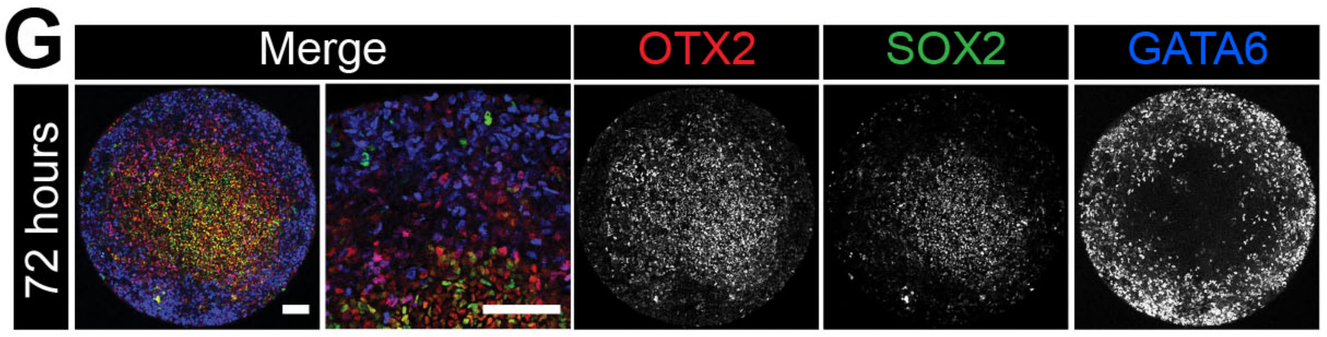


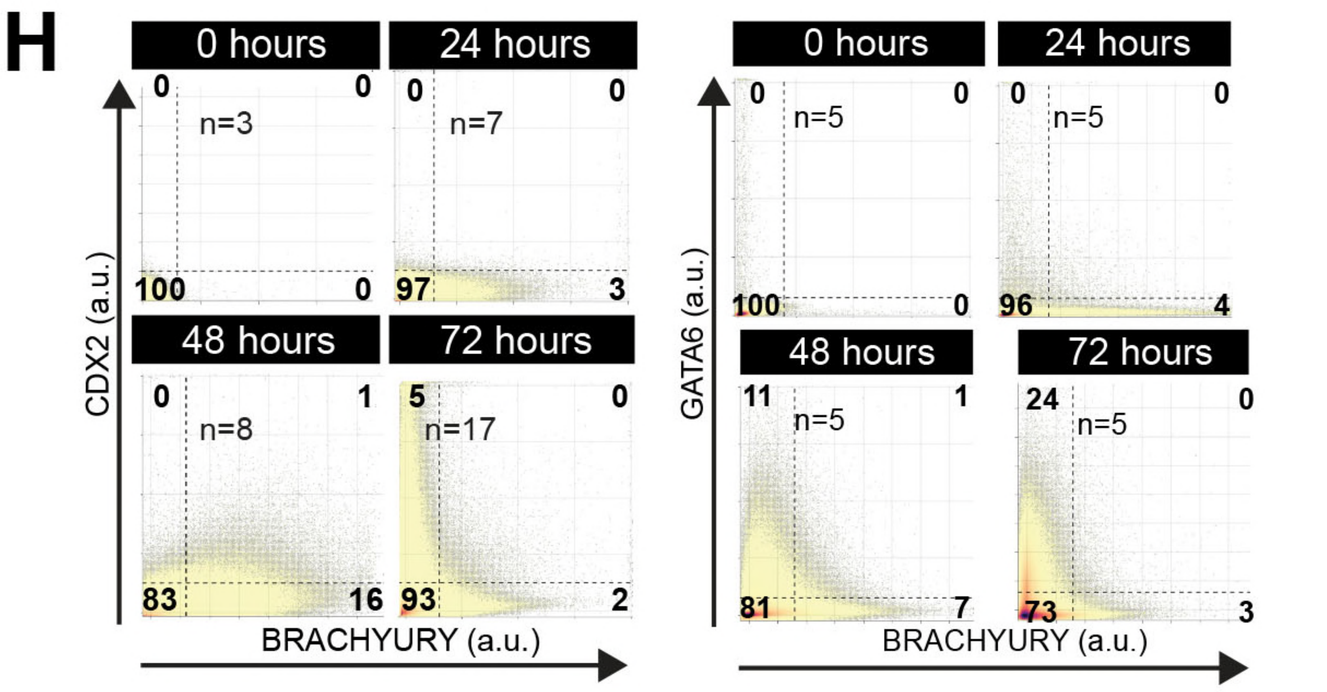


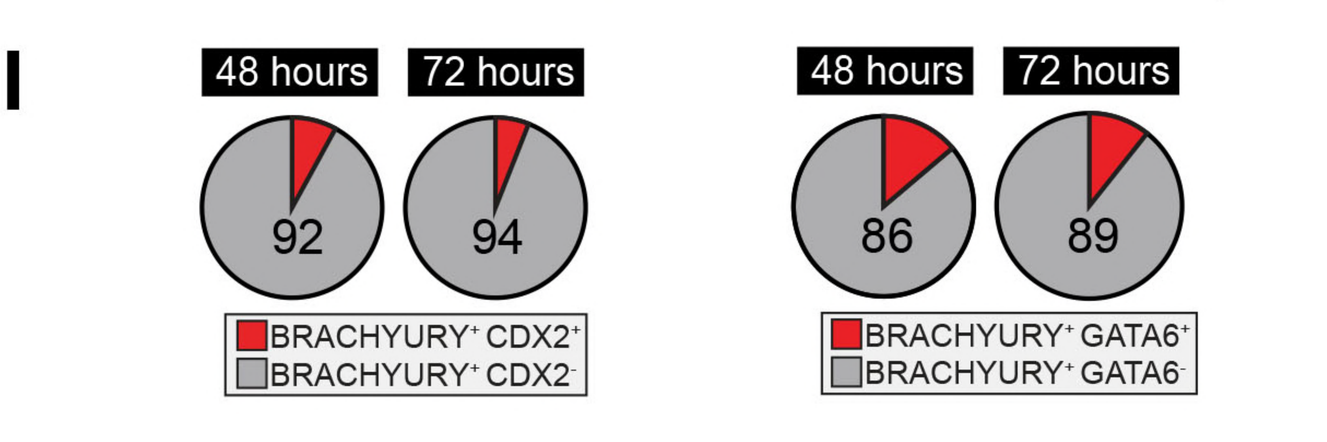












