Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate 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 (150 Words)

During gastrulation epiblast cells exit pluripotency as they specify and spatially arrange the three germ layers of the embryo. Similarly, human pluripotent stem cells (PSCs) 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, with 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 identities, including definitive endoderm. By contrast, epiblast stem cells, a 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 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, acting at the proximal, posterior of the embryo. The resulting 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 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]. Standard differentiation protocols generate cell fates in a spatially disorganized manner that is incomparable with *in vivo* gastrulation. However, it was recently it was shown that, when human embryonic stem cells (hESCs) were differentiated within geometrically uniform, circular micropatterns, they reproducibly patterned cell fates with radial symmetry [7-9]. Based on a limited number of markers,

hESC micropatterned colonies were suggested to give rise to a central ectoderm population followed by concentric circular territories of mesoderm, endoderm, and an outer trophectoderm layer [7]. These findings revealed the capacity of the BMP, Wnt and Nodal signaling pathways to collectively organize cell fates. The scalability and reproducibility of this assay coupled with the ease of genetically modifying PSCs, the ability to manipulate culture conditions and 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 cell identities.

Here we adapted the micropattern-based system to defined medium conditions to precisely dissect signaling requirements, and to mouse PSCs for which *in vivo* reference points are accessible to assign cell fates. We first converted mouse ESCs to epiblast-like cells (EpiLCs), the *in vitro* counterpart of the Epi of the early pre-gastrulation embryo [10]. Mouse EpiLCs seeded onto circular micropatterned surfaces formed a simple epithelial morphology in a flat-disc geometry. By all markers examined, these cells were identical to the Epi of the E5.5-E6.0 embryo. When exposed to gastrulation-promoting factors, micropatterned EpiLCs underwent an EMT and recapitulated organized germ layer differentiation 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 differentiation under these conditions was specific to EpiLCs. Under the same conditions neither ESCs, corresponding the pre-implantation Epi, nor epiblast stem cells (EpiSCs), corresponding to the gastrulating Epi (Figure 1A), demonstrated robust cell

fate patterning. Hence, the mouse micropattern system offers a defined and quantitative tool to functionally assess the spectrum of described mouse pluripotent states [11].

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 (distal) PS showed no activity. We hypothesized that by modulating the signals provided to mouse PSCs we could recapitulate the proximal-distal environments operative *in vivo* and generate distinct regional identities *in vitro*. Exposure of micropatterned EpiLCs to posterior signals, BMP, FGF, ACTIVIN (NODAL) and WNT, promoted an EMT and acquisition of posterior Epi, PS, embryonic and extraembryonic mesoderm identities. When BMP was removed, emulating the anterior PS environment (in which FGF, ACTIVIN and WNT are acting), anterior Epi, anterior PS and/or AxM and DE cell types were formed.

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 DE and anterior PS/AxM 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 most developed mammalian genetic model.

Results

Micropatterned EpiLCs correspond to the pre-gastrulation epiblast

The pluripotent state is a continuum spanning from Epi cell specification in the pre-implantation blastocyst (at approximately E3.5) to differentiation at gastrulation which initiates at E6.25 [11] (Figure 1A). Prior to the onset of gastrulation (E5.5-E6.0), 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) [11, 12]. 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 identified EpiLCs as the closest *in vitro* counterpart of the formative Epi [10, 12-14]. We sought to determine whether this correlation held when EpiLCs were grown on micropatterned surfaces. EpiLCs were generated as described [10] 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 [15, 16]. While EpiLCs are generated on a Fibronectin substrate [10], 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 POU5F1 (OCT4), SOX2, NANOG and OTX2, as well as the post-implantation Epi marker, POU3F1 (OCT6) (Figure 1B,D) [17-21]. Neither the Epi of the pre-gastrula

embryo nor EpiLCs expressed KLF4, a key regulator of naïve pluripotency (Figure 1B,D) [22]. 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 in rare cells that we interpret as having spontaneously differentiated.

Micropatterned EpiLC colonies formed an epithelial monolayer (marked by CDH1, also referred to as E-CADHERIN) (Figure 1E - figure supplement 1A). Cell density within the epithelium increased by nearly a factor of four in 24 hours (Figure 1 - figure supplement 1B). 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 1 – figure supplement 1B). However, *in vivo* development can also proceed when the Epi cell density is experimentally decreased [23].

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 from the colony center to the colony edge (Figure 1 – figure supplement 1C) [7, 9]. 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-E6.0.

EpiLCs exposed to FGF/ACTIVIN/Wnt signaling form a PS-like population

At the onset of gastrulation (E6.25-E6.75), the anterior and posterior of the Epi can be distinguished by the expression of specific markers. SOX2 is elevated within the anterior Epi, while NANOG is restricted to the posterior Epi (Figure 2A-C – figure supplement 1A)

[19, 21, 24-26]. The PS emerges in the proximal posterior Epi and is marked by BRACHYURY expression (Figure 2A-C – figure supplement 1A) [27].

When micropatterned EpiLC colonies were cultured with FGF2 and ACTIVIN A for 72 hours, BRACHYURY was induced at the colony periphery (Figure 2 - figure supplement 1B-D), reminiscent of PS formation. In the presence of a small molecule inhibitor of Wnt signaling, XAV939 [28], BRACHYURY expression was abolished and SOX2 was homogeneously expressed at high levels (Figure 2 - figure supplement 1B-D). BMP inhibition had no obvious effect on BRACHYURY (data not shown). Hence in micropatterned colonies, as in EpiSC cultures [29-33], BRACHYURY expression was dependent on endogenous Wnt signaling. This suggested that FGF and ACTIVIN support a SOX2-high anterior Epi-like state, while FGF and ACTIVIN combined with endogenous WNT trigger a reduction in SOX2 levels, as in the posterior Epi, and PS gene expression at the colony periphery (Figure 2 - figure supplement 1E, Table 1). The later germ layer markers GATA6, SOX17 and CDX2 (Figure 2 - figure supplement 1C) were not expressed under these conditions indicating that additional factors were required to stimulate further differentiation.

Posteriorization of EpiLCs after 24 hours exposure to FGF/ACTIVIN/WNT/BMP

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. The extraembryonic signals include WNT3 produced by the VE, and BMP4 produced by the extraembryonic ectoderm (ExE) [2, 34] (Figure 2D). For the mouse micropattern differentiation we utilized pluripotent EpiLCs, corresponding to the Epi cells of the embryo, and thus the system likely lacked the neighboring extraembryonic cell types and the signals that they provide. We therefore asked whether supplying these signals

exogenously could mirror the *in vivo* signaling environment and 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 – figure supplement 2A-C).

After 24 hours of differentiation, the EpiLC micropatterned colonies gave rise to two populations – a central population expressing the Epi markers POU5F1, SOX2 and NANOG, and an outer population expressing the PS marker BRACHYURY (Figure 2F,G). From 0 to 24 hours, SOX2 levels were reduced approximately 2-fold within the central population (Figure 2F-H – figure supplement 2D). Conversely, there was no change in NANOG expression (Figure 2H – figure supplement 2D). Hence, a NANOG-positive, SOX2-low state emerged as is present in the posterior Epi of the embryo [35] (Figure 2A-C – figure supplement 2D). 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 2 – figure supplement 2E), which was also observed in cells within the PS *in vivo* (Figure 2A) and may mark an Epi-PS transition state. At this time, no later germ layer-associated markers (GATA6, SOX17, CDX2, FOXA2) were expressed (Figure 2F,H). Hence, the first 24 hours of micropattern differentiation with exogenous BMP and WNT generated populations resembling the posterior Epi and emerging PS of the mouse embryo at approximately E6.25-E6.75, as observed in the presence of FGF, ACTIVIN and endogenous WNT (Figure 2 – figure supplement 1B-D).

Identification of marker signatures to track germ layer differentiation

As gastrulation proceeds, an increasingly complex array of populations are specified – 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 3 – figure supplement 1A). To identify these cell states *in vitro* in the absence of the spatial and temporal context of the embryo, we sought to establish marker signatures of the cell types present in gastrula stage embryos. To do this, we performed immunostaining for a panel of factors on gastrulating wild-type mouse embryos, and supplemented our observations with published data (Supplemental Table 1 and Additional Data Resource).

At E6.5-E7.5, the anterior Epi expresses high levels of SOX2 and OTX2 while the posterior Epi expresses low levels of SOX2 and high levels of NANOG (Figure 2A-C, Figure 3 – figure supplement 1B). In agreement with a recent spatial transcriptional analysis of gastrulating mouse embryos [36], our immunostaining data suggested that posterior Epi cells may also express low levels of OTX2 (Figure 3 – figure supplement 1B). At E7.0-E7.5, a fraction of distal posterior Epi cells begin to express FOXA2 (Figure 3A) [37], some of which also express OTX2 [36, 38].Then, by E7.75, CDX2 is expressed throughout the posterior Epi (Figure 3B) [39]. By E7.75-E8.0, SOX2 continues to be expressed at high levels in the anterior neurectoderm and at low levels in the posterior, while NANOG is no longer observed within the Epi (see Additional Data Resource) [25].

Throughout gastrulation, BRACHYURY is expressed by cells within the PS (Figure 2A, 3A,B – figure supplement 1C,F). The first cells to leave the posterior Epi and exit the PS at E6.5-E6.75 coexpress BRACHYURY and GATA6 (Figure 3 – figure supplement 1C).

Over time, these cells adopt distinct mesodermal and DE identities. Cells that exit the posterior PS and move proximally into the extraembryonic region generate the extraembryonic mesoderm. The extraembryonic mesoderm forms structures involved in the exchange of materials between the embryo and the mother, including the allantois and yolk sac. Additionally, the extraembryonic mesoderm is a source of hematopoietic progenitors and factors associated with early hematopoiesis, such as *Sox17* and *Hhex*, are expressed within the allantois [40-43]. At E7.75-E8.0, the extraembryonic mesoderm can be subdivided into the allantois core domain (ACD) - expressing BRACHYURY, SOX17, CDX2, FOXF1, the allantois outer mesenchyme (AOM) - expressing SOX17, CDX2, GATA6, FOXF1 and the yolk sac mesoderm (YSM) - expressing GATA6 and FOXF1 (Figure 3C,D – figure supplement 1D) [39, 43-47]. The allantois is reported to express *Gata6* at this time [48], although at the protein level GATA6 was not evident until slightly later stages. At E8.5, CDX2, SOX17 and GATA6 are expressed throughout the allantois (Additional Data Resource). Over time, cells of the ACD contribute to the AOM to support allantois elongation [49].

Cells that originate from the posterior PS and move in an anterior direction around the embryo will form the embryonic mesoderm. At E7.5-E8.0, based on protein expression, we could distinguish two populations of embryonic mesoderm, which we refer to here 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 – figure supplement 1B) [44, 48]. Mesoderm 2 cells, which left the PS later and so were more posterior, expressed BRACHYURY and FOXF1 (Figure 3B-D) [46]. To note, GATA6 and FOXF1 showed a degree of overlap within the region between Mesoderm 1 and 2 (Figure 3D).

Over time, the PS extends distally within the cup-shaped mouse embryo and cells that emanate from the anterior PS give rise to the DE and AxM. Although BRACHYURY is expressed along the length of the PS, at the anterior PS a fraction of cells coexpress BRACHYURY and FOXA2 while others express only FOXA2 (Figure 3A) [37]. Both DE and AxM cells express FOXA2, OTX2, *Gsc* and *Hhex*, although *Gsc* and *Hhex* may be present only transiently within the AxM (Figure 3A – figure supplement 1E) [40, 50-53]. Additionally, DE cells express SOX17 (Figure 3 – figure supplement 1E) and GATA6 [15, 44]. AxM cells also express BRACHYURY (Figure 3 – figure supplement 1F) [27].

Utilizing these *in vivo* marker signatures (Figure 3 – supplemental table 1), we sought to assign identities to the cell populations arising at 48-72 hours of micropattern differentiation, and consequently map the *in vitro* differentiation to *in vivo* development.

FGF/ACTIVIN/WNT/BMP triggers spatially organized posterior fate specification

The first 24 hours of micropattern differentiation in the presence of FGF, ACTIVIN, BMP and WNT resulted in EpiLC posteriorization and the emergence of a PS-like population (Figure 2F,G). At 48 and 72 hours of micropattern differentiation, further spatially organized germ layer differentiation was observed. Three colony domains (central, mid and outer concentric circles) were evident based on marker expression (Figure 2 – figure supplement 2A-C) and the populations within these domains were largely conserved between time points.

At both 48 and 72 hours, the colony center continued to express NANOG and low levels of POU5F1 and SOX2 as in the posterior Epi (Fig, 2F-H – figure supplement 2D). Additionally, at 72 hours, OTX2 and FOXA2 were observed within the colony center (Figure 3E – figure supplement 1G) analogous to the expression of these markers within

the mid to distal region of the posterior Epi at E7.0-E7.5 (Figure 3A) [38]. While OTX2 and SOX2 are also highly expressed within the *in vivo* anterior Epi (Figure 3 – figure supplement 1B), the additional expression of the posterior-restricted marker NANOG, as well as FOXA2, within the colony center (Figure 2A – figure supplement 2D, Figure 3A,E) suggested that these cells were more similar to posterior than anterior Epi. *In vivo*, CDX2 expression is induced within the posterior Epi at E7.5-E7.75 (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 likely correlated to the posterior Epi later than E7.0, but prior to E7.75.

At 24 hours, BRACHYURY marked a PS-like population at the colony periphery (Figure 2F-H). However, at 48 and 72 hours BRACHYURY-positive cells were observed more centrally, within the mid micropattern domain (Figure 2F-H). This inwards shift of BRACHYURY suggested either a wave of gene expression propagating throughout the colony or an inward movement of BRACHYURY-expressing cells. *In vivo*, BRACHYURY is expressed by cells within the PS, but is additionally present in the first cells emanating from the PS – where it is coexpressed with GATA6, and in the extraembryonic mesoderm ACD – where it is coexpressed with CDX2 and SOX17. While we observed cells that coexpressed BRACHYURY and GATA6 or CDX2/SOX17 (discussed later), the majority of BRACHYURY-positive cells did not express these markers, and hence likely corresponded to PS (Figure 3 – figure supplement 1H,I).

The outermost micropattern domain comprised several distinct cell populations with expression signatures reminiscent of embryonic and extraembryonic mesoderm. A small fraction of cells coexpressed GATA6 and BRACHYURY, as in the first cells leaving the PS (Figure 3 – figure supplement 1C,J,K). Additionally, we observed cells that

coexpressed GATA6 and OTX2, as in Mesoderm 1 (Figure 3B – figure supplement 1B,G). While *in vivo* we could also discern a second population of embryonic mesoderm (Mesoderm 2) that expressed FOXF1 and BRACHYURY (Figure 3B-D), within the micropatterned colonies FOXF1 was restricted to the colony periphery in a spatially distinct domain from BRACHYURY (Figure 2H, 3H). Hence Mesoderm 2 was likely not generated under these *in vitro* differentiation conditions.

Within the same outer colony domain, we identified populations resembling extraembryonic mesoderm cell types. CDX2 was expressed within the outer domain from 48 hours of differentiation (Figure 2F-H, 3F). At this time, almost all CDX2-positive cells coexpressed BRACHYURY and SOX17 (Figure 2F, 3F,G – figure supplement 1J). *In vivo*, coexpression of BRACHYURY, SOX17 and CDX2 is first observed within extraembryonic mesoderm cells of the ACD (Figure 3 – figure supplement 1D). While BRACHYURY, SOX17 and CDX2 are also all expressed within cells of the hindgut at later stages of development (E8.5) [54], this micropattern population did not express additional hindgut markers such as FOXA2 [55-57] (Figure 3E). Hence BRACHYURY/SOX17/CDX2-positive cells correlated most strongly to the extraembryonic mesoderm ACD.

At 72 hours, the BRACHYURY/SOX17/CDX2 population was no longer observed. CDX2 and SOX17 continued to be coexpressed (Figure 3F,G) but these cells now lacked BRACHYURY expression (Figure 2F, 3F,G – figure supplement 1J). At 72 hours, CDX2-positive cells also expressed FOXF1, another marker found within the extraembryonic mesoderm (Figure 3D). Hence, SOX17/CDX2/FOXF1-positive cells likely corresponded to the AOM, suggesting a temporal progression of ACD cells to an AOM state, as *in vivo*. We also observed a rarer population of cells that coexpressed FOXF1 and GATA6

(Figure 3H, yellow arrowhead) as in the YSM or embryonic mesoderm positioned between Mesoderm 1 and 2 (Figure 3D). While both embryonic (GATA6) and extraembryonic (SOX17/CDX2/FOXF1) mesoderm-like populations were present within the outer micropattern domain, they tended to exist within discrete clusters (Figure 3G).

