Cell-state transitions and collective cell movement generate an endoderm-like region in gastruloids

Ali Hashmi[1], Sham Tlili[1], Pierre Perrin[1], Alfonso Martinez-Arias[2,3], Pierre-François Lenne[1*]

1. Aix-Marseille University, CNRS, UMR 7288, IBDM, Turing Center for Living Systems, Marseille, France
2. Department of Genetics, University of Cambridge, United Kingdom
3. Systems Bioengineering, DCEXS , Universidad Pompeu Fabra, Doctor Aiguader, 88, ICREA , Pag Lluis Companys 23 , Barcelona, Spain

[*]Corresponding author: pierre-francois.lenne@univ-amu.fr

Abstract

Shaping the animal body plan is a complex process that involves the spatial organization and patterning of the different germ layers. Recent advances in live imaging have started to unravel the cellular choreography underlying this process in mammals, however, the sequence of events transforming an unpatterned cell ensemble into structured territories is largely unknown. Here, using gastruloids -3D aggregates of mouse embryonic stem cells- we study the formation of one of the three germ layers, the endoderm. We show that the endoderm is generated from an epiblast-like homogeneous state by a three-step mechanism: (i) a loss of E-cadherin mediated contacts in parts of the aggregate leading to the appearance of islands of E-cadherin expressing cells surrounded by cells devoid of E-cadherin, (ii) a separation of these two populations with islands of E-cadherin expressing cells flowing toward the aggregate tip, and (iii) their differentiation into an endoderm population. During the flow, the islands of E-cadherin expressing cells are surrounded by cells expressing T-Brachyury, reminiscent of the process occurring at the primitive streak. Consistent with recent in vivo observations, the endoderm formation in the gastruloids does not require an epithelial-to-mesenchymal transition, but rather a maintenance of an epithelial state for a subset of cells coupled with fragmentation of E-cadherin contacts in the vicinity, and a sorting process. Our data emphasize the role of signaling and tissue flows in the establishment of the body plan.

Introduction

Gastrulation is a crucial stage of animal development that results in the laying down of the body plan through the specification of the germ layers – the endoderm, the mesoderm and the ectoderm – and their organization with reference to three orthogonal axes (Wolpert et al., 2006; Solnica-Krezel and Sepich, 2012; Williams and Solnica-Krezel, 2017). Defects in gastrulation result in premature death or birth defects (Ferrer-Vaquer and Hadjantonakis, 2013). In most vertebrates, gastrulation entails the reorganization of a large mass of cells that has been generated in a burst of cell divisions and the process can be easily followed and manipulated in the lab as the embryos develop externally. Mammalian embryos are different, in that an increase in the cellular mass is concomitant with gastrulation (Snow, 1977); the fact that they develop in the uterus of the mother makes this coupling and the process of gastrulation difficult to study. For this reason, our understanding of mammalian gastrulation is poor relying on mutant analysis and description of patterns of gene expression (Anderson and Ingham, 2003; Artzt, 2012). Recent live imaging of the process (McDole et al., 2018) has added an important dimension to our understanding of the event, but this is low throughput and involves time consuming analysis. Gastrulation is, essentially, a cellular process that needs to be investigated at the cellular level and it would be useful to complement the *in vivo* studies with alternatives *in vitro*.

The advent of in vitro models of mammalian development provides several advantages (Fu et al., 2021) Firstly, unlike embryos, such systems are amenable to systematic perturbations. The system size and its composition can be conveniently modulated by changing the type and the initial number of cells seeded. Secondly, environmental cues - not limited to morphogen concentration (Heemskerk et al., 2019), such as geometric confinements and properties of extra-cellular matrix – can be finely tuned. Lastly, such systems are amenable to live imaging at single cell resolution to follow cell fates and cell dynamics.

Gastruloids offer the opportunity to enrich our understanding of the agents that may influence gastrulation/embryogenesis – particularly the role of environment, forces and signaling. Upon exposure to external cues and morphogens, gastruloids show an intrinsic capability to polarize, elongate and pattern in a manner that is reminiscent of the early embryo (van den Brink et al., 2014; Turner et al., 2017). Strikingly, 3D gastruloids exhibit a multiaxial organization and spatiotemporal patterning of Hox genes (Beccari et al., 2018) with signatures of somitogenesis (van den Brink et al., 2020; Veenvliet et al., 2020). From *in-vitro* assays, it was shown that combinations of embryonic and extraembryonic stem cells is needed to generate a pro-amniotic cavity (Harrison et al., 2017) and gastrulating embryo-like structures (Sozen et al., 2018). Moreover, by exploiting aggregates, the role of mechanics - for instance a relationship between the physical contact of cells with their confinement and the expression of the mesoderm transcription factor Brachyury (T-Bra) - has been revealed (Sagy et al., 2019) . Besides tridimensional (3D) systems, planar/2D gastruloids confined on micropatterns have emphasized the role of the boundary conditions, shapes (Blin et al., 2018) and scaling parameter on the patterning of the different germ layers (Warmflash et al., 2014; Etoc et al., 2016; Morgani et al., 2018; Xue et al., 2018).

However, how a genetically unpatterned and morphologically symmetric 3D ESCs aggregate transforms into a spatially organized and patterned array of germ layers is largely unknown. In particular, how cellular movements, their rearrangements and patterning coordinate has not been explored. In embryo, the widespread notion is that the endoderm is formed from a single pool of cells that firstly become mesenchymal, then migrate and a proportion of which eventually re-epithelialize (Williams et al., 2012).

This notion has been recently challenged by in vivo observations that show that the definitive endoderm is formed in absence of an epithelial to mesenchymal transition (EMT) (Scheibner et al., 2021).

Here we dissect the formation of an endoderm-like region in 3D gastruloids, focusing on the spatio-temporal deployment and cell state transitions that yield the aforesaid region; specifically, starting from an ensemble of epithelial cells, we follow the cell-state transitions that govern the mesendodermal-like progenitors by looking at expression of the transcription factor Brachyury (T-Bra) and E-cadherin (E-cad). We find that an endoderm-like region is established from a distinct pool of cells, different from the mesoderm. Furthermore, we identify tissue scale flow that localizes the progenitors of the endoderm at a pole. Lastly, we find that a heterogeneity of cellular junction tension could be responsible for segregating the endoderm-like region from the rest of the aggregate via a cell-sorting mechanism.

Results

A pole of E-cadherin expression emerges at the tip of the aggregate

To induce a pluripotent, post-implantation epiblast state of the mouse ESCs (Hayashi et al., 2011), we placed the cells in Activin and FGF throughout the protocol (Figure 1A). A portion of the aggregates were exposed to the Wnt agonist Chiron (Chi) as the combination with Activin and Chi tends to promote endoderm formation (Figure 1A) (Morgani et al., 2018; Engert et al., 2013). We investigated the aggregate morphologies and gene expression patterns following a pulse of Chi (lasting a day from the 2nd day post plating).

Aggregates exposed to Chi lose their spherical morphology and acquire a teardrop shape (Figure 1B, day 4, Video 1). We find that 30/37 or ~ 80% aggregates qualify as elongated with an elongation threshold set to 0.125 (see also Video 2). In contrast 15/20 or 75 % of control aggregates remained spherical (Figure 1b, day 4). At the pointed end of the Chi exposed aggregates, we observed an E-cadherin (E-cad) rich region, which was absent in the control aggregates (Figure 1C). The E-cad pole emerges from an initially homogeneous distribution of E-cad (Figure 1D, Video 3). Likewise, a T-Brachyury (T-Bra) rich pole, clearly visible at the tip of the aggregate at day 4, emerges from a spatially homogeneous distribution at day 3.

Time-lapse imaging shows distinct time evolution of global E-cad and T-Bra signals. While E-cad is expressed from the early time points (Figure 1D, day 2), T-Bra expression follows exposure to Chi (Figure 1D and Video 4). The Chi pulse triggers a decay in E-cad level (mean intensity, Figure 1E and Video 5). T-Bra level, maximal just before removal of Chi, gradually decreases thereafter (Videos 2 and 6).

Most aggregates show a single pole of T-Bra (35/64), some display multiple poles (18/64) and 11/64 do not polarize (Video 2). We then asked whether the polarized expression of E-cad precedes, follows or is concomitant with the deformation of the aggregates. Using descriptors of aggregate's shape and polarization of fluorescence signal (see methods), we found that E-cad and T-Bra polarize prior to the onset of tip formation (Figure 1F, G), suggesting a possible role of E-cad and T-Bra in the aggregate's shape polarization.

E-cadherin expression defines an endoderm-like region in the gastruloid

To resolve the spatial distribution of E-cad and T-Bra expressing cells at the single cell level, aggregates exposed to Chi were stained at day 4. As shown in Figure 2A, a group of E-cad expressing cells is localized at the tip, spatially segregated from, and surrounded by T-Bra expressing cells. Furthermore, we observed that this core expresses Sox17 (Figure 2A, inset, Figure 2 – figure supplement 1) and Foxa2 (Figure 2 - figure supplement 1, Video 7) – markers of the endoderm. We confirmed that Sox2 is present earlier but absent at 5 days-pp while Sox17 is retained (Figure 2 – figure supplement 2). These observations lead us to conclude that the group of E-cad expressing cells demarcates an endoderm-like region in the gastruloids. Figure 2B shows a histogram of the various cell populations at the tip of aggregates. The core of E-cad/Sox17 expressing cells and the T-Bra expressing cells constitute the majority of the cell population near the tip (Figure 2B). Nevertheless, a minority outside the core expresses either Sox17 alone or in conjunction with T-Bra. To obtain some insight into the supracellular tissue organisation, we further determined the local neighbourhood of the different cell populations (Figure 2C). The E-cad/Sox17 expressing cells that form the core maintain a relatively larger simply-connected-component (126 of 131 Ecad+/Sox17+ cells) and a homotypic neighbourhood (111 E-cad+/Sox17+ cells maintain an Ecad+/Sox17+ majority neighbourhood), which implies that they form a tightly packed group of cells. In contrast, the T-Bra expressing cells form multiple simply-connected-components (20 components) suggesting that they are sparsely distributed around the core. The E-cad+/Sox17+ cells thus form a compact region that is flanked by sparse T-Bra + cells and few other cell types that span widely about the tip (Figure 2D and Figure 2 - figure supplement 3 for two other aggregates).

Endoderm is established from an epiblast-like population

To probe how the endoderm-like region is established in gastruloids, we performed time-lapse imaging of the aggregates at single cell resolution. At the early stages E-cad expression pervades the entire aggregate; It is present at every cell-cell contact forming a quite homogeneous network of junctions. However, this network undergoes a gradual fragmentation (Figure 3A and Videos 8-11). The level of E-cad decreases temporally while a small fraction is retained, which eventually localizes near the aggregate's tip. Within our immunostainings, we found consistently the presence of some E-cad/T-Bra cells that express lower levels of E-cad compared to their neighbours devoid of T-Bra (Figure 3 – figure supplement 1, 92 hrs-pp). In view of earlier studies (Turner et al., 2014; Fernando et al., 2010), a gradual repression of E-cad via increasing levels of T-Bra may possibly account for the fall of the E-cad level in the junctional network. Furthermore, in immunostainings at 3 days-pp we notice only a small fraction of cells – distributed in a salt and pepper manner - that express Foxa2 (Figure 3- figure supplement 2). Therefore, at 72 hrs-pp, we do not find an organized group of cells that can potentially yield the endoderm. Additionally, at this stage the cell population in the aggregates can be broadly classified into three categories: a class of cells that express both E-cad and T-Bra and two classes that express either one of the two markers (Figure 3B, left). Additionally, the E-cad+/T-Bra- cells are also found to be the ones expressing Sox2, akin to an epiblast. The E-cad/T-Bra expressing cells constitute most of the cell population (~70 % of the total) and are surrounded by the T-Bra expressing cells (Figure 3B, right). The cells expressing only E-cad – present in the minority - are present mainly within the core of the E-cad/T-Bra expressing cells (Figure 3B, right, inset). A spatial connectivity graph shows that the cells in the core (E-cad+/T-Bra+) are found to be tightly packed – with 70 out of 81 of E-cad+/T-Bra+ cells forming the largest connected component - and maintain a homotypic neighbourhood (72 E-cad+/T-Bra+ cells maintain an E-cad+/T-Bra+ neighbourhood) (Figure 3b,

right). Yet again the cells expressing only T-Bra are sparsely wrapped around the core (13 components for 31 T-Bra+ cells) and therefore, have a relatively heterotypic neighbourhood (12 T-Bra+ cells have an E-cad+/T-Bra+ majority neighbourhood). Spatial distribution of the cell types (mean of n = 7 aggregates) at 72 hrs-pp depicts that the cells expressing only E-cad are mostly present within the E-cad+/T-Bra+ cell islands, whereas the T-Bra+ cells are situated at the periphery of the aggregate (Figure 3C). Interestingly, at later time-points but before the aggregates are fully polarized, we observed an "archipelago" of E-cad expressing cells separated by T-Bra expressing cells (Figure 3D, 92 hrs-pp, Figure 3 – figure supplement 1). At day 4, the E-cad expressing cells exhibit low levels of N-cadherin and Snail1, supporting the absence of an EMT transition for the E-cad expressing cells (Figure 3 – figure supplements 3 and 4). These data are consistent with recent observations in the mouse (Scheibner et al., 2021). Additionally, the islands of E-cad expressing cells were found to be positive for the pluripotency markers Oct4 (Figure 3D). The Oct4 expression correlates inversely with the T-Bra expression, suggesting pluripotency and an epiblast-like nature of these cells (Figure 3E). Altogether our data suggests that the endoderm-like region - identified at the end of day 4 - originates from islands of pluripotent and epithelial cells expressing E-cadherin.