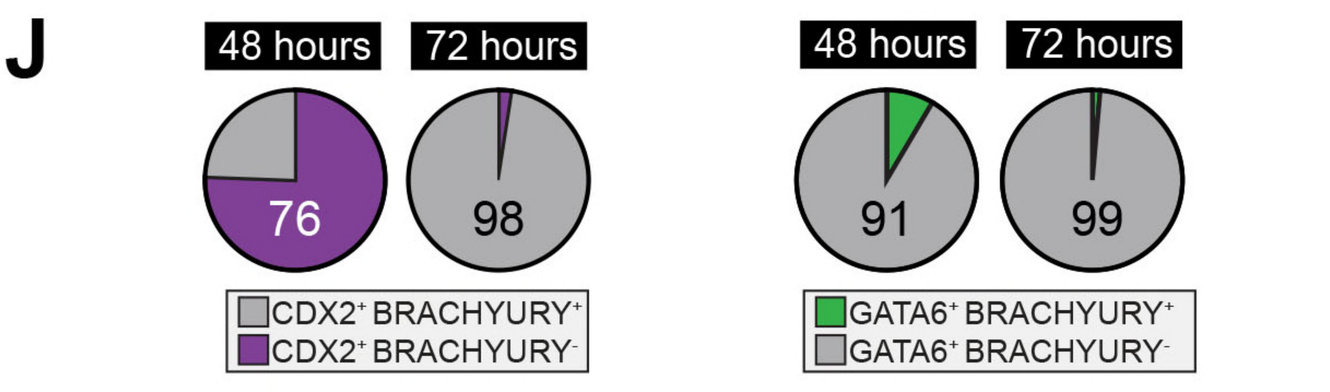
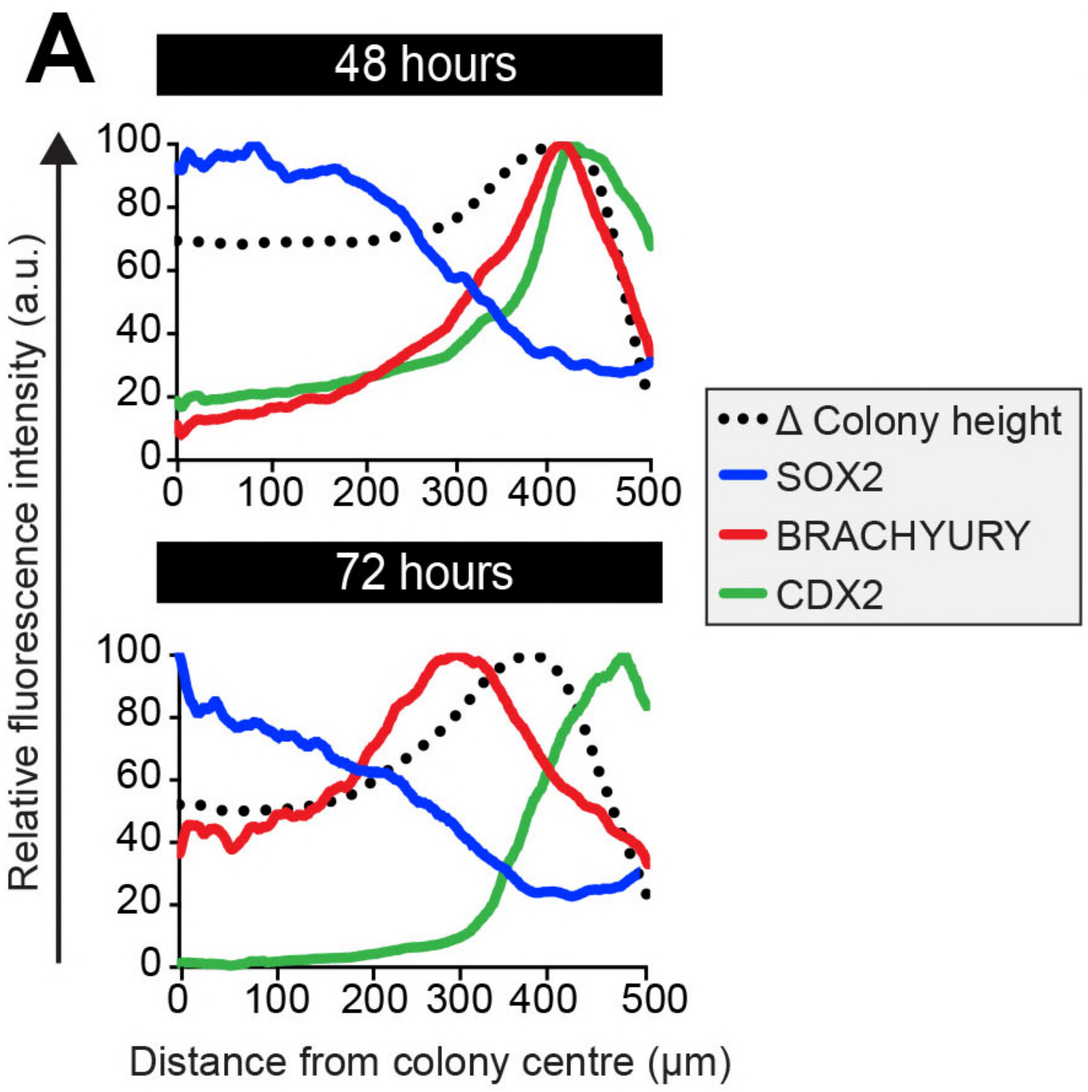
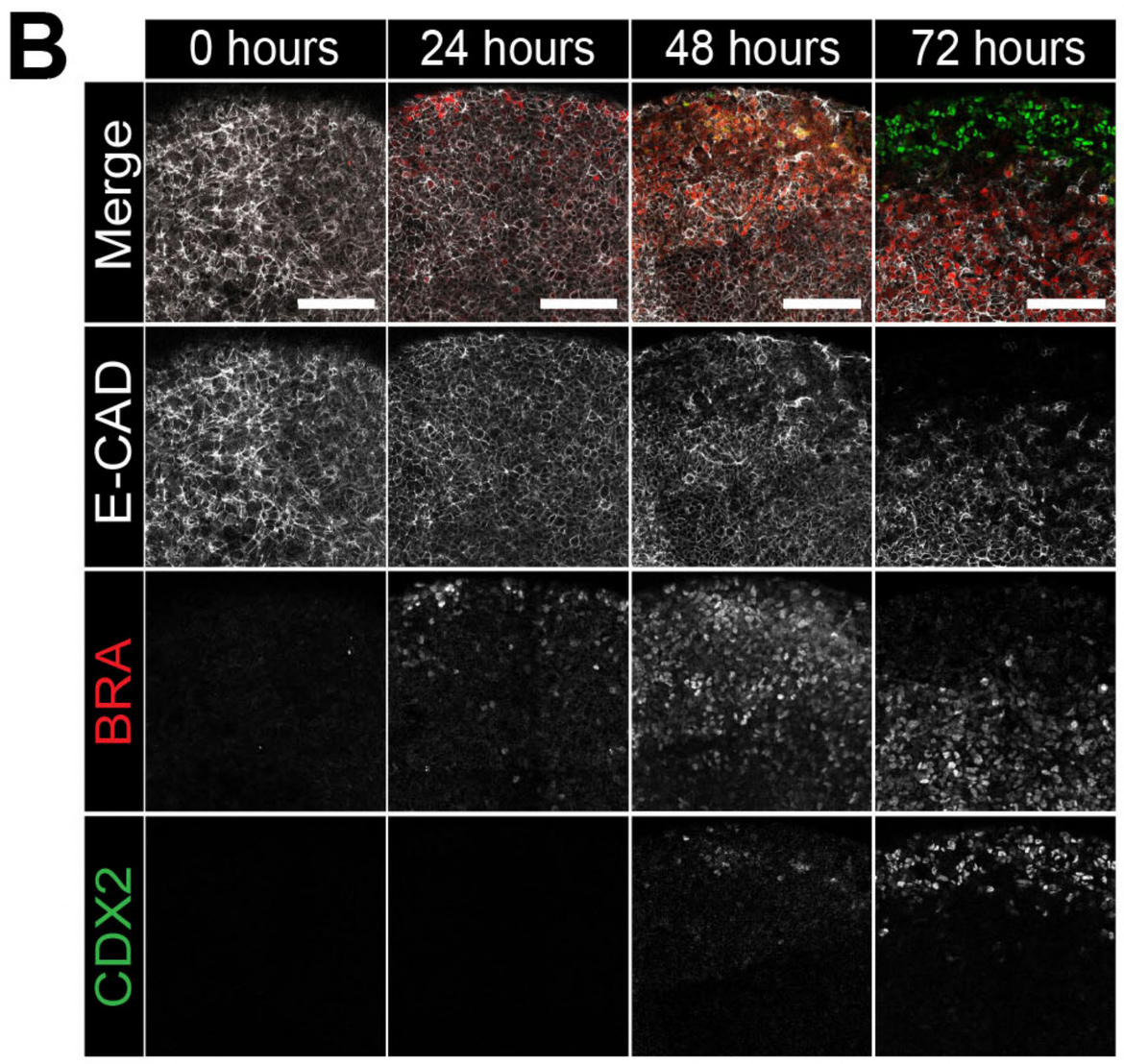
