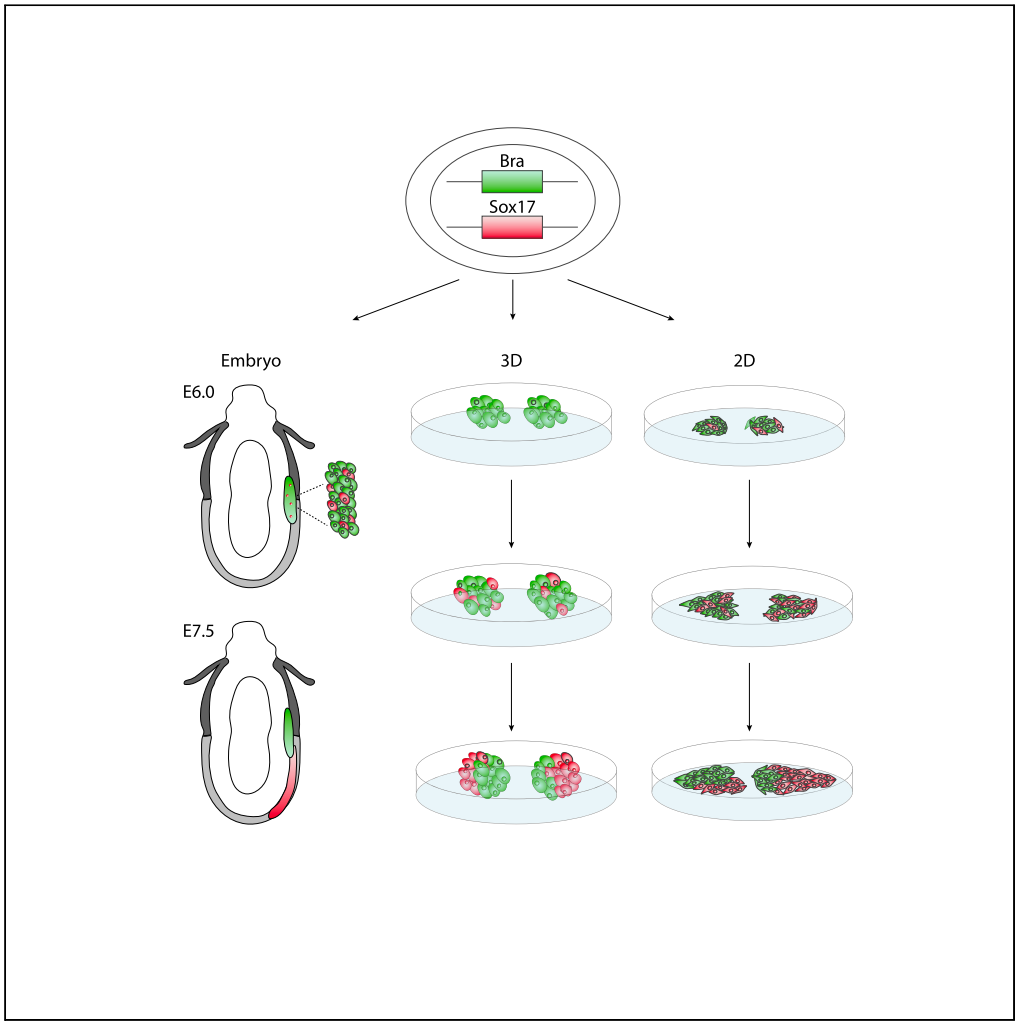
iScience

CellPress
OPEN ACCESS

Article

Emergence and patterning dynamics of mouse-definitive endoderm



Maayan Pour, Abhishek Sampath Kumar, Naama Farag, ..., Lars Wittler, Alexander Meissner, Iftach Nachman

iftachn@tauex.tau.ac.il

Highlights

Sox17 onsets in a few isolated cells within Bra-expressing population

Sox17 onset followed by expansion and self-sorting

Final number of Sox17+ cells does not depend on self-sorting or cell movement

The DE segregation pattern is similar in *in vivo* and in 2D, 3D *in vitro* systems

Pour et al., iScience *25*, 103556
January 21, 2022
https://doi.org/10.1016/j.isci.2021.103556

iScience

CellPress
OPEN ACCESS

Article

Emergence and patterning dynamics of mouse-definitive endoderm

Maayan Pour,[1] Abhishek Sampath Kumar,[2] Naama Farag,[1] Adriano Bolondi,[2] Helene Kretzmer,[2] Maria Walther,[2] Lars Wittler,[3] Alexander Meissner,[2,4,5] and Iftach Nachman[1,6,*]

SUMMARY

The segregation of definitive endoderm (DE) from bipotent mesendoderm progenitors leads to the formation of two distinct germ layers. Dissecting DE commitment and onset has been challenging as it occurs within a narrow spatio-temporal window in the embryo. Here, we employ a dual Bra/Sox17 reporter cell line to study DE onset dynamics. We find Sox17 expression initiates *in vivo* in isolated cells within a temporally restricted window. In 2D and 3D *in vitro* models, DE cells emerge from mesendoderm progenitors at a temporally regular, but spatially stochastic pattern, which is subsequently arranged by self-sorting of Sox17 + cells. A subpopulation of Bra-high cells commits to a Sox17+ fate independent of external Wnt signal. Self-sorting coincides with upregulation of E-cadherin but is not necessary for DE differentiation or proliferation. Our *in vivo* and *in vitro* results highlight basic rules governing DE onset and patterning through the commonalities and differences between these systems.

INTRODUCTION

Early in mouse embryonic development, epiblast cells undergo gastrulation, the process that derives the three primary germ layers. Mesendoderm progenitors from the epiblast egress through the extending primitive streak (PS) and segregate to two distinct lineages of the body, mesoderm and definitive endoderm (DE) (Tam and Loebel). The decision of mesendoderm progenitor cells to become DE is thus one of the earliest developmental decisions during embryogenesis. Ingressing epiblast cells switch on their DE identity, turning on DE regulators such as Sox17 and downstream endodermal genes, while moving ventrally and laterally and eventually intercalating with the visceral endoderm layer (Burtscher and Lickert, 2009, reviewed in Rivera-Perez and Hadjantonakis, 2014). In this process, they segregate from future mesodermal cells, eventually resulting in distinct germ layers with the correct ratio of cell numbers. Several signaling pathways, such as Wnt and Nodal/Activin, are known to play dominant roles in driving this decision (Robertson, 2014; Singh et al., 2012; Tam and Loebel, 2007; Yoney et al., 2018). Wnt signaling plays an essential function in mesendoderm specification by controlling the expression of Nodal and its co-receptor Cripto during gastrulation (Norris and Robertson, 1999; Tam and Loebel, 2007). Mesendoderm progenitors initially start from a Wnt-high region (the proximal-posterior PS), and committed DE cells migrate toward the anterior side. Nodal/Activin signaling forms a proximal-distal axis at the posterior region of the embryo while Nodal inhibitors are expressed at the anterior region (Arnold and Robertson, 2009; Yamamoto et al., 2004). How cells become DE or anterior mesoderm within the same physical position in the embryo, where they are exposed to the same signaling environment is unclear. It is not clear whether local signaling gradients or stochastic epigenetic or expression state differences drive the decision of specific cells to become DE, nor at what point cells commit to their DE fate.

One of the challenges to study DE decision in the mouse embryo is the lack of unique markers, as most of the endoderm markers are also expressed in the primitive endoderm (PrE) or visceral endoderm (VE). For example, Sox17, one of the earliest DE regulators, is expressed in PrE, VE, DE, and later endodermal tissues in the developing embryo (Kanai-Azuma et al., 2002; Niakan et al., 2010). Recently, it has been shown that epiblast-derived DE cells and VE cells converge not only spatially (through egression and intercalation) but also transcriptionally, both contributing to endodermal structures, such as the gut endoderm (Ibarra-Soria et al., 2018; Nowotschin and Hadjantonakis, 2018; Viotti et al., 2014). Therefore, in order to decipher the onset and dynamics of the epiblast-derived DE, there is a need to selectively track these cells and distinguish them from extra-embryonic derived endoderm. Even then, it is hard to tease apart the contribution of

[1]School of Neurobiology, Biochemistry and Biophysics, Department of Biochemistry and Molecular Biology, Tel Aviv University, Tel Aviv 6997801, Israel

[2]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany

[3]Department of Developmental Genetics, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany

[4]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

[5]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[6]Lead contact

*Correspondence: iftachn@tauex.tau.ac.il

https://doi.org/10.1016/j.isci.2021.103556

other crucial factors such as cellular localization, signaling gradient, and expression profile, as these are all coupled in the developing embryo. *In vitro* 2D and 3D stem-cell systems have been used in recent years to model early stages of embryonic development, including mesendoderm progression (Beccari et al., 2018; Boxman et al., 2016; Harrison et al., 2017; Rivron et al., 2018; Sozen et al., 2018; Turner et al., 2017; van den Brink et al., 2014; Warmflash et al., 2014). These provide accessible systems, amenable to manipulation and decoupling of different external conditions. The comparative view of commonalities and differences between such systems and the embryo provide a promising opportunity to dissect basic rules of cell fate decisions and early patterning (Morgani et al., 2018). Such systems can be used to quantitatively study the specific dynamics of mesoderm (ME) and DE segregation.

Here, we set out to study the onset dynamics and progression of the epiblast-derived Sox17+ cell population (DE), using a newly engineered dual Bra-GFP (Brachyury), Sox17-RFP reporter cell line, combining *in vivo* and *in vitro* observations. Dual-reporter embryos derived through tetraploid embryo complementation experiments reveal that the DE emerges in few isolated cells within the mesendoderm spread population. To better understand the factors driving the DE fate decision, we quantitatively dissected its dynamics in complementary 2D and 3D *in vitro* contexts. In both systems, DE cells arise from within the mesendoderm population in a temporally synchronized manner, but in a spatially stochastic "salt-and-pepper" pattern. Single-cell segmentation reveals that both *in vivo* and *in vitro* Sox17 cells are Bra-low at the onset time. This is followed by expansion via cell divisions and a self-sorting phase of the Sox17+ cells, leading to aggregation (2D) or lumenogenesis (3D) as the *in vitro* counterpart to intercalation in the embryo. This self-sorting precedes high E-cadherin expression shown both at the mRNA and at the protein levels, and is not essential for DE differentiation or proliferation. Analyzing the temporal dependencies of DE differentiation on Wnt and Nodal signaling, we find that a small subpopulation of high-expressing Bra cells is already committed to their Sox17+ fate independent of external Wnt signal. The robust properties we observe in the different *in vitro* contexts shed new light on principles underlying DE dynamics in-vivo.