During the micropattern differentiation, multiple DE-associated markers were expressed, namely FOXA2, SOX17, GATA6 and OTX2. *In vivo* these markers are coexpressed within DE cells (Figure 3 – figure supplement 1E) while in the micropattern differentiation FOXA2, SOX17 and GATA6 were expressed in a mostly mutually exclusive manner (Figure 3E – figure supplement 2A), hence they marked separate non-DE populations. Furthermore, BRACHYURY and FOXA2 were expressed within distinct micropattern domains (mid and central respectively) (Figure 2F,G, 3E) suggesting that AxM cells were not present. To further validate these conclusions, we assessed the expression of the DE and AxM markers *Gsc* and *Hhex* using a $Gsc^{GFP/+}$; $Hhex^{RedStar/+}$ dual reporter ESC line [58]. After 72 hours of differentiation, $Hhex^{RedStar}$ was observed at the outer colony edge but Gsc^{GFP} was not expressed (Figure 3 – figure supplement 2B). The expression of $Hhex^{RedStar}$ and SOX17, without FOXA2 and Gsc^{GFP} expression, confirmed the absence of DE and AxM fates and likely indicated the presence of hematopoietic progenitors that arise from the allantois [40, 59] and

Therefore, *in vitro*, a combination of BMP, WNT, ACTIVIN (NODAL), and FGF 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). However, fates arising from the anterior PS including DE and AxM were not formed under these conditions.

Micropattern differentiation involves a TGFβ-regulated EMT

One of the primary hallmarks of gastrulation is an EMT, involving downregulation of the epithelial marker CDH1 (E-CADHERIN) and upregulation of the mesenchymal marker CDH2 (N-CADHERIN) in cells ingressing through the PS (Figure 4A,B). Epi cells that do not undergo an EMT differentiate into neurectoderm, while those that undergo an EMT emanate from the PS and acquire mesoderm or DE identities [3, 60]. 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, 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-positive PS and CDX2-positive AOM populations (Figure 4- figure supplement 1A). Cells at the border of the BRACHYURY-expressing region downregulated the epithelial marker CDH1 and upregulated the mesenchymal marker CDH2 (Figure 4D,E – figure supplement 1B). As *in vivo*, the outer CDH2 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 4 – figure supplement 1C-E), a transcriptional repressor that regulates the gastrulation EMT [61]. At 48 hours, CDH2-positive cells emerged at the base of the colony, beneath the CDH1-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 CDH1 (Figure 4D,E).

Various signaling pathways including Wnt, FGF and TGFβ regulate EMT in development and cancer [60, 62, 63]. In particular, the role of TGFβ signaling through SMAD2/3 has been well characterized [64]. Mice with null mutations in *Smad2/3* or *Nodal* do not gastrulate and lack normal mesoderm structures [65-68]. 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 CDH1 and accumulated at the edge of colonies (Figure 4G,H – figure supplement 1F). Furthermore, they failed to downregulate SOX2 and did not differentiate, evidenced by the lack of BRACHYURY, GATA6, CDX2 or SOX17 expression (Figure 4I – figure supplement 1F). 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.

Colony diameter is a critical factor involved in patterning

It was previously shown that, in the micropattern system, hESCs give rise to the broadest spectrum of cell fates when differentiated within colonies of 500-1000 μm diameter [7, 9]. When the micropattern diameter was decreased, the outer cell fate

domains were preserved while the inner cell fates were lost, suggesting an edge-sensing input into the differentiation [7, 9]. We asked whether colony diameter also affected cell fate specification and patterning of mouse PSCs.

At 72 hours of mouse PSC differentiation on micropatterns of 1000 µm diameter, three concentric domains (Regions A-C) could be defined with respect to SOX2, BRACHYURY and CDX2 expression (Figure 2F, 5A). The most central domain was Epi-like and predominantly expressed SOX2 (Region A), the mid domain was PS-like and predominantly expressed BRACHYURY (Region B), and the outermost region (Region C) comprised both CDX2-positive extraembryonic mesoderm and GATA6-positive embryonic mesoderm cells (Figure 5A). While SOX2 levels were highest within the colony center, it was expressed at reduced levels throughout all domains (Figure 5A,B).

We noted that the temporal order of differentiation was maintained across the different colony diameters analyzed - 500, 225, 140 and 80 µm (Figure 5C-E). BRACHYURY-expressing PS-like cells were observed at 24 hours, followed by BRACHYURY/CDX2 coexpression within ACD-like cells at 48 hours and the emergence of cells expressing CDX2 but not BRACHYURY, as in the AOM, at 72 hours (Figure 5C-D). However, the spatial organization of cell fates after 72 hours of differentiation was dependent on colony diameter (Figure 5B). Within 500 µm diameter colonies, the outer extraembryonic mesoderm population marked by CDX2 was maintained but, in contrast to colonies of 1000 µm diameter, BRACHYURY-positive PS cells were positioned within the colony center in place of the SOX2 only Epi-like population (Figure 5B-E – figure supplement 1).

At even smaller micropattern diameters of 80-140 µm, colonies were comprised predominantly of CDX2 or SOX2 –positive cells with almost no BRACHYURY

expression observed (Figure 5B-E – figure supplement 1). CDX2 and SOX2 marked distinct, apparently randomly positioned, clusters of cells (Figure 5B-E – figure supplement 1). Within 1000 μm diameter colonies, the outer micropattern domain (Region C) was comprised of both embryonic (GATA6) and extraembryonic (SOX17/CDX2) mesoderm populations (Figure 3F-H). Both embryonic and extraembryonic mesoderm populations were also observed within smaller diameter colonies whereby CDX2 and SOX17 -positive cells were present within distinct domains from cells that expressed GATA6 (Figure 5F). It should be noted that, on smaller diameter micropatterns, the width to height ratio of colonies was altered such that 80-140 μm diameter colonies generated taller, embryoid body-like aggregates. Over time, these three-dimensional structures exhibited morphological asymmetries (Figure 5F), which may explain the loss of radial symmetry in marker expression. Colonies of 225 μm diameter patterned cell fates in a manner intermediate to that observed in colonies of 500 μm and 80-140 μm diameter, with some BRACHYURY cells still observed (Figure 5C-E – figure supplement 1).

Taken together, these data show that micropattern diameter is a critical factor in determining mouse PSC fate specification and patterning. While the non-uniform, three-dimensional geometry of smaller colonies made data difficult to interpret, the loss of central populations within 500 μm diameter colonies indicates that, like human PSCs [7], mouse PSCs may specify fates as a function of distance from the colony edge.

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-E6.75), BMP4 is expressed by the ExE and later (E7.5-E8.0) by the allantois and chorion [69] (Figure 6- figure supplement 1A), and acts on adjacent tissues.

In E6.5-E6.75 embryos, BMP signaling (marked by pSMAD1/5/8) was active at low levels within the proximal, but not distal, Epi and elevated within cells of the PS and embryonic and extraembryonic mesoderm (Figure 6 – figure supplement 1B). At this stage, pSMAD1/5/8 levels correlated with BRACHYURY expression (Figure 6A,B – figure supplement 1B,C). From E7.0 onwards, as the PS extended, pSMAD1/5/8 was observed within the posterior PS but not anterior PS, consistent with anterior cells being positioned furthest from the ExE source of BMP4 (Figure 6 – figure supplement 1D,E). Furthermore, pSMAD1/5/8 was observed in embryonic Mesoderm 1 but not BRACHYURY-positive Mesoderm 2 cells and consequently, the correlation with BRACHYURY expression was lost (Figure 6 – figure supplement 1E-G).

At 0 hours of micropattern differentiation, nuclear pSMAD1/5/8 was observed at low levels throughout colonies (Figure 6C), corresponding to the low BMP signaling activity within the proximal embryonic Epi at E6.5-E6.75 (Figure 6 – figure supplement 1B). From 24-72 hours of micropattern differentiation, nuclear pSMAD1/5/8 was elevated at the colony edge within the PS, embryonic and extraembryonic mesoderm cell fate domains (Figure 6C-E). At 24 and 48 hours, the majority of pSMAD1/5/8-positive cells expressed BRACHYURY but by 72 hours, the fraction of BRACHYURY/pSMAD1/5/8 - positive cells was significantly reduced (Figure 6C – figure supplement 1H,I). This likely

corresponded to the presence of nuclear pSMAD1/5/8, but not BRACHYURY, within Mesoderm 1 cells of the embryo (Figure 6 – figure supplement 1E,F).

These data revealed that, as with micropattern differentiated hESCs [7-9], signaling activity exhibits radial dependence. Furthermore, the cell types identified within the *in vitro* micropattern system experienced a comparable BMP signaling history to their *in vivo* counterparts, with low BMP signaling activity present within posterior Epi-like cells and elevated activity within the posterior PS, embryonic Mesoderm 1 and extraembryonic mesoderm populations. *In vivo*, the distal Epi and anterior PS were devoid of BMP activity (Figure 6 – figure supplement 1B,D,E) but, in the presence of FGF, ACTIVIN, BMP and WNT, a comparable signaling niche that lacked BMP activity was not observed within the micropatterned colonies.

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 [8, 9]. 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 6 – figure supplement 2A,B). Under these conditions, CDX2, GATA6 and SOX17 were expressed at the outer colony edge indicating that mesoderm differentiation was unaffected (Figure 6 – figure supplement 2C,D). However, the BRACHYURY expression territory was expanded throughout the center of the colony (Figure 6 – figure supplement 2E,F), recapitulating the expansion of BRACHYURY expression in CHIR-cultured embryos [33]. 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

BMP, WNT, ACTIVIN and FGF directed micropattern 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 replicate this signaling niche and create an environment permissive to specify anterior, but not posterior fates. EpiLCs were plated onto micropatterns and differentiated for 72 hours with FGF2, ACTIVIN A, BMP4 and WNT3A (referred to as +BMP), FGF2, ACTIVIN A and WNT3A (referred to as -BMP), or FGF2, ACTIVIN A and WNT3A with a small molecule inhibitor of BMP signaling, DMH1 [70] (referred to as BMPi) (Figure 7A). In +BMP 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 7B-E). In medium conditions lacking BMP (-BMP), the absence of BMP signaling activity was confirmed by lack of nuclear pSMAD1/5/8 (Figure 7B,E). Under these conditions, the domain of extraembryonic mesoderm, marked by CDX2, was lost (Figure 7C,E). Instead we observed elevated SOX17 in outer cells, which was now robustly coexpressed with FOXA2 (Figure 7D-F) representing DE (Figure 3 – figure supplement 1E). We also observed a separate population of outer cells that coexpressed FOXA2 and BRACHYURY (Figure 7F,G), likely representing cells within the anterior PS, node or AxM (Figure 3A, yellow arrowheads – figure supplement 1F).

To further investigate the anterior cell fates formed in the absence of BMP, we differentiated *Gsc*$^{GFP/+}$; *Hhex*$^{RedStar/+}$ ESCs for 72 hours in the presence of FGF, ACTIVIN and WNT. While in the presence of BMP4, *Hhex*^{RedStar} but not *Gsc*^{GFP} was expressed

(Figure 3 – figure supplement 2B), in the absence of BMP, we observed FOXA2, BRACHYURY and *Gsc*^{GFP} expression at the edge of micropatterned colonies from 24 hours of differentiation, followed by *Hhex*^{RedStar} expression at 48 hours (Figure 7 – figure supplement 1A-C). The number of FOXA2, BRACHYURY, *Gsc*^{GFP} and *Hhex*^{RedStar} -expressing cells increased over time. The majority of *Hhex*^{RedStar}-positive cells coexpressed *Gsc*^{GFP} and FOXA2 (Figure 7 – figure supplement 1D,E) a signature of both DE and AxM. However, as we observed little overlap between *Gsc*^{GFP}, *Hhex*^{RedStar} and the AxM marker BRACHYURY within individual cells (Figure 7 – figure supplement 1F,G), *Gsc*^{GFP}/*Hhex*^{RedStar}/FOXA2 coexpression likely represented DE. The FOXA2/BRACHYURY-coexpressing cells observed in the absence of BMP (Figure 7F,G) may correspond to a subpopulation of anterior PS cells (Figure 3A) or alternatively AxM cells that have downregulated *Gsc* and *Hhex*. Global transcriptional analysis may be required to resolve these possibilities. We also frequently observed cells that coexpressed *Gsc*^{GFP} and FOXA2 but not *Hhex*^{RedStar} (Figure 7 – figure supplement 1D), which may represent a BRACHYURY negative anterior PS-like state (Figure 3A) [43].

In conditions lacking BMP signaling activity, SOX2 levels were elevated relative to those in the presence of BMP (Figure 7B-E). This suggested that central cells represent a more anterior Epi state. To determine whether this was the case, we utilized a *Sox1*^{GFP} fluorescent reporter ESC line [71]. *Sox1*^{GFP} marks early neurectoderm specification from cells within the anterior Epi [71]. We differentiated *Sox1*^{GFP} ESCs as described in Figure 7A, 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 7 – figure supplement 2A,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 7 – figure supplement 2B,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 7 – figure supplement 2B,C). This agrees with the later embryonic expression of OTX2 within the DE (Figure 7 – figure supplement 2D). Hence, removing BMP from the (FGF, ACTIVIN and WNT) growth factor cocktail promoted differentiation towards anterior Epi, DE and anterior PS and/or AxM fates (Figure 7H, Table 1).

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

EpiSCs, maintained under standard FGF and ACTIVIN (F/A) culture conditions [72, 73], correlate to later embryonic stages than EpiLCs do [10]. While EpiLCs represent the pre-gastrulation Epi, EpiSCs are similar to the Epi during gastrulation and express markers associated with the anterior PS [74]. We therefore asked whether EpiSCs demonstrated a distinct differentiation capacity from EpiLCs in the context of the micropattern system.

EpiSC9 cells [75] were cultured in defined medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A. EpiSCs were plated onto the micropatterns as described for EpiLCs and differentiated in the same manner - for 72 hours in the presence or absence of BMP (Figure 8A). In the presence of BMP, differentiated EpiSC colonies showed an elevated expression of lineage-associated markers, including BRACHYURY, GATA6, FOXA2 and SOX17, at the colony periphery but lacked obvious spatial organization within more central regions (Figure 8B,C). GATA6 and SOX17/FOXA2 expression represented a DE fate while BRACHYURY was expressed at the outer colony edge in the same domain as FOXA2 corresponding to anterior PS or AxM cell types. SOX2-expressing cells were also present, likely representing an Epi-like state (Figure 8B). Under these conditions,

EpiSCs generated few CDX2-positive cells indicating a significant reduction in the formation of extraembryonic mesoderm (Figure 8B). In the absence of BMP, GATA6, FOXA2 and SOX17 were expressed more uniformly throughout the colonies (Figure 8B,C). Hence, unlike EpiLCs, EpiSCs specified anterior cell fates both in the presence and absence of BMP.

We then utilized published microarray data from Hayashi et al [10], comparing the pre-gastrulation E5.75 Epi, EpiLCs and EpiSCs, to ask what may underlie this difference in the micropattern differentiation of EpiLCs and EpiSCs. As previously described [10, 74] the E5.75 Epi, EpiLCs and EpiSCs all express high levels of the pluripotency marker *Pou5f1,* but EpiSCs also express high levels of the anterior markers *Foxa2* and *Sox17* (Figure 8D). Furthermore, EpiSCs show a marked increase in the expression of the BMP pathway inhibitor *Chordin*, the Wnt pathway inhibitor *Dkk1*, and the Nodal pathway inhibitor *Lefty2* (Figure 8D) that, in this context, may render EpiSCs unresponsive to the BMP posteriorization signal.

Discussion

We have developed a robust, quantitative and scalable micropattern protocol promoting the organized differentiation of mouse EpiLCs, the *in vitro* counterpart of the pre-gastrulation pluripotent Epi of the embryo [10]. In response to FGF, ACTIVIN (NODAL), BMP and WNT, the critical gastrulation-inducing signals acting in the mouse embryo [3], EpiLCs grown 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 marker expression but

also of cell position) and micropatterns allowed us to link the *in vitro* differentiation to *in vivo* developmental time and space. In the absence of the spatial and temporal information of the embryo, we defined a cohort of 15 markers (SOX2, POU5F1, NANOG, SOX1, OTX2, BRACHYURY, CDX2, GATA6, SOX17, FOXA2, FOXF1, CDH1, CDH2, SNAIL, pSMAD1/5/8) that allowed us to distinguish between cell fates such as anterior versus posterior Epi, or extraembryonic mesoderm versus trophectoderm and DE, as these cell types express many common factors. This emphasizes the necessity of expression signatures, rather than individual markers, to accurately assign cell fates *in vitro*.

During 72 hours of differentiation, micropatterned colonies advanced from an E5.5 pluripotent Epi-like state to comprising an array of populations present in the embryo just prior to E7.75 (Figure 9A). Hence, under these culture conditions, *in vitro* cellular differentiation was slower than *in vivo* development. Conceivably, further manipulation of the timing, levels and combination of signaling factors provided to EpiLCs, as well as the extracellular matrix composition and stiffness of the substrate on which cells are maintained to more closely mimic that of the embryo, could alter the rate of differentiation and support the specification of cell fates emerging at later gastrulation stages.

At 72 hours, micropatterned colonies could be divided into 3 spatially distinct domains (central, mid and outer) (Figure 9A). Cells within the colony center showed minimal BMP signaling and expressed posterior Epi markers. PS markers were initially expressed at the periphery, but over time were observed more centrally. This was accompanied by an EMT and the emergence of outer mesenchymal cells, plausibly emanating from the PS-like region. The outer domain displayed elevated BMP activity and contained multiple

populations including allantois and yolk sac extraembryonic mesoderm and early embryonic mesoderm.