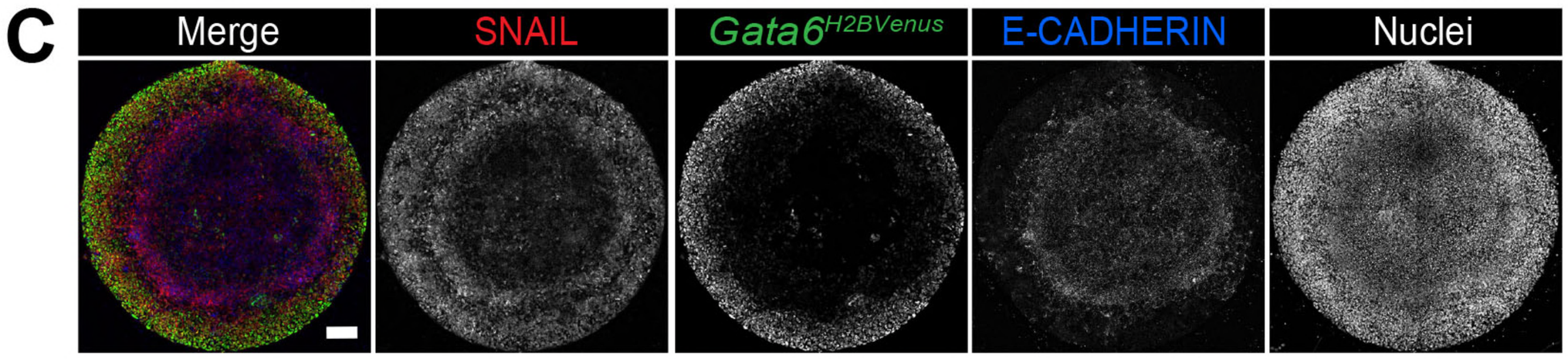
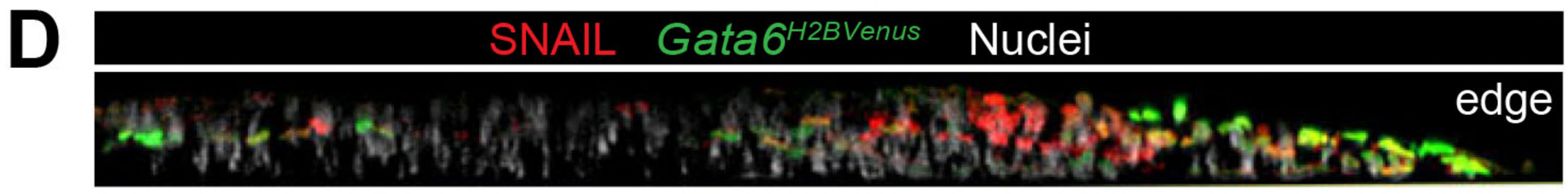


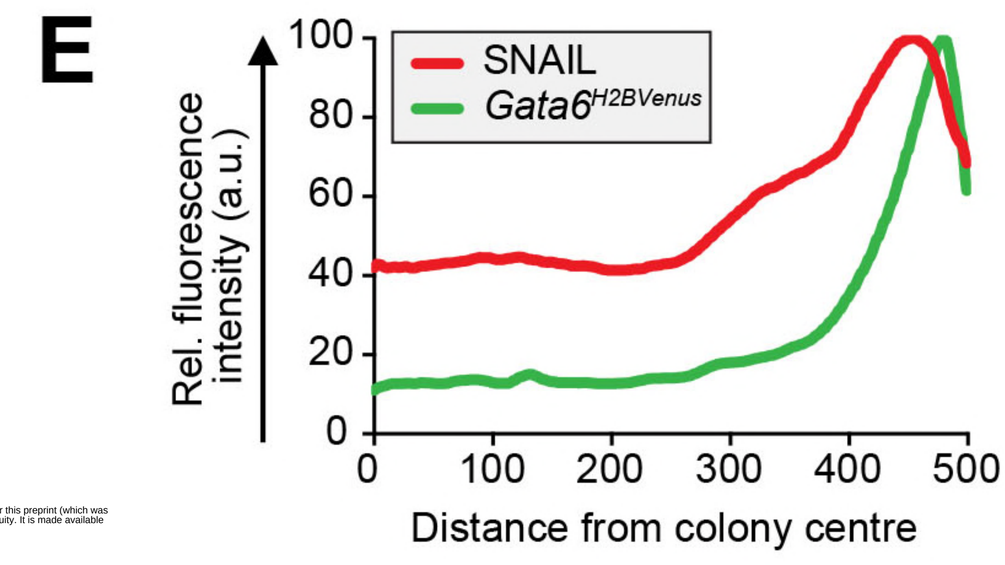
Figure S4, Morgani et al.