RESULTS

Dual live marker cell line demonstrates sporadic onset pattern of Sox17 in the primitive streak within Bra-low cells

To study the onset of mesendoderm-derived DE and the relative patterning of DE and mesendoderm progenitors, we have inserted an RFP (mStrawberry) reporter downstream of the Sox17 ORF in an E14 Brachyury-GFP mESC line, resulting in a dual-reporter Sox17-RFP/Bra-GFP cell line (Figures 1A and 1B) (Fehling et al., 2003). As single-cell RNA sequencing data from E6.5 and E7.5 embryos confirm that embryonic expression of Bra and Sox17 mark ME and DE, respectively (Figure S1A) (Nowotschin et al., 2019), we expected the dual-reporter line to serve as a faithful reporter for these two germ layers. We generated transgenic mouse embryos from the reporter cells by tetraploid complementation such that all embryonic tissues (e.g. primitive streak and mesendoderm progenitors) are derived from the diploid transgenic mESC line, while extra-embryonic tissues (including visceral endoderm) are from the wild-type tetraploid donor (Burtscher and Lickert, 2009) (Figure 1C). We imaged these embryos between gastrulation initiation (E6.5) and early organogenesis (E8.5) to monitor the onset and expansion pattern of Sox17-RFP within the embryonic tissues. As expected, Bra-GFP was expressed in the primitive streak, node, and notochord, while Sox17-RFP was expressed in the developing DE, and later along the gut tube (Figures 1D, 1E, and S3A). Immunostaining at equivalent time points shows Sox17 expression also in the extra-embryonic tissues, confirming that in the tetraploid aggregation embryos Sox17-RFP cells are limited to the DE (Figure S2). Consistent with previous observations, the Sox17-RFP+ population expands during migration laterally toward the anterior side, and is spread over the whole distal side of the epiblast as distinct separated cells by E7.5 (Viotti et al., 2014). These cells intercalate with the cells of the visceral endoderm and undergo epithelialization, and by E8.5 all embryo-derived Sox17+ cells are localized along the gut tube (Figures 1E and S3A). In gastrulating embryos, the first Sox17-RFP+ cells arise within the Bra-GFP+ cells at the anterior part of the PS (Figure 1D). Unlike the Bra-GFP + cell population which is adjoined, Sox17 initially arises in a few isolated cells (Figure 1D). The first Sox17-expressing cells arise within the Bra-expressing population but at this point show lower Bra-GFP expression, which is also associated with the location in the primitive streak (Figures 1D and 1F(black points), Figures S3B andS3C). The Sox17 onset is followed by further segregation from the Bra-positive cells, as these cells downregulate their GFP expression as RFP levels accumulate and differentiation continues (Figure 1F). This phenomenon is confirmed at the RNA level, by single-cell RNA sequencing showing that a subset of E6.5 cells expresses both Bra and Sox17 at a lower level than single positive cells at E7.5 (Figure S1B left). This segregation is accompanied by

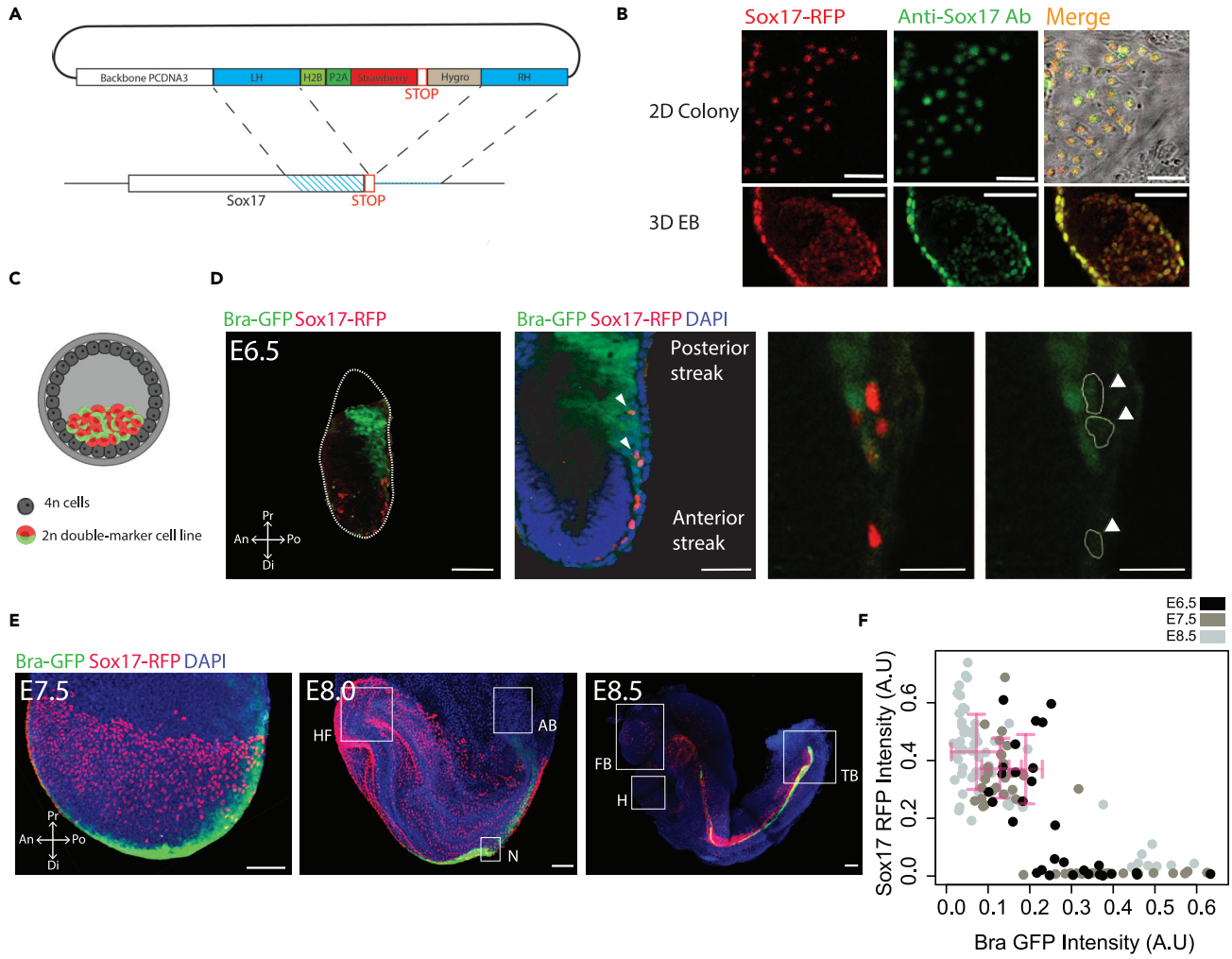


Figure 1. Sox17 emerges in few isolated Bra-low cells in the mouse tetraploid embryos

(A) CRISPR knockin of the H2B-RFP cassette in the endogenous Sox17 gene design.

(B) Overlap between Sox17-RFP live marker (left) and Sox17 antibody immunostaining (middle) to validate the live marker cell line in a 2D colony (top row) and in a 3D EB (bottom row). Scale bar 50 μm.

(C) A schematic of a tetraploid-aggregation embryo at the blastocyst stage (created with BioRender.com).

(D) Single Z-slice image within E6.5 embryo shows the emergence of Sox17+ cells within the Bra+ cells in the primitive streak. Bra-GFP levels are low in RFP-high cells, marked by white contour and arrows. Scale bar 50 μm.

(E) Images of maximum intensity projections of E7.0–E8.5 tetraploid embryos with the double marker cell line, shows the expression and resolution of Bra+/Sox17+ cells in the gastrulating embryos. Scale bar 100 μm. n = 3.

(F) Bra-GFP and Sox17-RFP intensity levels in single cells within the E6.5 (black), E7.5 (dark gray), and E8.5 (light gray) embryos. Cross represents the mean ± standard deviation of GFP and RFP intensities for the RFP-high cells (intensity higher than 0.15) for each embryo. RFP-high cells show lower levels of Bra compared with the RFP-low ones. As development progresses, RFP-high cells show lower GFP intensity (n > 30 cells per embryo, each channel was normalized to its background and maximal intensity, see STAR Methods). HF, head fold; N, notochord; AB, allantois bud; FB, forebrain; H, heart; TB, tail bud

E-cadherin (E-cad) upregulation by the Sox17 cells (Figure S1B right). These observations raise questions about how the DE population segregates from the mesendoderm, and what mechanisms set its resulting pattern.

Dynamics of Sox17 emergence from bra-expressing cells are conserved between 2D and 3D *in vitro* systems

To better understand the rules governing onset and expansion of Sox17 within mesendoderm cells in a robust and quantitative manner, we differentiated our reporter ES cells in 2D colonies or 3D embryoid bodies (EBs) in defined medium (N2B27). While 2D culture allows uniform exposure to external signals