In contrast to most gastrulating viviparous mammalian embryos, which exhibit a flat-disc geometry, rodents including the mouse are cup-shaped. A conceptual flattening of the cell fate arrangement within the mouse embryo [76], could not fully recapitulate the organization of cell types observed within the flat-disc micropatterns. Therefore the most evident correspondence between embryonic and micropattern cell fates was signaling history. However, while all cells within the outer micropattern domain experienced high levels of BMP signaling, both embryonic and extraembryonic mesoderm fates were specified. It is therefore unclear whether additional morphogens distinguish embryonic and extraembryonic mesoderm, or if factors such as three dimensional growth, migration and extracellular matrix composition or substratum stiffness dictate fate. Extension of the micropattern system to different geometries, morphogens and inhibitors should resolve these questions.

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 [77] 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. The identity of these inhibitors represents an open question to be elucidated in future studies.

In vivo, localized signaling from the extraembryonic tissues, notably the anterior visceral endoderm (AVE), induces molecular asymmetries within the bilaterally symmetrical Epi leading to anterior-posterior axis establishment [78]. As the described micropattern system does not contain extraembryonic VE or ExE cells, and the disc-shaped colonies are also morphologically symmetrical, there is no apparent chemical or physical source of symmetry-breaking. Therefore, micropattern-differentiated PSCs might generate radially symmetric cell fate domains and be reminiscent of mutant embryos with defects in AVE specification or positioning [79-82], that lack the endogenous source of symmetry breaking. Interestingly, a mouse PSC-based three-dimensional differentiation system involving embryoid body-like aggregates, demonstrated asymmetric lineage marker expression in the absence of extraembryonic cell types [83]. However, as these structures are not geometrically uniform, the polarized marker expression likely stems from initial morphological asymmetries and there is still no evidence of spontaneous symmetry breaking within geometrically uniform structures.

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 [78], and WNT3 on the posterior [84] side of the embryo, whereas the Epi and its derivatives express WNT3, LEFTY2 and DKK1 [36, 85]. 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 [86-89], 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, anterior rather than posterior cell fates were specified (Figure 7B). These data revealed 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 survey signaling conditions and systematically screen interactions between individual genes and pathways.

While a spectrum of mouse PSC states have been captured *in vitro* [11], their comparative functional capacities and relation to the embryo is largely unknown. The micropattern system represents a quantifiable means to test the differentiation potential of PSC states and cell lines under defined conditions. Here we observed that EpiLCs patterned either posterior or anterior cell fates in the presence or absence of BMP respectively. Conversely, although EpiSCs can contribute to all germ layers in chimaera assays [90], within the micropattern system they predominantly generated DE cells and exhibited minimal self-organization. The limited capacity of EpiSCs to pattern in isolation may stem from their elevated expression of signaling inhibitors [10]. Alternatively, exogenous FGF, ACTIVIN, BMP and WNT may not be sufficient to induce the expression of secondary factors required for patterning and posterior cell fate specification of EpiSCs. Further manipulation of the micropattern differentiation conditions may give novel insights into the unique requirements of these different pluripotent cell states (Figure 1A) for organized cell fate specification.

Due to a paucity of data on gastrulating human embryos, cell fates arising during hESC micropattern differentiation can only been predicted [7, 9]. The mouse micropattern differentiation provides the essential missing link between *in vitro* gastrulation models in mouse and human, and *in vivo* mouse development. The identification of an extraembryonic mesoderm population within mouse, but not human micropatterns prompts an analysis of the human system with equivalent marker combinations under comparable serum-free medium conditions containing both BMP4 and WNT3A to determine whether populations such as extraembryonic mesoderm can be generated, or whether in human, extraembryonic mesoderm, as has been shown for the amnion [91], does not arise from the Epi at gastrulation. Human and mouse embryos are of a similar size (Figure 8C,D), their corresponding *in vitro* PSCs undergo micropattern differentiation within equivalent diameter colonies and specify cell fates as a function of distance from the colony edge, suggesting that these species use common mechanisms to regulate cell fate specification and tissue patterning. The further correlation of mouse and human *in vitro* micropattern data, in the context of different pluripotent starting states, and corroborated with *in vivo* data from mouse embryos should yield insights into the conserved and divergent mechanisms regulating fundamental aspects of early mammalian development.

Materials and Methods

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
gene ()				
strain, strain background ()	Crl:CD1 (ICR)		RRID:IMSR_CRL:22	CD1 Mus musculus wild-type outbred mouse
genetic reagent ()				
cell line ()	ES-E14	[92]	RRID:CVCL_C320	Embryonic stem cell line: Mus musculus
cell line ()	ES-R1	[93]	RRID:CVCL_2167	Embryonic stem cell line: Mus musculus
cell line ()	*Sox17*$^{GFP/+}$	[94]		Embryonic stem cell line: Mus musculus
cell line ()	*T*$^{GFP/+}$	[95]		Embryonic stem cell line: Mus musculus
cell line ()	46C cell line (*Sox1*^{GFP})	[71]	RRID:CVCL_Y482	Embryonic stem cell line: Mus musculus
cell line ()	*Gsc*$^{GFP/+}$; *Hhex*$^{RedStar/+}$	[58]		Embryonic stem cell line: Mus musculus
cell line ()	EpiSC9	[75]		Epiblast stem cell line: Mus musculus
transfected construct ()				
biological sample ()				
antibody	anti-BRACHYURY	R&D Systems	Cat# AF2085, RRID:AB_2200235	1:200
antibody	anti-CDH1	Sigma-Aldrich	Cat# U3254, RRID:AB_477600	1:500
antibody	anti-CDH2	Santa Cruz Biotechnology	Cat# sc-7939, RRID:AB_647794	1:300
antibody	anti-CDX2	BioGenex	Cat# AM392, RRID:AB_2650531	1:200
antibody	anti-DsRed	Clontech Laboratories, Inc.	Cat# 632496, RRID:AB_10013483	1:500
antibody	anti-FOXA2	Abcam	Cat# ab108422, RRID:AB_11157157	1:500
antibody	anti-FOXF1	R&D Systems	Cat# AF4798, RRID:AB_2105588	1:200
antibody	anti-GATA6	R&D Systems	Cat# AF1700, RRID:AB_2108901	1:100
antibody	anti-GATA6	Cell Signaling Technology	Cat# 5851, RRID:AB_10705521	1:500
antibody	anti-GFP	Aves Labs	Cat# GFP-1020, RRID:AB_10000240	1:500
antibody	anti-KLF4	R&D Systems	Cat# AF3158, RRID:AB_2130245	1:200
antibody	anti-NANOG	Thermo Fisher Scientific	Cat# 14-5761-80, RRID:AB_763613	1:200
antibody	anti-NANOG	Cosmo Bio Co	Cat# REC-RCAB0002PF, RRID:AB_567471	1:500

antibody	anti-OTX2	R&D Systems	Cat# AF1979, RRID:AB_2157172	1:500
antibody	anti-POU3F1	Millipore Sigma	MABN738	1:100
antibody	anti-POU5F1	Santa Cruz Biotechnology	Cat# sc-5279, RRID:AB_628051	1:100
antibody	anti-pSMAD1/5/8	a gift from Dr. Ed Laufer, Columbia University, New York, NY	N/A	1:200
antibody	anti-SNAIL	R&D Systems	Cat# AF3639, RRID:AB_2191738	1:100
antibody	anti-SOX2	Thermo Fisher Scientific	Cat# 14-9811-82, RRID:AB_11219471	1:200
antibody	anti-SOX17	R&D Systems	Cat# AF1924, RRID:AB_355060	1:200
recombinant DNA reagent				
sequence-based reagent				
peptide, recombinant protein				
commercial assay or kit				
chemical compound, drug				
software, algorithm	Ilastik	http://ilastik.org/	RRID:SCR_015246	3-D Nuclear mask generation
other				

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) [92], R1 (129/Sv background) [93], *Sox17*$^{GFP/+}$ (R1, 129/Sv background) [94], *T*$^{GFP/+}$ (E14.1, 129/Ola background, also known as GFP-Bry) [96] and *Sox1*^{GFP} (E14Tg2a background, also known as 46C) [71], *Gsc*$^{GFP/+}$ *; Hhex*$^{RedStar/+}$ (E14Tg2a background) [58]. ESCs were

routinely cultured on 0.1% gelatin coated tissue culture grade plates (Falcon) in serum and LIF medium as previously described [97]. 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% confluence 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) [75]. EpiSCs were cultured under standard conditions as previously described [73], in defined, serum-free N2B27 medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A. EpiSCs were passaged upon reaching approximately 80% confluence 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, ESCs were converted to a transient EpiLC state as previously described [10]. First, 10 cm plates were coated with 16 µg/ml of Fibronectin for 1 hour at room temperature followed by two washes with PBS. ESCs 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 (ROCKi, 10 µM Y-27632) for the first 2 hours after plating, to reduce apoptosis [98, 99]. 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, after which time cells were highly confluent and cell death was observed. 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 (+BMP) or with FGF2, ACTIVIN A and WNT3A without BMP4 (-BMP) or FGF, ACTIVIN and WNT with the addition of 2 µM DMH1 (Sigma) (BMPi).

Immunostaining, 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 minutes 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 Key Resources Table. 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 [100], an interactive image classification software. Using this information, a 3D 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 [101], 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 artifacts. To generate plots of radial immunostaining fluorescence intensity, each voxel within a colony was assigned a distance from the colony center. The fluorescence intensity for each marker was measured per voxel and then the average fluorescence intensity of voxels at a particular radial position (binned into discrete radial bands) was calculated for each colony. The average radial fluorescence intensity across multiple colonies was then calculated. To display the expression of multiple different markers across the radii of colonies on the same scale, the relative level of each marker was quantified by normalizing to the highest level of expression (shown as 100) either across a time-course or within an individual time-point. Spatial patterning across multiple colonies was also demonstrated by generating average colony images for individual markers where each segmented cell was represented as a dot whose color indicates its fluorescence in the specified channel.

Coexpression analysis was carried out on a voxel level, i.e. the fluorescence level of each marker within a single voxel was calculated and plotted. For genes that were not expressed, or only expressed at low levels, at the start of the differentiation, gates could be drawn based on the fluorescence at 0 hours and used to quantify the percentage of total voxels expressing a particular marker at later time points.

Nuclear density measurements in micropatterns and embryos

The number of nuclei per 100 μm was quantified for 0 hours and 24 hours of micropattern differentiation utilizing the colony side view (z-axis) from confocal images acquired using a 40x objective at 0.5 μm interval steps. The number of nuclei was quantified across the entire width of the colony at 10 distinct positions and the average number of nuclei per 100 μm distance were calculated. For E5.5 embryos, the number of nuclei per 100 μm was quantified on sagittal confocal optical sections based on the number of nuclei within a sagittal optical section of the epiblast and the distance around the epiblast within the same section, manually measured using ImageJ software. For E6.5 the same was done using confocal images of transverse cryosections. Only cells within the epiblast were counted. Five embryos at E5.5 and five at E6.5 were analyzed in this manner.

Inter-nuclear distance was manually measured using ImageJ software. A line was drawn 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 Key Resources Table) 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; Josh Brickman for *Gsc*$^{GFP/+}$ *; Hhex*$^{RedStar/+}$ and Austin Smith for *Sox1*^{GFP} mESCs ; 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 under the supervision of JN and AKH. 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).

Competing interests

The authors have no competing interests to declare.

Ethics

Animal experimentation: All mice used in this study were maintained in accordance with the guidelines of the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee (IACUC) under protocol number 03-12-017 (PI Hadjantonakis).

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. Tewary, M., et al., *A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning.* Development, 2017.
9. Etoc, F., et al., *A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization.* Developmental Cell, 2016. **39**(3): p. 302-315.
10. 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.
11. 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.
12. Smith, A., *Formative pluripotency: the executive phase in a developmental continuum.* Development, 2017. **144**(3): p. 365-373.
13. Kalkan, T., et al., *Tracking the embryonic stem cell transition from ground state pluripotency.* Development, 2017.
14. 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).
15. 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.
16. Bedzhov, I. and M. Zernicka-Goetz, *Self-Organizing Properties of Mouse Pluripotent Cells Initiate Morphogenesis upon Implantation.* Cell, 2014. **156**(5): p. 1032-1044.
17. 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**.
18. Acampora, D., L.G. Di Giovannantonio, and A. Simeone, *Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition.* Development, 2013. **140**(1): p. 43-55.

19. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes & Development, 2003. **17**(1): p. 126-140.
20. Rosner, M.H., et al., *A Pou-Domain Transcription Factor in Early Stem-Cells and Germ-Cells of the Mammalian Embryo.* Nature, 1990. **345**(6277): p. 686-692.
21. Hart, A.H., et al., *Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human.* Developmental Dynamics, 2004. **230**(1): p. 187-198.
22. Guo, G., et al., *Klf4 reverts developmentally programmed restriction of ground state pluripotency.* Development, 2009. **136**(7): p. 1063-9.
23. 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.
24. Morkel, M., et al., *Beta-catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation.* Development, 2003. **130**(25): p. 6283-94.
25. Osorno, R., et al., *The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression.* Development, 2012. **139**(13): p. 2288-2298.
26. Di-Gregorio, A., et al., *BMP signalling inhibits premature neural differentiation in the mouse embryo.* Development, 2007. **134**(18): p. 3359-69.
27. Wilkinson, D.G., S. Bhatt, and B.G. Herrmann, *Expression pattern of the mouse T gene and its role in mesoderm formation.* Nature, 1990. **343**(6259): p. 657-9.
28. Huang, S.M., et al., *Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling.* Nature, 2009. **461**(7264): p. 614-20.
29. Kim, H., et al., *Modulation of beta-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal.* Nat Commun, 2013. **4**: p. 2403.
30. Kurek, D., et al., *Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells.* Stem Cell Reports, 2015. **4**(1): p. 114-28.
31. Tsakiridis, A., et al., *Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors.* Development, 2014. **141**(6): p. 1209-21.
32. Wu, J., et al., *An alternative pluripotent state confers interspecies chimaeric competency.* Nature, 2015. **521**(7552): p. 316-21.
33. 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.
34. Robertson, E.J., *Dose-dependent Nodal/Smad signals pattern the early mouse embryo.* Semin Cell Dev Biol, 2014. **32**: p. 73-9.
35. Hoffman, J.A., C.I. Wu, and B.J. Merrill, *Tcf7l1 prepares epiblast cells in the gastrulating mouse embryo for lineage specification.* Development, 2013. **140**(8): p. 1665-75.
36. 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.
37. Burtscher, I. and H. Lickert, *Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo.* Development, 2009. **136**(6): p. 1029-1038.
38. Engert, S., et al., *Wnt/beta-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse.* Development, 2013. **140**(15): p. 3128-3138.

39. Deschamps, J. and J. van Nes, *Developmental regulation of the Hox genes during axial morphogenesis in the mouse.* Development, 2005. **132**(13): p. 2931-2942.
40. Thomas, P.Q., A. Brown, and R.S. Beddington, *Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors.* Development, 1998. **125**(1): p. 85-94.
41. Bedford, F.K., et al., *Hex - a Novel Homeobox Gene Expressed during Hematopoiesis and Conserved between Mouse and Human.* Nucleic Acids Research, 1993. **21**(5): p. 1245-1249.
42. Sakamoto, Y., et al., *Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos.* Biochemical and Biophysical Research Communications, 2007. **360**(3): p. 539-544.
43. Burtscher, I., et al., *The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development.* Genesis, 2012. **50**(6): p. 496-505.
44. Freyer, L., et al., *A loss-of-function and H2B-Venus transcriptional reporter allele for Gata6 in mice.* Bmc Developmental Biology, 2015. **15**.
45. Choi, E., et al., *Dual lineage-specific expression of Sox17 during mouse embryogenesis.* Stem Cells, 2012. **30**(10): p. 2297-308.
46. Peterson, R.S., et al., *The winged helix transcriptional activator HFH-8 is expressed in the mesoderm of the primitive streak stage of mouse embryos and its cellular derivatives.* Mechanisms of Development, 1997. **69**(1-2): p. 53-69.
47. Fleury, M., et al., *FOXF1 inhibits hematopoietic lineage commitment during early mesoderm specification.* Development, 2015. **142**(19): p. 3307-20.
48. Morrisey, E.E., et al., *GATA-6: A zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm.* Developmental Biology, 1996. **177**(1): p. 309-322.
49. 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.
50. Wu, T., A.K. Hadjantonakis, and S. Nowotschin, *Visualizing endoderm cell populations and their dynamics in the mouse embryo with a Hex-tdTomato reporter.* Biology Open, 2017. **6**(5): p. 678-687.
51. Rodriguez, T.A., et al., *Distinct enhancer elements control Hex expression during gastrulation and early organogenesis.* Developmental Biology, 2001. **234**(2): p. 304-316.
52. Belo, J.A., et al., *The prechordal midline of the chondrocranium is defective in Goosecoid-1 mouse mutants.* Mechanisms of Development, 1998. **72**(1-2): p. 15-25.
53. Sasaki, H. and B.L.M. Hogan, *Differential Expression of Multiple Fork Head-Related Genes during Gastrulation and Axial Pattern-Formation in the Mouse Embryo.* Development, 1993. **118**(1): p. 47-59.
54. 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.
55. Monaghan, A.P., et al., *Postimplantation Expression Patterns Indicate a Role for the Mouse Forkhead/Hnf-3 Alpha,Beta and Gamma Genes in Determination of the Definitive Endoderm, Chordamesoderm and Neuroectoderm.* Development, 1993. **119**(3): p. 567-578.