Tissue flows and a local sorting segregate islands of E-cadherin expressing cells

Our observations raise the possibility that the islands of E-cad+ cells sort and form a compact group of cells. To test this possibility, we analyzed cell movements prior to the establishment of the aggregate tip. We imaged the aggregates with inverted two-photon microscopy adding Sulforhodamine B (SRB) to the medium which enables to visualize cell contours and extra-cellular space. We perform optical flow (See method) on the SRB signal (Figure 4A-B, Videos 8-14) and reconstruct the coarse-grained velocity field within the tissue. The flow field clearly shows that there is a directed flow of cells towards the future region that will form the tip of the gastruloid (Figure 4A-B, Figure 4- figure supplement 1, and Videos 15-17). E-cad- cells flow backward at the periphery, away from the tip, which creates a physical separation between the two populations: E-cad+ and E-cad- cells (Videos 8-10). islands of E-cad+ cells are highly dynamic during their movement towards the tip (Video 11). The directed flow and recirculation are not observed in absence of Chi pulse (Figure 4- figure supplement 1).

Despite dynamic cell rearrangements and shear in the bulk of the aggregate, the islands of E-cad expressing cells are not fully fragmented, implying that physical forces maintain their cohesion (Video 11). The E-cad expressing cells in islands have a significantly smaller contact angle with the surrounding cells than with each other (Figure 4c, inset). Assuming that the cell-cell contacts are determined by interfacial tension (Supplementary material), we measured the angle of contacts between E-cad expressing cells within the islands and with neighbouring cells (Figure 4C). They correspond to a significant higher tension at the outer boundary of islands at days 3 and 4 (Figure 4C, Figure 4 – figure supplement 2). Such difference could account for the maintenance of E-cad+ cell islands during their transport as well as their segregation at the tip, where E-cad+ cells would sort from the T-Bra+ cells to form a large endoderm domain.

Discussion

In contrast to classical gastruloids, when aggregates of ESCs are grown in Activin and FGF, they go through a clear phase similar to the postimplantation epiblast. In this state, suppression of Wnt signalling leads to

anterior neural development (Girgin et al., 2021) but, as we show here, exposure to Wnt signalling leads to fates associated with primitive streak, as evidenced by expression of T-Bra and the emergence of an endoderm-like region. We also observe directional and coordinated movements of cell ensembles that mimic some aspects of mammalian gastrulation, in particular the epithelial organization of the gastruloid at the time of Chi application. Under Chi exposure, that mimics the Wnt/β-catenin signaling, the pluripotent epiblast-like cells (with the exception of a few cells) experience a massive wave of T-Bra expression and a loss of E-cad mediated contacts (Figure 4D). The cells near the periphery specifically exhibit high levels of T-Bra and low levels of E-cad expression (Figure 3 B, C). A similar observation has been made on 2D micropatterns in which a T-Bra wave emanates and travels inwards, from the boundary where E-cad is absent (Martyn et al., 2019). Loss of E-cad contacts in parts of the aggregate leads to the appearance of pluripotent E-cad+ cell islands (Sox2 or Oct4 positive), that are enveloped by T-Bra expressing cells (Figure 3, B, D, Figure 3: figure supplement 1, and Figure 4 D). All the while a coordinated motion of the cells within the aggregate arrange the islands of the E-cad/Sox2/Oct4 expressing cells at one pole of the aggregate. This pole protrudes outwards to form the tip. The E-cad islands subsequently segregate at the tip and are found to express markers characteristic of the endoderm (Figure 2, Figure 4D).

While we cannot rule out the possibility that the mechanogenetic trajectory followed by cells in the gastruloids may be different from that in the embryo, our data are consistent with recent observations in the embryo (Scheibner et al., 2021) that challenged the prevailing view that endoderm forms from a mesendodermal primodium. In the mouse (Scheibner et al., 2021) and in the gastruloids (our work and (Vianello and Lutolf, 2021)) the endodermal progenitors do not undergo an Epithelial-Mesenchymal transition (EMT) and emerge from a different pool of cells (E-cad+ /T-Bra-) than the mesodermal-like precursor cells (T-Bra+). Vianello and Lutolf further reported that within 7 days the endoderm primordium in the gastruloids give rise to different expression territories corresponding to anterior foregut, midgut and hindgut (Vianello and Lutolf, 2021), suggesting that what we observe in vitro follows closely the events in vivo. Altogether the in vivo and in vitro data point to a mechanism where the endoderm tissues originate from the epiblast via a trajectory that differs from that of the mesodermal cells.

A surprising feature of endoderm specification in gastruloids (see also (Pour et al., 2019; Vianello and Lutolf, 2021)) is the appearance of endodermal cells as small islands that coalesce into a continuous tissue. These observations are compatible with what is observed in mouse embryos where the emergence of endodermal progenitors in small islands of cells has been overlooked (see Figure 3 in (Probst et al., 2021)). Furthermore, a similar mode of endoderm specification has been reported in zebrafish, where it is the predominant mode (Montero et al., 2005; Nair and Schilling, 2008; Warga and Nusslein-Volhard, 1999). The clear manifestation in the gastruloids might reflect the suggestion that ancestral modes are emphasized in this model (Steventon et al., 2021).

The segregation of endodermal-like cells in the gastruloids is a consequence of a wave of cell state transitions, a global flow and surface tension heterogeneity. Whether the same mechanism exists at the primitive streak in the mouse embryo remains to be explored. The mechanism described here cannot be reduced to a simple cell-sorting mechanism based on differential adhesion. Different cellular mechanisms are at stake: advection by tissue flow, cell sorting by cellular motions, but also cell division, cell death, and cell differentiation that constantly modify the number of constituent cells and their pattern. How these different cell events lead to the eventual tissue pattern will require further study. Our work emphasizes the coupling between biochemical signaling and tissue flows during endoderm formation, and perhaps

more generally during the morphogenesis of germ layers. Furthermore, it illustrates the potential of in vitro models of early embryogenesis to decouple and study these events.

Materials and Methods

1. Cell Culture and Gastruloids

A complete description of the culture conditions and the protocol for making gastruloids is present in (Baillie-Johnson et al., 2015). Nevertheless, a twofold modification was made to the original protocol.

a) Unless explicitly mentioned approximately 50-60 cells were seeded per well to generate smaller gastruloids.
b) The gastruloids were exposed to a 12.5 ng/ml of bFGF (RD system ref: SUN11602) and 25 ng/ml of Activin A (RD system ref: 338-AC-050) throughout its development

* We formed our aggregates in low-adherence condition in 96 well plates (Costar ref: 7007) and microwell plates (400 um wells, SUNBIOSCIENCE ref: Gri3D).

List of cell lines used:
1) E-cad-GFP/Oct4-mCherry (generous gift from the lab of Joshua Brickmann, Centre for Stem Cell Biology, University of Copenhagen, Denmark)
2) T-Bra-GFP/NE-mKate2 (generous gift from the lab of Guillaume Blin, CRG, University of Edinburgh, Scotland)
3) E14Tg2a.4 mouse Embryonic Stem Cells (mESC) (MMRRC, University of California Davis, US).

All cell lines were tested free from mycoplasma contamination by qPCR.

2. Immunostaining

Immunostaining was performed by fixing the aggregates (4% PFA in PBS with 1 mM $CaCl_2$ and 0.49 mM $MgCl_2$ overnight at 4°C), followed by washing and blocking for 1 hour in PBSFT buffer (PBS with 10% FBS, 0.2% Triton X-100). Aggregates were incubated with primary antibodies, washed, and incubated again with secondary antibodies. The incubations were done overnight at 4°C.

Mentioned below is a list of primary antibodies used in this study:

Antibodies	Species	Reference	Provider	Dilution
Brachyury	Goat	AF2085	RD system	1:80
E-cadherin	Rat	M108	Takara	1:200
anti-GFP	Chicken	GFP-1020	Aves	1:1000
Sox17	Goat	AF1924	RD system	1:80

FoxA2	Goat	sc-6554	Santa-Cruz	1:50
CDX2	Rabbit	MA514494	Thermo	1:200
Gata6	Goat	AF1700	RD system	1:100
Sox2	Mouse	Sc-365823	Santa Cruz	1:100
Snail	Goat	AF3639	RD system	1:50

* We used Alexa dyes (ThermoFisher) against the primary antibodies

3. Live and fixed imaging

Epifluorescence time-lapse imaging of gastruloids was achieved in a chamber - maintained at 37°C, 5% CO_2 with a humidifier - using an AxioObserver inverted microscope (Carl Zeiss) with a 20x LD Plan-Neofluar 0.4 NA Ph2 objective. A white LED source (Laser2000, Kettering UK) was used for illumination. A second setup with similar specifications was also utilized for long-term imaging.

Fixed gastruloids were loaded on MatTek dishes (MatTek corporation, ref: P35G-1.5-14-C) and imaged using a confocal microscope (Zeiss LSM 880) with a 40x 1.2 NA W Korr objective.

Biphoton imaging of gastruloids was achieved in a chamber- maintained at 37°C, 5% CO2 with a humidifier – using a Zeiss 510 NLO (Inverse - LSM) with a femtosecond laser (Laser Mai Tai DeepSee HP) (900 nm) with a 40x/1.2 C Apochromat objective. 72h old Gastruloids were transferred from 96 well plates to either MatTek dishes (MatTek corporation, ref: P35G-1.5-14-C) or to microwell plates (400 um wells, SUNBIOSCIENCE ref: Gri3D). To visualize inter-cellular space, we added 32 µL/mL of a Sulforhodamin B solution (1 mg of Sulforhodamin B powder, 230162 Aldrich per mL of medium) to the medium. We imaged sub-volumes of gastruloids (10 slices z-stacks with a z-interval of 2µm) near their midplane during a total time between 10h and 24h (time interval between two frames between 2 and 4 minutes depending on the experiment). We imaged the GFP (T-Bra-GFP/NE-mKate2 and E-cadherin-GFP/Oct4-mCherry cell lines) and the Sulforhodamin B channels using a non-descanned detector.

4. Cryosection

Aggregates were embedded in Optimal Cutting Temperature (OCT) after fixing in 4% Paraformaldehyde (PFA) and then flash frozen in liquid nitrogen. Slices of 25 and 30 µm were manually cut from the resin with a cryostat (Leica CM3050 S). The slices were mounted on Superfrost® Plus slides before proceeding with immunostaining.