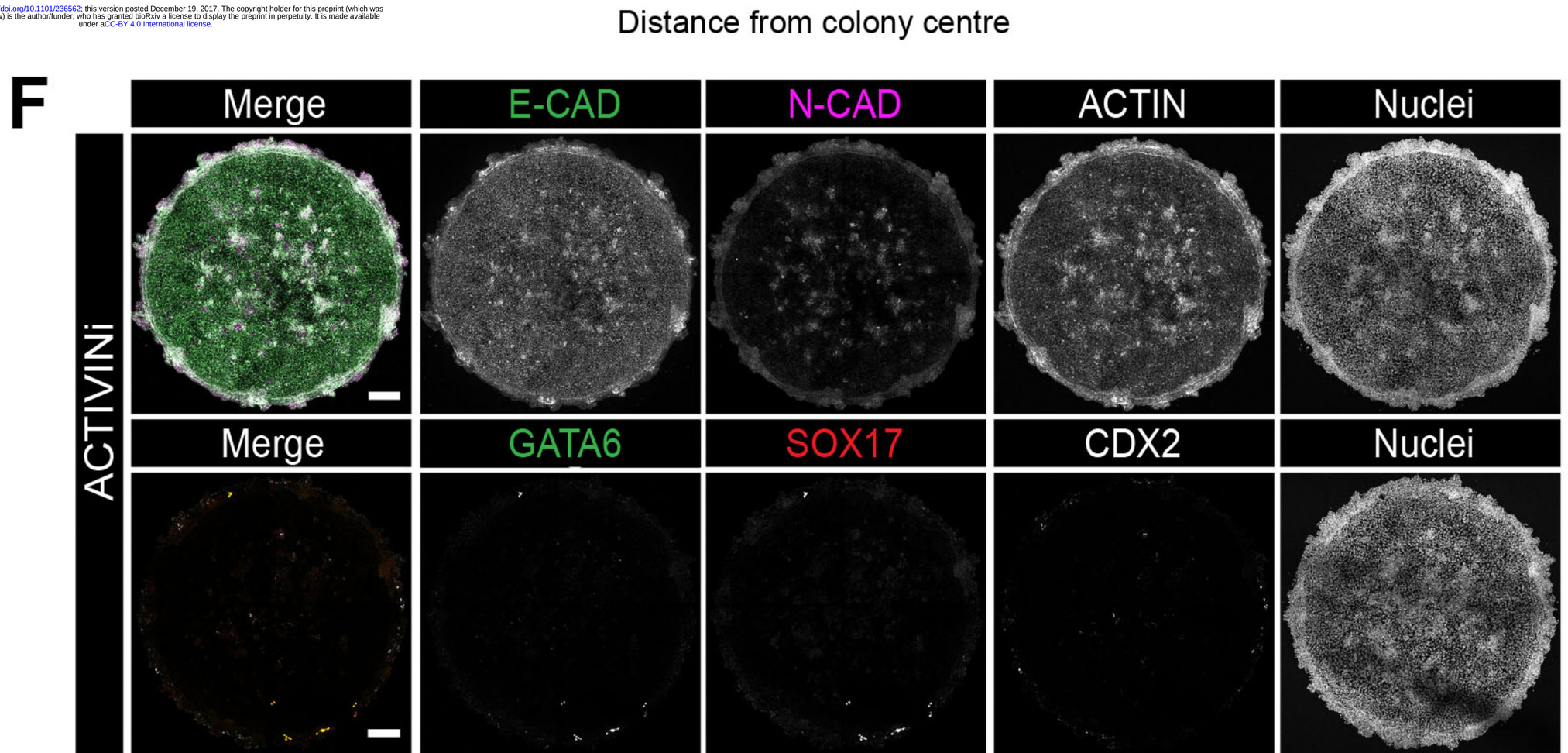
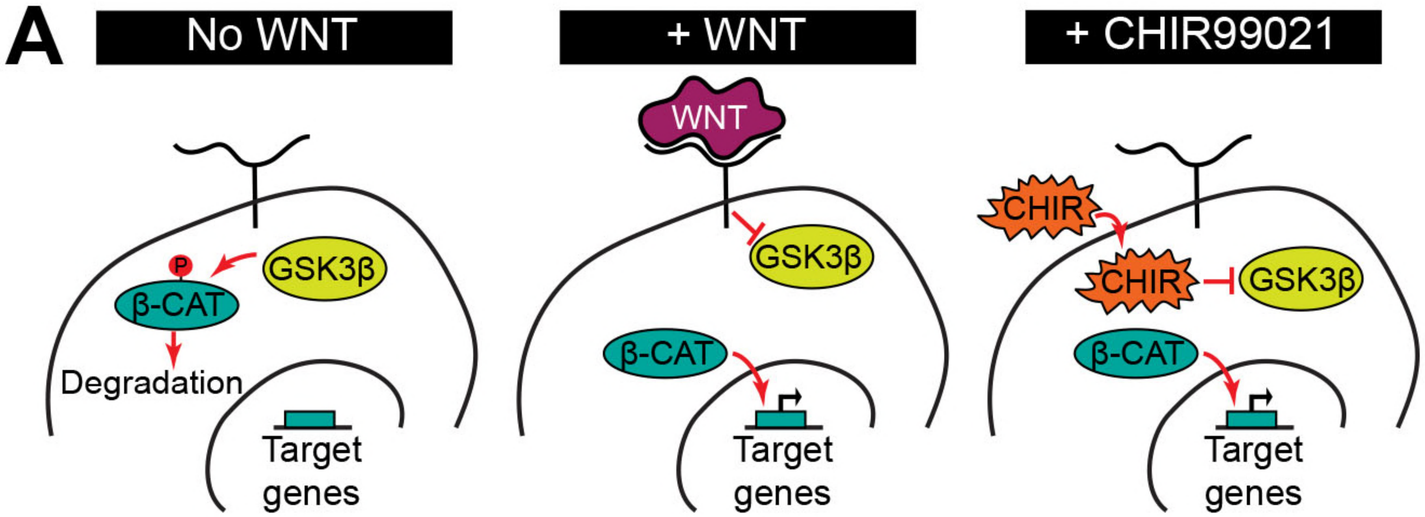
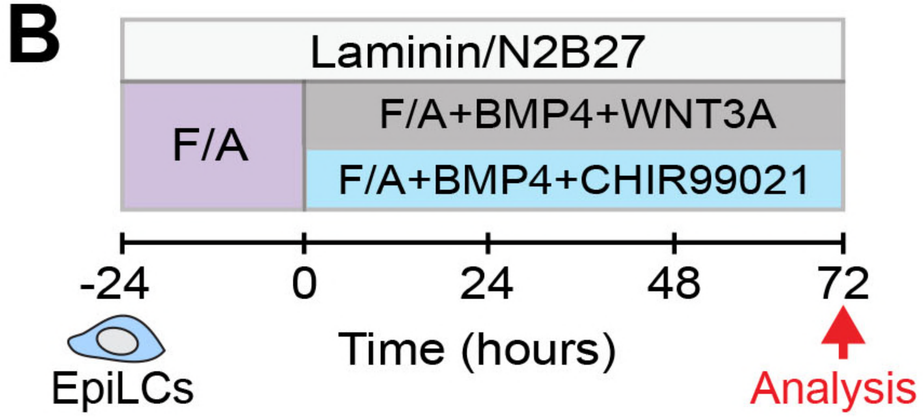