56. Ang, S.L., et al., *The Formation and Maintenance of the Definitive Endoderm Lineage in the Mouse - Involvement of Hnf3/Forkhead Proteins.* Development, 1993. **119**(4): p. 1301-1315.
57. Altaba, A.R.I., et al., *Sequential Expression of Hnf-3-Beta and Hnf-3-Alpha by Embryonic Organizing Centers - the Dorsal Lip/Node, Notochord and Floor Plate.* Mechanisms of Development, 1993. **44**(2-3): p. 91-108.
58. Villegas, S.N., et al., *PI3K/Akt1 signalling specifies foregut precursors by generating regionalized extra-cellular matrix.* Elife, 2013. **2**: p. e00806.
59. Crompton, M.R., et al., *Identification of a novel vertebrate homeobox gene expressed in haematopoietic cells.* Nucleic Acids Res, 1992. **20**(21): p. 5661-7.
60. 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.
61. 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.
62. Kang, Y.B. and J. Massague, *Epithelial-mesenchymal transitions: Twist in development and metastasis.* Cell, 2004. **118**(3): p. 277-279.
63. 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.
64. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-72.
65. 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.
66. Nomura, M. and E. Li, *Smad2 role in mesoderm formation, left-right patterning and craniofacial development.* Nature, 1998. **393**(6687): p. 786-90.
67. 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.
68. 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.
69. 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.
70. 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.
71. Ying, Q.L., et al., *Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture.* Nature Biotechnology, 2003. **21**(2): p. 183-186.
72. 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.
73. Brons, I.G., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-5.
74. 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.
75. 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.

76. Behringer, R.R., et al., *A flattened mouse embryo: leveling the playing field.* Genesis, 2000. **28**(1): p. 23-30.
77. 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.
78. 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).
79. 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.
80. 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).
81. 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.
82. 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.
83. van den Brink, S.C., et al., *Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells.* Development, 2014. **141**(22): p. 4231-42.
84. 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.
85. 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.
86. 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.
87. Mishina, Y., et al., *Multiple roles for activin-like kinase-2 signaling during mouse embryogenesis.* Developmental Biology, 1999. **213**(2): p. 314-326.
88. 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.
89. 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.
90. Huang, Y.L., et al., *In Vivo Differentiation Potential of Epiblast Stem Cells Revealed by Chimeric Embryo Formation.* Cell Reports, 2012. **2**(6): p. 1571-1578.
91. 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.
92. Hooper, M., et al., *HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells.* Nature, 1987. **326**(6110): p. 292-5.
93. 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.
94. 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.

95. Abe, M. and M.C. Naski, *Regulation of sprouty expression by PLCgamma and calcium-dependent signals.* Biochem Biophys Res Commun, 2004. **323**(3): p. 1040-7.
96. 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.
97. Morgani, S.M., et al., *Totipotent embryonic stem cells arise in ground-state culture conditions.* Cell Rep, 2013. **3**(6): p. 1945-57.
98. 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.
99. Watanabe, K., et al., *A ROCK inhibitor permits survival of dissociated human embryonic stem cells.* Nat Biotechnol, 2007. **25**(6): p. 681-6.
100. 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.
101. 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-E6.0 embryos. Yellow dashed line demarcates Epi. Scale bars, 25 µm. Non-nuclear anti-BRACHYURY/CDX2/POU3F1 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/POU3F1; n=14 BRACHYURY/FOXA2. BRA, BRACHYURY.

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/SOX2 VE fluorescence likely 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). Orange dashed lines delineate regions of interest. **C.** Quantification (5 sections /embryo /stage) of SOX2 and NANOG in manually selected (panel B) anterior (A) and posterior (P) Epi of E6.5-E6.75 and E7.0-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. Upper 2 panels represent a merge of the markers shown below. 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, POU5F1/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 BRACHYURY/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). Orange dashed lines delineate regions of interest. **E,F,H.** MIPs of immunostained micropatterns. High magnification shows region at the colony edge. Yellow arrowhead H marks GATA6/FOXF1 cell. Scale bars, 100 µm. **G.** High magnification of colony edge. Outer domain represents a mixture of populations often organized in clusters, highlighted by 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.** Schematic diagram summarizing the cell identities observed 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.

Figure 5. Smaller diameter colonies pattern in the same order of events but lose central populations. A. EpiLCs were differentiated with FGF2 and ACTIVIN A (F/A), BMP4 and WNT3A as described in Figure 2E. Confocal optical section of a representative 1000 µm diameter colony after differentiation. Dashed circles define 3 regions of distinct marker expression, shown at higher magnification in adjacent panels. While SOX2 is expressed quite broadly, regions were defined based on the marker that was predominantly expressed. Region A (central) = SOX2 (blue), Region B (intermediate) = BRACHYURY (red), Region C (outer) = CDX2 (green). Scale bars, 25 µm. **B.** Schematic diagram showing the changing marker expression in colonies of different diameters. **C-F.** Representative confocal maximum intensity projections of colonies at 0, 24, 48 and 72 hours after addition of BMP4 and WNT3A to F/A medium.

Images show colonies of 500 µm, 225 µm, 140 µm and 80 µm diameter. Scale bars, 100 µm.

Figure 6. BMP signaling in micropatterns and embryos correlates with embryonic and extraembryonic mesoderm fates. A. Transverse cryosection of immunostained embryo in figure supplement 1B. Scale bar, 50 µm. **B.** 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 Epi. Quantification was carried out on 3 cryosections per embryo. N, number of cells. Data represents mean fluorescence intensity +/- S.D. normalized to Hoechst fluorescence. **C.** 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. **D.** Depiction of spatial patterning across multiple colonies. Each dot represents a single cell. **E.** 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.

Figure 7. 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.** Schematic diagram summarizing the cell fates observed after 72 hours *in vitro* differentiation under conditions described in A and corresponding *in vivo* cell types at E7.75-E8.0. 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 anterior primitive streak or axial mesoderm cells.

Figure 8. Epiblast stem cells undergo definitive endoderm differentiation in the presence or absence of BMP. A. Epiblast stem cells (EpiSCs) of the EpiSC9 line [75] 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/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 [10]. Data shown is from amplified RNA samples and represents the mean +/- S.D for two independent replicates.

Figure 9. 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-E6.5. 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-E7.25 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, SOX17, *Gsc* and *Hhex* representing definitive endoderm (DE), FOXA2 and BRACHYURY representing anterior PS (A-PS) or AxM and or FOXA2 and *Gsc* likely representing A-PS. 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-E6.0) early (E6.5-E6.75), mid- (E7.0-E7.5) and late (E7.75-E8.0) 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 the schematic diagram. N, number of embryos. **D.** Box plots showing human embryonic disc measurements compiled from human embryo data collections. Abnormal embryo data were excluded. D = embryonic day. Carnegie stage 6a, pre-streak; 6b, early streak; 7, early-mid gastrulation.

Figure Supplement Legends

Figure 1 – figure supplement 1. Micropatterned EpiLC colonies begin as an epithelial monolayer that increases density over time. A. Confocal optical section of a region of an immunostained EpiLC micropatterned colony at 0 hours. Scale bar, 50 µm. **B.** Quantification of nuclei/100 µm and distance between nuclei (in µm) *in vitro* after 0 hours and 24 hours of micropattern differentiation and *in vivo* at embryonic day (E) 5.5 and 6.5. For nuclei/100 µm, the number of nuclei across the width of a micropatterned colony at 10 different positions was quantified. *In vivo* quantification was carried out on sections of 5 different embryos at E5.5 and 6.5. For quantification of inter-nuclear distance, 0 and 24 hours of micropattern differentiation: n=150, E5.5: n=128 (5 embryos), E6.5: n=189. Quantification details can be found in the methods section. **C.** Schematic diagram of quantification of radial fluorescence intensity of immunostaining on micropatterns.

Figure 2 – figure supplement 1. FGF, ACTIVIN and endogenous WNT induce a primitive streak state. A. Confocal sagittal optical section of an early streak embryo. Non-nuclear anti-BRACHYURY VE fluorescence likely represents non-specific binding. PS, primitive streak; A, anterior; P, posterior; Pr, proximal; D, distal. Scale bars, 50 µm. **B.** EpiLCs were generated as in Figure 1C. EpiLCs were plated overnight onto Laminin-coated micropatterned surfaces (-24 hours) in N2B27 medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A (F/A). Cells were cultured for a further 72 hours in either N2B27 with F/A or F/A with 10 µM of XAV939 (Wnt signaling inhibitor, WNTi). **C.** Representative confocal maximum intensity projections of immunostained 1000 µm diameter colonies after 72 hours in the conditions in panel D. Scale bars, 100 µm. **D.** Quantification of SOX2 and BRACHYURY immunostaining voxel fluorescence intensity, in arbitrary units (a.u.), from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies. Control: n=32, WNTi: n=7. **E.** Schematic diagram depicting cell fates generated after 72 hours of *in vitro* micropattern differentiation with FGF2 and ACTIVIN A (+XAV) or FGF2, ACTIVIN A and endogenous Wnt signaling and corresponding *in vivo* cell types.

Figure 2 – figure supplement 2. Robust micropattern differentiation of EpiLCs. A. Confocal maximum intensity projection (MIP) showing micropatterned colonies of 1000,

500, 250, 140 and 80 µm diameter. Scale bar, 1000 µm. **B.** Left panels show confocal MIP images of 5 independent 1000 µm diameter colonies (C1-5). Middle panels show quantification of SOX2, BRACHYURY and CDX2 immunostaining voxel fluorescence intensity, in arbitrary units (a.u.), from colony center (0) to edge (500). Data relative to maximum voxel intensity for each marker. Right panel shows average radial profile of these 5 colonies. **C.** While much of the work in this study utilized E14 ESCs, comparable patterning was observed with other mouse ESC lines. Quantification of immunostaining voxel fluorescence intensity of differentiated $T^{GFP/+}$ and $Sox17^{GFP/+}$ cell lines, in arbitrary units (a.u.), from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies relative to maximum voxel intensity for each marker. For $T^{GFP/+}$ cells, GATA6, SOX2: n=10, BRACHYURY, CDX2, T^{GFP}: n=11. For $Sox17^{GFP/+}$ cells, SOX2: n=5, BRACHYURY, CDX2, $Sox17^{GFP}$: n=6, GATA6: n=5. **D.** Confocal MIPs of immunostained colonies differentiated as in Figure 2E. Scale bars, 100 µm. **E.** High magnification confocal image of the colony edge after 24 hours in conditions defined in Figure 2E. Yellow arrowheads mark cells coexpressing BRACHYURY and SOX2.

Figure 3 – figure supplement 1. Populations arising during *in vivo* and *in vitro* differentiation. A. Schematic diagram showing cell types arising during mouse gastrulation. Upper panels show sagittal sections. Lower panels show transverse sections. Dotted lines demarcate approximate position of transverse sections. A, anterior; P, posterior; Pr, proximal; D, distal; R, right; L, left. Cell fates color-coded as in legend. Epi, epiblast; ExM, extraembryonic mesoderm; YSM, yolk sac mesoderm; AOM, allantois outer mesenchyme; ACD, allantois core domain; TE, trophectoderm; VE, visceral endoderm. **B-F.** Confocal maximum intensity projections (MIP), sagittal optical sections and transverse cryosections of gastrulating mouse embryos. Dashed lines mark transverse plane. Non-nuclear anti- BRACHYURY/CDX2 VE fluorescence represents non-specific binding. PS, primitive streak; M1, Mesoderm 1; M2, Mesoderm 2; Meso, mesoderm; Epi, epiblast; ExE, extraembryonic ectoderm; ExM, extraembryonic mesoderm; ACD, allantois core domain; ML, midline, A, anterior; P, posterior; Pr, proximal; D, distal. Scale bars, 50 µm. **D.** Images of transverse cryosections through posterior extraembryonic region of the allantois of the embryo. **E.** Yellow arrowhead marks SOX17/FOXA2 -coexpressing definitive endoderm cell. **G.** MIP of immunostained micropatterned colony. High magnification views shows a region at the colony edge. Scale bars, 100 µm. **H.** Quantification of marker coexpression by voxel. Each dot

indicates the fluorescence intensity of a single voxel, in arbitrary units (a.u.). Color represents voxel density within the plot. Gates were defined based on the 0 hour time point where BRACHYURY, CDX2 and GATA6 were not expressed. Numbers within each quadrant represent the % of voxels within gate, rounded to the nearest whole number. N, number of colonies. **I,J.** Pie charts illustrating % of BRACHYURY-positive cells that coexpress CDX2 or GATA6 (I) and % of CDX2 or GATA6-positive cells coexpressing BRACHYURY (J). Percentages shown within the largest fraction. **K.** High magnification confocal image of a region of the outer colony edge after 48 hours of differentiation under the conditions described in Figure 2E. Yellow arrowheads highlight GATA6/BRACHYURY-coexpressing cells. Scale bars, 50 µm.

Figure 3 – figure supplement 2. Micropatterns in the presence of BMP express SOX17 and *Hhex* but not *Gsc*. EpiLCs were differentiated as in Figure 2E, with FGF2, ACTIVIN A, BMP4, WNT3A for 72 hours. **A.** High magnification confocal optical section of a region of the colony edge after 72 hours of differentiation with FGF2, ACTIVIN A, WNT3A, BMP4. **B.** Confocal MIPs of differentiated and immunostained $Gsc^{GFP/+}$;$Hhex^{RedStar/+}$ dual reporter cells. Scale bars, 100 µm.

Figure 4 – figure supplement 1. Cells undergo an epithelial to mesenchymal transition during gastrulation and *in vitro* differentiation. A. Quantification of immunostaining voxel fluorescence intensity (in arbitrary units, a.u.) of SOX2, BRACHYURY and CDX2 from colony center (0) to edge (500) as well as relative colony height. Data represents average voxel intensity or change in colony height across multiple colonies relative to the maximum value for each marker. **B.** Confocal optical sections of the outer colony edge. Scale bars, 100 µm. **C.** MIPs of immunostained 72 hour colonies. Scale bars, 100 µm. **D.** Images of z-axis profile from colony center (left) to edge (right). **E.** Quantification of immunostaining. Voxel fluorescence intensity was measured from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies (n=9) relative to maximum voxel intensity for each marker. **F.** EpiLCs were plated onto micropatterns overnight in N2B27 with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A (F/A). The following day medium was changed to 12 ng/ml FGF2, 50 ng/ml BMP4, 200 ng/ml WNT3A and 10 µM SB431542 (ACTIVINi). Confocal maximum intensity projections of immunostained colonies after 72 hours of differentiation. Scale bars, 100 µm. BRA, BRACHYURY.

Figure 5 – figure supplement 1. Patterning of cell fates is altered at different colony diameters. EpiLCs were differentiated as previously described in Figure 2E on different diameter micropatterns. Images show confocal maximum intensity projections of immunostained colonies after 72 hours of differentiation. Three representative examples are shown for each colony diameter. Scale bars, 100 µm.

Figure 6 – figure supplement 1. BMP signaling is active in the posterior primitive streak, embryonic and extraembryonic mesoderm. A. Schematic diagram depicting the sources of BMP (pink) within gastrulating embryos, described in [69]. At the start of gastrulation (E6.5-E6.75), BMP4 is expressed by the extraembryonic ectoderm and later (E7.5-E8.0) in the allantois, amnion and chorion. **B,D,E.** Representative confocal images of immunostained gastrulating embryos showing BMP signaling activity based on nuclear localization of pSMAD1/5/8. Images represent maximum intensity projections (MIP), sagittal optical sections and transverse cryosections. Transverse section 'b' from panel B is shown in Figure 6A. Non-nuclear anti-BRACHYURY VE fluorescence represents to non-specific binding. AVE, anterior visceral endoderm; Epi, epiblast; ExE, extraembryonic ectoderm; VE, visceral endoderm; A-PS, anterior primitive streak; Al, allantois; M1, mesoderm 1; M2, mesoderm 2; A, anterior; P, posterior; Pr, proximal; D, distal; R, right; L, left; BRA, BRACHYURY; E-CAD, E-CADHERIN; pS1/5/8, pSMAD1/5/8. Scale bars, 50 µm. Dashed lines mark transverse plane. Dashed boxes outline regions in higher magnification in lower panels. In F, white bracket demarcates the posterior primitive streak and red bracket the anterior primitive streak. **C,G** Scatter plots showing the levels of pSMAD1/5/8 and BRACHYURY in arbitrary units (a.u.) within single mesoderm cells of early streak (E6.5-E6.75 - C) and early headfold (E7.75-E8.0 - G) embryos. Each dot represents a single cell. Linear regression curves were fitted to the points (red line). **F.** Quantification of pSMAD1/5/8 and BRACHYURY fluorescence intensity in arbitrary units (a.u.) in early headfold (E7.75-E8.0) embryos. Cells within the epiblast, primitive streak (PS) and Mesoderm 1 and 2 (Meso1/2) were manually selected on confocal images of transverse cryosections using ImageJ software as shown in right-hand panel. The PS was identified as BRACHYURY-positive cells at the embryo posterior. Mesoderm cells were separated into 2 categories, those that had migrated further anteriorly and expressed low BRACHYURY and high pSMAD1/5/8 (Meso1) and those closest to the PS expressing high BRACHYURY (Meso2). Quantification was carried out on 3 cryosections per embryo. N, number of cells. Data represents mean

fluorescence intensity normalized to the fluorescence intensity of Hoechst nuclear stain +/- S.D. **H.** Quantification of marker coexpression by voxel. Each dot indicates the fluorescence intensity of a single voxel, in arbitrary units (a.u.). Color represents voxel density within the plot. Gates were defined based on the 0 hour time point where BRACHYURY and CDX2 were not expressed and pSMAD1/5/8 was expressed only at low levels. Numbers within each quadrant represent the % of voxels within gate, rounded to the nearest whole number. N, number of colonies. **I.** Pie charts illustrating the % of pSMAD1/5/8-positive cells that coexpress BRACHYURY or CDX2. Percentages shown within the largest fraction.