5. Relative surface tension

Under the assumption that the final shape and configuration of the cells are largely dominated by surface tension, we can compute the relative tension between the E-cadherin/non-E-cadherin (γ_{+-}) and the E-cadherin/E-cadherin expressing cells (γ_{++}) as the reciprocal of the sum of the cosines of the two angles subtended at the interface of the E-cadherin/non-E-cadherin cells (Figure 4C).

$$\left(\frac{\gamma_{+-}}{\gamma_{++}}\right) = \frac{1}{cos\ \theta_1\ + cos\ \theta_2}$$

6. Image Processing and Data Analysis

triangulation/Probability Map

The gastruloids were initially manually segmented using brightfield images or a combination of brightfield and the fluorescence signal. An artificial convolutional neural network, UNET (Ronneberger et al., 2015) was also employed to yield a high throughput pipeline for generating binary masks of aggregates with brightfield images. The UNET was trained on an augmented dataset (n = 390 and n = 264 images for aggregates generated from Ecad/Oct4 line and T-Bra/NE-mKate2 lines respectively) and then applied onto the time series aggregate images to obtain binarized masks. Any imperfections in the generated masks were manually corrected in FIJI/ImageJ. The mean signal intensity of the aggregates was computed by averaging the pixel intensities within the binary masks.

The overall elongation of the aggregates at each time point was quantified by determining the ratio of the eigenvalue pairs of the inertia tensor of the binary mask. This is equivalent to finding the ratio of the major to the minor axes of the best-fit ellipse to the mask. Furthermore, to precisely capture the onset of the tip formation we relied on Fourier-Elliptic Descriptors[3](FED) as a shape descriptor (Kuhl and Giardina, 1982). The aggregate perimeter was defined by a set of n-harmonics/ellipses (n = 1,2,3 …) – each associated with four coefficients - such that the difference of area between the aggregate mask and the morphology obtained from the harmonics was less than 2 %. The onset of tip formation was determined using the 2nd FED harmonic at each time point. The corresponding coefficients were transformed to the lengths of the major and minor axes of the ellipse. The magnitude of the second harmonic was then taken to be the product of the lengths of the two axes, which is a proxy for the extent of tip formation and shape polarization.

The polarisation of the fluorescence was defined as the maximum of contrast between the two-halves of the aggregate. It was calculated in two-steps: first the aggregate mask was chopped into two halves about the eigenvector of the inertia matrix, which passes through the centroid of the mask. The contrast was then calculated as (I_max-I_min)/(I_max+I_min) where I_max and I_min represent the intensities of the brighter and less bright halves, respectively. The contrast maximum was then determined by angular scanning by repeatedly rotating the cutting plane by 5°.

The different cell-types in the confocal image-stacks were manually identified and tagged via *Cell Counter* – a built-in FIJI plugin. Corrections were made to ensure that each cell was designated a unique identity. The positions of the marked cells were extracted for further processing. The centroids of the cell were extracted to create a neighbourhood graph and deduce spatial statistics. The cell populations at the aggregate tip (at 96 hrs-pp) are spatially binned to generate probability maps of the different cell types along the tip. Owing to the sphericity of the aggregates at 72 hrs-pp, cell positions are normalized to lie in a unit sphere. Aggregates are registered via their common origin and all the cell populations are pooled to create a spatial probability density map (using SmoothKernelDistribution function in Mathematica). A numerical integration of the density function yields the probability of finding the various cell types as a function of the distance from the origin.

To build the graph we generated a Delaunay triangulation (using the built-in DelaunayMesh in Mathematica) from the cell centroids and pruned any edges that were more than 12um in length (avg. size of a cell). The neighbourhood type of a cell is assigned by counting the cell types of all the adjoining cells and taking the most dominant cell-type as the neighbourhood type; for instance, if majority of the neighbouring cells, around an E-cad + / T-Bra + cell, are also E-cad + / T-Bra + then the cell has a homotypic neighbourhood, else the neighbourhood is designated as being heterotypic.

To measure velocity fields corresponding to Gastruloids polarization and elongation, we use the signal obtained from the Sulforhodamin B (SRB) channel in two-photon timelapses (which is present in all the parts of the Gastruloid in contrary to signals such as T-Bra GFP or E-Cadherin GFP). The SRB signal corresponds to (i) cell contours, (2) dead cells and (3) bright spots of subcellular size moving rapidly corresponding probably to vesicles are visible. As dead cells and bright spots are much brighter than cell contours, we filter them with a thresholding operation, and perform the optic flow on the filtered image where only cell contours are visible after timelapse registration to remove rapid rotations and translations of the aggregate. We use a custom-made Matlab optic flow code based on the Kanade Lucas Tomasi (KLT) algorithm with a level 2 pyramid and a window size corresponding to a square of 2 cells by 2 cells (Lucas and Kanade, 1981).

Images in the timeseries data were registered using *MultiStackReg* - an ImageJ plugin - and custom Matlab scripts. The scripts for analyzing data, image processing, neural networks and computer vision techniques were implemented in Wolfram Mathematica 12 and Matlab. Analysis scripts are provided in the supplemental information of this paper.

Figure Legends

Figure 1. (A) Schematic of the experimental protocol; mESCs seeded onto a low-adherence substrate form aggregates that are exposed to N2B27 medium, FGF and Act throughout the experiment. In addition, a pulse of Chi is provided for 24 hrs - starting at 48 hrs-post plating. (B) The box-whisker plot shows the distribution of elongation for Chi + (n = 37) and the control (Chi -) aggregates (n = 20). The insets delineate the contours of the two classes of aggregates. (C) Box-whisker plot showing the distribution of E-cad polarization in the Chi + (n = 37) and Chi - aggregates (n = 20). The polarisation of the signal is defined by the maximum contrast between the two halves of the aggregate. The insets display E-cad intensity within the aggregates. (D) Snapshots of aggregates (first column composed of E-cad-GFP/Oct4-mCherry and the second column composed of T-Bra-GFP/NE-mKate2 and E14Tg2a.4) exposed to Chi pulse at different developmental stages (scale bars for the columns represent 100 µm). The T-Bra-GFP/NE-mKate2 aggregate was generated from ~ 200 cells. (E) Temporal changes in the mean E-cad intensity profile between 48 and 92 hrs for aggregates (Ecad-GFP/Oct4-mCherry) exposed to Chi (n=7 aggregates) and the control (n=7 aggregates). (F) The plot illustrates the polarization of E-cad (n = 14 aggregates) with respect to the deformation of the aggregates between 72 and 92 hrs; the thick and the dashed lines represent the mean and the mean ± SD respectively. (G) T-Bra polarization (n = 11 aggregates) with respect to the aggregate deformation between 72 and 96 hrs; the thick and the dashed lines represent the mean and the mean ± SD respectively.

Figure 2. (A) aggregate stained at 96 hrs-pp for E-cad and T-Bra; the inset shows the endoderm-like cells at the tip expressing both E-cad and Sox17, surrounded by T-Bra + cells. (B) the bar chart delineates the population fraction of the different cell populations present near the tip (n = 3 aggregates) with the error bars indicating the SD. (C) Spatial connectivity graph (Delaunay Mesh) of the different cell populations in an aggregate fixed at 96 hrs-pp (D) Spatial distribution of the different cell types near the tip of a representative aggregate (stained at 96 hrs-pp), its projection and the respective population fractions as a function of the distance from the tip of the aggregate.

Figure 2 - figure supplement 1. The islands of E-cad+ cells express Foxa2 and Sox17 (immunostaining performed at 96 hrs-pp).

Figure 2 – figure supplement 2. At 120 hrs-pp, E-cad + cells are found to still express high levels of Sox17. Sox2 being an early marker of the endoderm is absent at this stage.

Figure 2 – figure supplement 3. Spatial distribution of the different cell types near the tip of two representative aggregates at 96 hrs-pp

Figure 3. (A) Time lapse images showing a gradual fragmentation and loss of E-cad network at 73, 80, 85 and 92 hrs-pp; scale bar represents 100 µm (B) (left) The three different cell populations in the aggregate are marked at 72 hrs-pp: namely cells that express only T-Bra, cells that express both T-Bra and E-cad, and cells that express only E-cad. (right) the fraction of the total population of cells for the three classes are provided in the histogram (n = 7 aggregates) with the SD shown by the error bars; inset shows the spatial distribution of the cells in one such aggregate (right) Spatial connectivity graph is generated for the aggregate at 72 hrs-pp. (C) Spatial distribution of the population fraction of the different cell types as a function of the radial distance from the aggregate centre. (D) Islands of E-cad/Oct4 expressing cells surrounded by T-Bra + cells are present within the aggregate fixed at 92 hrs-

pp; the inset zooms on a few such islands. (E) Pixel intensity correlation plot between Oct4 and T-Bra (n = 6 aggregates).

Figure 3 - figure supplement 1. (upper panel) white arrows mark the junctions between cells that are expressing both E-cad and T-Bra. Notice that these cells do not express Sox2 and express E-cad only weakly compared to the cells that express E-cad/Sox2. (lower panel) Islands of E-cad expressing cells also express Sox2 before the aggregates are fully polarized and an endoderm is established.

Figure 3 - figure supplement 2. At 72 hrs-pp, few Foxa2 positive cells are found within the tissue which expresses in large majority T-Bra.

Figure 3 - figure supplement 3. Staining for E-cadherin and N-cadherin in aggregates at 96 hrs-pp.

Figure 3 - figure supplement 4. Immunostaining of E-cadherin and Snail in Gastruloids at 66h, 72h and 90h.

Figure 4. (A-B) Cellular movements are quantified during the aggregate elongation using two-photon imaging of the aggregates with Sulforhodamine B (SRB) added in the medium, which enables to visualize cell contours. Optic flow measurements are performed on the SRB signal and the obtained velocity fields overlapped on the GFP signal corresponding to (A) E-cad and (B) T-Bra. (C) Box-whisker plot shows the distribution of the ratio of the cell-junction tension heterogeneity between E-cad expressing cells (+), and between E-cad expressing cells and their neighbours (-) at 72 hrs-pp (n = 28 junctions) and 96 hrs-pp(n = 54 junctions); 25th and 75th quartiles together with the minimum and maximum values are shown; the insets to the left depict regions on the E-cad+ cell islands where the angle measurements were made (D) Cartoon depicting the different steps leading to the formation of an endoderm-like region within the gastruloid. Upon addition of Chi, majority of the E-cad expressing cells begin to express T-Bra . The transient expression of T-Bra represses E-cad expression in T-Bra +/E-cad +. Formation of islands of E-cad expressing cells surrounded by T-Bra+ cells; these islands are transported to one end of the aggregate via a tissue-scale movement. Islands of E-cad expressing cells segregate to give rise to the endoderm.

Figure 4 - figure supplement 1. Cell collective movements are quantified in both aggregates cultured in normal conditions (with a chiron pulse between 48h-72h pp) and in non-pulsed aggregates. Mosaic aggregates are used (100 cells wild-type and 20 cells with T-Bra::GFP and nuclear membrane mKate) which enables to perform optic flow on the sparse nuclear membrane signal obtained with epifluorescence microscopy. Here the velocity fields (averaged over two hours) are represented on the T-Bra::GFP signal. Collective anisotropic flows associated with a re-circulation of cells from the tip to the back are only observed in the chiron pulse case.

Figure 4 - figure supplement 2. Examples of E-cad islands used to measure relative tensions between junctions inside an E-cad island and the non-E-cad cells

Videos legends

Video 1.

Gastruloid polarisation between 72h and 96h imaged with bright-field microscopy.

Video 2.

Gastruloids polarisation in a 96 wells plate between 72 h and 100 h after formation (mosaic aggregates made of 100 cells E14 + 20 cells T-Bra::GFP) imaged with epifluorescence. In magenta, aggregates were pulsed with chiron between 48 hours and 72 hours and have a T-Bra expression and polarization. In blue, control Gastruloids that were not pulsed and do not exhibit T-Bra expression.

Video 3.

E-cadherin::GFP signal evolution during Gastruloid polarization between 72h and 96h imaged with epifluorescence.