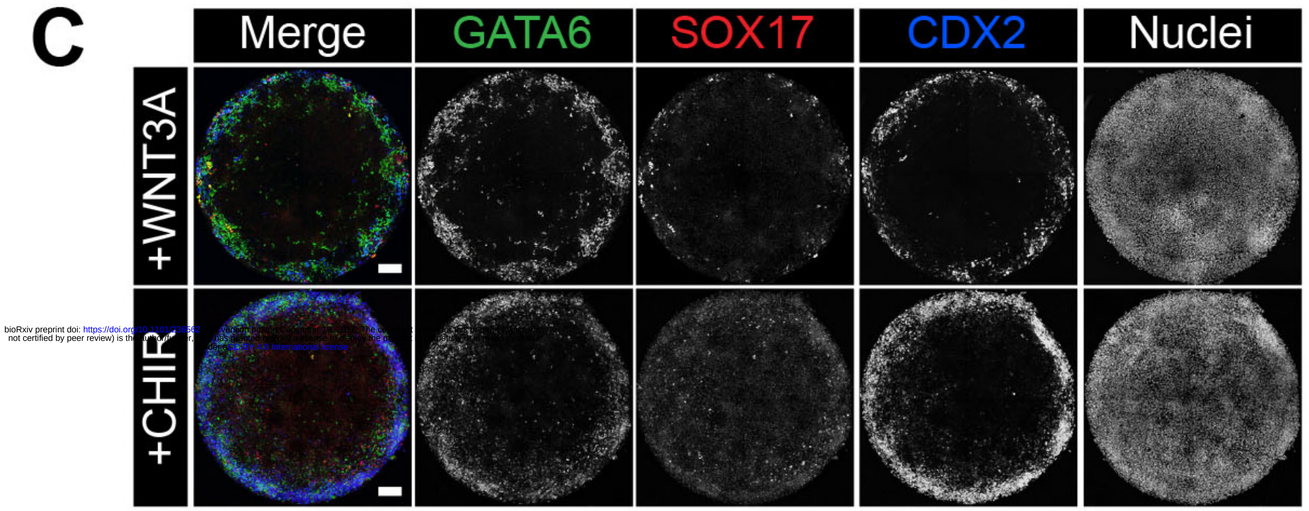


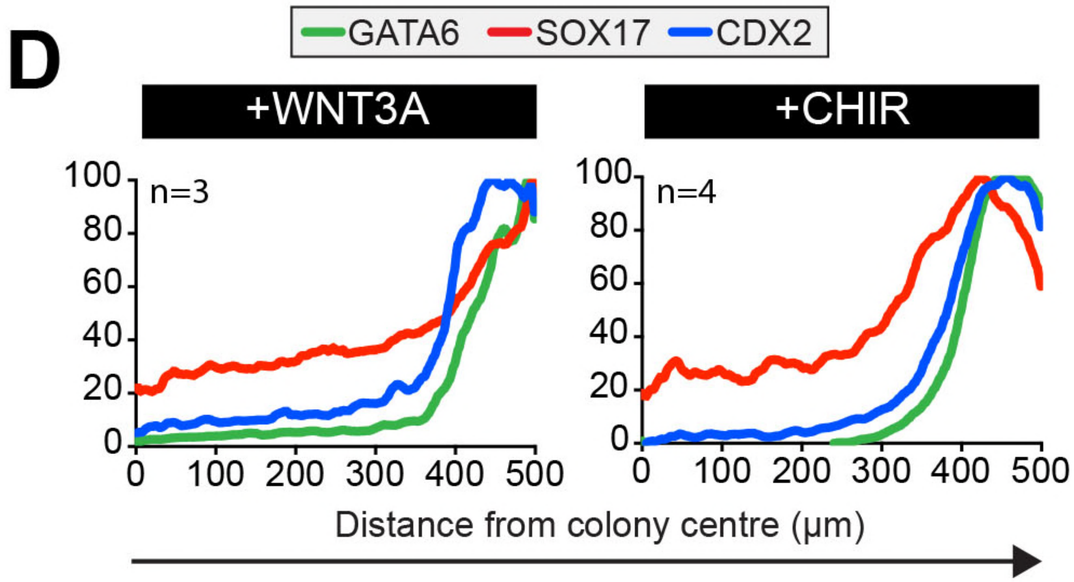
Figure S5, Morgani et al.