Figure 6 – figure supplement 2. Bypassing the WNT receptor alters spatial patterning. A. Simplified schematic of the Wnt pathway. In the absence of WNT receptor binding (left panel), GSKβ phosphorylates β-catenin, targeting it for degradation. β-catenin target genes are inactive. When WNT binds to the receptor (middle panel), factors downstream of the receptor are activated and inhibit GSKβ activity. β-catenin is not phosphorylated and not degraded, but instead translocates to the nucleus and activates target genes. CHIR99021 (CHIR) is a small molecule inhibitor of GSKβ. CHIR directly inactivates GSKβ independent of WNT receptor activation (right panel). As with WNT receptor binding, inactive GSKβ cannot phosphorylate β-catenin hence target genes are activated. **B.** Schematic diagram of differentiation conditions. EpiLCs were generated as described in Figure 1C. EpiLCs were plated overnight onto Laminin-coated micropatterns (-24 hours) in N2B27 medium with 12 ng/ml FGF2 and 20 ng/ml ACTIVIN A (F/A). Various conditions were then used for further differentiation, F/A with BMP4 (50 ng/ml) and WNT3A (200 ng/ml) (Control) or F/A, BMP4 and 3 μM CHIR99021, a small molecule inhibitor of GSK3. Colonies were analyzed after 72 hours differentiation. **C,E.** Confocal maximum intensity projections of immunostained colonies after 72 hours of differentiation. Scale bars, 100 μm. **D,F.** Quantification of immunostaining voxel fluorescence intensity, in arbitrary units (a.u.), from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies.

Figure 7 – figure supplement 1. Anterior primitive streak and definitive endoderm populations are formed in the absence of BMP. *Gsc*$^{GFP/+}$;*Hhex*$^{RedStar/+}$ dual reporter ESCs were differentiated as in Figure 7A, with FGF2, ACTIVIN A, WNT3A for 72 hours. **A,B.** Confocal MIPs of immunostained colonies. Scale bars, 100 μm. **C.** Quantification of

immunostaining voxel fluorescence intensity in arbitrary units (a.u.), from colony center (0) to edge (500). Data represents average voxel intensity across multiple colonies (n=10/time point) relative to maximum voxel intensity for each marker. **D,F.** High magnification confocal optical sections of a region of the colony edge after 72 hours of differentiation. **D.** Red arrowheads mark FOXA2/*Gsc*^{GFP}/*Hhex*^{RedStar} -expressing cells. Cyan arrowheads mark FOXA2/*Gsc*^{GFP} -expressing cells. **F.** Yellow arrowheads mark *Gsc*^{GFP}/*Hhex*^{RedStar} –expressing cells. White arrowheads mark *Gsc*^{GFP}/BRACHYURY –expressing cells. Scale bars, 25 µm **E,G.** Quantification of marker coexpression by voxel. Each dot indicates the fluorescence intensity of a single voxel, in arbitrary units (a.u.). Color represents voxel density within the plot. Gates were defined based on the 0 hour time point when FOXA2, BRACHYURY, *Gsc*^{GFP} and *Hhex*^{RedStar} were not expressed. Numbers within each quadrant represent the % of voxels within gate, rounded to the nearest whole number. n=10 colonies/time point.

Figure 7 – figure supplement 2. An anterior epiblast/neurectoderm population is formed in the absence of BMP. *Sox1*$^{GFP/+}$ 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) or F/A, WNT3A (-BMP). Colonies were analyzed after 72 hours differentiation. **A,B.** MIPs of immunostained 72 hour colonies. *Sox1*$^{GFP/+}$ cells were stained with an anti-GFP antibody. 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 relative to maximum voxel intensity for each marker. n=10. **D.** Confocal maximum intensity projections (MIP) and transverse cryosections of an E8.0 mouse embryo. Dashed lines mark transverse plane. Al, allantois; NE, neurectoderm; ADE, anterior definitive endoderm; A, anterior; P, posterior; Pr, proximal; D, distal; L, left; R, right. Scale bars, 100 µm.

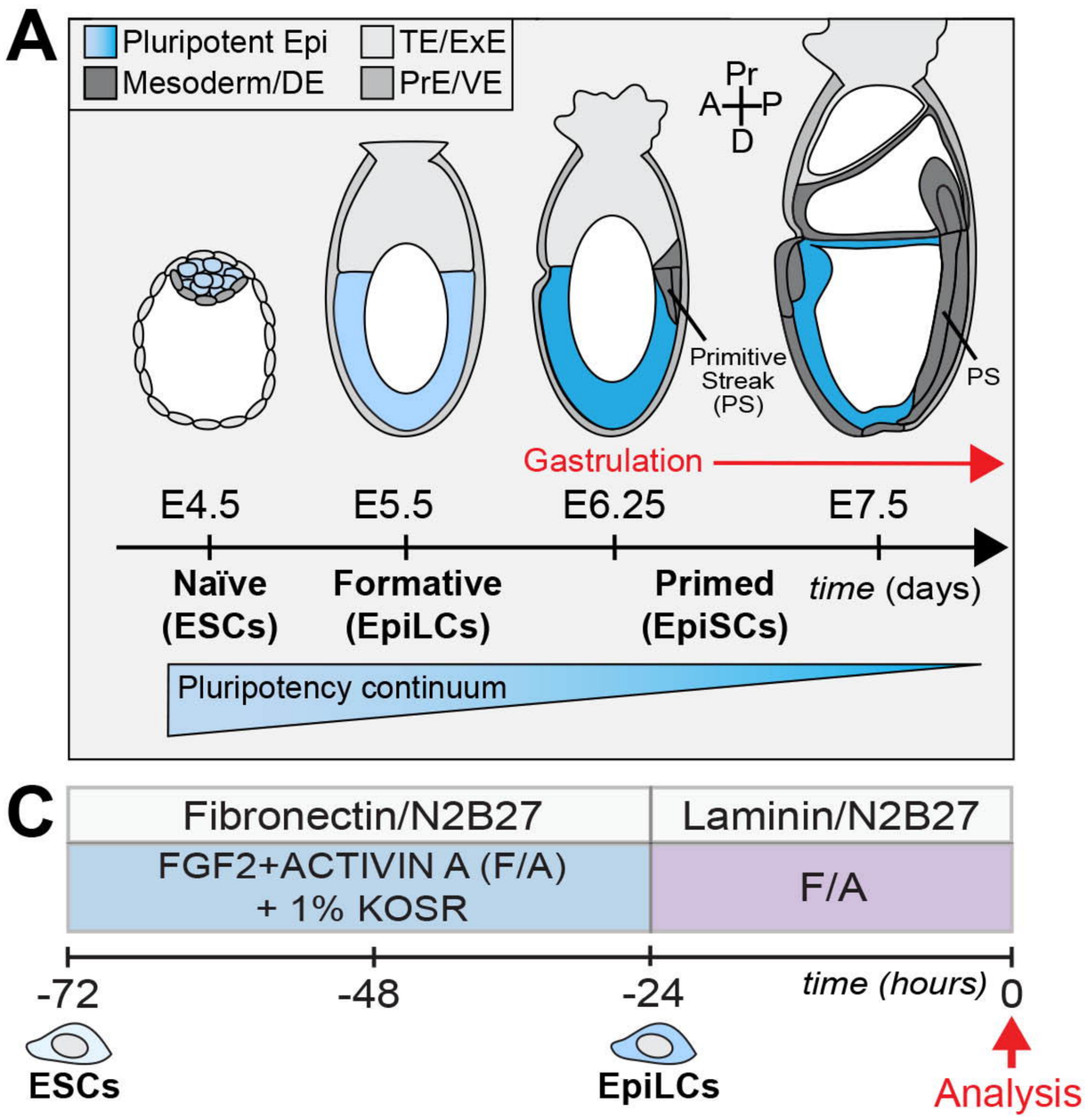
Supplemental Table 1. Summary of expression data for a panel of factors used for cell fate assignments in this study. Where data was not generated in this study, appropriate literature references are provided – listed in bold lowercase letters, full references provided below the table. Where primary data is included in this study, the reference is given in the following format – (Main text figure (Fig._) – figure supplement (S_): reference for figures within additional data resource). Cell fates are color-coded as

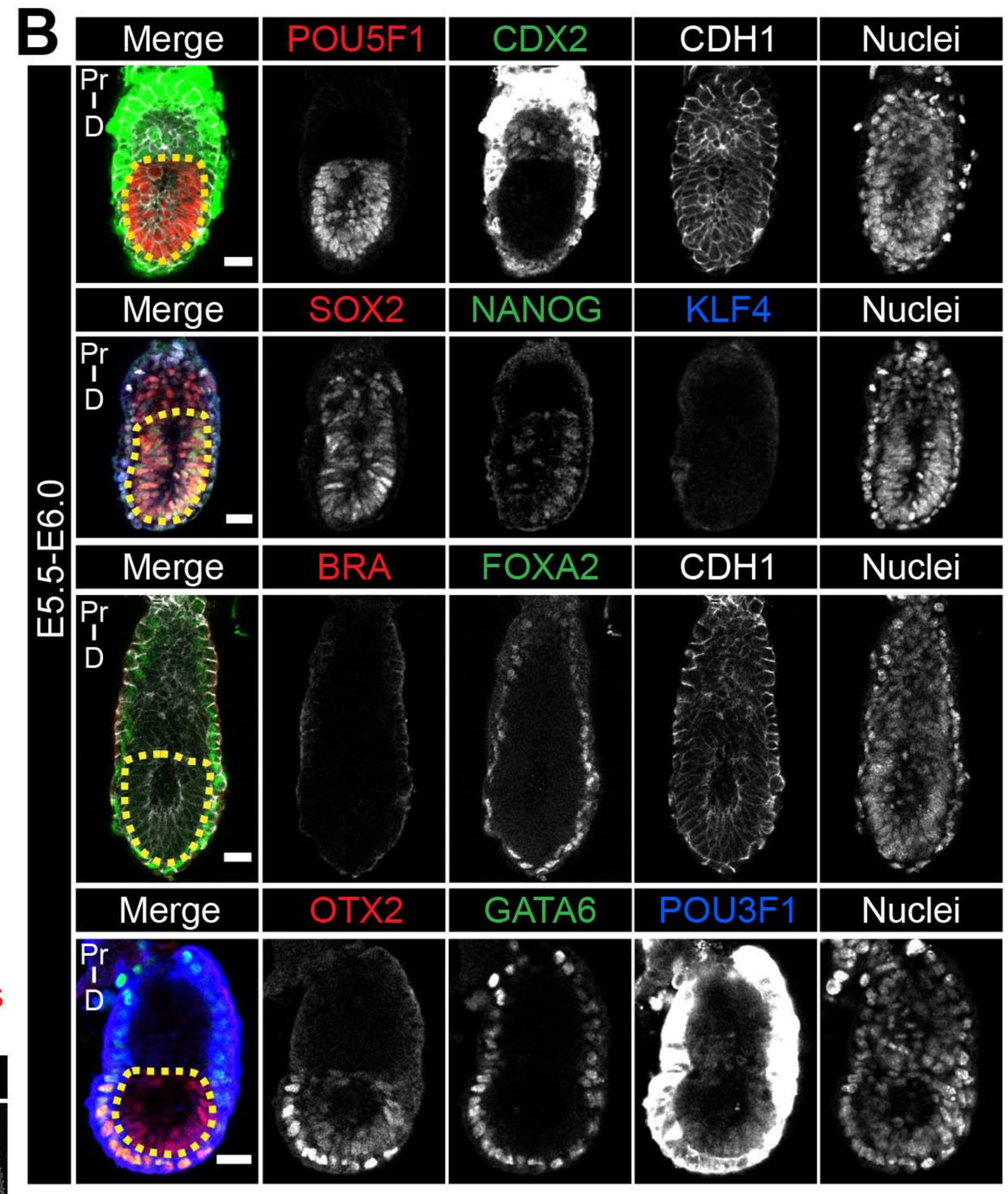
in Figure 3 – figure supplement 1A and in table legend. At each gastrulation stage, each cell type was classified as 1 = factor expressed, 0.5 = factor expressed heterogeneously or at low levels, 0 = not expressed (white), ? = expression unclear from immunostaining and exact localization is unclear/unknown from published literature. FOXF1 is expressed in the embryonic mesoderm and also reported to be expressed within the PS at E7.0-E7.5 but PS expression is unclear. *Gsc* and *Hhex* are both expressed in the AxM although their expression has been reported to be transient in some studies and the exact timing of expression is not clear. To note, some cells of the PS coexpressed BRACHYURY, SOX2 and NANOG, likely a transition state between Epi and PS before SOX2 and NANOG are downregulated. It was unclear whether GATA6 was absent or expressed only at low levels in the ACD and early AOM. Certain markers were expressed in some but not all embryos at a particular stage – likely indicating the onset of expression, for example SOX17 within the ExM at E7.0-E7.5. ******* At E7.0-E7.5 both DE and mesoderm cells arise from the anterior PS and are spatially intermixed hence it is difficult to identify DE cells based only on localization. At this time DE therefore refers to presumptive DE cells arising from the anterior PS based on known marker expression.

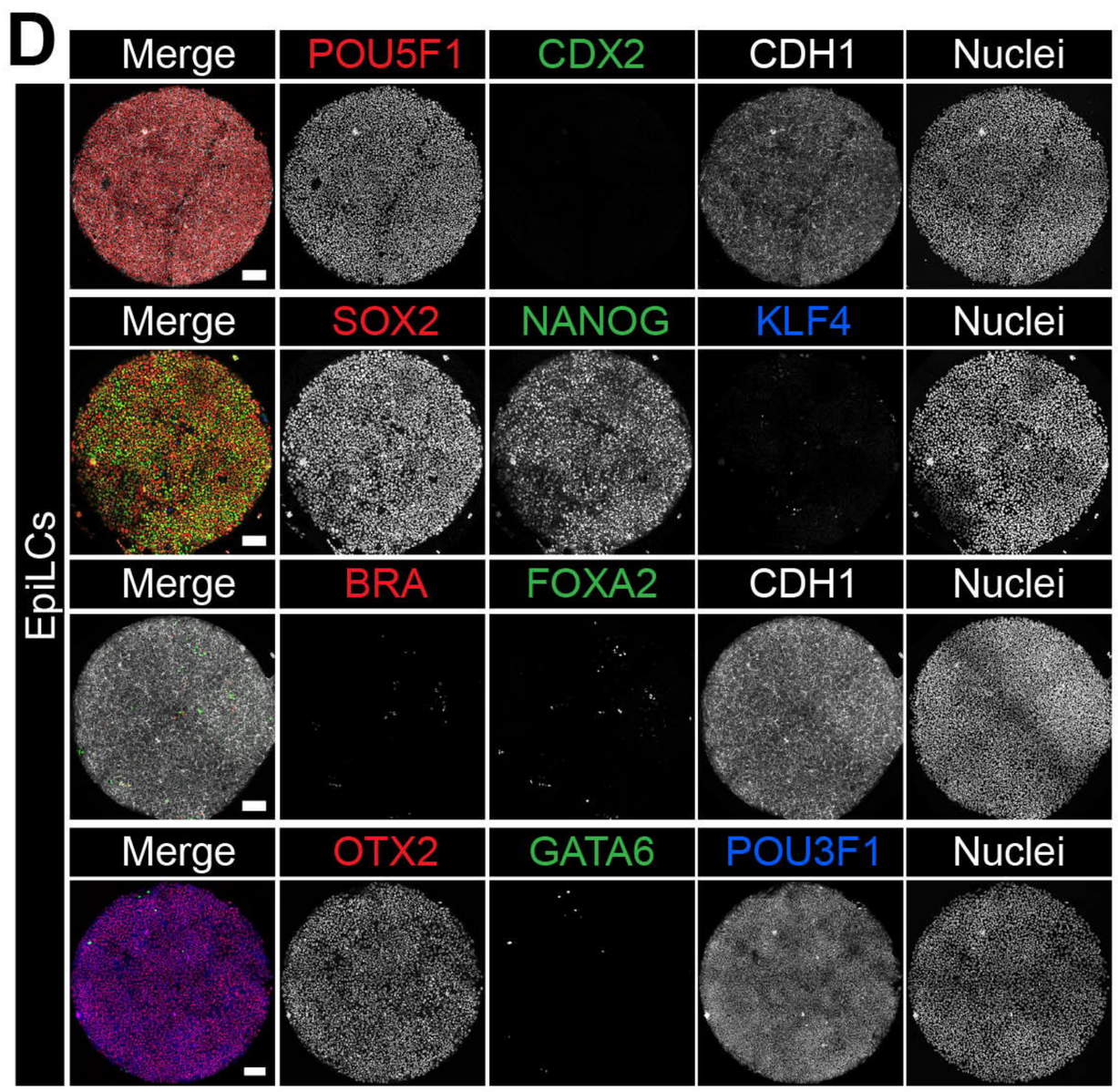
Table 1. Summary of cell fates arising in the presence of various signaling factors. Cell fates generated after 72 hours of mouse micropattern differentiation with described cytokine combinations. It should be noted that cells do not detect these signals homogeneously. WNT inhibition (WNTi) refers to XAV treatment; ACTIVIN inhibition (ACTIVINi) refers to the absence of ACTIVIN and SB431542.