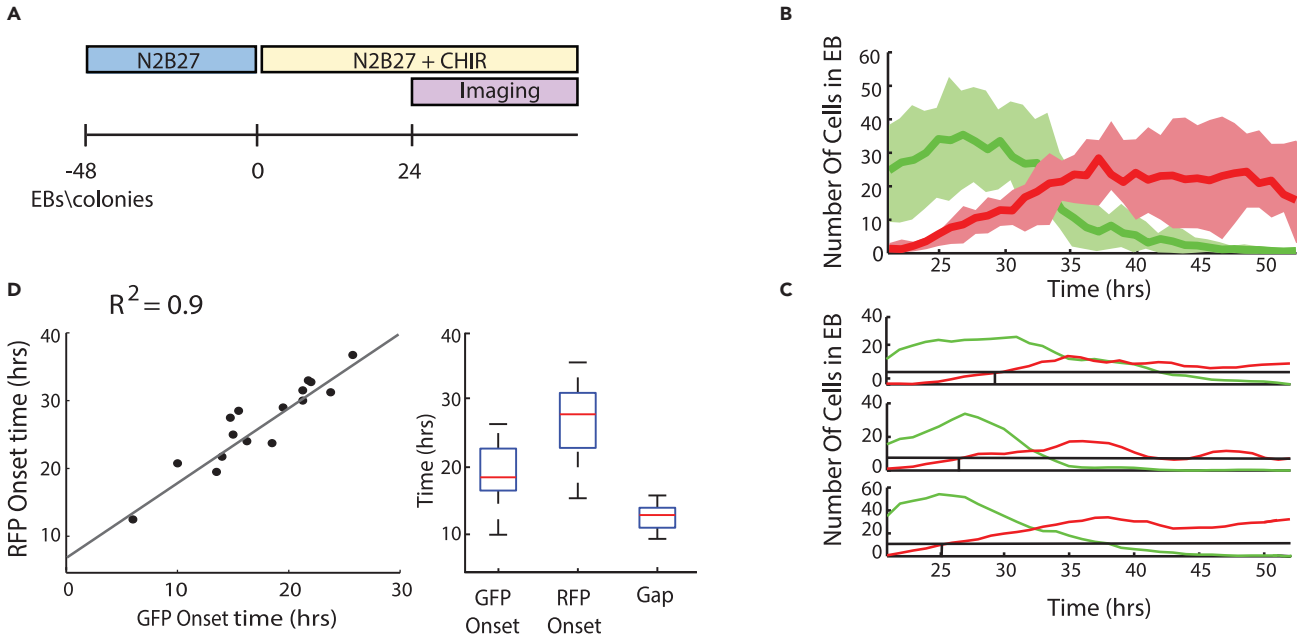


Figure 2. A fixed temporal pattern between Brachyury and Sox17 onsets in 3D cultures

(A) Experiment design. ESCs were grown in N2B27 medium as 2D colonies or as EBs. After 48 h, CHIR was added to the medium (marked as 0 h). Imaging started at ~24 h.

(B and C) Number of Bra-GFP (green) cells or number of Sox17-RFP (red) cells per EB over time. Data are represented as mean ± standard deviation of 10 EBs from two independent experiments (B). Three individual EBs as an example for the coupling between Bra-GFP peak and Sox17-RFP onset. Later GFP peak time show later RFP onset time (C).

(D) Bra-GFP onset time versus Sox17-RFP onset time in EBs. Each dot represents one EB (n = 18) (left). Data are represented as mean ± standard deviation of onset times and the gap between them (right).

from the media, 3D cultures allow cell–cell interactions that resemble more the *in vivo* structure (Kapalczynska et al., 2018). At 48 h, CHIR (a Wnt signaling activator) was added, as these conditions have been shown before to induce both mesoderm and endoderm (Turner et al., 2017; van den Brink et al., 2014). We designated the time of CHIR addition as 0 h (Figure 2A). We first verified the identity and origin of Sox17-RFP+ cells in the 2D and 3D cultures. Under pluripotency conditions, only in 2D colonies, but not in 3D embryoid bodies, do we observe some peripheral RFP+ cells, which have been previously attributed to extraembryonic endoderm-fated cells (Niakan et al., 2010) (Figure S4A). Staining for Foxa2 results in a full overlap between Sox17-RFP+ cells and Foxa2 expression by 48 h, supporting the notion that Sox17+ cells at this stage represent DE (Figure S4B top). Similar overlap was observed in the E6.5 embryos (Figure S4B bottom). Bulk RNA sequencing at 48 h after CHIR induction on sorted populations (FACS) shows that Bra-GFP+ cells express mesodermal markers and Sox17-RFP+ cells express definitive endoderm markers, providing validation for the use of Bra and Sox17 in the double-marker cell line as markers for ME and DE, respectively (Figure S5). We next tracked the origin of DE cells in the 2D and 3D cultures. The GFP and RFP signals were quantified over time in each EB/colony (see STAR methods). We found that in embryoid bodies, Sox17-RFP emerged at a highly uniform temporal window from Bra-GFP positive cells (Figure 2B). The first Sox17-RFP cells are observed 27 ± 2.5 h (n = 13, from 2 independent experiments) after addition of CHIR, 13 ± 4 h after the onset of Bra-GFP, and 2 ± 1.5 h before the peak of Bra-GFP expansion. This pattern is highly reproducible in individual EBs, with high correlation between the onset time of the two genes (Figures 2C and 2D), suggesting they are part of a sequential differentiation process. Tracking individual cells at high temporal resolution in 2D differentiating colonies reveals similar temporal expression pattern even at the single-cell level, where cells turn off their Bra-GFP expression during (or before) Sox17-RFP ramp-up (Figure S6A). Following the number of Sox17-RFP cells per colony over time in 2D culture, shows resembling dynamics between these two very different experimental systems (Figure S6B). The robustness of the relative pattern to the differences between the 2D and 3D setups and between independent experiments suggests this sequential process is not dependent on cell interactions with the environment or other cells.

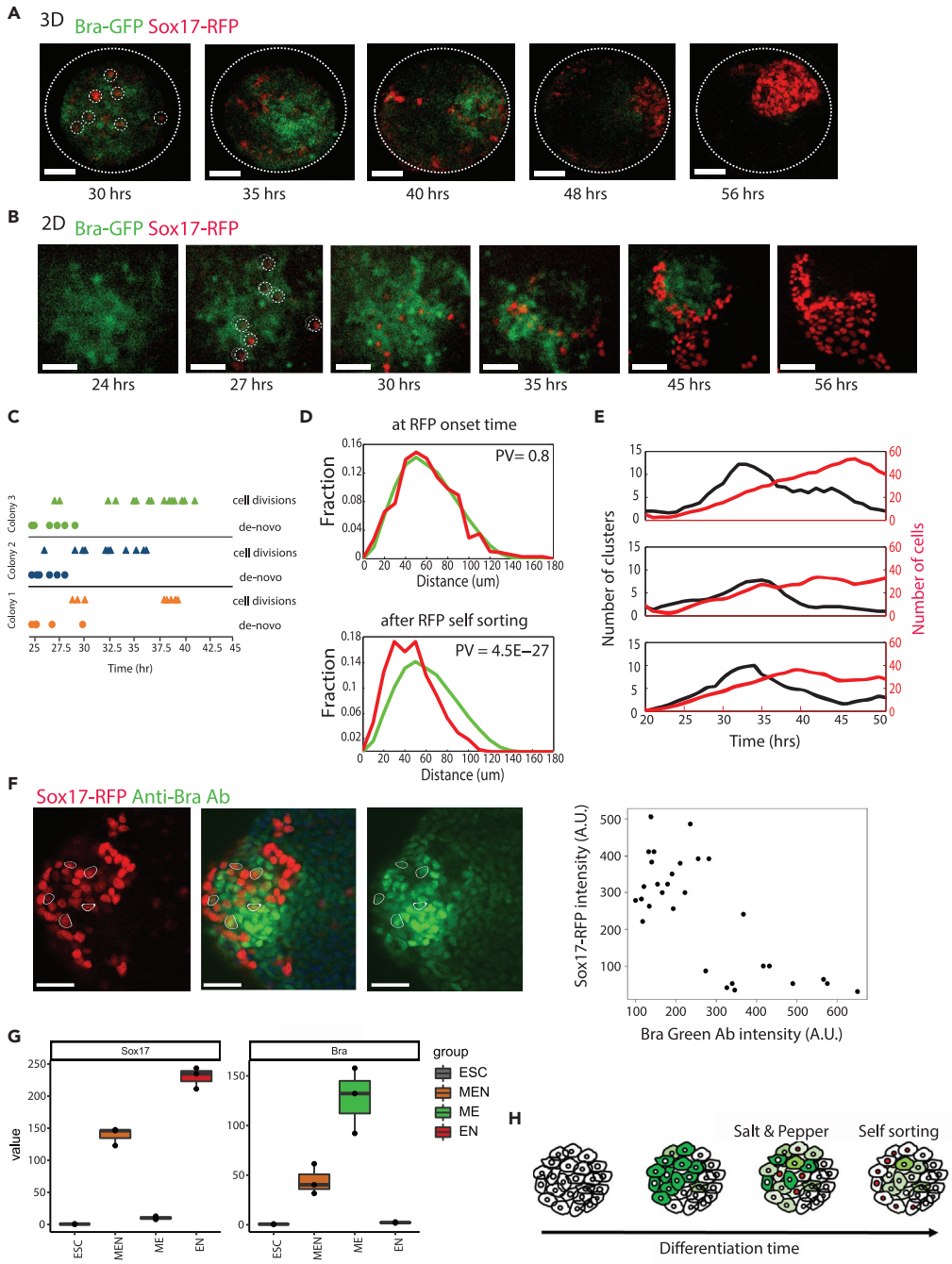


Figure 3. Sox17 onsets in a salt-and-pepper pattern followed by self-sorting in 2D and 3D cultures

(A and B) Sox17-RFP cells emerge within the Bra-GFP population and expand to form an RFP only locus in EBs (A) or in 2D colonies (B). Scale bar 50 μm.

(C) De novo expression and cell divisions in three differentiating 2D colonies, from 24 h after CHIR addition to 45 h (image every 15 min). Circle - a new cell that turned on Sox17-RFP expression; Triangle - a cell division.

(D) The distribution of distances between RFP+ cells (red) compared with distances between all cells (green). The two distributions are similar at the time of Sox17-RFP onset (P = 0.815, Kolmogorov-Smirnov test) (top). At 56 h of differentiation, the two distributions differ significantly (P = 4.46 × 10^{-27}), after RFP cells have clustered together (bottom).

(E) Number of Sox17-RFP cells (red line, right Y axis) and number of clusters (black line, left Y axis) as an indicator for the self-sorting process over time. While the number of total Sox17-RFP cells keeps rising, the number of clusters starts to

Figure 3. *Continued*
decrease due to merging of cell clusters. Clusters were automatically determined as isolated surfaces (see STAR Methods).
(F) Sox17-RFP (red) and anti-Brachyury staining (green) in a 48 h EB. Dual-positive cells are outlined in white. Right: anti-Bra staining and Sox17-RFP intensity levels in single cells. RFP-high cells show lower level of Bra compared with the RFP-low ones. (n = 30).
(G) mRNA expression levels of Sox17 and Bra of undifferentiated ESCs and differentiated EBs, 48 h post CHIR addition, sorted to Bra-GFP+ (ME), Bra-GFP+ Sox17-RFP+ (MEN), and Sox17-RFP+ (DE) populations, show that Sox17 and Bra are each expressed at the RFP+ and GFP+ single-positive populations, respectively, while the double-positive population shows lower mRNA levels of both markers. Data are represented as mean ± standard deviation (n = 3).
(H) A model of the dynamics observed for Sox17 onset and expansion.