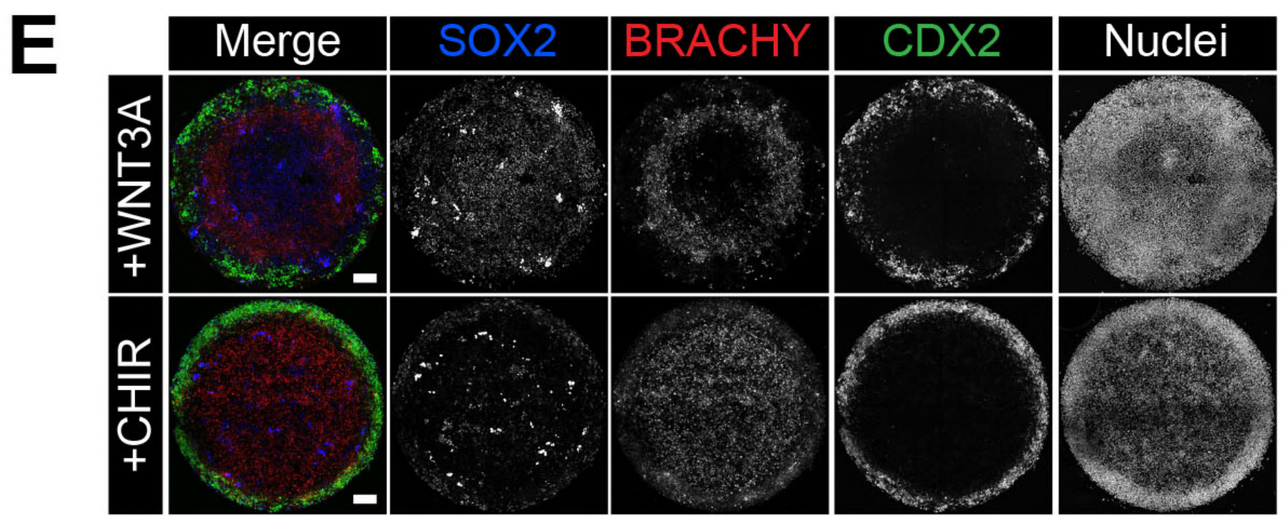
Figure S6, Morgani et al.











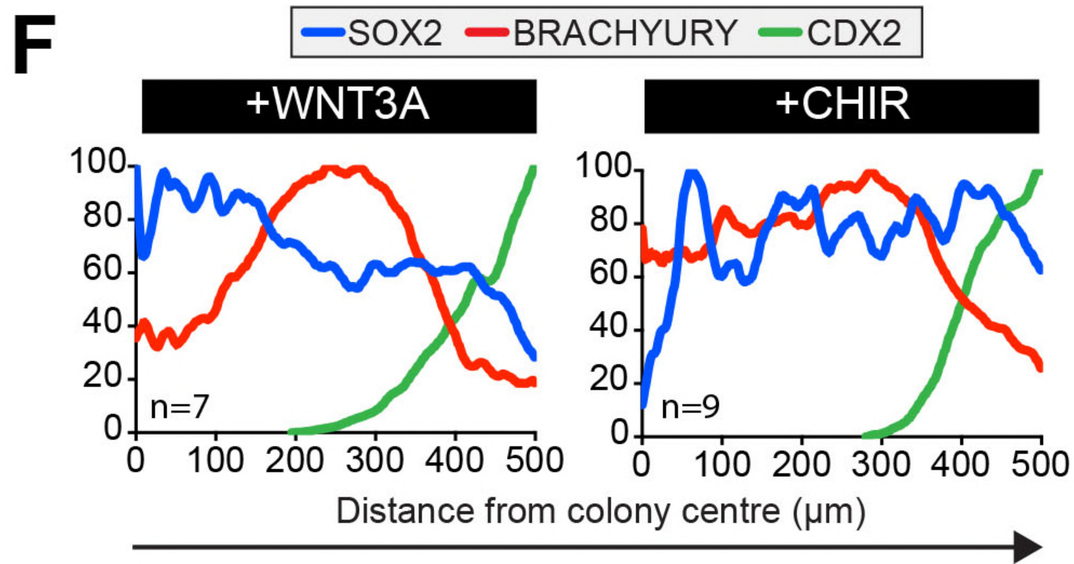