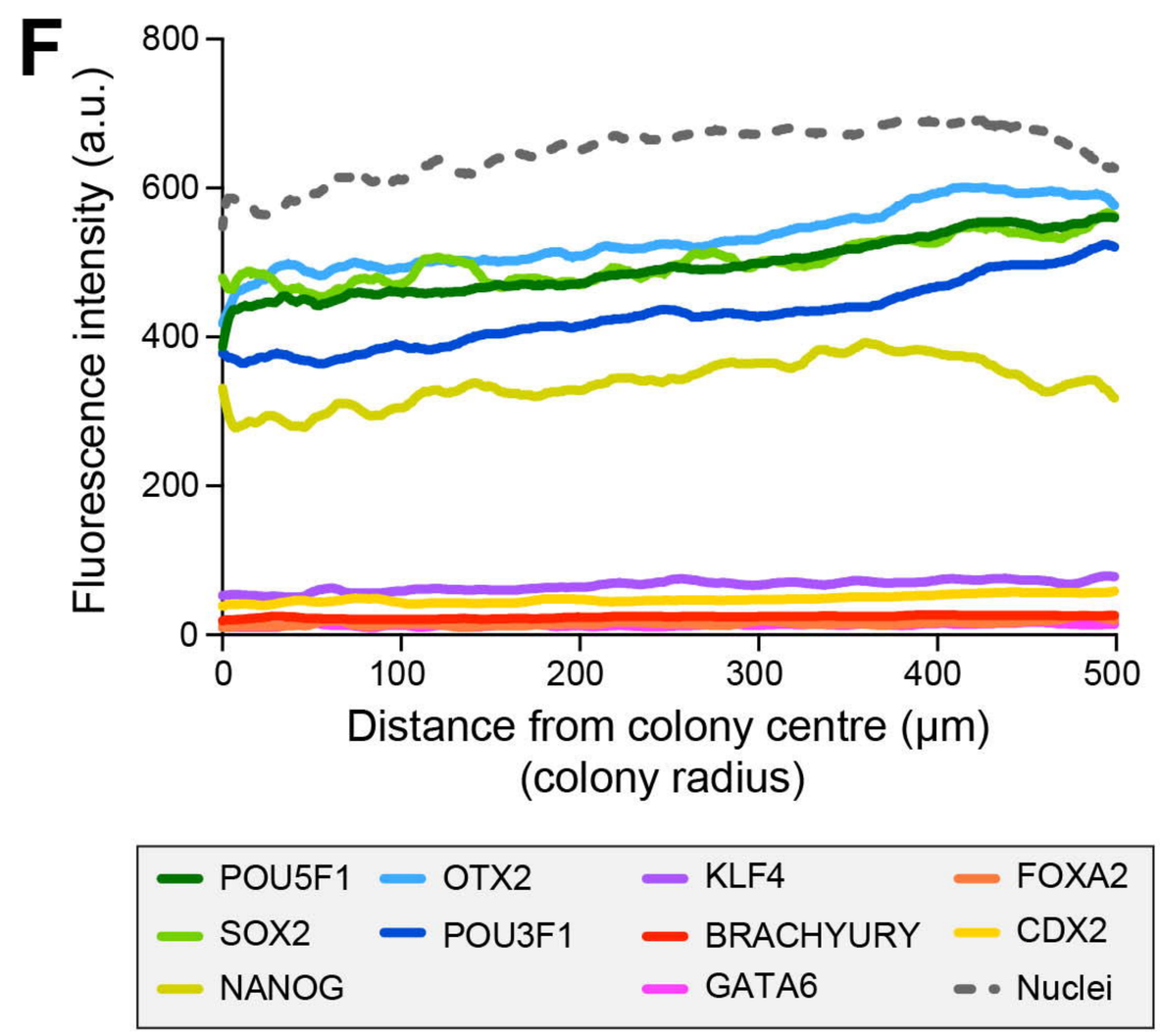
Signaling pathways	Outcome
FGF, ACTIVIN, (WNTi)	Epiblast
FGF, ACTIVIN, endogenous WNT	Epiblast PS
FGF, ACTIVIN, BMP, WNT	Posterior epiblast PS Extraembryonic mesoderm Embryonic mesoderm
FGF, ACTIVIN, WNT	Anterior epiblast Definitive endoderm Axial mesoderm or Anterior PS
FGF, BMP, WNT (ACTIVINi)	Epiblast

Figure 1, Morgani et al.









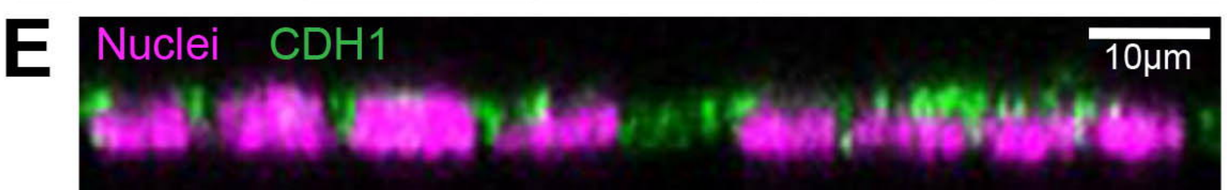


Figure 2, Morgani et al.

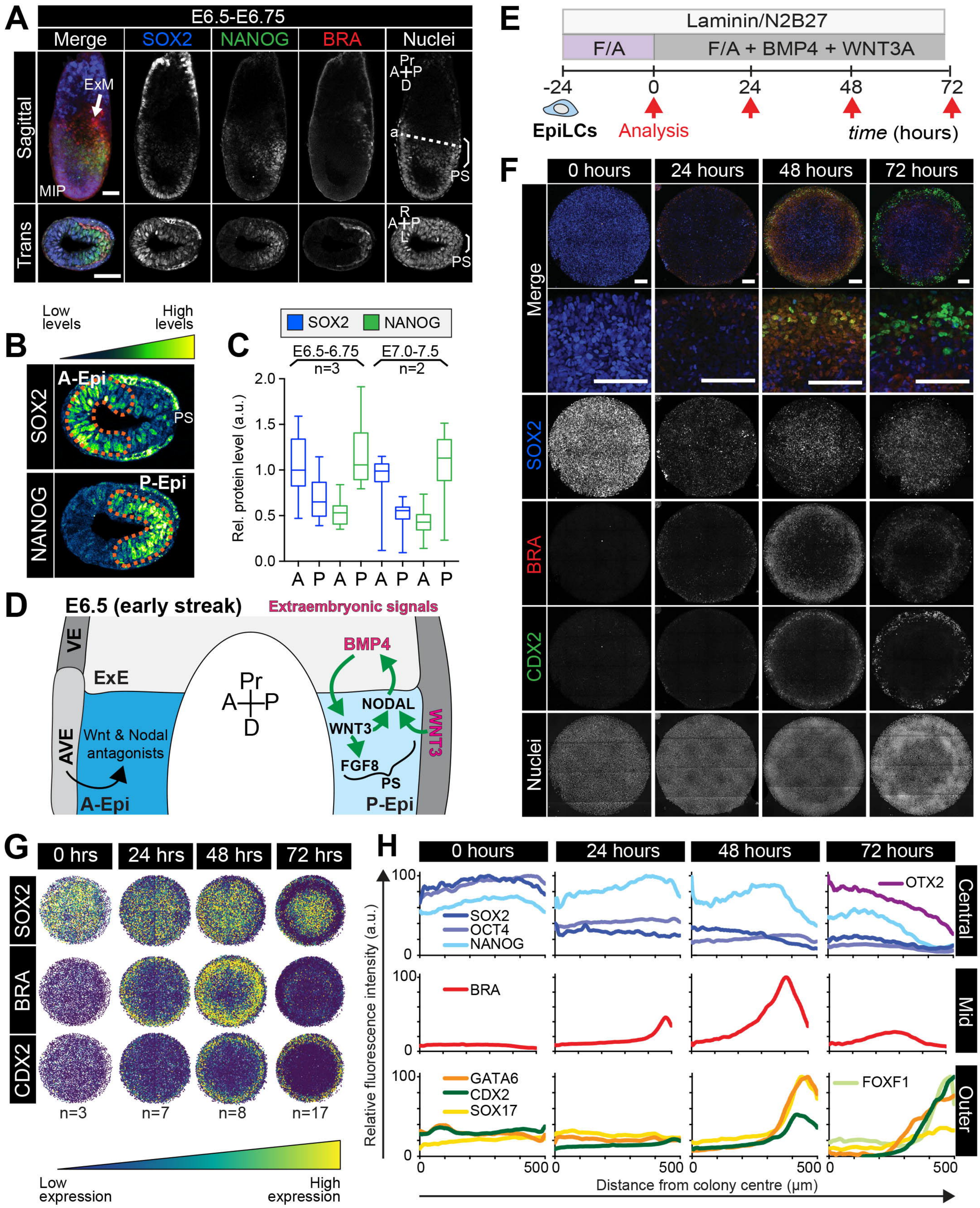


Figure 3, Morgani et al.

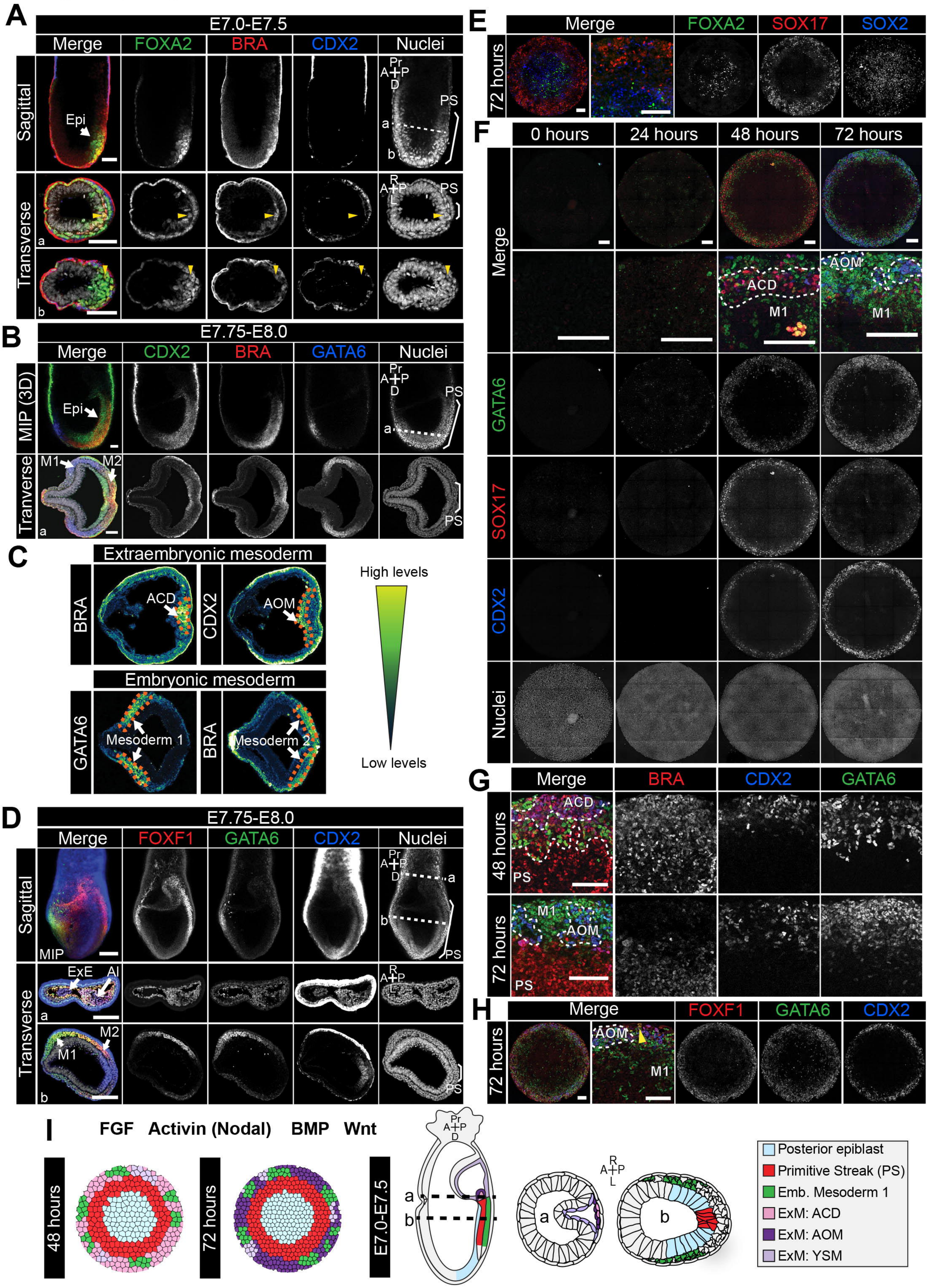


Figure 4, Morgani et al.

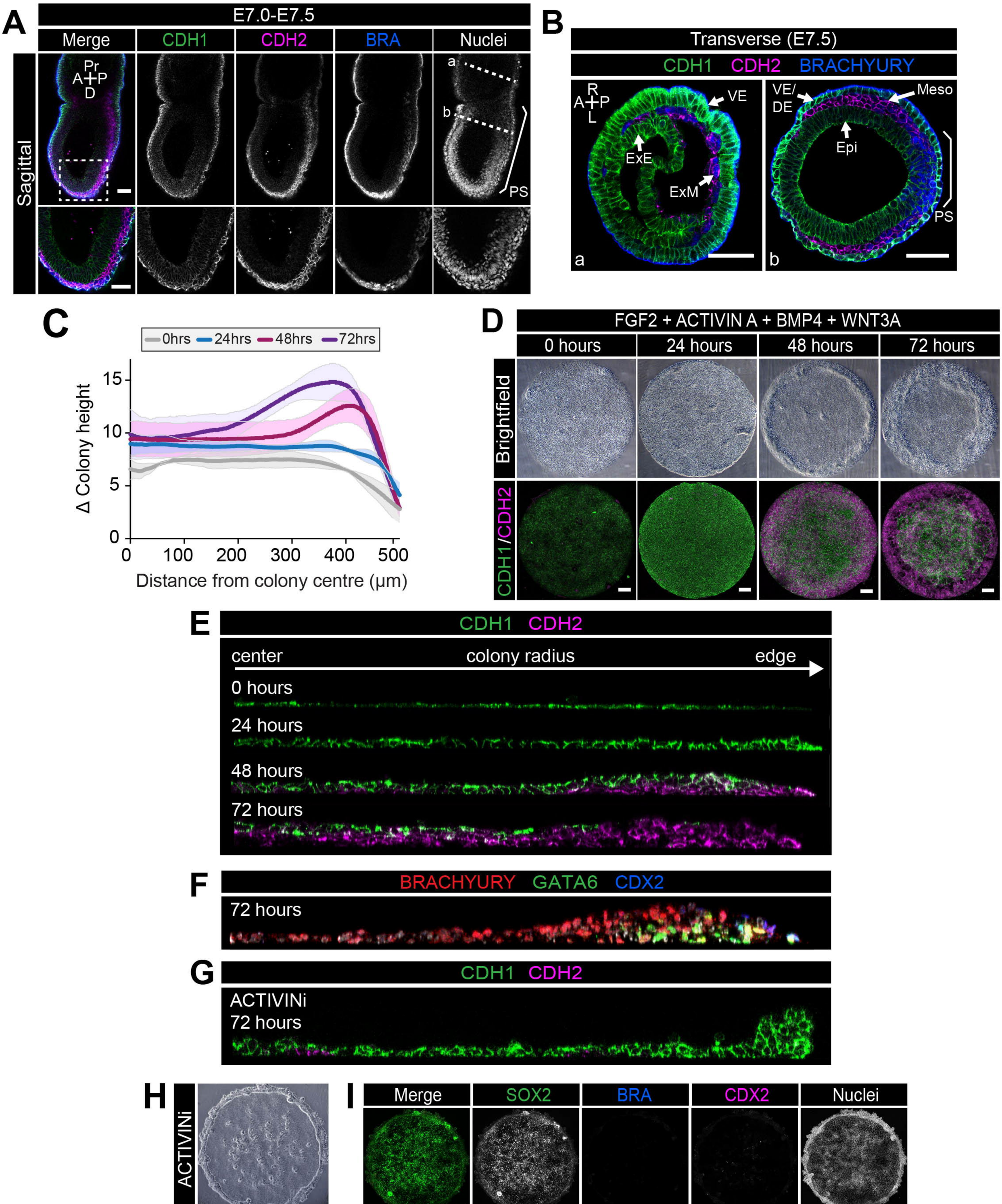


Figure 5, Morgani et al.