Sox17 emerges in a salt-and-pepper onset pattern followed by expansion through cell divisions and self-sorting in 2D and 3D *in vitro* systems

We next wanted to characterize the spatial pattern of Sox17 emergence. In embryoid bodies, Sox17-RFP+ cells initially emerge within the Bra-GFP+ population, close to the peak time of Bra expression. The spatial pattern of appearance is salt-and-pepper like: RFP+ cells are randomly distributed within the GFP+ population, with no clear spatial organization, similar to their early pattern in the embryo (Figure 1D). However, the cells begin to self-sort 13 ± 1.5 h after the appearance of Sox17, moving within the GFP+ population, eventually finding each other to form either internal lumens or an envelope at the outer shell of the EB (Figure 3A, Video S1). A similar spatial dynamic is observed in 2D colonies (Figure 3B and Video S2), where after sorting out of the GFP + area, the Sox17-RFP+ cells form distinct continuous patches. Tracking Sox17-RFP+ cells from their onset time (denovo expression) to the final sorted pattern in single 2D colonies revealed that this dynamic consisted of an early Sox17 onset phase, followed by an expansion phase through cell-division of Sox17+ cells, that take place up until ~40 h after CHIR induction. Finally, the resulted pattern obtained by self-sorting (Figure 3C). To quantify the amount of spatial bias at the onset of Sox17, we analyzed the distribution of distances between all pairs of Sox17-RFP+ cells (Figure 3D). At the onset time (when we spot 7–10 RFP+ cells within the EB), this distribution is indistinguishable from that of all cells in the EB volume ($p = 0.8$, two-sample Kolmogorov-Smirnov test), consistent with random locations and no spatial bias (Figure 3D top). In contrast, after 24–30 h, when RFP+ cells have self-sorted into consistent lumens/aggregates, the distance distribution of RFP+ cells significantly differs from that of all cells ($p = 4 \times 10^{-27}$) (Figure 3D bottom). To quantify the aggregation dynamics of the RFP+ cells in EBs, we performed a coarse segmentation to count the number of continuous cell clusters (Figure 3E and video S3). For example, each formed lumen is counted as a single cluster. Initially, the number of clusters increases with the number of cells. Once self-sorting begins, the number of clusters begins to decrease, as Sox17-RFP+ cells or clusters find each other and merge into larger clusters.

Our imaging data show that Sox17+ cells emerge within the Bra+ region in 2D, 3D, and embryos (Figures1-D,3A,and 3B), with a fixed temporal pattern. At early time points, in the embryo, Sox17+ cells co-express low levels of Bra (Figures 1D and 1F). Our EB system captures the same dynamics (Figure 3F). Furthermore, bulk RNA sequencing of sorted populations of Bra-GFP and Sox17-RFP single-positive, as well as double-positive cell populations, confirms this phenomenon at the RNA level, where the RFP and GFP single-positive populations show higher expression of Sox17 and Bra, respectively, than the double-positive population (Figure 3G). Tracking single cells in 2D reveals that Sox17 upregulation is concomitant with Bra down-regulation in individual cells (Figure S6A). Furthermore, it was previously shown that sorted Brachyury + cells uniquely give rise to Sox17 + cells in differentiating mESCs (Kubo et al., 2004). Taken together, the spatiotemporal pattern of these signals suggests an intermediate double-positive phase preceding DE segregation from the mesendoderm. To further verify the differentiation potential of this intermediate phase, we sorted the cells by FACS at different time points after Bra onset, which was shown to be a synchronizing event (Boxman et al., 2016) (Figure S7A). A double-positive cell population was detected 36 h after Bra onset. We plated GFP+/RFP-, GFP+/RFP+, and GFP-/RFP+ populations sorted at this time point, and imaged them 24 h after the plating. The GFP+/RFP- population showed no fluorescent signal, and displayed a mesenchymal morphology, suggesting a later mesodermal fate, and no switching to the DE state at later time points. In contrast, both the GFP+/RFP+ and GFP-/RFP+ populations showed only an RFP signal and a rounded morphology (Figure S7A). Sorting of the cells from the 2D differentiation protocol led to similar results (Figure S7B). Together, these results suggest that Sox17 is activated within a subset of the Bra-GFP+ cells in the studied *in vitro* models, within a narrow time window concomitant with down-regulation of Brachyury in these cells, similarly to its onset dynamics in the embryo (Figure 1). Commitment

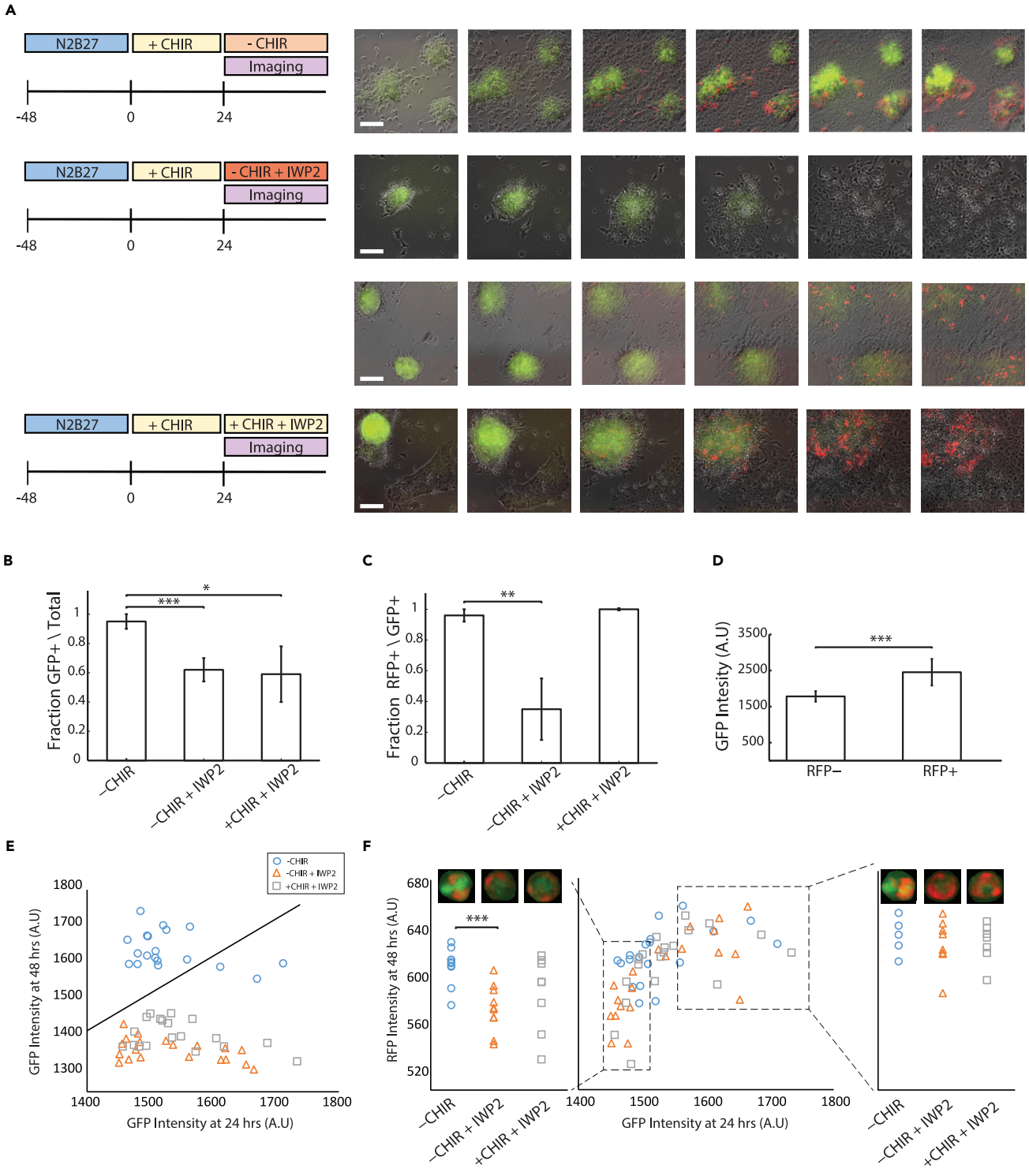


Figure 4. Sox17+ establishment is independent of Wnt in Bra-high colonies

(A) Time-lapse images of Bra and Sox17 expression over time under withdrawal of CHIR at 24 h (top), withdrawal of CHIR and blocking the Wnt pathway (IWP2) (middle), or blocking the Wnt pathway (IWP2) with CHIR in the medium (bottom).

(B) Fraction of colonies with GFP+ cells out of all the colonies for each condition. Data are represented as mean ± standard deviation.

(C) Fraction of colonies with RFP+ cells out of all colonies with GFP+ cells.

Figure 4. *Continued*
(D) GFP intensity at 24 h, in RFP+ and RFP- colonies (at 60 h).
(E) GFP mean intensity measured in individual EBs at 24 h (before treatment) vs. the intensity at 48 h (after treatment).
(F) GFP mean intensity measured in individual EBs at 24 h (before treatment) vs. RFP mean intensity at 48 h (after treatment). Treatment protocol is the same as in (A). n = 20 for each condition. ***P < 0.001. **P < 0.01. *P < 0.05

to the DE fate is obtained earlier than 36 h after Bra onset, and expansion of the Sox17-RFP+ population beyond this time point occurs through proliferation alone, followed by self-sorting (Figure 3H).

Temporal requirements of Wnt and Nodal/Activin signaling for Sox17 specification

In the embryo, Activin/Nodal signaling is required for endoderm differentiation (Camus et al., 2006; Kubo et al., 2004; Zhou et al., 1993). Timed interactions between Wnt/β-catenin and Nodal signaling allow anteroposterior polarization in gastruloids. Canonical Wnt TCF/beta-catenin was also shown to directly bind to the Sox17 promoter during endoderm differentiation (Engert et al., 2013). While both signaling pathways are required for Bra expression and axis formation, their roles in the segregation between ME and DE remain unclear (Turner et al., 2017).