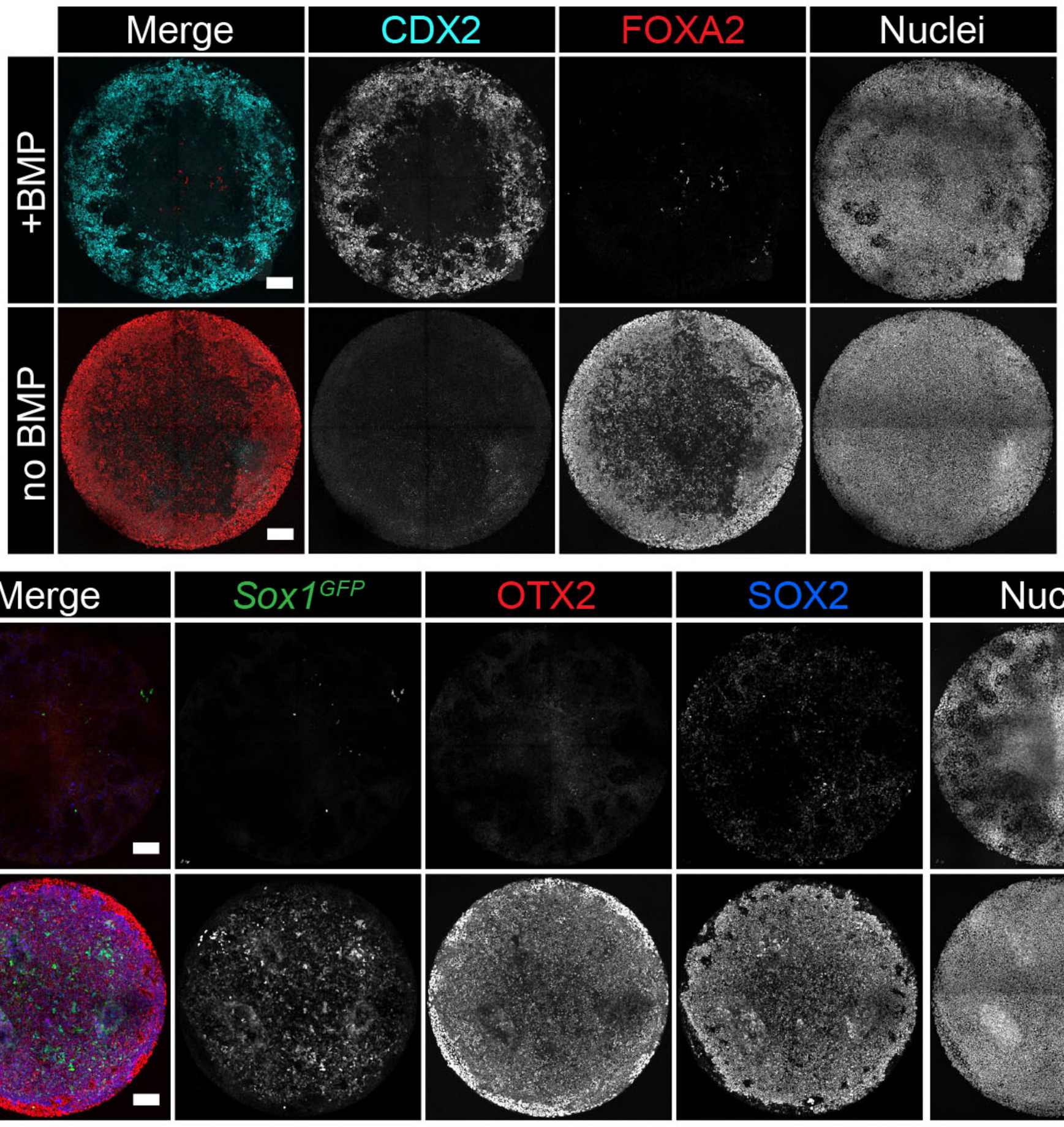
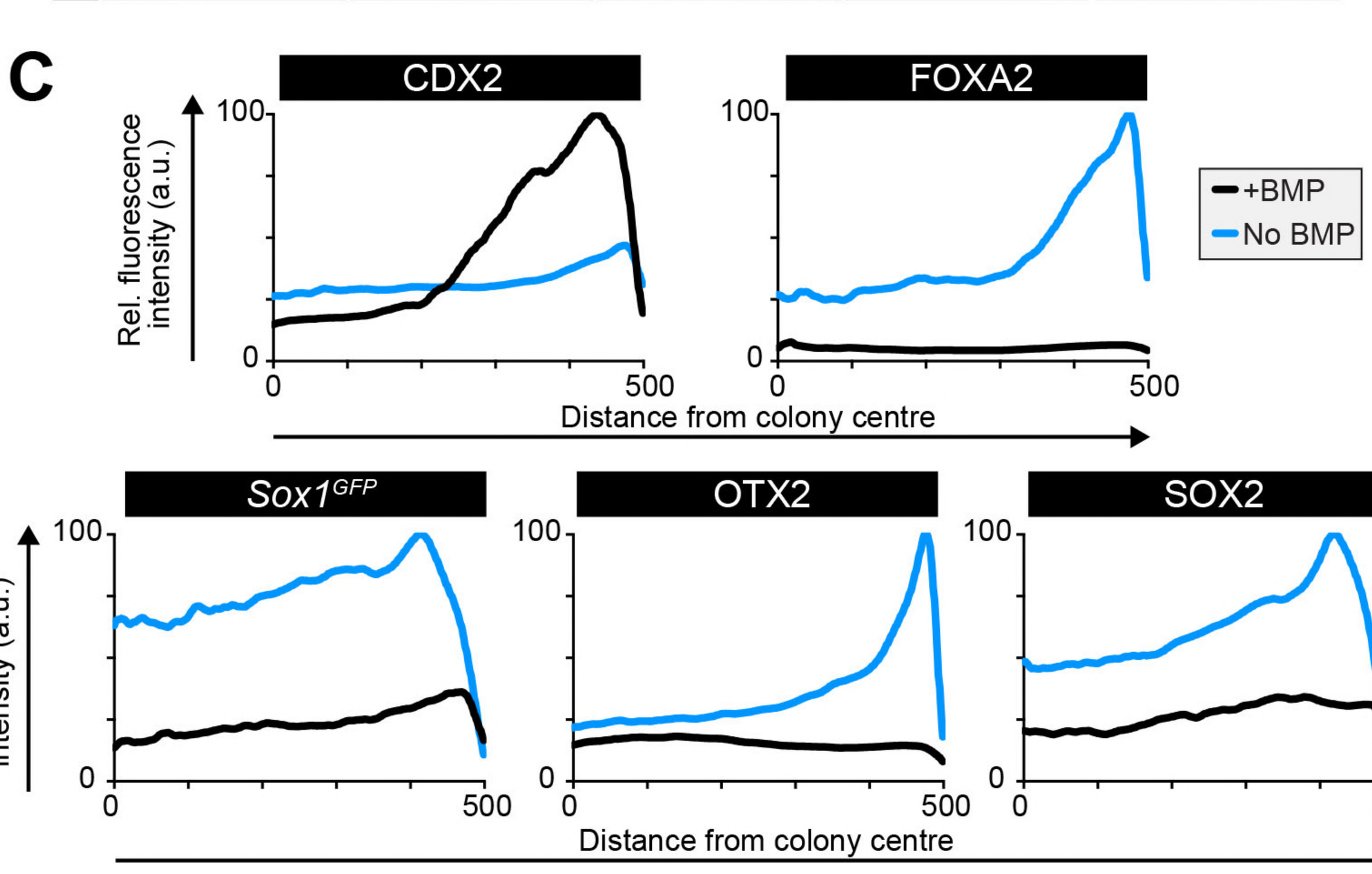


Figure S7, Morgani et al.





