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], 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*

[*]*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 different cell 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 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 state by a three-step mechanism: a release of islands of E-cadherin expressing cells, their flow toward the aggregate tip, and their segregation. Unlike the prevailing view, this mechanism does not require epithelial-to-mesenchymal transitions and vice-versa but rather a fragmentation, which is mediated by Wnt/β–catenin, and a sorting process. Our data emphasize the role of signaling and cell flows in the establishment of the body plan.

Introduction

Gastrulation is the 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[1–3]. Any defects or deviations during gastrulation will manifest as severe abnormalities in embryos, often resulting in premature death. 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 the increase in cellular mass is

concomitant with gastrulation[4]; 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[5,6]. Recent live imaging of the process[7] 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[8,9] provides several advantages: 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 - which are not limited to morphogen concentration[10], 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 and more broadly aggregates of embryonic stem cells (ESCs), 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[11,12]. Strikingly, 3D gastruloids exhibit a multiaxial organization and spatiotemporal patterning of Hox genes[13] with signatures of somitogenesis[14,15]. From *in-vitro* assays, it is evident that a combination of particular cell types is needed to generate lumen in the blastocyst[16] and establish an anterior-posterior axes prior gastrulation[17]. 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[18]. Besides tridimensional (3D) systems, planar/2D gastruloids confined on micropatterns have emphasized the role of the boundary conditions, shapes[19] and scaling parameter on the patterning of the different germ layers[20–23].

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[24]. On the other hand, despite some evidence there has been a longstanding debate whether the mesoderm and the endoderm originate from bipotent precursor cells[25].

Here we begin to address this question by studying the formation of an endoderm-like region in 3D gastruloids, focusing on the spatio-temporal deployment and state transitions of cells 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 track cell movements that localize 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

During the early stages of gastruloid development, cells lack an epithelial organization. This may be due to low levels of signals driving the exit from pluripotency. We placed ESCs in Activin and FGF throughout the protocol as it is done to transition cells to the post-implantation stage[26]. A portion of the aggregates were exposed to the Wnt agonist Chiron (Chi) as the combination of Activin and Chi tends to promote endoderm formation[22,27]. We investigated the aggregate morphologies and gene expression patterns following a pulse of Chi (lasting a day from day 2).

Aggregates exposed to Chi lose their spherical morphology and acquire a teardrop shape (Fig 1a, left, day 4, Movie 1a). We find that 30/37 or ~ 81% aggregates qualify as elongated with an elongation threshold set to 0.125. In contrast 15/20 or 75 % of control aggregates remained spherical (Figure 1b, right, day 4). At the pointed end of Chi exposed aggregates, we observed an E-cad rich region, which was absent in control (Figure 1b). The E-cad pole emerges from an initially homogeneous distribution of E-cad (as delineated in Figure 1c, Movie 1b). Likewise, a 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 that while E-cad is expressed from the early time points (Figure 1d, day 2), T-Bra is expressed following exposure to Chi. A sharp rise in T-Bra (at day 2.5, Movie 2) precedes the decay of E-cad (Fig 1d, Movies 3a, 3b, 4a and 4b). The E-cad level, minimal after removal of Chi, increases thereafter to reach a constant value. T-Bra level, maximal just before removal of Chi, gradually decreases thereafter (Fig 1d, Movies 5a and 5b). 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 accumulates at one pole prior to the onset of tip formation (Figure 1e, left). T-Bra polarizes as the aggregate deforms, implying a possible causal link between T-Bra polarization and the aggregate's shape (Figure 1e, right). A comparison between T-Bra and E-cad elongation - as a function of aggregate polarization - indicates that the E-cad polarization is more abrupt than T-Bra polarization, which remains more homogeneous at early stages (Supplementary Fig. 1). Our imaging data suggests two coordinated, albeit distinct, processes for T-Bra and E-cad patterning.