Video 4.

Brachyury::GFP signal apparition during the chiron pulse from 48h to 72h imaged with epifluorescence.

Video 5.

E-cadherin::GFP progressive decay during the chiron pulse from 48h to 72h imaged with epifluorescence.

Video 6.

Brachyury::GFP signal evolution during Gastruloid polarization between 72h and 96h imaged with epifluorescence.

Video 7.

Z-stack (one slice every 10 microns) of a 96 hours Gastruloid (made of 150 T-Bra::GFP cells) immunostained with FoxA2 and E-cadherin.

Video 8-10.

Two-photon timelapses of Gastruloid (E-cadherin::GFP) elongation and polarisation with E-cadherin GFP and sulforhodamine B tagging.

Video 11.

Two-photon rapid timelapse of a Gastruloid early phase of symmetry breaking (E-cadherin::GFP).

Video 12-14.

Two-photon timelapses of Gastruloid (T-Bra::GFP) elongation and polarisation with T-Bra and sulforhodamine B tagging.

Video 15. Velocity field obtained by optical flow on the sulforhodamine B signal of the movie 12 superimposed on the T-Bra signal.

Video 16. Velocity field obtained by optical flow on the sulforhodamine B signal of the movie 8 superimposed on the E-cadherin signal.

Video 17. Velocity field obtained by optical flow on the sulforhodamine B signal of the movie 12 superimposed on the sulforhodamine B signal.

Data and code availability

All the source data and scripts can be accessed at https://zenodo.org/record/5727050

Supplemental Information

Supplementary figures, movies and table (Supplementary File 1), giving the number of aggregates analyzed and number of independent experiments

Acknowledgements

The work is supported by the Leverhulme Trust (visiting professorship to PFL VP2-2015-022 and RPG-2018-356) and Agence Nationale de la Recherche ("Investissements d'Avenir", Labex INFORM Phd Grant to AH, ANR-11-LABX-0054 and ANR-16-CONV-0001 from Excellence Initiative of Aix-Marseille University - A*MIDEX and generic grant to PFL, ANR-19-CE13-0022). We also acknowledge the France-Bioimaging Infrastructure (ANR-10-INBS-04). We thank Simone Probst and Sebastian Arnold for their advice. We are grateful to Rosanna Dono for her suggestions and help, and Tina Balayo for assistance with tissue culture. We thank all members of the Lenne and Martinez Arias groups. We thank Guillaume Blin (CRG, University of Edinburgh) and Joshua Brickman (Centre for Stem Cell Biology, University of Copenhagen, Denmark) for sharing their cell lines.

Declaration of interests: the authors declare no competing interests.

References

Anderson KV, Ingham PW. 2003. The transformation of the model organism: a decade of developmental genetics. *Nat Genet* **33**:285–293. doi:10.1038/ng1105

Artzt K. 2012. Mammalian Developmental Genetics in the Twentieth Century. *Genetics* **192**:1151–1163. doi:10.1534/genetics.112.146191

Baillie-Johnson P, van den Brink SC, Balayo T, Turner DA, Martinez Arias A. 2015. Generation of Aggregates of Mouse Embryonic Stem Cells that Show Symmetry Breaking, Polarization and Emergent Collective Behaviour In Vitro. *J Vis Exp* e53252. doi:10.3791/53252

Beccari L, Moris N, Girgin M, Turner DA, Baillie-Johnson P, Cossy A-C, Lutolf MP, Duboule D, Arias AM. 2018. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**:272–276. doi:10.1038/s41586-018-0578-0

Blin G, Wisniewski D, Picart C, Théry M, Puceat M, Lowell S. 2018. Geometrical confinement controls the asymmetric patterning of brachyury in cultures of pluripotent cells. *Development (Cambridge, England)* **145**:dev166025. doi:10.1242/dev.166025

Engert S, Burtscher I, Liao WP, Dulev S, Schotta G, Lickert H. 2013. Wnt/β-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. *Development* **140**:3128–3138. doi:10.1242/dev.088765

Etoc F, Metzger J, Ruzo A, Kirst C, Yoney A, Ozair MZ, Brivanlou AH, Siggia ED. 2016. A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Dev Cell* **39**:302–315. doi:10.1016/j.devcel.2016.09.016

Fernando RI, Litzinger M, Trono P, Hamilton DH, Schlom J, Palena C. 2010. The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *J Clin Invest* **120**:533–544. doi:10.1172/JCI38379

Ferrer-Vaquer A, Hadjantonakis A-K. 2013. Birth Defects Associated with Perturbations in Pre-implantation, Gastrulation & Axis Extension: from Conjoined Twinning to Caudal Dysgenesis. *Wiley Interdiscip Rev Dev Biol* **2**:427–442. doi:10.1002/wdev.97

Fu J, Warmflash A, Lutolf MP. 2021. Stem-cell-based embryo models for fundamental research and translation. *Nat Mater* **20**:132–144. doi:10.1038/s41563-020-00829-9

Girgin MU, Broguiere N, Mattolini L, Lutolf MP. 2021. Gastruloids generated without exogenous Wnt activation develop anterior neural tissues. *Stem Cell Reports* **16**:1143–1155. doi:10.1016/j.stemcr.2021.03.017

Harrison SE, Sozen B, Christodoulou N, Kyprianou C, Zernicka-Goetz M. 2017. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science* **356**:eaal1810. doi:10.1126/science.aal1810

Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. 2011. Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. *Cell* **146**:519–532.

doi:10.1016/j.cell.2011.06.052

Heemskerk I, Burt K, Miller M, Chhabra S, Guerra MC, Liu L, Warmflash A. 2019. Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells. *eLife* **8**:e40526. doi:10.7554/eLife.40526

Kuhl F, Giardina C. 1982. Elliptic Fourier features of a closed contour. *Comput Graph Image Process*. doi:10.1016/0146-664X(82)90034-X

Lucas BD, Kanade T. 1981. An iterative image registration technique with an application to stereo visionProceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'81. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. pp. 674–679.

Martyn I, Brivanlou AH, Siggia ED. 2019. A wave of WNT signaling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. *Development* **146**. doi:10.1242/dev.172791

McDole K, Guignard L, Amat F, Berger A, Malandain G, Royer LA, Turaga SC, Branson K, Keller PJ. 2018. In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell* **175**:859-876.e33. doi:10.1016/j.cell.2018.09.031

Montero J-A, Carvalho L, Wilsch-Bräuninger M, Kilian B, Mustafa C, Heisenberg C-P. 2005. Shield formation at the onset of zebrafish gastrulation. *Development* **132**:1187–1198. doi:10.1242/dev.01667

Morgani SM, Metzger JJ, Nichols J, Siggia ED, Hadjantonakis A-K. 2018. Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. *Elife* **7**:1040. doi:10.7554/eLife.32839

Nair S, Schilling TF. 2008. Chemokine Signaling Controls Endodermal Migration During Zebrafish Gastrulation. *Science* **322**:89–92. doi:10.1126/science.1160038

Pour M, Kumar AS, Walther M, Wittler L, Meissner A, Nachman I. 2019. Emergence and patterning dynamics of mouse definitive endoderm. doi:10.1101/728642

Probst S, Sagar, Tosic J, Schwan C, Grün D, Arnold SJ. 2021. Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation. *Development* **148**. doi:10.1242/dev.193789

Ronneberger O, Fischer P, Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science. Cham: Springer International Publishing. pp. 234–241. doi:10.1007/978-3-319-24574-4_28

Sagy N, Slovin S, Allalouf M, Pour M, Savyon G, Boxman J, Nachman I. 2019. Prediction and control of symmetry breaking in embryoid bodies by environment and signal integration. *Development* **146**. doi:10.1242/dev.181917

Scheibner K, Schirge S, Burtscher I, Büttner M, Sterr M, Yang D, Böttcher A, Ansarullah null, Irmler M, Beckers J, Cernilogar FM, Schotta G, Theis FJ, Lickert H. 2021. Epithelial cell plasticity drives endoderm formation during gastrulation. *Nat Cell Biol* **23**:692–703. doi:10.1038/s41556-021-00694-x

Snow MHL. 1977. Gastrulation in the mouse: Growth and regionalization of the epiblast. *Development*

42:293–303. doi:10.1242/dev.42.1.293

Solnica-Krezel L, Sepich DS. 2012. Gastrulation: making and shaping germ layers. *Annu Rev Cell Dev Biol* **28**:687–717. doi:10.1146/annurev-cellbio-092910-154043

Sozen B, Amadei G, Cox A, Wang R, Na E, Czukiewska S, Chappell L, Voet T, Michel G, Jing N, Glover DM, Zernicka-Goetz M. 2018. Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat Cell Biol* **20**:979–989. doi:10.1038/s41556-018-0147-7

Steventon B, Busby L, Arias AM. 2021. Establishment of the vertebrate body plan: Rethinking gastrulation through stem cell models of early embryogenesis. *Developmental Cell* **56**:2405–2418. doi:10.1016/j.devcel.2021.08.012

Turner DA, Girgin M, Alonso-Crisostomo L, Trivedi V, Baillie-Johnson P, Glodowski CR, Hayward PC, Collignon J, Gustavsen C, Serup P, Steventon B, P Lutolf M, Arias AM. 2017. Anteroposterior polarity and elongation in the absence of extra-embryonic tissues and of spatially localised signalling in gastruloids: mammalian embryonic organoids. *Development (Cambridge, England)* **144**:3894–3906. doi:10.1242/dev.150391

Turner DA, Rué P, Mackenzie JP, Davies E, Martinez Arias A. 2014. Brachyury cooperates with Wnt/β-catenin signalling to elicit primitive-streak-like behaviour in differentiating mouse embryonic stem cells. *BMC Biol* **12**:63. doi:10.1186/s12915-014-0063-7

van den Brink SC, Alemany A, van Batenburg V, Moris N, Blotenburg M, Vivié J, Baillie-Johnson P, Nichols J, Sonnen KF, Martinez Arias A, van Oudenaarden A. 2020. Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**:405–409. doi:10.1038/s41586-020-2024-3

van den Brink SC, Baillie-Johnson P, Balayo T, Hadjantonakis A-K, Nowotschin S, Turner DA, Martinez Arias A. 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development (Cambridge, England)* **141**:4231–4242. doi:10.1242/dev.113001

Veenvliet JV, Bolondi A, Kretzmer H, Haut L, Scholze-Wittler M, Schifferl D, Koch F, Guignard L, Kumar AS, Pustet M, Heimann S, Buschow R, Wittler L, Timmermann B, Meissner A, Herrmann BG. 2020. Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* **370**:eaba4937. doi:10.1126/science.aba4937

Vianello S, Lutolf MP. 2021. In vitro endoderm emergence and self-organisation in the absence of extraembryonic tissues and embryonic architecture. doi:10.1101/2020.06.07.138883

Warga RM, Nusslein-Volhard C. 1999. Origin and development of the zebrafish endoderm. *Development* **126**:827–838. doi:10.1242/dev.126.4.827