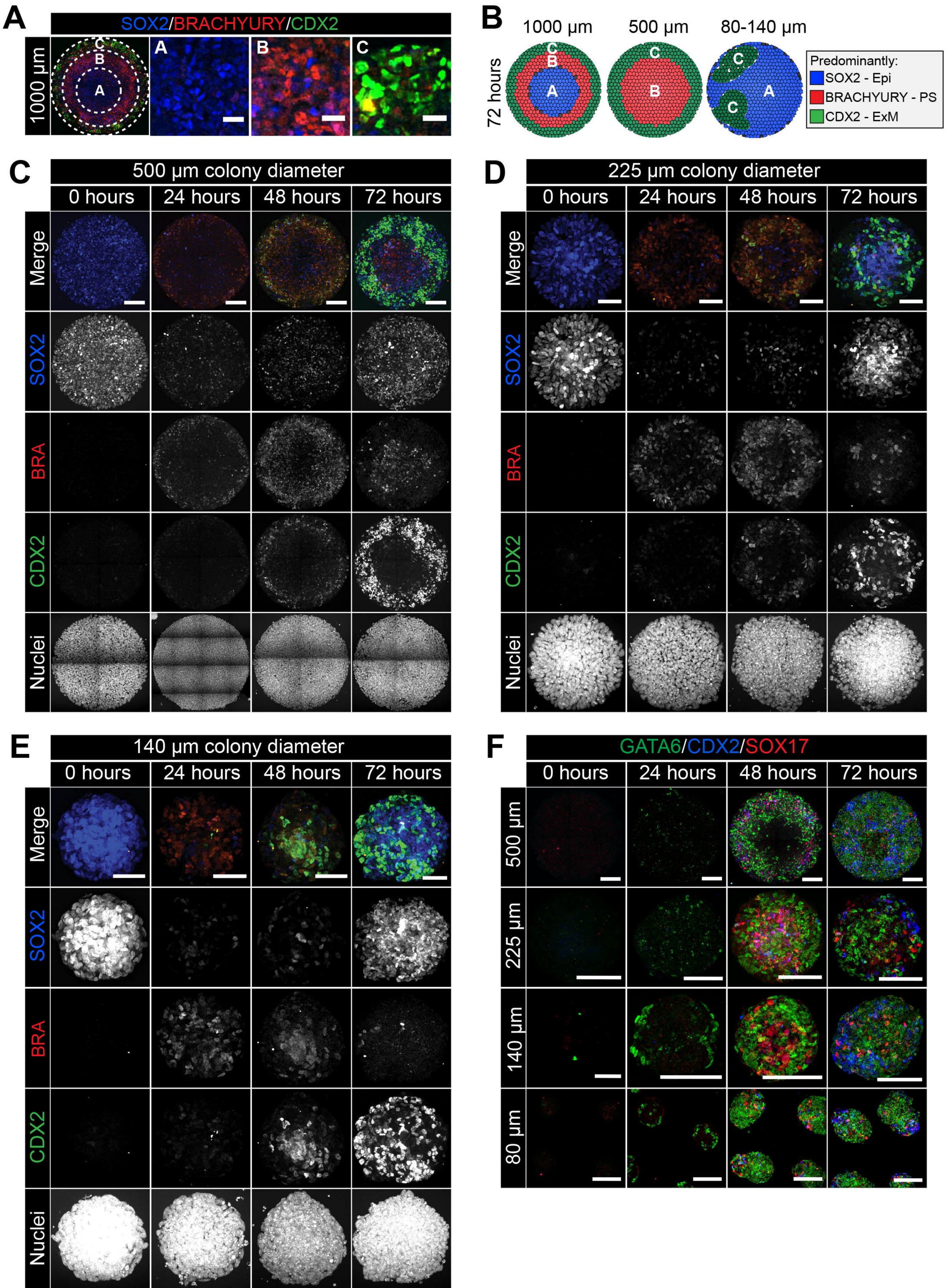


Figure 6, Morgani et al.

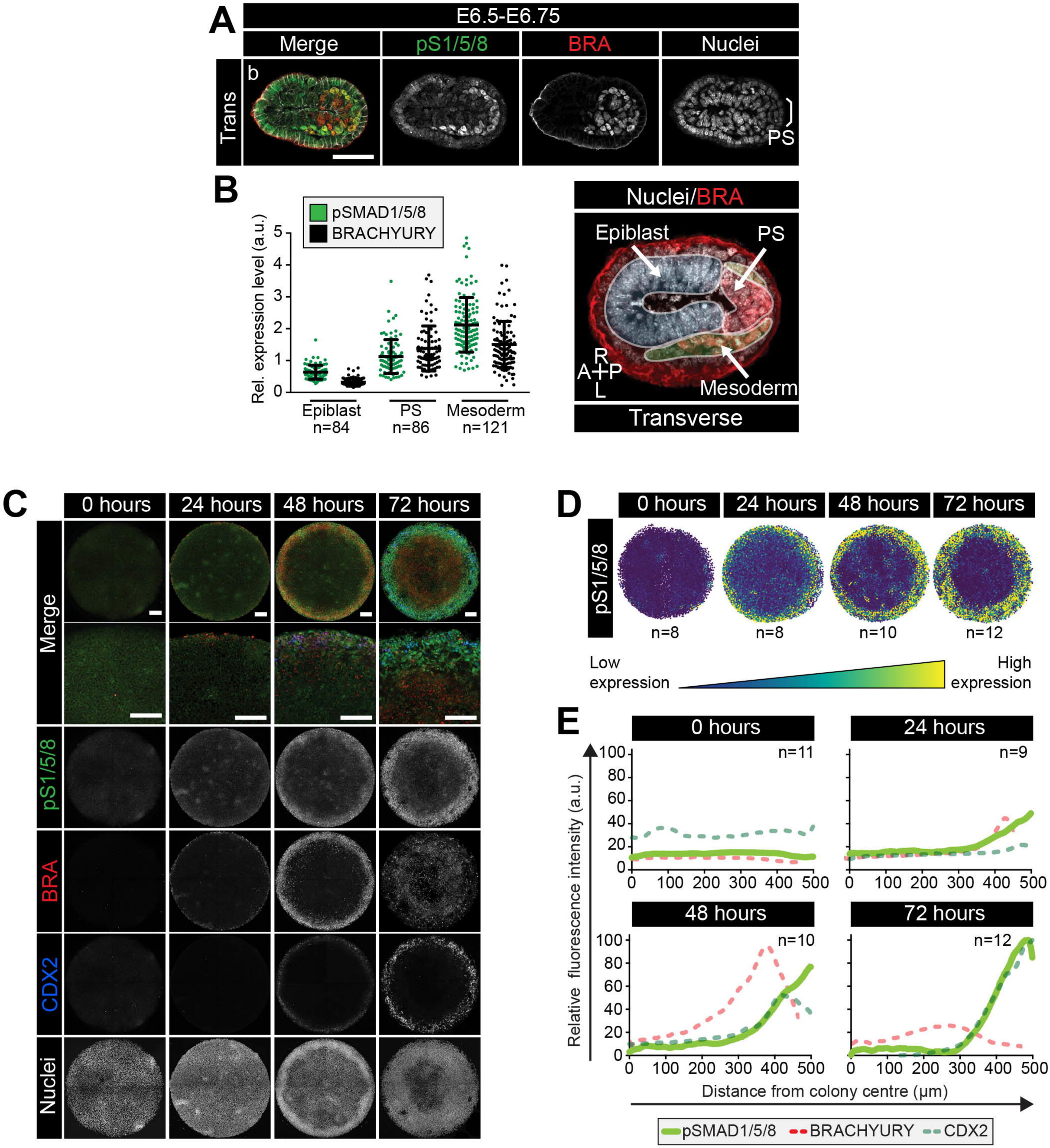
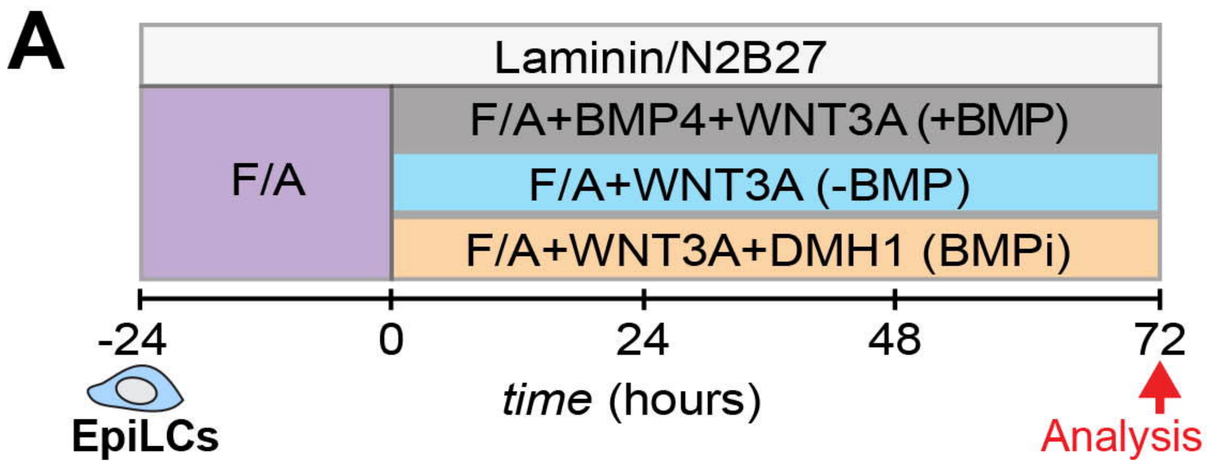
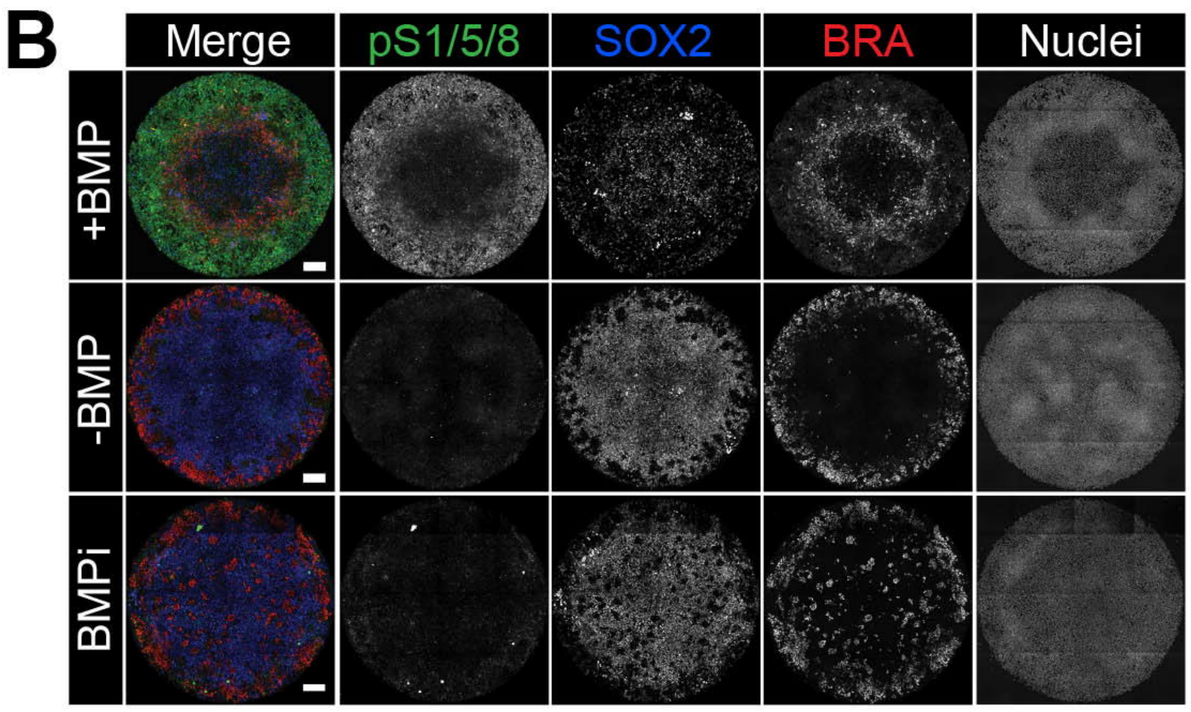
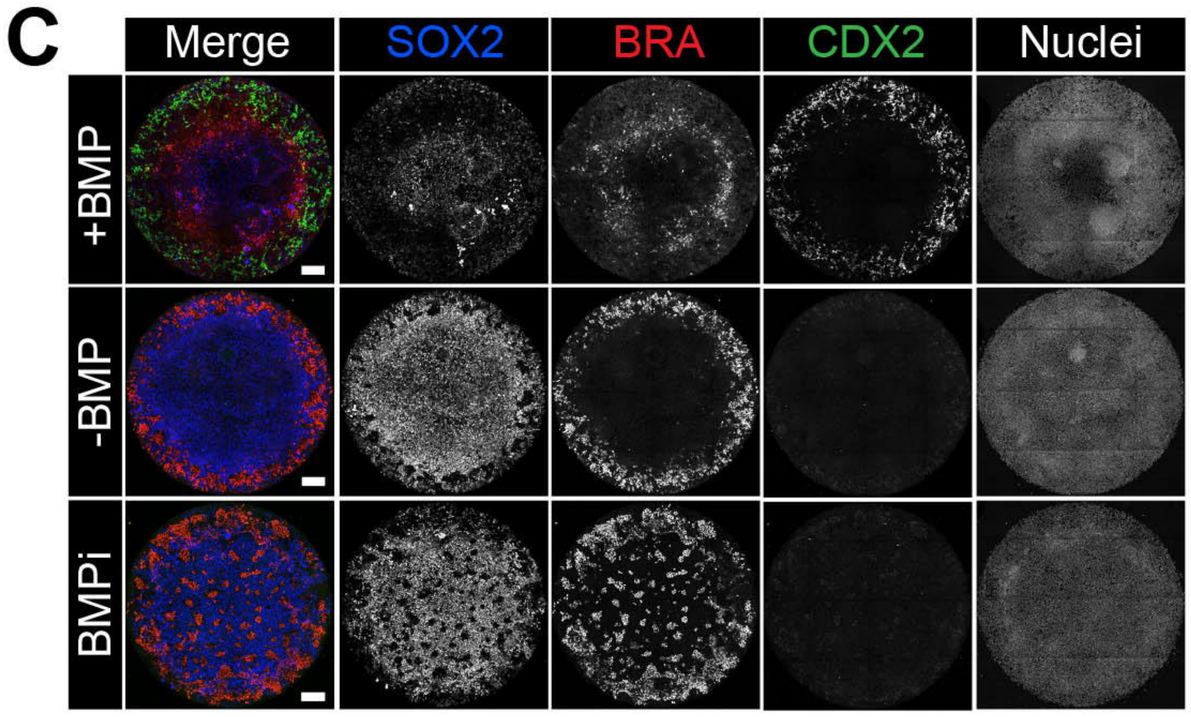
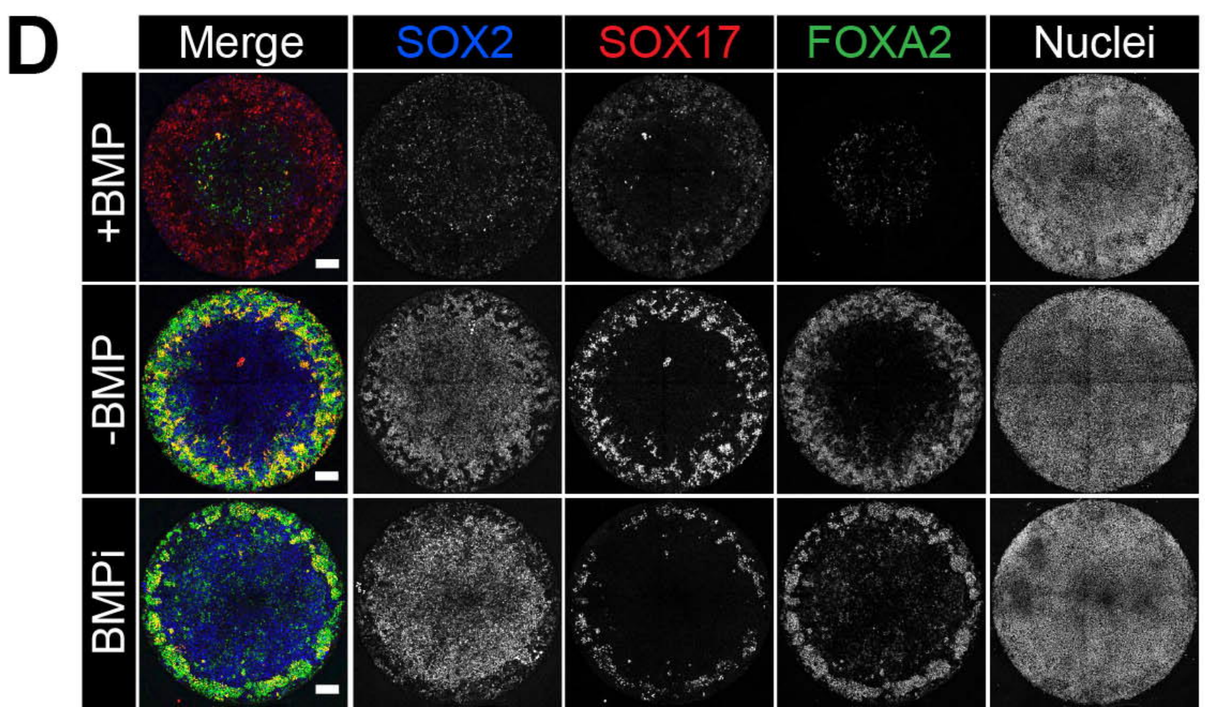