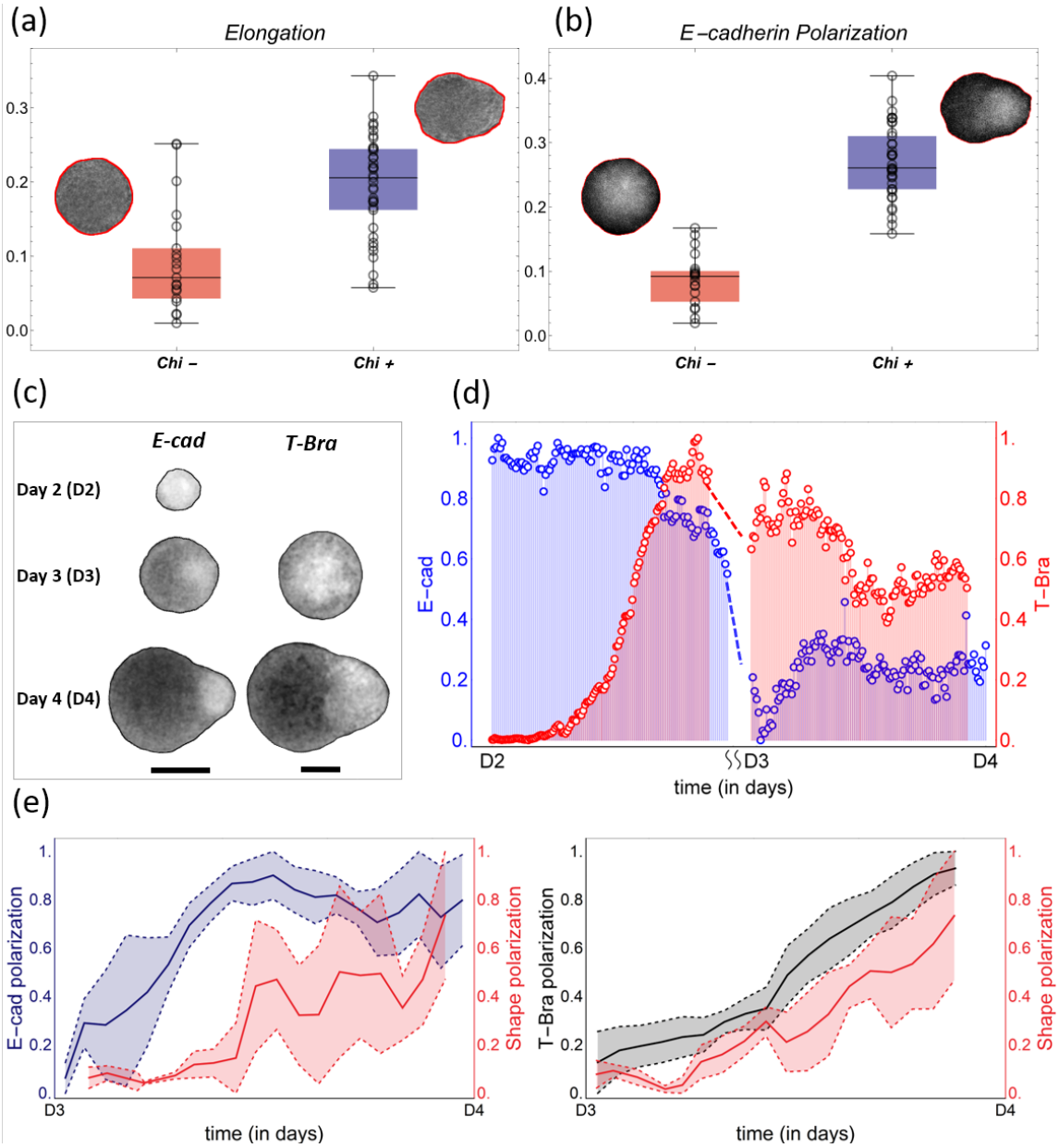

Figure 1: (a) The box-whisker plot shows the distribution of elongation for Chi + (n = 37) and the control (Chi -) aggregates (n = 20). The insets delineate the contours of the two classes of aggregates. (b) The box-whisker plot shows the distribution of *E-cad* polarization in the Chi + (n = 37) and Chi - aggregates (n = 20). The insets show the intensity profile of *E-cad* within the aggregates (c) 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*) at different developmental stages and exposed to Chi pulse (scale bars for the columns represent 100 um). The T-Bra-GFP/NE-mKate2 aggregate shown here was generated from ~ 200 cells (d) Mean temporal intensity profiles of *E-cad* and *T-Bra* between the addition of Chi (D2, 48 hrs post-plating (hrs-pp)), removal of Chi (D3 or 72 hrs-pp) and D4 (92 hrs-pp), n = 4 aggregates for *E-cad-GFP/Oct4-mCherry* and n = 3 aggregates for *T-Bra-GFP/NE-mKate2*. (e) The plots illustrate the polarization *of E-cad* (left panel, n = 4 aggregates) and *T-Bra* (right-panel, n = 3 aggregates) with respect to the deformation of the aggregate between D3 and D4; the thick and the dashed lines represent the mean and the mean ± SD respectively.

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 found that this core expresses Sox17 (Figure 2a, inset), Foxa2 (Supplementary Fig. 2) and Gata6 (Supplementary Fig. 3) – all distinct markers of the endoderm. We confirmed that Sox2 is present earlier but absent at 5 days-pp (Supplementary Fig. 4). These observations lead us to conclude that the group of E-cad expressing cells demarcates an endoderm-like region in the gastruloids. The layer of cells surrounding the core (T-Bra +/Sox17 -) might be the future presumptive mesoderm, as reported in 2D micropattern[20]. Figure 2b shows a histogram of the various cell population near the tip of an aggregate (n = 3). The core of E-cad/Sox17 expressing cells and the T-Bra expressing cells constitute the majority (Figure 2b and inset) of the cell population near the tip. 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 and a homotypic neighbourhood, which implies that they form a tightly packed group of cells. In contrast, the T-Bra expressing cells form multiple simply-connected-components suggesting that they are sparsely distributed around the core. The cells farther away from the tip express Gata6 – a marker of the cardiac crescent cells (Supplementary Fig. 3) – consistent with an earlier study[13].