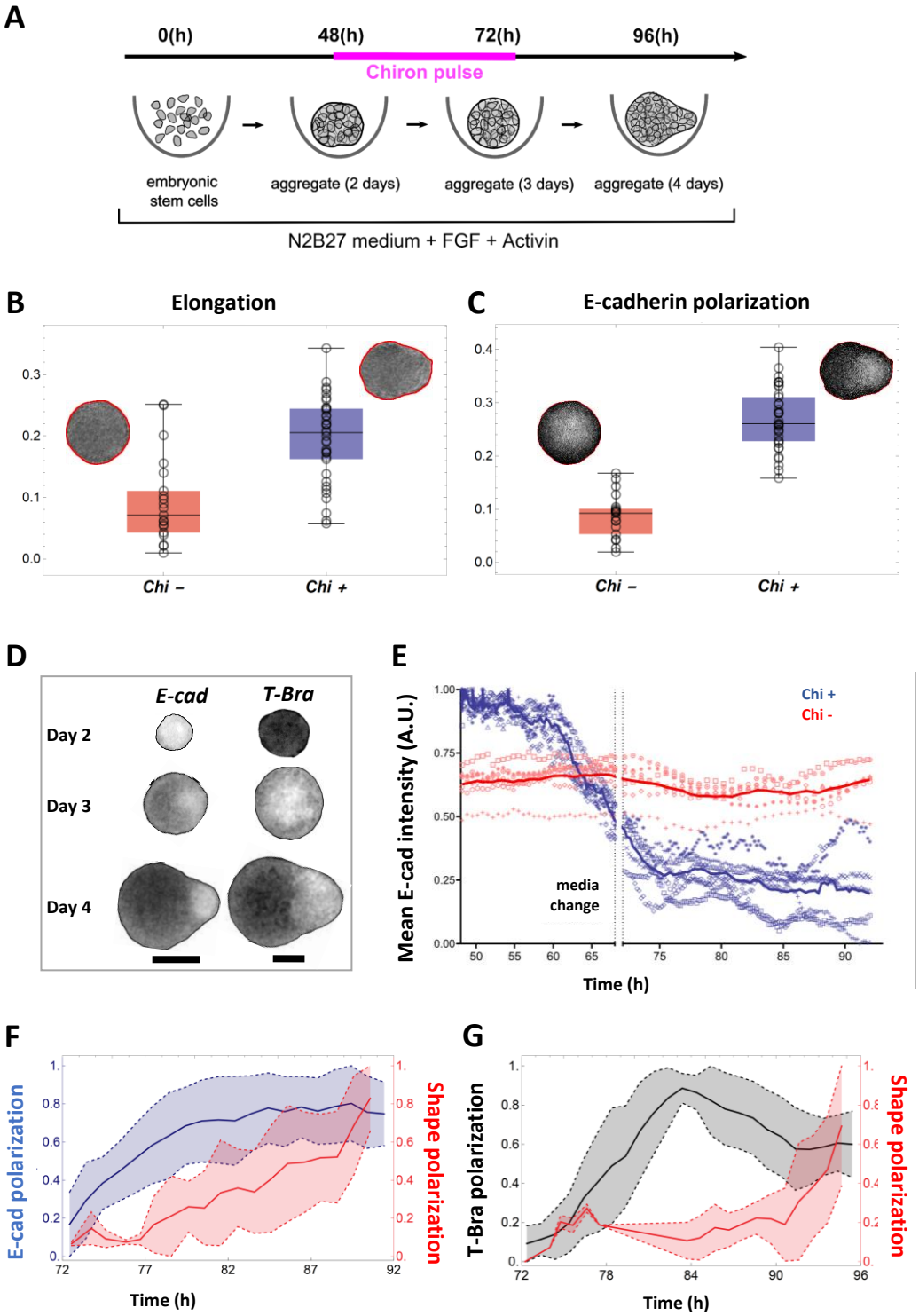
Warmflash A, Sorre B, Etoc F, Siggia ED, Brivanlou AH. 2014. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature methods* **11**:847–854. doi:10.1038/nmeth.3016

Williams M, Burdsal C, Periasamy A, Lewandoski M, Sutherland A. 2012. Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Developmental Dynamics* **241**:270–283. doi:10.1002/dvdy.23711

Williams ML, Solnica-Krezel L. 2017. Regulation of gastrulation movements by emergent cell and tissue interactions. *Current Opinion in Cell Biology*, Cell Dynamics **48**:33–39. doi:10.1016/j.ceb.2017.04.006

Wolpert L, Smith J, Jessell T, Lawrence P, Robertson E, Meyerowitz E. 2006. Principles of Development, 3rd ed. Oxford University Press, USA.

Xue X, Sun Y, Resto-Irizarry AM, Yuan Y, Aw Yong KM, Zheng Y, Weng S, Shao Y, Chai Y, Studer L, Fu J. 2018. Mechanics-guided embryonic patterning of neuroectoderm tissue from human pluripotent stem cells. *Nature Mater* **17**:633–641. doi:10.1038/s41563-018-0082-9


A
0(h)
48(h)
72(h)
96(h)
Chiron pulse
embryonic stem cells
aggregate (2 days)
aggregate (3 days)
aggregate (4 days)
N2B27 medium + FGF + Activin
B
Elongation
Chi –
Chi +
C
E-cadherin polarization
Chi –
Chi +
D
E-cad
T-Bra
Day 2
Day 3
Day 4
E
Mean E-cad intensity (A.U.)
Chi +
Chi -
media change
Time (h)
F
E-cad polarization
Shape polarization
Time (h)
G
T-Bra polarization
Shape polarization
Time (h)

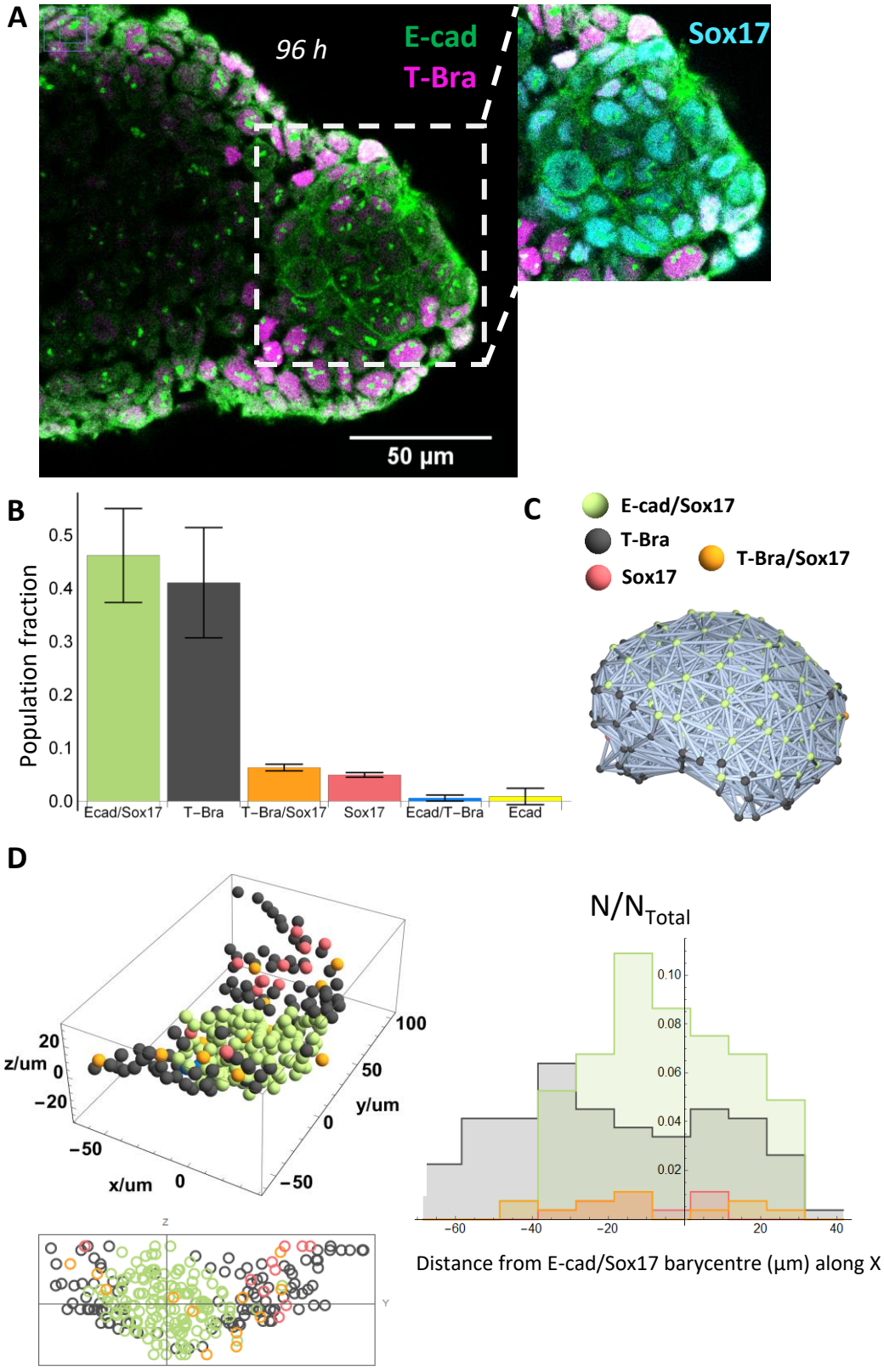

A
96 h
E-cad
T-Bra
Sox17
50 µm
B
Population fraction
0.5
0.4
0.3
0.2
0.1
0.0
Ecad/Sox17
T−Bra
T−Bra/Sox17
Sox17
Ecad/T−Bra
Ecad
C
E-cad/Sox17
T-Bra
Sox17
T-Bra/Sox17
D
z/um
20
0
−20
x/um
−50
0
100
50
y/um
0
−50
N/N_Total
0.10
0.08
0.06
0.04
0.02
−60
−40
−20
20
40
Distance from E-cad/Sox17 barycentre (µm) along X