In our system, though external inducing signals (e.g. CHIR) are supplied as a uniform field, cells still respond variably, either due to differential activation and response of signaling pathways or alternative chromatin states. To explore the interplay and temporal role of the Wnt and Nodal signaling pathways in DE specification under these conditions, we tested the effect of signal perturbations on the differentiation process. It was shown that a 24-h pulse of Wnt activation is sufficient in gastruloids for the induction of Bra expression and at least a low level of Sox17 (Turner et al., 2017; van den Brink et al., 2014). We first tested and quantified the needed duration of Wnt induction in our system. Induction with CHIR for 24 h, or even 12 h, is sufficient for the expression of both Bra-GFP and Sox17-RFP in 2D colonies (Figure 4A top row; Figure S8A top). Replacing CHIR at 24 h with an inhibitor of endogenous Wnt signaling (using the porcupine inhibitor IWP2, thus inhibiting Wnt ligand secretion) reduced the number of Sox17 + colonies in a GFP-intensity dependent manner: GFP-high colonies later yielded Sox17-RFP+ cells, while GFP-low colonies failed to do so (Figure 4A middle rows, b, c, and d). This effect is abolished when IWP2 is added 12 h after CHIR, leading to no Sox17-RFP colonies (Figure S8). When IWP2 is added after 24 h while maintaining CHIR, all the Bra-GFP+ colonies give rise to Sox17-RFP populations, while Bra-GFP- colonies remain negative (Figure 4A bottom row, B, C). Similar treatment conditions on 3D EBs result in a similar pattern. Wnt inhibition blocks additional Bra expression, leading to lower Bra-GFP levels at the end of the experiment, compared to control (Figure 4E). Moreover, the Bra-GFP levels at the start of treatment serve as a proxy to the final RFP levels (Figure 4F). In Bra-GFP-low EBs, the Wnt inhibition significantly reduces the Sox17-RFP expression ($p = 0.0001$) (Figure 4F left) while Bra-GFP-high EBs show no difference in Sox17-RFP expression between treatment and control (Figure 4F right). Taken together, these results suggest that a subpopulation of Bra+ cells gain their potential to differentiate to a Sox17+ state shortly after Bra onset. Cells that are specified for DE differentiation do not depend on the activation of Wnt signaling anymore, and can differentiate under its inhibition at this point. Bra+ cells that are not specified for activation of Sox17 at this point will fail to further differentiate to a Sox17+ state under Wnt inhibition.

It was previously shown that treatment with Activin allows DE differentiation and Sox17 expression (van den Brink et al., 2014; Yasunaga et al., 2005). In our system, the Wnt signaling activation as an external field is sufficient for Sox17 emergence. We therefore wanted to dissect the role of Wnt signaling and Activin\Nodal signaling on the onset of Sox17, using a different double-reporter cell line with Wnt (SuTOP-CFP) and Nodal/Activin (AR8-RFP) pathway reporters. Under our standard differentiation protocol (Figure 2A), CHIR leads to activation of the Activin\Nodal pathway (Figure S9A(Turner et al., 2017)). When replacing CHIR with Activin as an external signal, the onset timing of the Sox17 cells was delayed (Figure S9B). When replacing CHIR at 24 h with the Activin\Nodal pathway inhibitor SB431542 (SB), the potential of Bra-GFP+ colonies to differentiate to Sox17-RFP+ cells was reduced by ~3-fold and was not rescued by the addition of CHIR (Figures S9C andS9D). Adding SB with CHIR from the beginning results in no mesodermal nor endodermal differentiation (Figure S9E), similar to what was shown *in vivo* with a Nodal null mutant (Camus et al., 2006). These results suggest that both Wnt and Activin signaling pathways are needed for the initial differentiation of Bra, and commitment for DE emergence. While Wnt signaling pathway is crucial for Bra differentiation and Wnt signaling inhibition in mature Bra-GFP+ cells does not prevent Sox17 expression, Activin\Nodal is required for the expression of Sox17 throughout the differentiation process. Our findings in the *in vitro* systems suggest that Wnt activation is needed for the onset

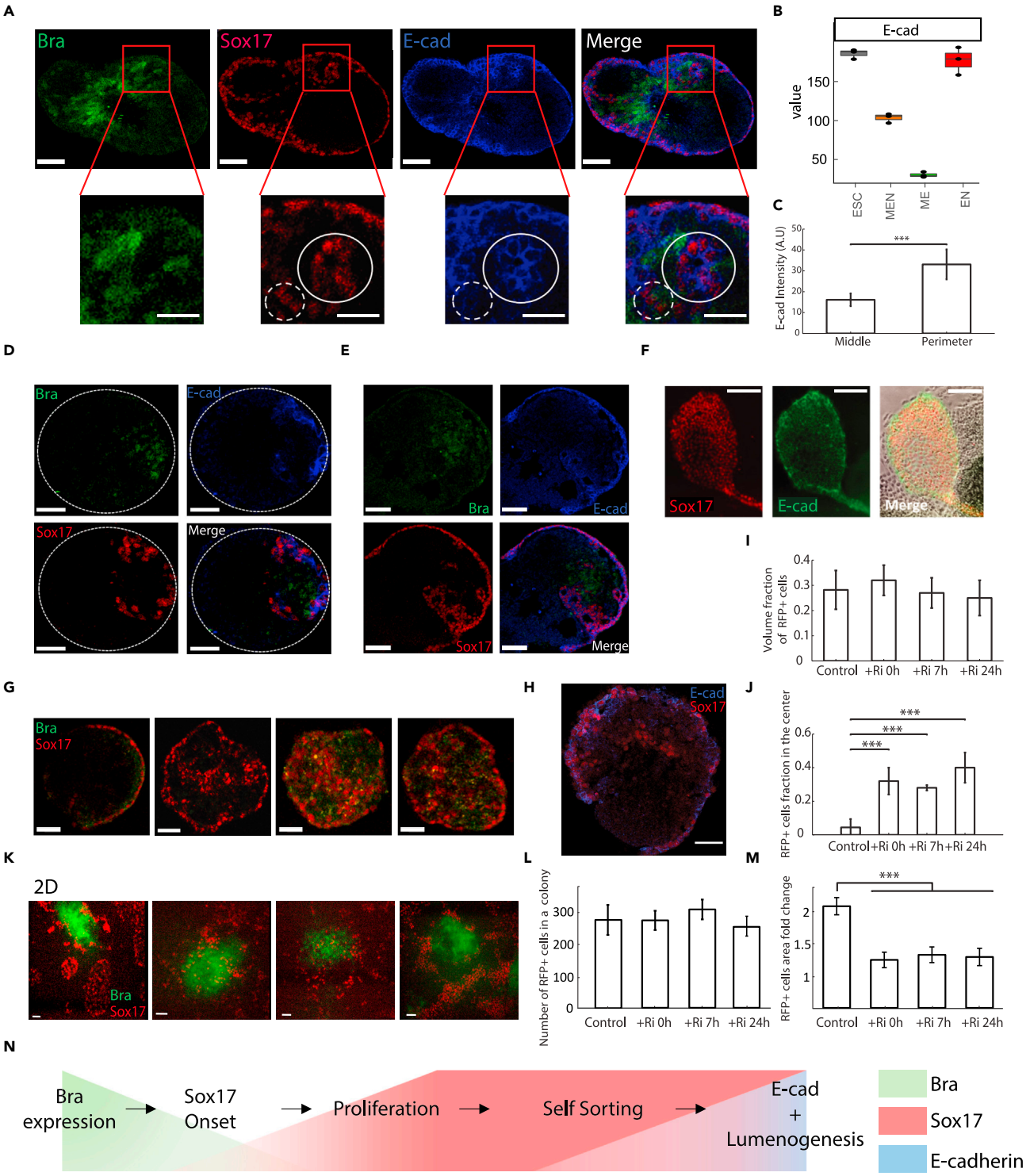


Figure 5. Self-sorting is not required for establishment of Sox17+ population, and is concomitant with E-cad upregulation

(A) E-cad immunostaining at 48 h of differentiation. E-cad is restricted to the Sox17 cells and not the Bra expressing cells. The highest E-cad expression is obtained in the Sox17 cells surrounding the outer layer of the EB or forming an interior lumen. Scale bar 50 μm. Bottom row - Zoom-in in the same representative EB. E-cad expression level depends on Sox17 cells location and aggregation stage. Scale bar 25 μm.

(B).TPM values of the E-cad expression by undifferentiated ESCs and differentiated EBs, 48 h post CHIR addition, sorted to Bra-GFP+ (ME, green), Bra-GFP+ Sox17-RFP+ (MEN, orange), and Sox17-RFP+ (DE, red) populations, as was measured by bulk RNA sequencing. Double-positive cells have lower E-cad mRNA levels compared with EN cells. Data are represented as mean ± standard deviation. (n = three independent differentiation and sorting assays).

Figure 5. *Continued*

(C) E-cad intensity in the cells in the middle of the EB is significantly lower than that on the perimeter of the EB (n = 10, from two independent experiments, $P = 10^{-5}$).

(D) E-cad is restricted to the Sox17 cells on the outer layer of the EB. Scale bar 50μm.

(E and F) E-cad expression of Sox17 cells in 72 h differentiated EBs (E) or 2D colonies (F). Scale bar 50 μm and 100 μm, respectively.

(G) Representative EBs at 65 h into differentiation, without (left) or with addition of ROCK inhibitor (Y-27632, 20 μM) at 0 h, 7 h, or 24 h after the addition of CHIR (Left to right). Scale bar 50 μm.

(H) An EB subjected to ROCK inhibition at 48 h, stained for E-cad (blue).

(I) Number of Sox17-RFP cells per EB, normalized to the EB volume, for different timings of ROCK inhibitor addition to the medium.

(J) Fraction of Sox17+ cells in the middle of the EB out of all Sox17 cells, for different ROCK inhibitor addition timings. (n = 8).

(K) Representative 2D colonies at 48 h into differentiation, without (left) or with addition of ROCK inhibitor (Y-27632, 20 μM) at 0 h, 7 h or 24 h after the addition of CHIR (Left to right). Scale bar 100 μm.

(L) Number of RFP+ cells in a colony. No significant difference between conditions (n = 6 colonies per condition, sampled from three independent wells for each condition).

(M) Fold change of the RFP+ covered area for each colony between 42 h and 48 h of differentiation.

(N) A summarized model of the differentiation pattern dynamics. ***P < 0.001. **P < 0.01. *P < 0.05

phase but not for the expansion and self-sorting phase (Figure 3C), and recapitulate the observations from embryos, where Wnt-high posterior environment allows DE onset, followed by migration to anterior, Wnt-low regions and intercalation of the cells.