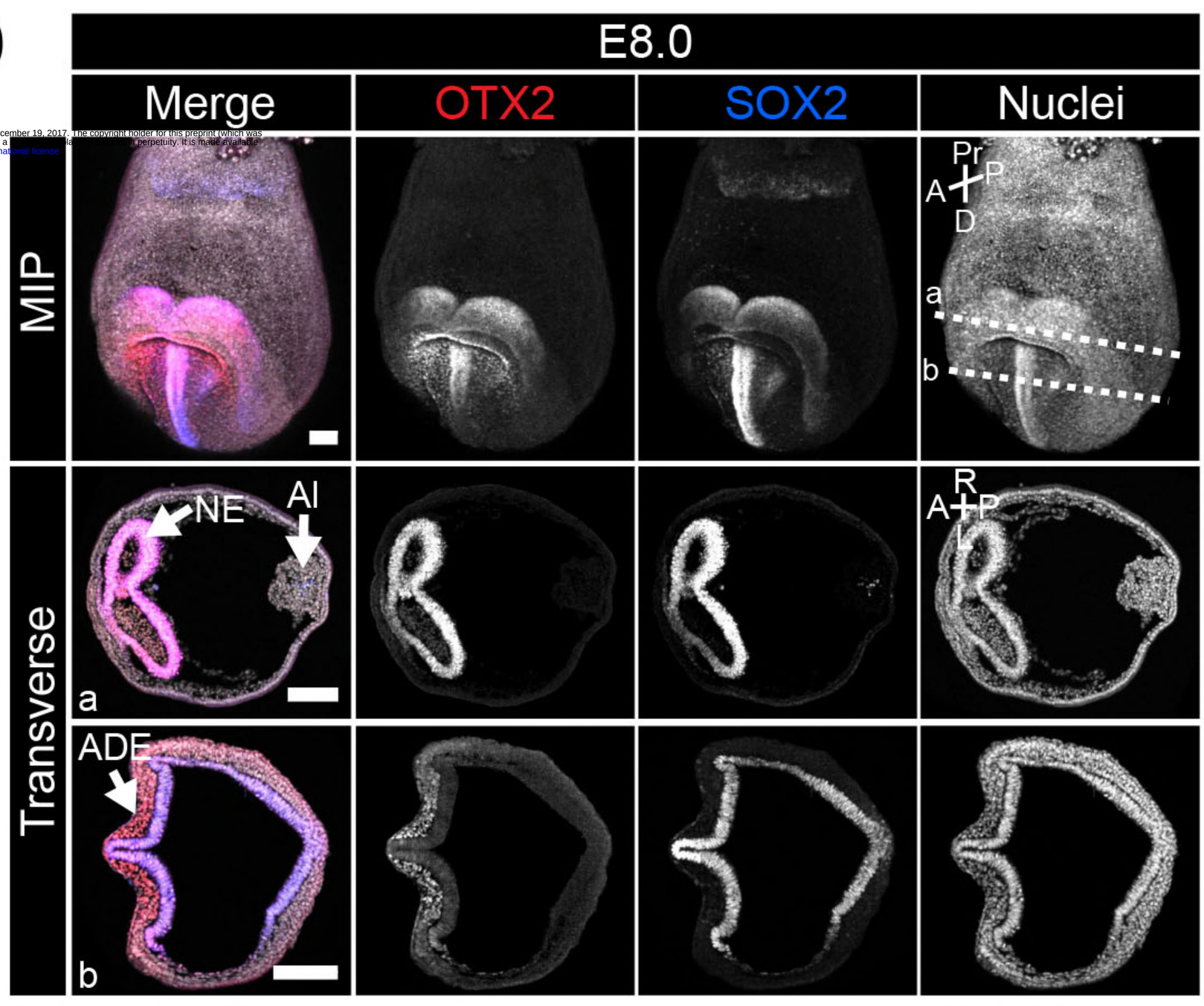
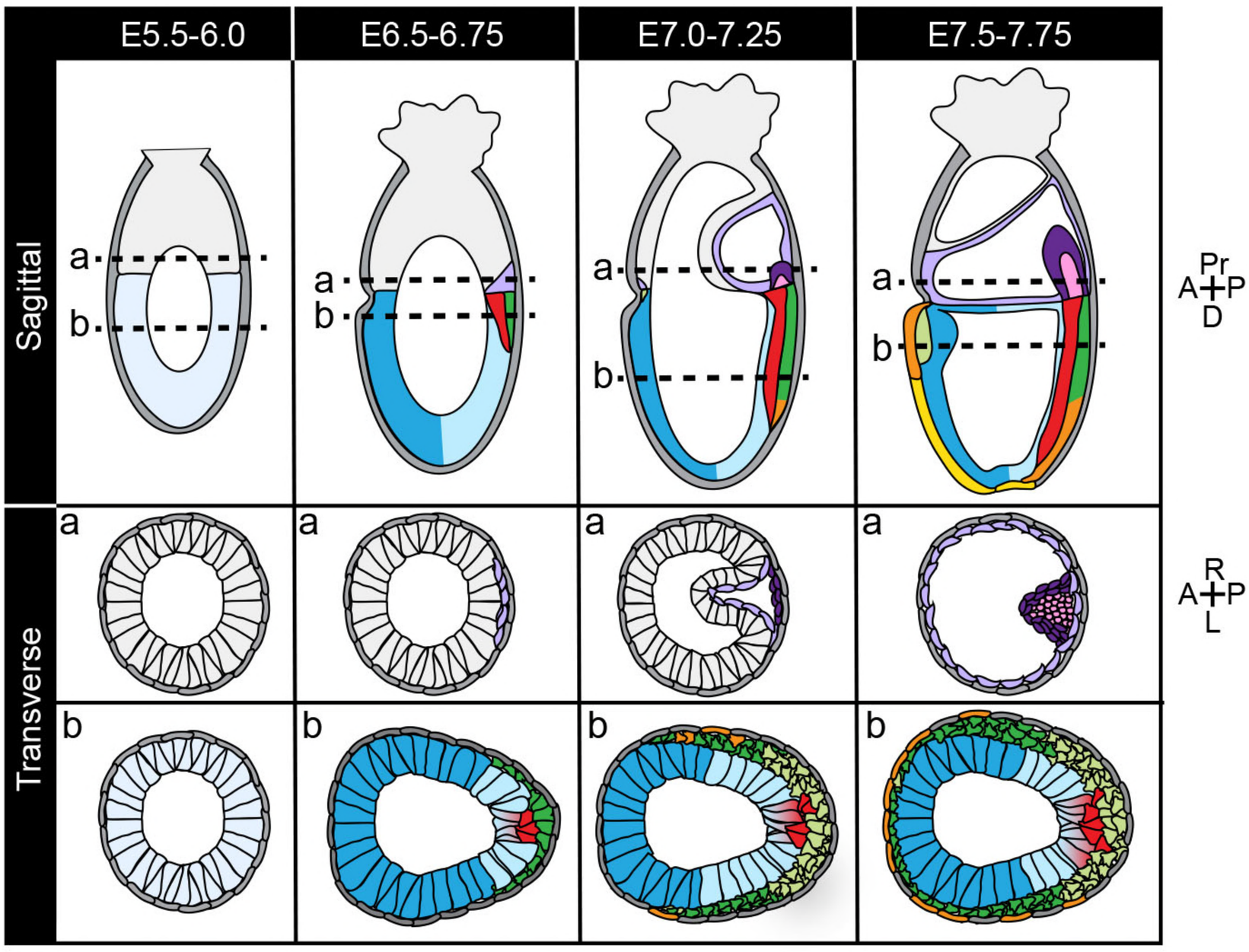


Figure S8, Morgani et al.

A



B