Figure 2-figure supplement 1

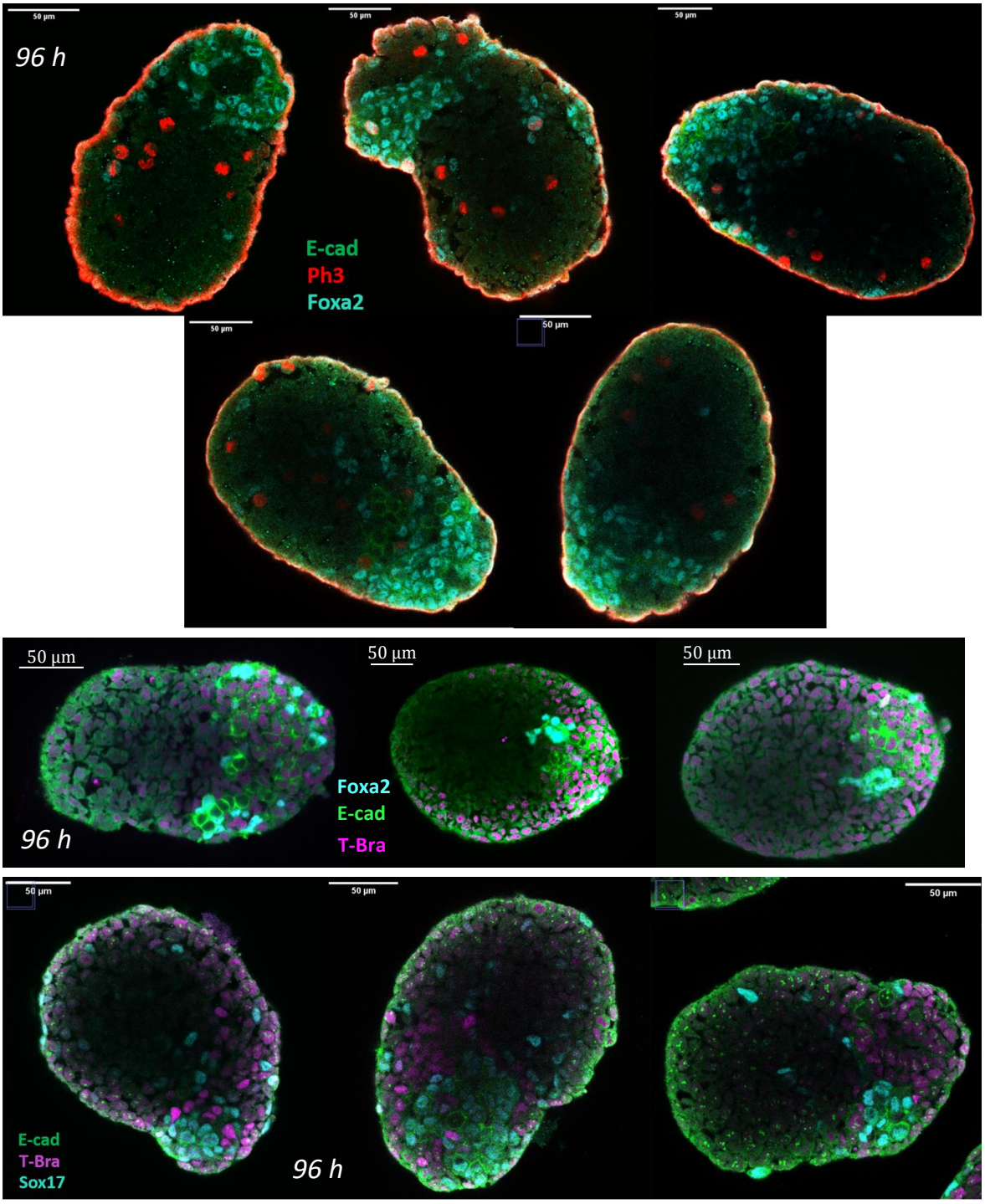
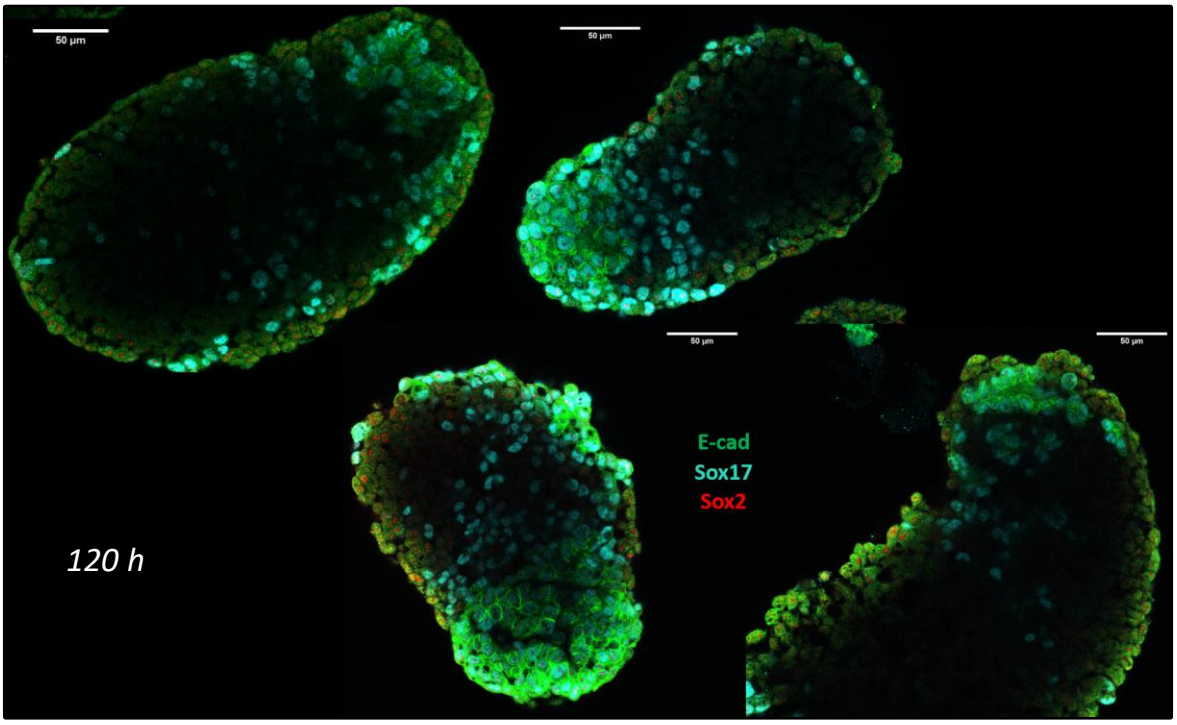
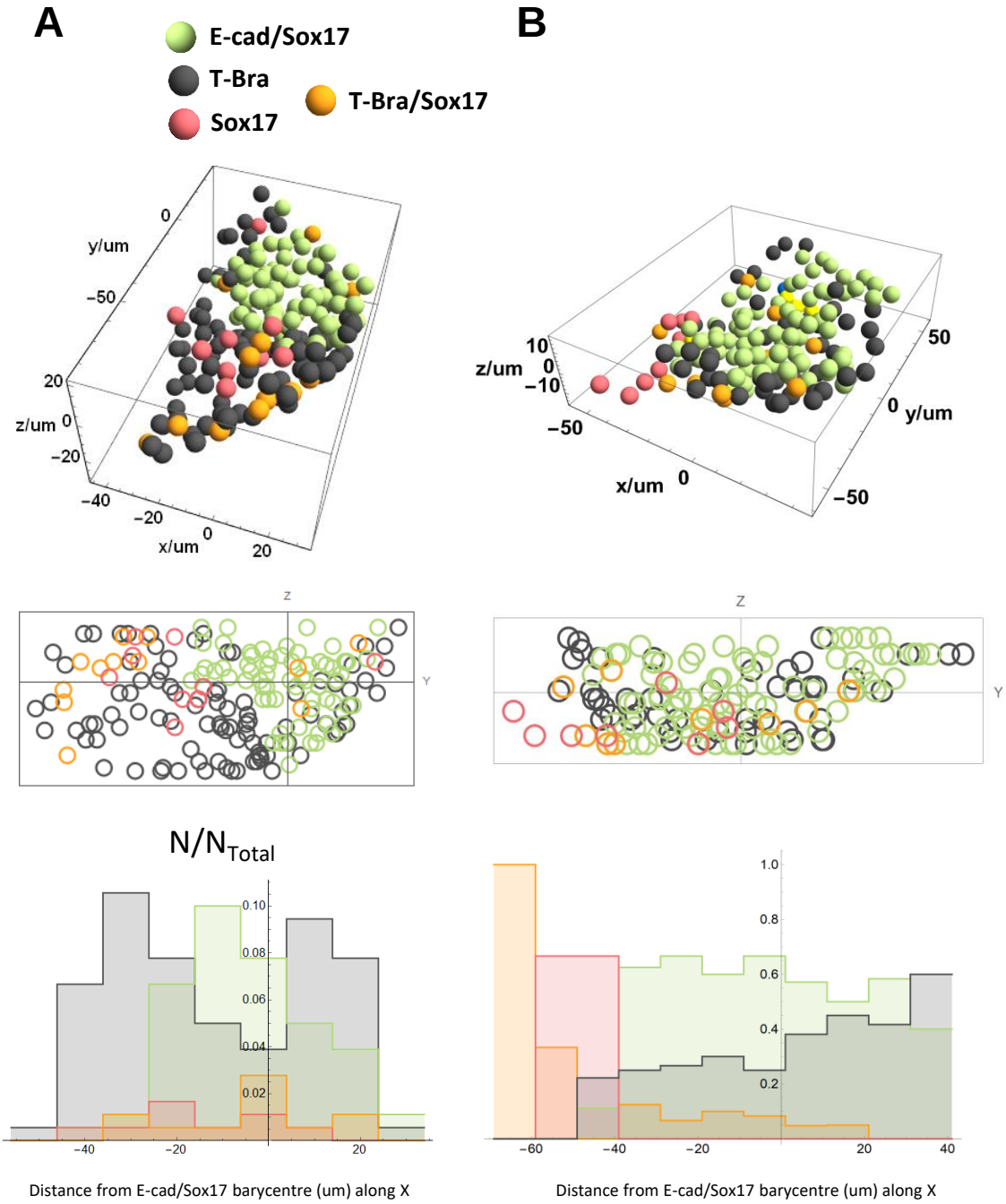


Figure 2-figure supplement 2




A
E-cad/Sox17
T-Bra
Sox17
T-Bra/Sox17
B
y/um
z/um
x/um
z/um
y/um
x/um
N/N_Total
Distance from E-cad/Sox17 barycentre (um) along X
Distance from E-cad/Sox17 barycentre (um) along X

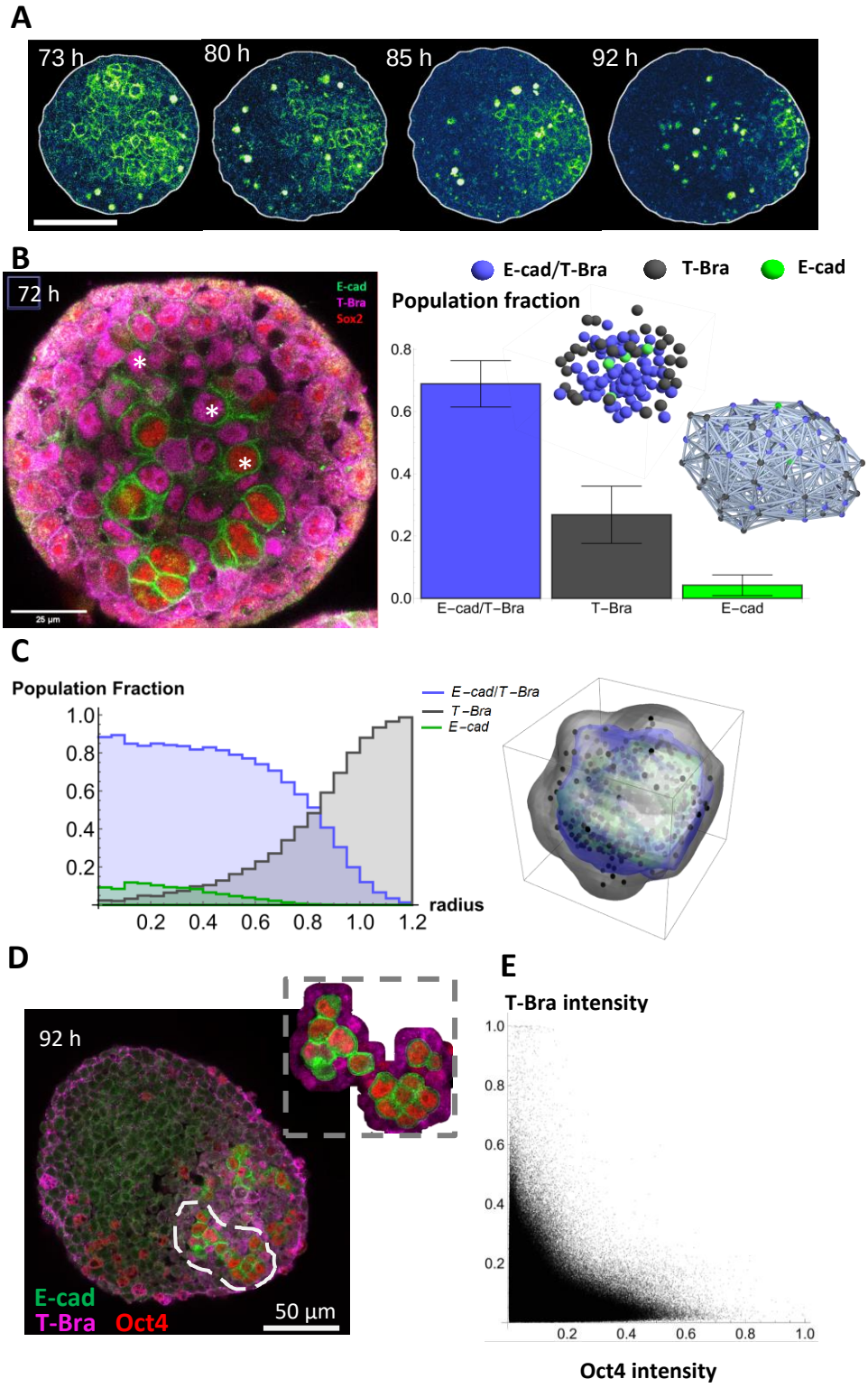

A
73 h
80 h
85 h
92 h
B
72 h
E-cad
T-Bra
Sox2
25 µm
E-cad/T-Bra
T-Bra
E-cad
Population fraction
0.8
0.6
0.4
0.2
0.0
E−cad/T−Bra
T−Bra
E−cad
C
Population Fraction
1.0
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1.0
1.2
radius
E−cad/T−Bra
T−Bra
E−cad
D
92 h
E-cad
T-Bra
Oct4
50 µm
E
T-Bra intensity
1.0
0.8
0.6
0.4
0.2
0.2
0.4
0.6
0.8
1.0
Oct4 intensity