Sox17+ cells expand through self-sorting in correlation with E-cadherin upregulation

The differentiation of mesendoderm cells to DE is known to be associated with reactivation of E-cadherin (Acloque et al., 2009; Kalluri and Weinberg, 2009). Single-cell RNA sequencing confirms that Sox17+ DE cells at E7.5 upregulate E-cad expression (Figure S1). We hypothesized that the self-sorting movements of the Sox17-RFP+ cells are correlated with the mesenchymal to epithelial transition (MET) and E-Cadherin (E-cad) expression. It has been shown on mutant mice that Sox17 is required for DE cell egression, and that E-cad was differentially expressed in egressing DE cells compared with cells that fail to egress (Ferrer-Vaquer et al., 2010; Viotti et al., 2014). In our systems, differentiating EBs show low expression of E-cad by Bra-GFP+ cells, and upregulation of E-cad expression shows low expression during DE differentiation both at the protein and at the RNA levels (Figures 5A and 5B). Moreover, Sox17-RFP+ cells that have already migrated to the outer layer of the EB or have formed a confined lumen show the highest E-cad expression (Figure 5A). Sox17-RFP+ unsorted cells show low to no E-cad expression (Figures 5A and 5C, Video S4). Outer-layer EB cells that are Sox17-RFP-, do not express E-cad (Figure 5D). Moreover, at a later stage, when all Sox17-RFP cells are sorted, the high E-cad expression is seen in all Sox17-RFP+ cells and is limited to Sox17-RFP+ cells only (Figures 5E and 5F; Video S5).

These results suggest that Sox17+ cells dispersed in the EB start to accumulate E-cad, and possibly the low concentration of E-cad expression is sufficient to initiate self-sorting. The aggregation of the Sox17+ cells results in an upregulation and higher expression of E-cad. To test the effect of cell movement and resulting aggregation on the final distribution of Sox17-RFP+ vs. Sox17-RFP- populations, we blocked cell movement by adding the Rho-associated protein kinase inhibitor Y-27632 (ROCKi) at different time points during differentiation. As expected, the addition of ROCKi reduced significantly the self-sorting of the Sox17-RFP+ cells, resulting in a higher fraction of these cells in the interior of the EB at 65h compared to the control, where these interior cells are not forming lumens nor express E-cad (Figures 5G, 5H, and 5J). The final fraction of Sox17-RFP+ cells in each EB was not altered by the addition of ROCKi (Figure 5I). Cell-movement inhibition by adding ROCKi at different time points during differentiation in 2D colonies, led to similar outcome. While the treatment inhibits the cell movement, measured by the fold change of the RFP+ covered area per colony between 42 and 48 h (Figures 5K and 5M), the number of RFP+ cells per colony was not affected by the treatment (Figures 5K and5L). Taken together, our results show that Sox17 onset within the Bra+ population is followed by proliferation and self-sorting dynamics of these cells (Figure 3G). Self-sorting, aggregation, and increase in the expression of E-cad occur after cells already express Sox17, and are not required for Sox17 onset and proliferation (Figure 5N).

DISCUSSION

Here, we explored the onset and expansion dynamics of mesendoderm-derived DE cells. Through tetraploid complementation experiments with a dual live marker, we were able to uniquely follow epiblast-derived DE in the embryo. The first Sox17-RFP+ cells arise in a dispersed pattern in the epiblast, though

limited to the anterior PS, followed by expansion and migration. We then demonstrated in both 2D and 3D *in vitro* cultures, that Sox17+ cells emerge in a salt-and-pepper pattern within the Bra+ cell population. In this sense, the *in vitro* conditions expand the Sox17-competent region compared to the embryo, while maintaining the disperse pattern. This expansion may be attributed to more uniform signals sensed by the cells in the *in vitro* systems, compared to the embryo, where localized sources of Wnt and Nodal signaling result in spatial gradients of these signals. This onset pattern is in sharp contrast to the strongly localized pattern of Brachyury onset (Boxman et al., 2016; ten Berge et al., 2008; van den Brink et al., 2014). Our evidence from *in vivo* and *in vitro* systems, as well the single-cell tracing suggests that most of the Sox17+ cells were Bra+ cells that downregulated their Bra expression shortly before or around Sox17 onset. The discrepancy from recent *in vivo* reports may be attributed to the longer expression memory provided by the GFP reporter, compared to measurements at the mRNA level that found little overlap between these genes (Probst et al., 2021). The ordered pattern of Sox17+ cells is then obtained through self-sorting of the cells. The transition from stochastic to self-sorted pattern is emphasized in the *in vitro* systems due to the spatial expansion of the Sox17-competent region. Our quantitative results in embryoid bodies suggest that in the embryo, mesendoderm progenitors, while egressing through the primitive streak, stochastically commit to definitive endoderm, and then through cadherin-mediated self-sorting, get spatially sorted apart from mesoderm cells (Kwon et al., 2008; Viotti et al., 2014). In the embryo, the sparse DE cells egress at this stage and intercalate into the VE layer, while in the *in vitro* systems, in the absence of VE, they cluster together to form lumens or an envelope layer (Burtscher and Lickert, 2009). The signaling gradient patterns and signal accessibility in different regions differ greatly between the 2D and 3D setups (Etoc et al., 2016). The robustness of both the relative temporal pattern between Bra and Sox17 and the spatial stochasticity in DE onset in these two model types therefore suggests its difference in internal molecular state that determines which cells will adopt a DE fate, as opposed to difference in external signaling.

In the Embryo, Nodal and Wnt gradients are formed as the primitive streak progresses, allowing pattern formation during gastrulation (Norris and Robertson, 1999). *In vitro*, it was shown that a 24-h pulse of CHIR (with or without Activin) induces EB gastrulation-like events, including Bra and Sox17 expression (van den Brink et al., 2014). Moreover, it was shown that Activin induces DE differentiation and Sox17 expression in Bra+ cells (Gadue et al., 2006; Kubo et al., 2004). Here, we showed that even a short Wnt activation is sufficient to induce Bra, and shortly after, a small subpopulation of the Bra+ cells is already specified to their Sox17+ fate independently of external Wnt activation. The correlation we find between Sox17 activation potential under Wnt inhibition and the colony or EB GFP level, suggests that self-sustained Wnt signaling (which is coupled to Bra through positive feedback (Evans et al., 2012; Yamaguchi et al., 1999)) may provide the needed Wnt activation in the future Sox17+ population. Alternatively, the separation we observe between Sox17 onset phase and a later expansion phase through cell division, suggests that Wnt signaling may not be required for the latter, consistent with DE cell anterior migration and expansion in the embryo, away from Wnt signaling sources. Our results further suggest that a disperse local activation of the Activin\Nodal pathway may partially explain this potentiated subpopulation, but cell-to-cell variation in expression or chromatin state may also account for the differential DE potential. A better understanding of the specific chromatin and expression determinants of this potential will provide a comprehensive picture of this spatially stochastic cell fate decision.

We find that *in vitro*, E-cadherin expression increases in parallel to the self-sorting of Sox17+ cells. This is consistent with previous observations in the embryo, where high E-cadherin is observed in DE cells only after they have egressed and intercalated into the VE (Ferrer-Vaquer et al., 2010; Viotti et al., 2014). Recently, it was shown that the segregation of DE cells from the primitive streak is facilitated by epithelial cell plasticity rather than a full EMT-MET cycle, resulting in E-cadherin expressing intercalated DE cells (Scheibner et al., 2021). The sparse, spatially stochastic appearance of DE cells, along with the late induction of E-cadherin may facilitate the intercalation process, preventing premature formation of DE cell clusters or lumens within the epiblast, such as the ones formed in EBs. Once self-sorting is completed in the EBs or 2D colonies, E-cadherin expression is unique to the Sox17+ populations. Limiting cell movement, by inhibition of ROCK, an important controller of actin microfilaments, did not affect the differentiation to Sox17+ cells or their proliferation, regardless of the timing of inhibition. It was previously shown that ROCK inhibition and cytoskeletal reorganization can mediate activation of Wnt signaling and the induction of EMT, and can affect cell differentiation *in vitro* (Galli et al., 2012; Korostylev et al., 2017; Maldonado et al., 2016). While ROCK inhibition early in differentiation may induce EMT and mesendoderm differentiation and by that enhance Sox17 expression later on, the emergence of a similar fraction of Sox17+ cells

when ROCKi is added later in differentiation indicates that Sox17 expression does not require cell movement. The order of Sox17 onset, E-cadherin upregulation, and cell movement suggests the activation of Sox17 does not depend on the latter two, but may drive these processes. Further research is needed to understand the mechanisms for delayed E-cadherin upregulation and accumulation. Nonetheless, our *in vivo* and *in vitro* results highlight basic rules governing DE onset and patterning through the commonalities and differences between these systems. Importantly, the *in vitro* system allows us to propose a sequential model of events leading to this developmental pattern.

Limitation of the study

Although we showed the temporal and spatial dynamics of the DE and ME segregation and its dependencies on Wnt signaling and cell movement, a detailed characterization of the first few cells that will become DE is still needed. In addition, some of the cell-movement aspects of the described model are yet to be verified *in vivo*. Finally, we used mouse cells and embryos for our experiments. This dynamic is yet to be confirmed in human cells.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - cell culture
  - Sox17-RFP live marker
  - Differentiation assay
  - Live imaging
  - Immunohistochemistry
  - Aggregation of mouse embryonic stem cells and animal procedures
  - Embryo isolation, staging and imaging
  - Flow cytometry
  - Bulk RNA sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Segmentation and image analysis
  - Statistical analysis
  - Single-cell RNA sequencing analysis of E6.5 and E7.5 mouse embryos data from Nowotschin et al (2019)

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103556.

ACKNOWLEDGMENTS

We thank Z.D. Smith and A. Arczewska for help with CRISPR knockin design and implementation, and G.S. Franca for graphic design. This work was supported by the Israel Science Foundation (1665/16), the Good Food Institute research program, the New York Stem Cell Foundation, the NIH (P01 GM099117 and P50 HG006193) as well as the Max Planck Society.

AUTHOR CONTRIBUTIONS

MP and IN designed and conceived the study. MP, ASK, and NF performed experiments. LW and MW assisted with tetraploid aggregation and embryo isolation. AB and HK performed RNA-seq experiments and analysis, respectively. MP performed the data analysis. IN and AM oversaw the project. MP and IN wrote the paper with contributions from ASK and AM.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 11, 2021
Revised: October 21, 2021
Accepted: December 1, 2021
Published: January 21, 2022

REFERENCES

Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. J. Clin. Invest. *119*, 1438–1449.