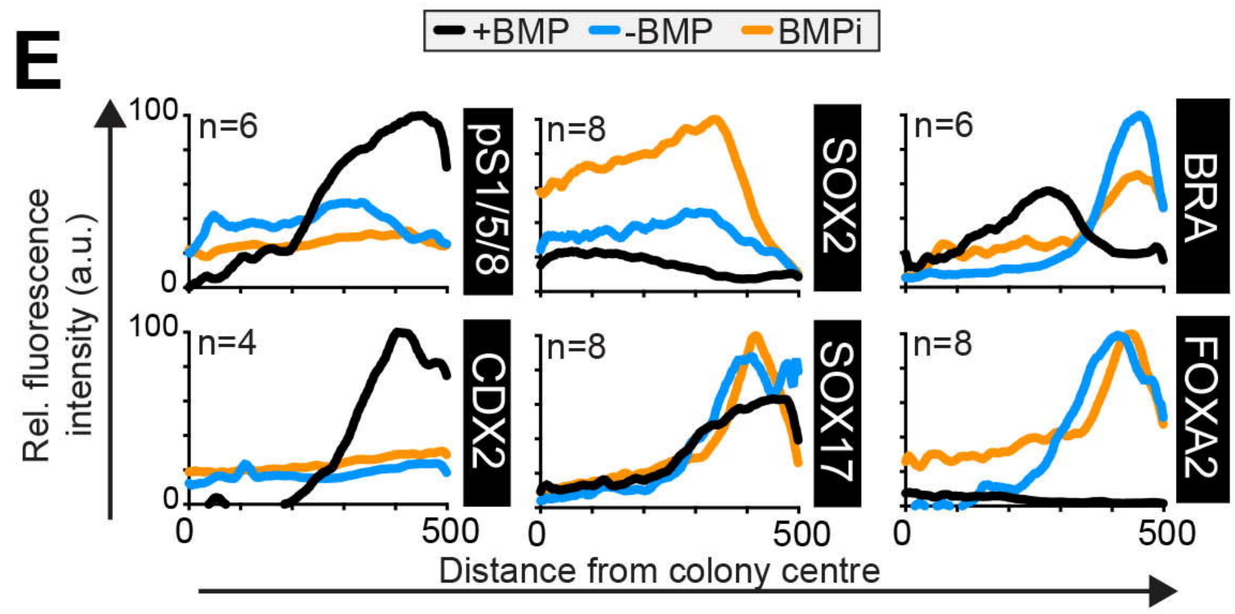
Figure 7, Morgani et al.

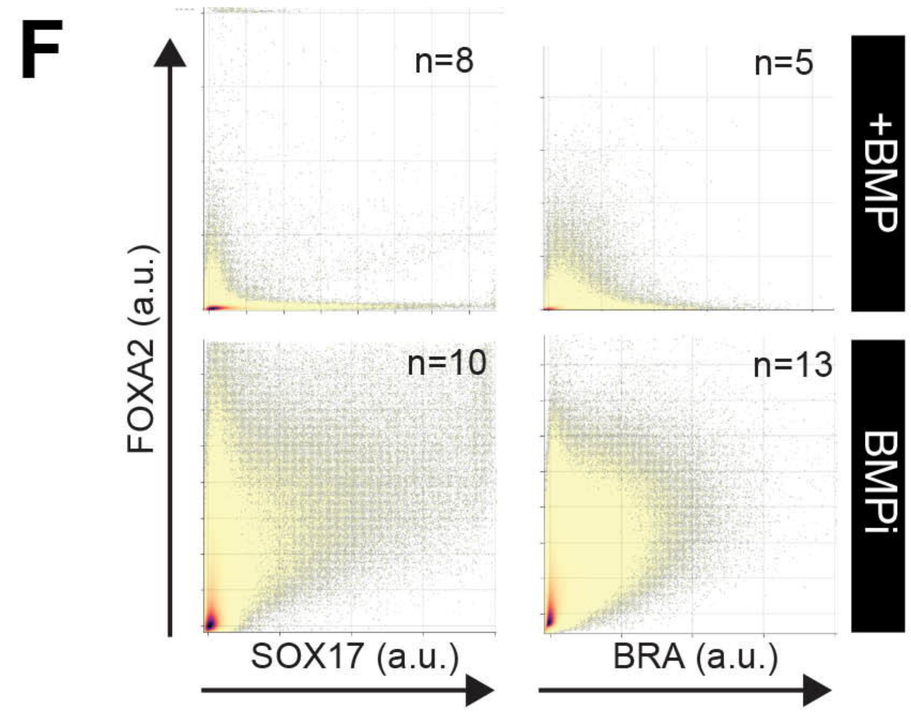


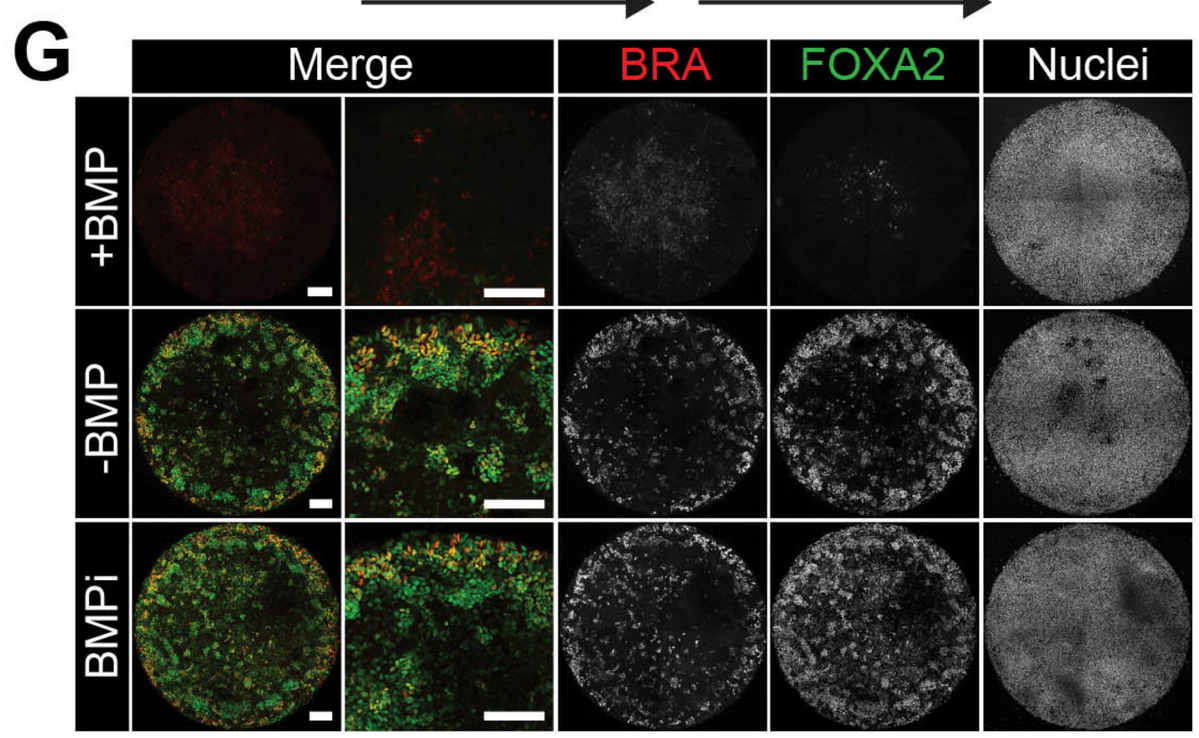












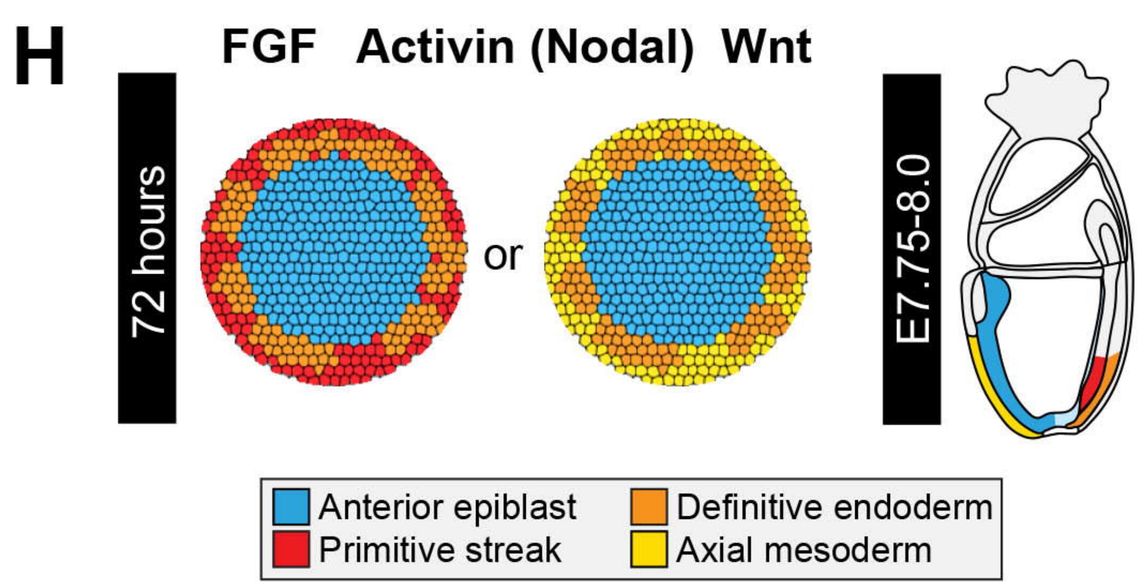


Figure 8, Morgani et al.

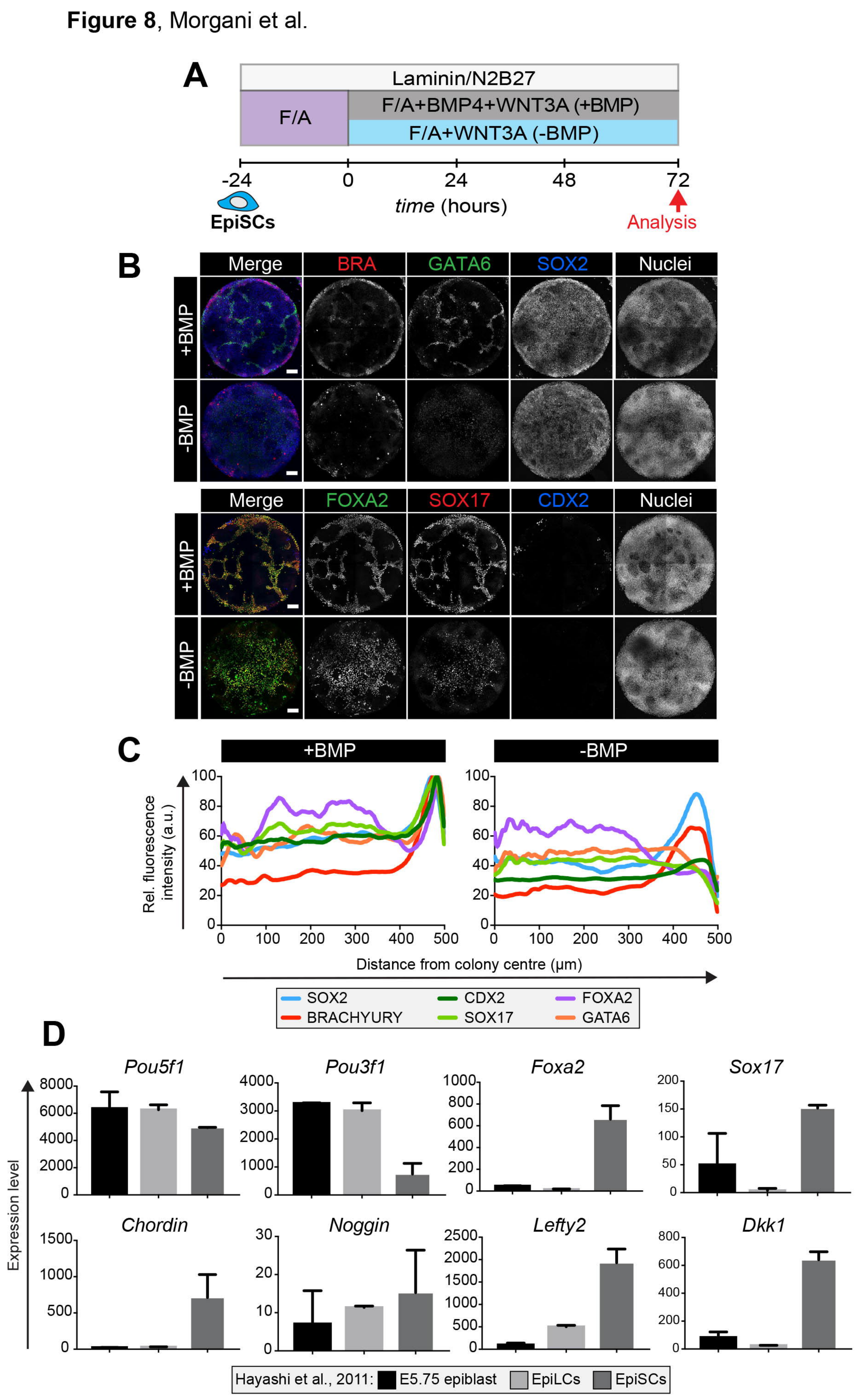
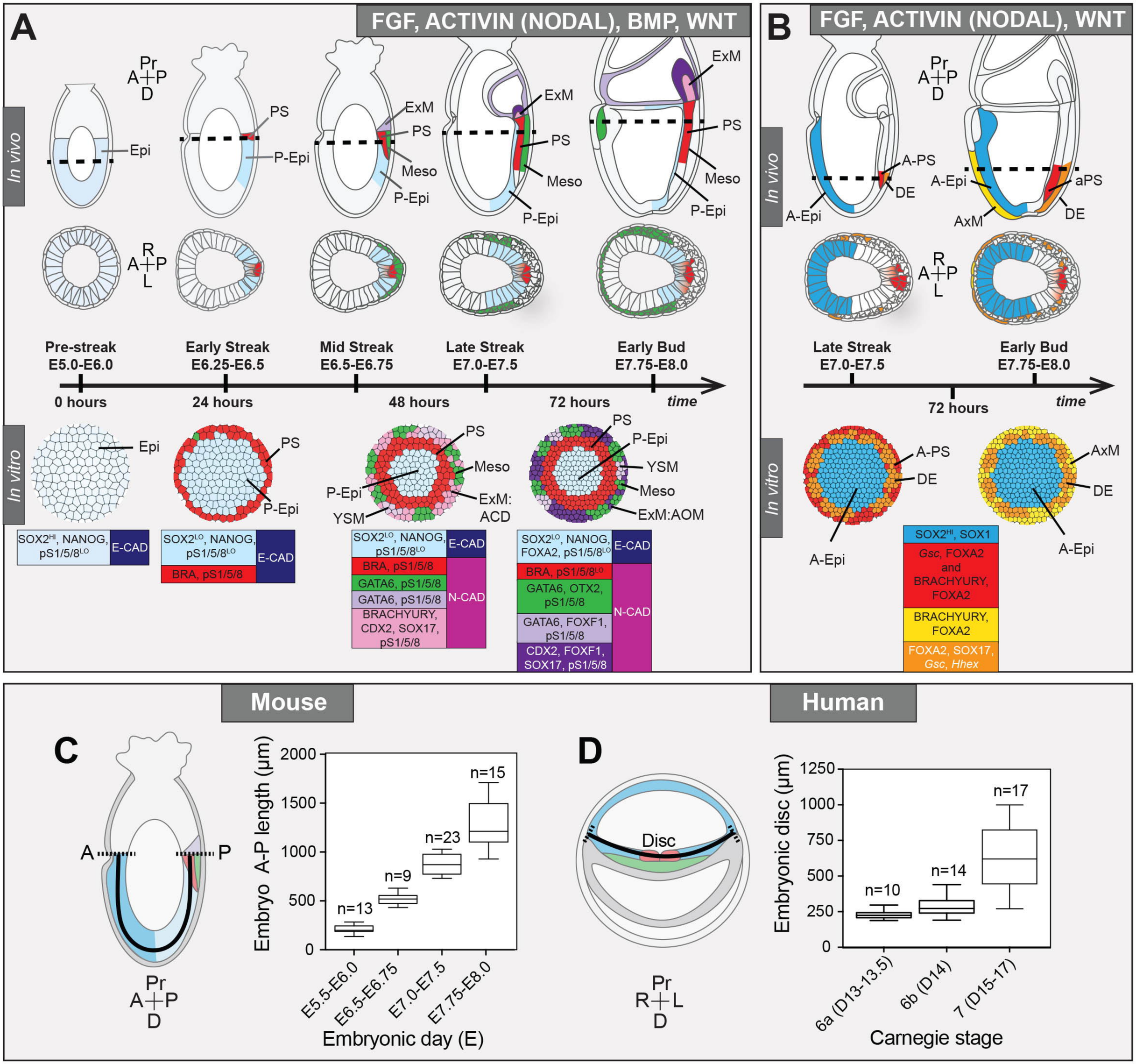



Figure 9, Morgani et al.