Figure 3-figure supplement 1

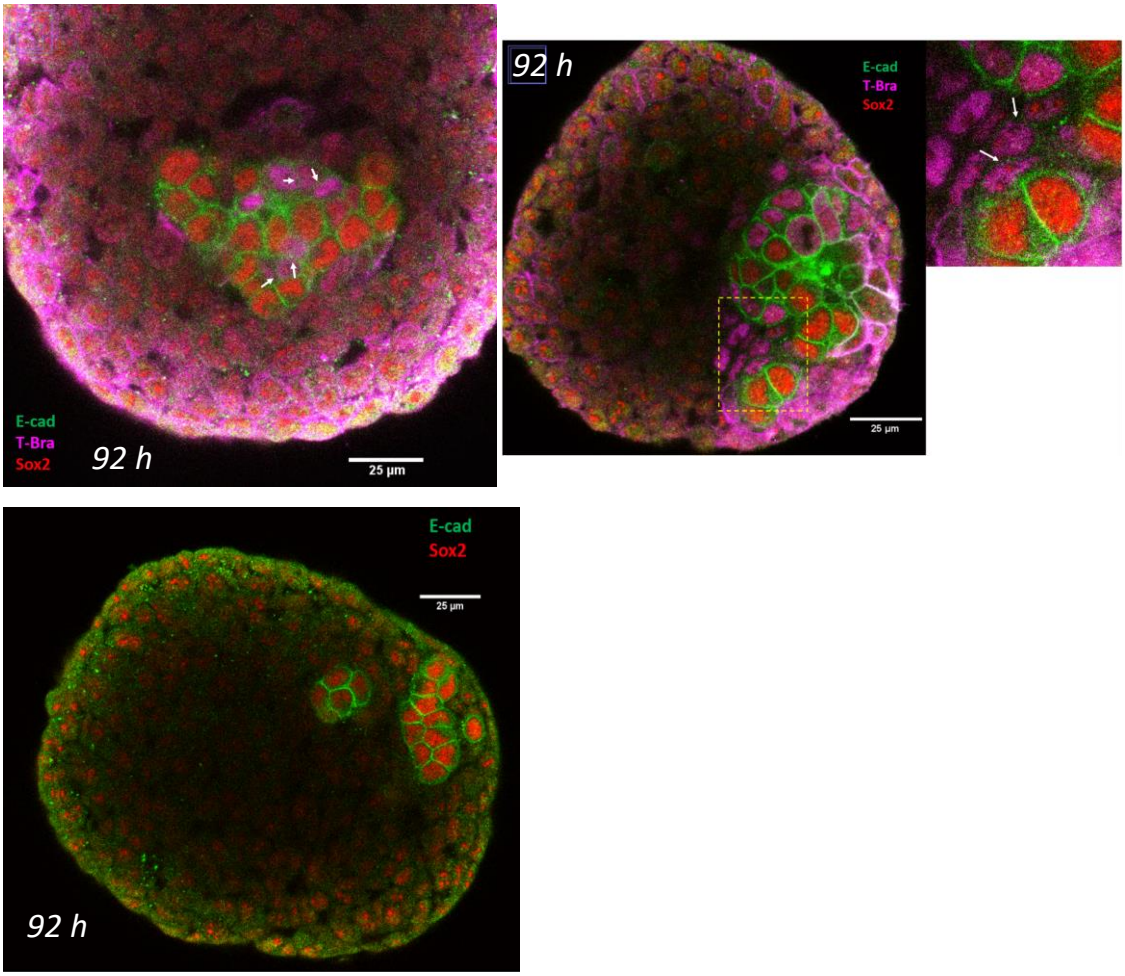


Figure 3-figure supplement 2

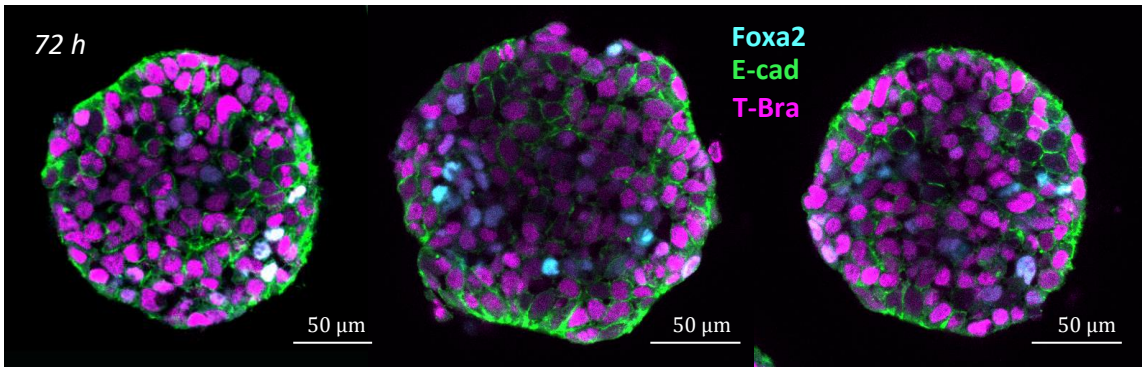
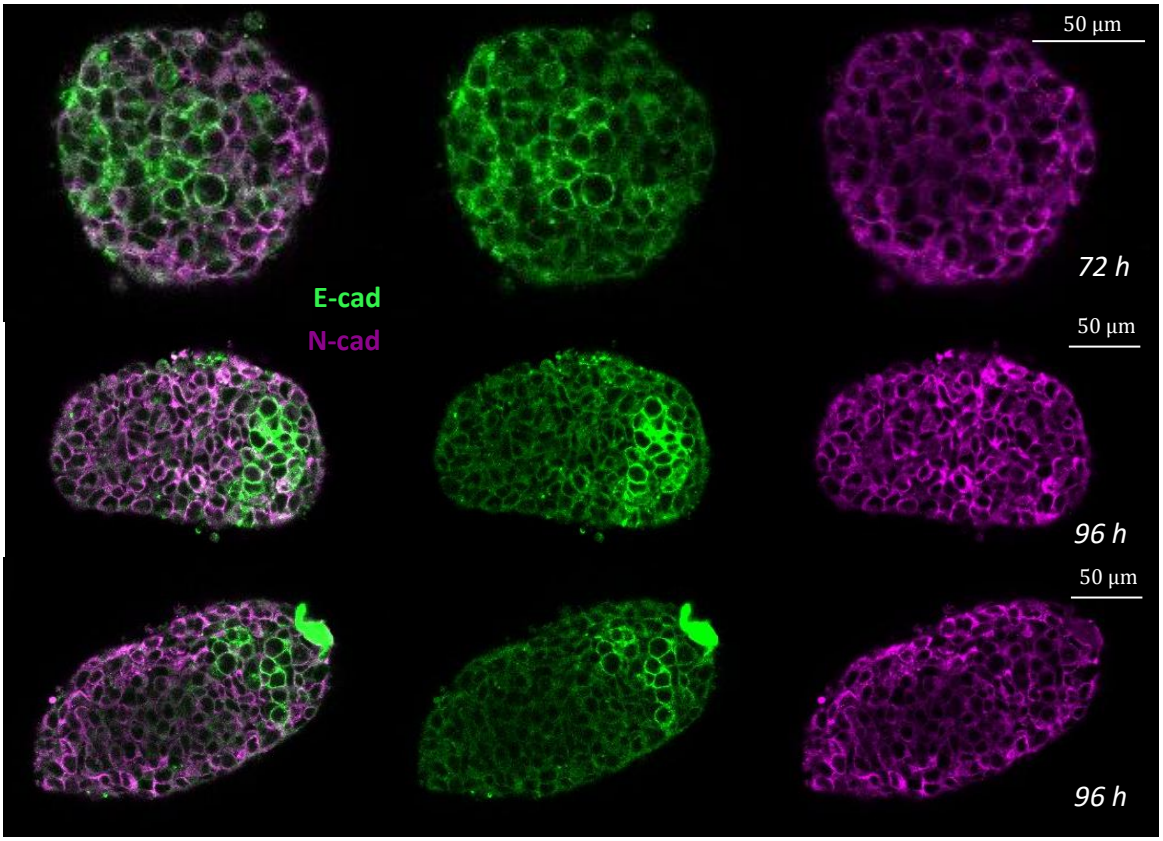
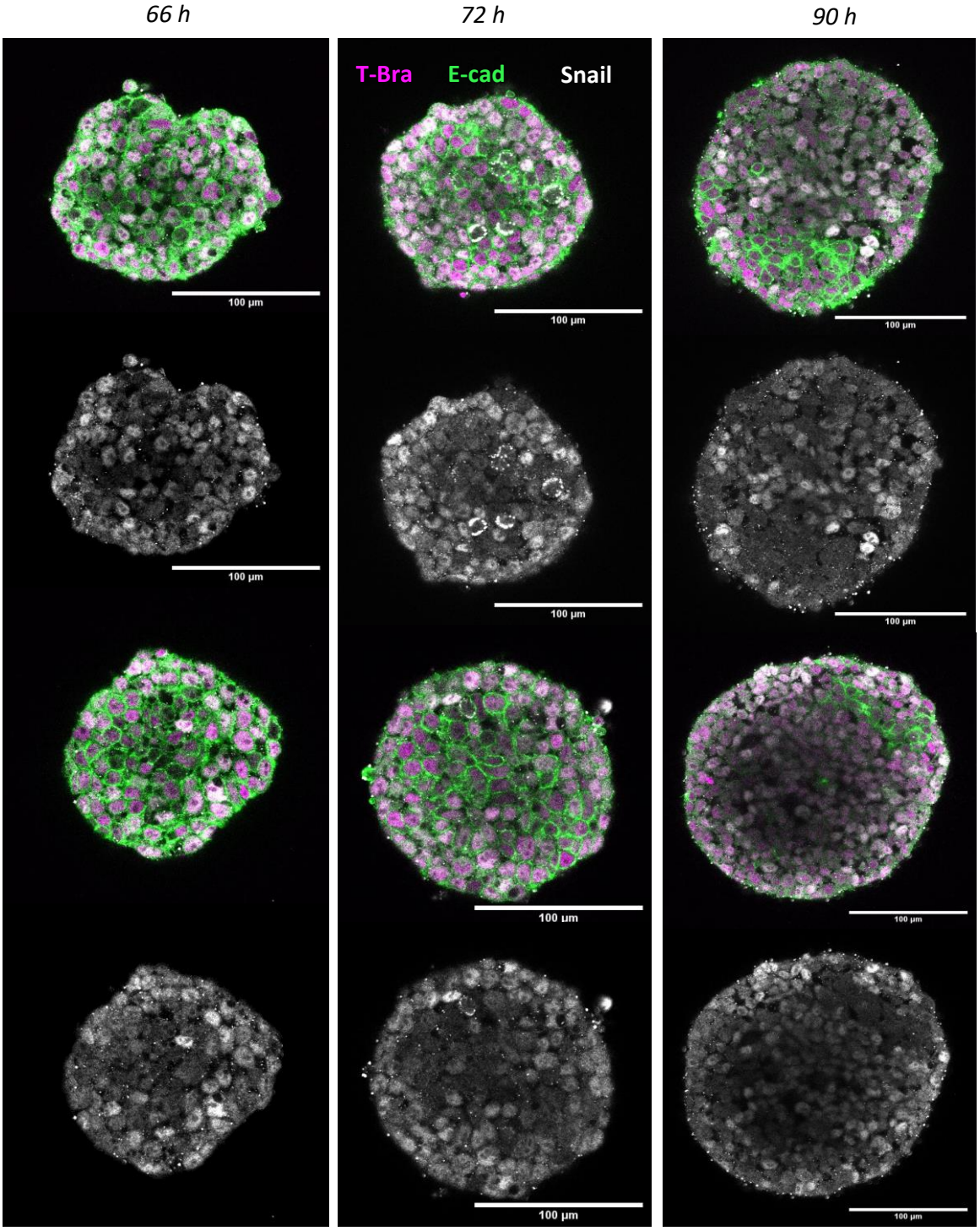