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

Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A.C., Lutolf, M.P., Duboule, D., and Arias, A.M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. Nature *562*, 272–276.

Boxman, J., Sagy, N., Achanta, S., Vadigepalli, R., and Nachman, I. (2016). Integrated live imaging and molecular profiling of embryoid bodies reveals a synchronized progression of early differentiation. Sci. Rep. *6*, 31623.

Burtscher, I., and Lickert, H. (2009). Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. Development *136*, 1029–1038.

Camus, A., Perea-Gomez, A., Moreau, A., and Collignon, J. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. Dev. Biol. *295*, 743–755.

Engert, S., Burtscher, I., Liao, W.P., Dulev, S., Schotta, G., and Lickert, H. (2013). Wnt/beta-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. Development *140*, 3128–3138.

Etoc, F., Metzger, J., Ruzo, A., Kirst, C., Yoney, A., Ozair, M.Z., Brivanlou, A.H., and Siggia, E.D. (2016). A balance between secreted inhibitors and Edge sensing controls gastruloid self-organization. Dev. Cell *39*, 302–315.

Evans, A.L., Faial, T., Gilchrist, M.J., Down, T., Vallier, L., Pedersen, R.A., Wardle, F.C., and Smith, J.C. (2012). Genomic targets of Brachyury (T) in differentiating mouse embryonic stem cells. PLoS One *7*, e33346.

Fehling, H.J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., and Kouskoff, V. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. Development *130*, 4217–4227.

Ferrer-Vaquer, A., Viotti, M., and Hadjantonakis, A.K. (2010). Transitions between epithelial and mesenchymal states and the morphogenesis of the early mouse embryo. Cell Adh.Migr. *4*, 447–457.

Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc. Natl. Acad. Sci. U S A *103*, 16806–16811.

Galli, C., Piemontese, M., Lumetti, S., Ravanetti, F., Macaluso, G.M., and Passeri, G. (2012). Actin cytoskeleton controls activation of Wnt/beta-catenin signaling in mesenchymal cells on implant surfaces with different topographies. Acta Biomater. *8*, 2963–2968.

Harrison, S.E., Sozen, B., Christodoulou, N., Kyprianou, C., and Zernicka-Goetz, M. (2017). Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. Science *356*, eaal1810.

Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., Ladopoulos, V., Scialdone, A., Jorg, D.J., Tyser, R.C.V., Calero-Nieto, F.J., Mulas, C., Nichols, J., et al. (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. Nat. Cell Biol. *20*, 127–134.

Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J. Clin. Invest. *119*, 1420–1428.

Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P., and Hayashi, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. Development *129*, 2367–2379.

Kapalczynska, M., Kolenda, T., Przybyla, W., Zajaczkowska, M., Teresiak, A., Filas, V., Ibbs, M., Bliźniak, R., Luczewski, L., and Lamperska, K. (2018). 2D and 3D cell cultures - a comparison of different types of cancer cell cultures. Arch. Med. Sci. *14*, 910–919.

Korostylev, A., Mahaddalkar, P.U., Keminer, O., Hadian, K., Schorpp, K., Gribbon, P., and Lickert, H. (2017). A high-content small molecule screen identifies novel inducers of definitive endoderm. Mol. Metab. *6*, 640–650.

Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., and Keller, G. (2004). Development of definitive endoderm from embryonic stem cells in culture. Development *131*, 1651–1662.

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

Maldonado, M., Luu, R.J., Ramos, M.E., and Nam, J. (2016). ROCK inhibitor primes human induced pluripotent stem cells to selectively differentiate towards mesendodermal lineage via epithelial-mesenchymal transition-like modulation. Stem Cell Res. *17*, 222–227.

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

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes Dev. *24*, 312–326.

Norris, D.P., and Robertson, E.J. (1999). Asymmetric and node-specific nodal expression patterns are controlled by two distinct cis-acting regulatory elements. Genes Dev. *13*, 1575–1588.

Nowotschin, S., and Hadjantonakis, A.K. (2018). Lights, camera, action! Visualizing the cellular choreography of mouse gastrulation. Dev. Cell *47*, 684–685.

Nowotschin, S., Setty, M., Kuo, Y.Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature *569*, 361–367.

Probst, S., Tosic, J., Schwan, C., Grun, D., and Arnold, S.J. (2021). Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation. Development *148*, dev193789.

Rivera-Perez, J.A., and Hadjantonakis, A.K. (2014). The dynamics of morphogenesis in the early mouse embryo. Cold Spring Harb. Perspect. Biol. *7*, a015867.

Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.C., Korving, J., Vivie, J., Truckenmuller, R.K., van Oudenaarden, A., van Blitterswijk, C.A., and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. Nature *557*, 106–111.

Robertson, E.J. (2014). Dose-dependent Nodal/Smad signals pattern the early mouse embryo. Semin. Cell Dev. Biol. *32*, 73–79.

Scheibner, K., Schirge, S., Burtscher, I., Büttner, M., Sterr, M., Yang, D., Böttcher, A., Irmler, M., Beckers, J., Cernilogar, F.M., et al. (2021). Epithelial cell plasticity drives endoderm formation during gastrulation. Nat. Cell Biol. *23*, 692–703.

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

Sozen, B., Amadei, G., Cox, A., Wang, R., Na, E., Czukiewska, S., Chappell, L., Voet, T., Michel, G., Jing, N., et al. (2018). Self-assembly of embryonic and two extra-embryonic stem cell types into

gastrulating embryo-like structures. Nat. Cell Biol. *20*, 979–989.

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

ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., and Nusse, R. (2008). Wnt signaling mediates self-organization and axis formation in embryoid bodies. Cell Stem Cell *3*, 508–518.

Turner, D.A., Girgin, M., Alonso-Crisostomo, L., Trivedi, V., Baillie-Johnson, P., Glodowski, C.R., Hayward, P.C., Collignon, J., Gustavsen, C., Serup, P., et al. (2017). Anteroposterior polarity and elongation in the absence of extra-embryonic tissues and of spatially localised signalling in gastruloids: mammalian embryonic organoids. Development *144*, 3894–3906.

van den Brink, S.C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.K., Nowotschin, S., Turner, D.A., and Martinez Arias, A. (2014). Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. Development *141*, 4231–4242.

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

Warmflash, A., Sorre, B., Etoc, F., Siggia, E.D., and Brivanlou, A.H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. Nat. Methods *11*, 847–854.

Yamaguchi, T.P., Takada, S., Yoshikawa, Y., Wu, N., and Mcmahon, A.P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes Dev. *13*, 3185–3190.

Yamamoto, M., Saijoh, Y., Perea-Gomez, A., Shawlot, W., Behringer, R.R., Ang, S.L., Hamada, H., and Meno, C. (2004). Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. Nature *428*, 387–392.

Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L.M., Nishikawa, S., Chiba, T., Era, T., and Nishikawa, S. (2005). Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. Nat. Biotechnol. *23*, 1542–1550.

Yoney, A., Etoc, F., Ruzo, A., Carroll, T., Metzger, J.J., Martyn, I., Li, S., Kirst, C., Siggia, E.D., and Brivanlou, A.H. (2018). WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids. Elife *7*, e38279.

Zhou, X., Sasaki, H., Lowe, L., Hogan, B.L., and Kuehn, M.R. (1993). Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. Nature *361*, 543–547.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Sox17	R&D Systems	AF1924; RRID:AB_355060
Anti-brachyury	R&D Systems	AF2085; RRID:AB_2200235
Anti-E-cadherin	Cell Signaling	24E10; RRID:AB_2291471
Biological samples		
E14 Bra-GFP mouse ESCs	provided by Dr. Gordon Keller	
AR8-RFP\SuTOP-CFP mouse ESCs	provided by Dr. Palle Serup	
Chemicals, peptides, and recombinant proteins		
Knockout DMEM	Gibco	10,829–018
Fetal bovine serum	Biological Industries	040071A
Penicillin/streptomycin	Biological Industries	03-033-1B
Glutamine	Biological Industries	03-020-1A
Non-essential amino acids	Biological Industries	01-340-1B
β-mercaptoethanol	Sigma	D2650
LIF	Merck-Millipore	ESG1106
Hygromycin	A.G Scientific	H-1012-PBS
Puromycin	Sigma	P9620
N2	Gibco	17502-048
B27 supplement	Gibco	17504-044
Gelatin	Sigma	G2500-100G
CHIR99021	Biovision	1677
PD0325901	Santa Cruz biotechnology	sc- 205427
Knockout serum replacement	Thermo Fisher Scientific	10828028
IWP2	ApexBio	A3512-10
Activin	Peprotech	120-14E
SB-431542	Sigma	S4317
Y-27632	ChemScene LLC	CS-0878
PBS	Sigma	D1408
4% paraformaldehyde	Merck-Millipore	104005
Triton ×100	Sigma	X100-100M
Critical commercial assays		
Xfect transfection reagent	Takara	Cat. # 631,318
Oligonucleotides		
fwd: 'TCTTGGAAGGCGTTGACCTTGGC'	This study	upstream of the SOX17 stop codon, forward primer
rev: 'GGTCTACTATTGCAACTACCCCGACATT'	This study	upstream of the SOX17 stop codon, reverse primer
fwd: 'TGACGGTTGCCGACCCGACCT'	This study	downstream of the SOX17 stop codon, forward primer
rev: 'CGCAGGTGAAGACACAAACCAGCATCC'	This study	downstream of the SOX17 stop codon, reverse primer
gRNA 'GCAACTACCCCGACATTTGA'	This study	Sox17 Guide RNA

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pSpCas9(BB)-2A-Puro (PX459)	Zhang laboratory	Addgene plasmid # 62988
Software and algorithms		
ImageJ	Schneider et al. (2012)	https://imagej.nih.gov/ij/
FlowJo	FlowJo	V 10
MATLAB	MathWorks	R2018a
Imaris	Oxford Instruments	8.3
Deposited data		
RNA-seq	This study	GEO accession: GSE189224

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Dr. Iftach Nachman (iftachn@tauex.tau.ac.il)

Materials availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact.

Data and code availability

The accession number for RNA-seq data reported in this paper is GEO: GSE189224.