Figure 3-figure supplement 3




66 h
72 h
90 h
T-Bra
E-cad
Snail
100 µm

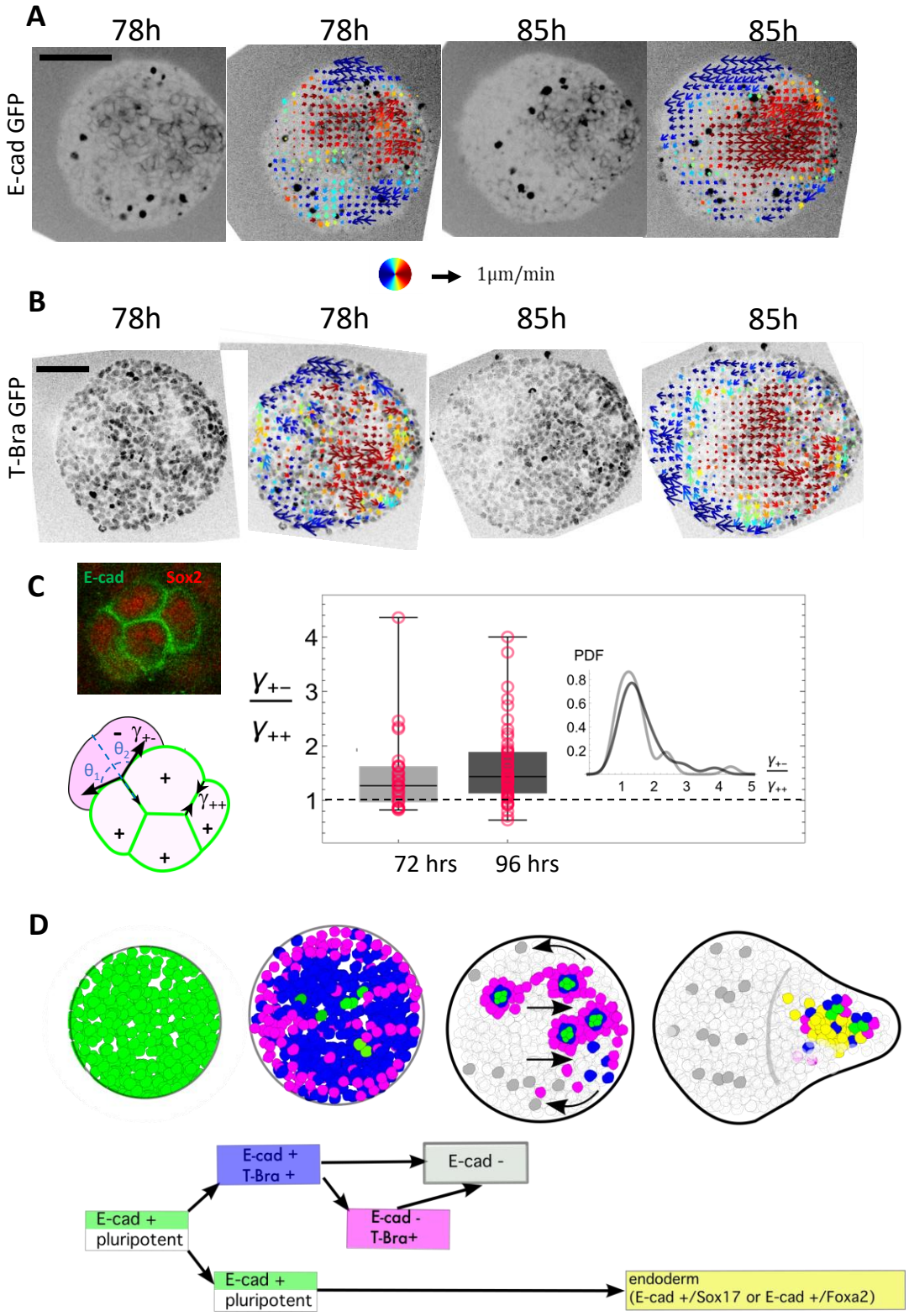

A
78h
78h
85h
85h
E-cad GFP
1μm/min
B
78h
78h
85h
85h
T-Bra GFP
C
E-cad
Sox2
PDF
72 hrs
96 hrs
D
E-cad +
T-Bra +
E-cad -
E-cad -
T-Bra+
E-cad +
pluripotent
E-cad +
pluripotent
endoderm
(E-cad +/Sox17 or E-cad +/Foxa2)

Figure 4-figure supplement 1

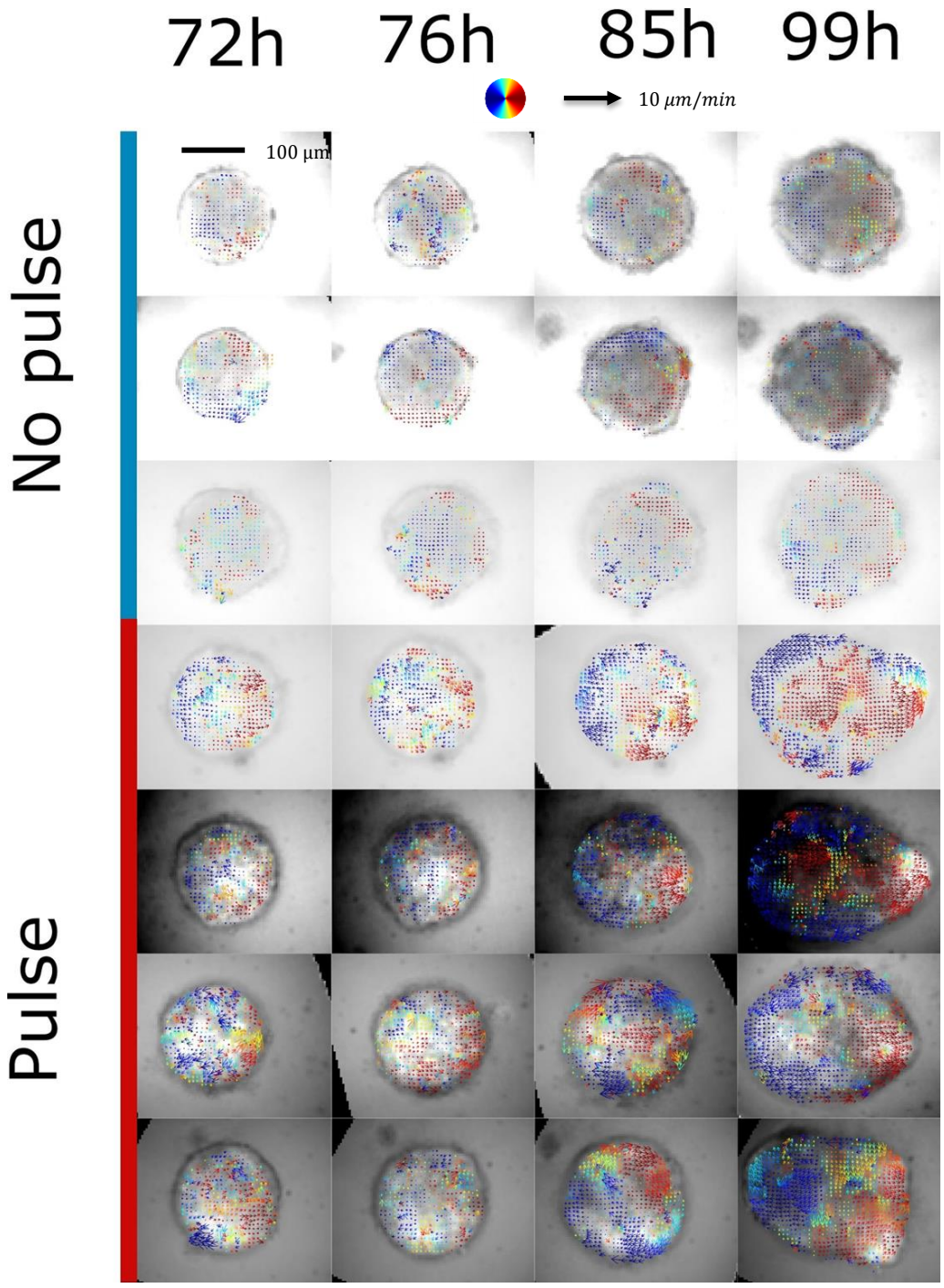
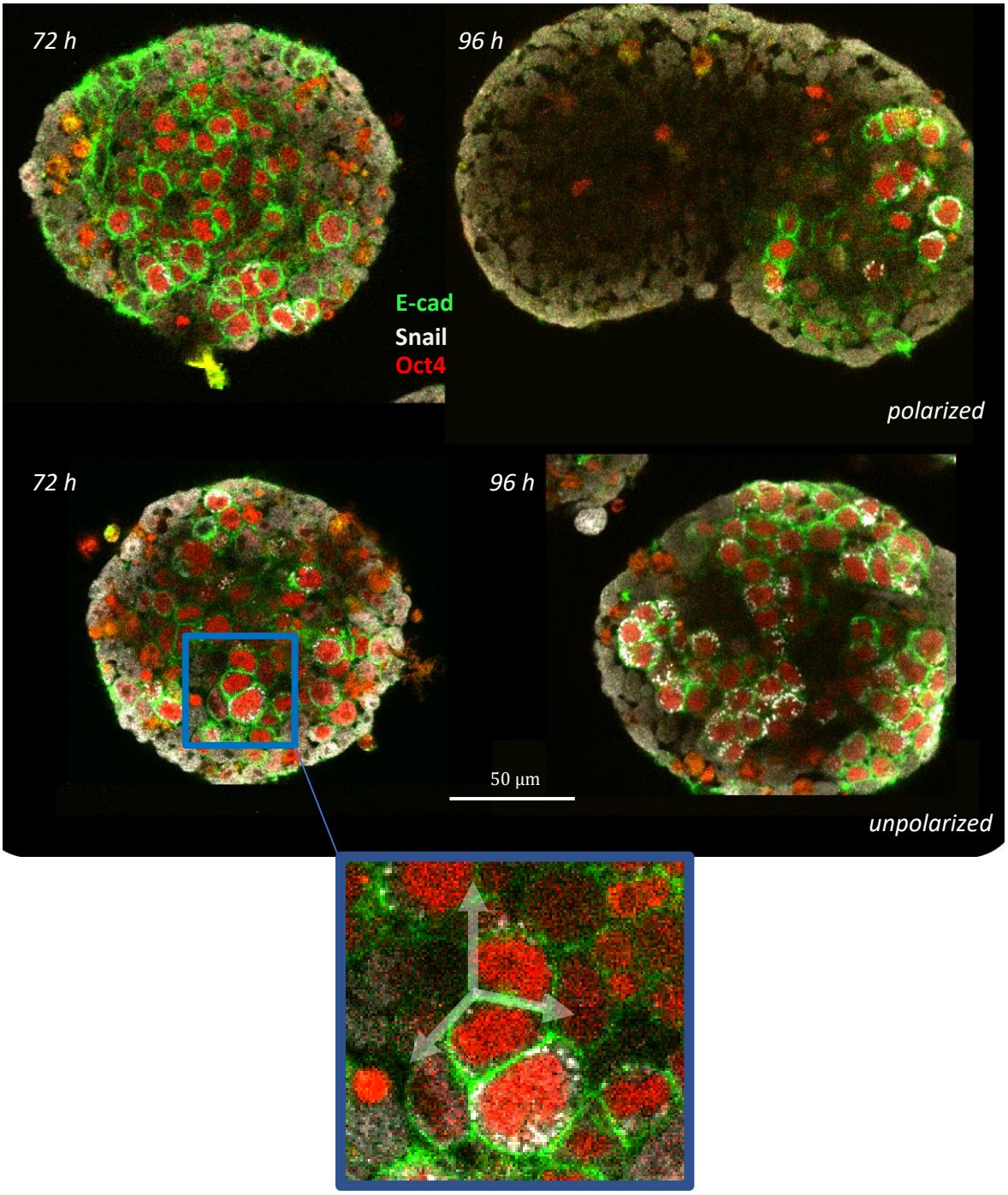



Figure 4-figure supplement 2