METHOD DETAILS

cell culture

Mouse ESCs were cultured on gelatin surface using standard conditions on irradiated primary mouse embryonic fibroblasts and knockout DMEM containing 15% fetal bovine serum, 50 μg/mL penicillin/streptomycin, 2 mM L Glutamine, 100 μM non-essential amino acids, 80 μM β-mercaptoethanol and 10^3 U/mL LIF. Cell lines used are E14 Bra-GFP mouse ESCs (kindly provided by Dr. Gordon Keller), AR8-RFP\SuTOP-CFP mouse ESCs (kindly provided by Dr. Palle Serup).

Sox17-RFP live marker

A sox17 homology donor (SOX17-HomDon) plasmid was generated, containing (i) a left homology arm containing a 1-kb sequence immediately upstream of the SOX17 stop codon (fwd: 'TCTTGGAAGGCGTT GACCTTGGC'; rev: 'GGTCTACTATTGCAACTACCCCGACATT'); (ii) a P2A-H2B- mStrawberry cassette; (iii) a floxed Hygromycin selection cassette (loxP-PGK-Hygro-loxP); and (iv) a right homology arm containing a 1-kb sequence immediately downstream of the SOX17stop codon (fwd: 'TGACGGTTGCCGACCC GACCT'; rev: 'CGCAGGTGAAGACACAAACCAGCATCC'. A single-guide RNA (sgRNA) recognizing a sequence near the stop codon of SOX17 ('GCAACTACCCCGACATTTGA') was cloned into a Cas9-nickase expression vector pSpCas9(BB)-2A-Puro (PX459) from the Zhang laboratory, Addgene plasmid # 62,988). This plasmid, together with the SOX17-HomDon plasmid, was transfected into E14 Bra-GFP cells using Xfect transfection reagent (Cat. # 631318, Takara). After seven days of antibiotic selection, a single clone was validated with Sanger sequencing from the genomic Sox17 locus. After Cre-floxing out of the Hygromycin selection marker, colonies were screened for negative selection, a single clone was validated by sequencing the Sox17 endogenous gene region again, confirming the Sox17 gene sequence followed by the P2A-H2B- mStrawberry cassette adjacent to the Sox17 stop codon. Primers used for verification before and after floxing were fwd: GATGGCACGGAATCCAACCAG (inside the Sox17 gene) and rev: GGAGCAAGTCCCTCAAGGCATT (downstream to the Sox17 stop codon). The final clone was expanded and frozen in aliquots for use in all experiments.

Differentiation assay

Cells were cultured in serum-free N2B27 media supplemented with LIF and and 2i (3 μM CHIR99021 and 1 μM PD0325901) for a minimum of 24 hours. Prior to 2D differentiation, ESCs were fully dissociated to single cells and seeded at low density (~12,000 cells/cm2) on gelatin-coated plates in LIF 2i medium supplemented with 5% KSR. Differentiation was induced a minimum of 18h later by rinsing ESCs thoroughly with 1X PBS and changing to N2B27 media without LIF 2i. For 3D Embryoid bodies' aggregation, single cells were transferred to a low-adherent culture dish, allowing cells to randomly aggregate together in small clumps, in N2B27 media without LIF 2i. In both 2D and 3D differentiation protocols, after 48hrs in in N2B27, media was changed to N2B27 media with CHIR99021 (3 μM).

For Wnt\Activin signaling pathways interplay experiments, cells were supplemented with CHIR99021 (3 μM) or IWP2 porcupine inhibitor (10μM) or Activin (20ng/mL) or SB-431542 Activin receptor inhibitor (10μM), as indicated in the relevant result. For movement inhibition, ROCK inhibitor (Y-27632, 20μM) was added at the described time points.

Live imaging

2D differentiation experiments were imaged using a Nikon TiE epi-fluorescence microscope equipped with a motorized XY stage (Prior) and taken within a connected 6 × 6 or 7 × 7 spatial range at 10× magnification in up to three fluorescent wavelengths and phase contrast using NIS Elements software.

3D differentiation experiments were imaged using a Zeiss LSM7 inverted two-photon microscope with a 20X/0.8NA air objective. Each EB was scanned at 3μm intervals along the *z* direction. Horizontal resolution was set to 512 × 512 pixels at approximately 0.6μm per pixel. GFP was excited at 920nm. RFP was excited at 760nm, Alexa 488, Alexa 405 and Alexa 594 were excited at 800nm. CFP was excited at 860 nm.

For both systems, acquisitions were taken every 30-60 minutes for 2-5 days. For live imaging, an Okolab incubation cage, maintaining 5% CO_2 and 37°C was used.

Immunohistochemistry

For immunostaining, cells were fixed in 4% paraformaldehyde (PFA) and immunostained for the indicated primary antibody at the concentration recommended by the manufacturer, at 4C O\N (unless indicated otherwise). For 2D colonies, cells were then washed 3 times in PBS and incubated with a secondary antibody for one hour at RT. 3D EBs were washed in four cycles in blocking buffer (PBS, 0.15% Triton x100, 7% FBS) for at least half an hour per cycle. A secondary antibody in blocking buffer was incubated o/n, followed by four wash cycle as described for the primary antibody. Antibodies used: Sox17 - R&D Systems AF1924 (3 hrs at RT); Bra - R&D Systems AF2085; E-cadherin - Cell Signaling 24E10.

Aggregation of mouse embryonic stem cells and animal procedures

Tetraploid complementation was performed with mESCs thawed from a frozen vial and cultured for three days on a layer of CD-1 feeders (Artus, J. & Hadjantonakis, A. K.). Aggregated embryos were re-transferred bi-laterally in a clutch of 10-15 embryos into uterine horn of CD-1 strain pseudopregnant foster females. All animal procedures were performed in accordance with the institutional, governmental and state regulations (LAGeSo Berlin, G0243/18).

Embryo isolation, staging and imaging

Embryos from the uterine decidua were dissected in cold 1X HBSS and the Reichert's membrane was carefully removed with a thin-drawn glass capillary. Since there is a natural variation in the development of embryos in the early stages of gastrulation which is more pronounced in tetraploid embryos, the embryos were staged and picked for analysis based on Theiler's nomenclature. The embryos were transferred through three washes in 1X PBS with 0.4% BSA and fixed overnight in 4% PFA at 4°C. The embryos were washed three times for 5 minutes each in 1X cold PBS with 0.1% TritonX-100. Nuclei were counterstained with 0.24μg/mL of DAPI in 1X PBS with 0.25% Triton X-100 at 4°C for 2 hours, followed by three washes with 1X PBS with 0.1% TritonX-100. Images were acquired with a Zeiss LSM880 AiryScan confocal microscope at 10X magnification with a z-stack of 2.73μm interval and averaging of 8 frames. Images were processed Imaris and Fiji.

Flow cytometry

Flow cytometry for differentiation kinetics of the reporters was performed on the BD Accuri C6 flow cytometer. 100,000 events were acquired and the cells were gated for size and singlets using forward (FSC-A/W) and side scatter (SSC-A/W). GFP and RFP was measured by excitation of blue and yellow-green laser and filtered through 530/30 and 610/20 band pass filter respectively. The gated data was exported and processed using FlowJo version 10.

Bulk RNA sequencing

For Bulk RNA sequencing experiment, EBs were differentiated for 48hrs after the addition of CHIR, and were collected and sorted to three populations - Bra-GFP+ (ME), Bra-GFP+ Sox17-RFP+ (MEN), and Sox17-RFP+ (DE). Undifferentiated ESC cell line were served as a control.

Three independent differentiation assays and sorting were performed and serve as three replicates, analyzed together. Paired-end reads were aligned with STAR v2.5.1 using default options. Gene counts were quantified using featureCounts v1.6.3. For featureCounts, fragments were counted at the gene-name level. Alignment and gene counts were generated against the GRCm38/mm10 genome assembly. STAR v2.5.1 and GRCm38/mm10 were included with CellRanger 3.0.2.

QUANTIFICATION AND STATISTICAL ANALYSIS

Segmentation and image analysis

Epi-fluorescence microscopy data for 2D colonies imaging was contrast enhanced in Fiji. The Sox17-RFP single cell segmentation and cell counting was done in Cell Profiler, for each colony separately. In order to average on colonies of different sizes, the cell number counted in each colony was normalized to its maximal value (the cell number counted in the time point of the highest cell count for each one of the colonies separately). The data was read and analyzed in MATLAB.

The Two-photon microscopy data for 3D EBs imaging was contrast enhanced in Fiji and then spot segmented in Imaris, where the integration radius was set to 7μm and the filter was set to sum intensity with a threshold of 110/255. For the GFP cytoplasmic signal, this resulted in a fixed cells per segmented point ratio, as previously shown (Boxman et al., 2016). The background is subtracted prior to segmentation. The segmented data was read and analyzed in MATLAB. In the analysis, the number of segmented spots was approximated to cell number and is equivalent to per colony intensity measured in 2D. Sox17 onset time was defined as the time point when 7 segmented RFP+ cells above intensity threshold were detected. For self-sorting analysis of the Sox17-RFP+ cells, 'Surface' Imaris tool were used. E- cad Alexa 405 staining intensity level was measured with Fiji.

Statistical analysis

For the analysis of Sox17-RFP+ cells onset pattern, segmented RFP cells and segmented GFP cells were analyzed for their location at Sox17 onset time point within the EBs. Images were cleaned from detected clumps of dead cells outside of the EB radius in order to prevent misleading distance measurements. The distances between every two RFP+ cells as well as between every two GFP+ cells were then calculated. The distributions of the distances were compared using two-sample Kolmogorov-Smirnov test under the null hypothesis that the two sets of distances are from the same continuous distribution. The results represent the analysis of 5 EBs of similar size.

For quantitative comparison of fluorescent intensities between two cell or colony groups, intensity measurements were taken using Fiji, and the significance P-value was calculated using two-tailed t-test in MATLAB. All experiments were repeated at least twice, with similar results. Results are reported from a single repeat, unless otherwise noted.

Single-cell RNA sequencing analysis of E6.5 and E7.5 mouse embryos data from Nowotschin et al (2019)

Publicly available data from Nowotschin et al (2019) [25] was downloaded and analyzed in R using Seurat V4 package. First, subsetting the data was preformed based on the Nowotschin et al (2019) annotations of day of embryonic development (time point, E6.5 and E7.5) and the developmental layer (cell type, Mes and DE), resulting in 6028 cells. The subsets of cells were than merged together to one Seurat object, were scaled and normalized together. Seurat was used to display dot plot and feature scatter plots of our genes of interest.