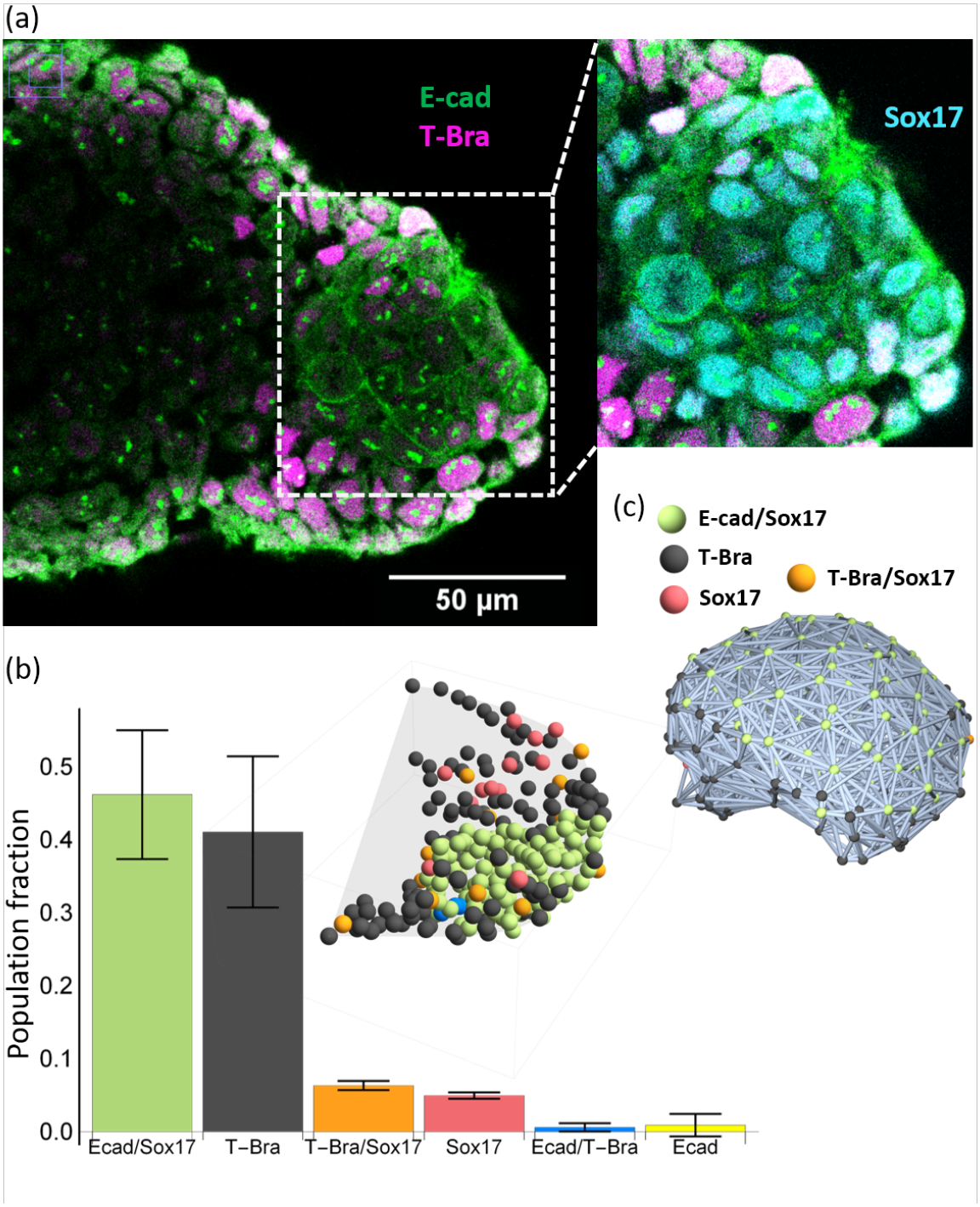

Figure 2:(a) aggregate stained at 96 hrs-pp for E-cad and T-Bra; the inset shows the endoderm-like cells at the tip expressing both E-cad and Sox17, surrounded by T-Bra + cells (b) the bar chart delineates the population fraction of the different cells populations present near the tip (n = 3 aggregates) with the error bars indicating the SD. The inset is the spatial distribution of the different cell types near the tip of a representative aggregate, fixed at 96 hrs-pp (c) spatial connectivity graph (Delaunay) of the different cell populations in an aggregate fixed at 96 hrs-pp

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. The E-cad is present at every cell-cell contact forming a quite homogeneous network of junctions. However, this network undergoes a gradual fragmentation (Figure 3a and Movie 6). 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 (Supplementary Fig. 5). In view of earlier studies[26,28], 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 Sox17 (Figure 3b). Therefore, at this stage 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 3c, left). The E-cad/T-Bra expressing cells constitute most of the cell population and are surrounded by the T-Bra expressing cells (Figure 3c, middle). The cells expressing only E-cad – present in the minority - are present mainly within the core of the E-cad/T-Bra expressing cells. It is these cells that are also found to be the ones expressing Sox2, akin to an epiblast (Supplementary Fig. 6). The cells in the core are found to be tightly packed and maintain a homotypic neighbourhood (Figure 3c, right). Yet again the T-Bra expressing cells are sparsely wrapped around the core and therefore, have a heterotypic neighbourhood. All the three classes of cells have the potential to undergo division (Supplementary Fig. 7). Interestingly, at later timepoints but before the aggregates are fully polarized, we observed an "archipelago" of E-cad expressing cells separated by T-Bra expressing cells (Figure 3d, left). Additionally, the islands of E-cad expressing cells were found to be positive for the pluripotency markers Sox2 (Supplementary Fig. 8) and Oct4. The Oct4 expression correlates inversely with the T-Bra expression, suggesting pluripotency and an epiblast-like nature of cells (see Figure 3d inset and Figure 3d, right). Altogether our data suggests that the endoderm-like region - identified at day 4 - originates from islands of E-cad expressing cells.

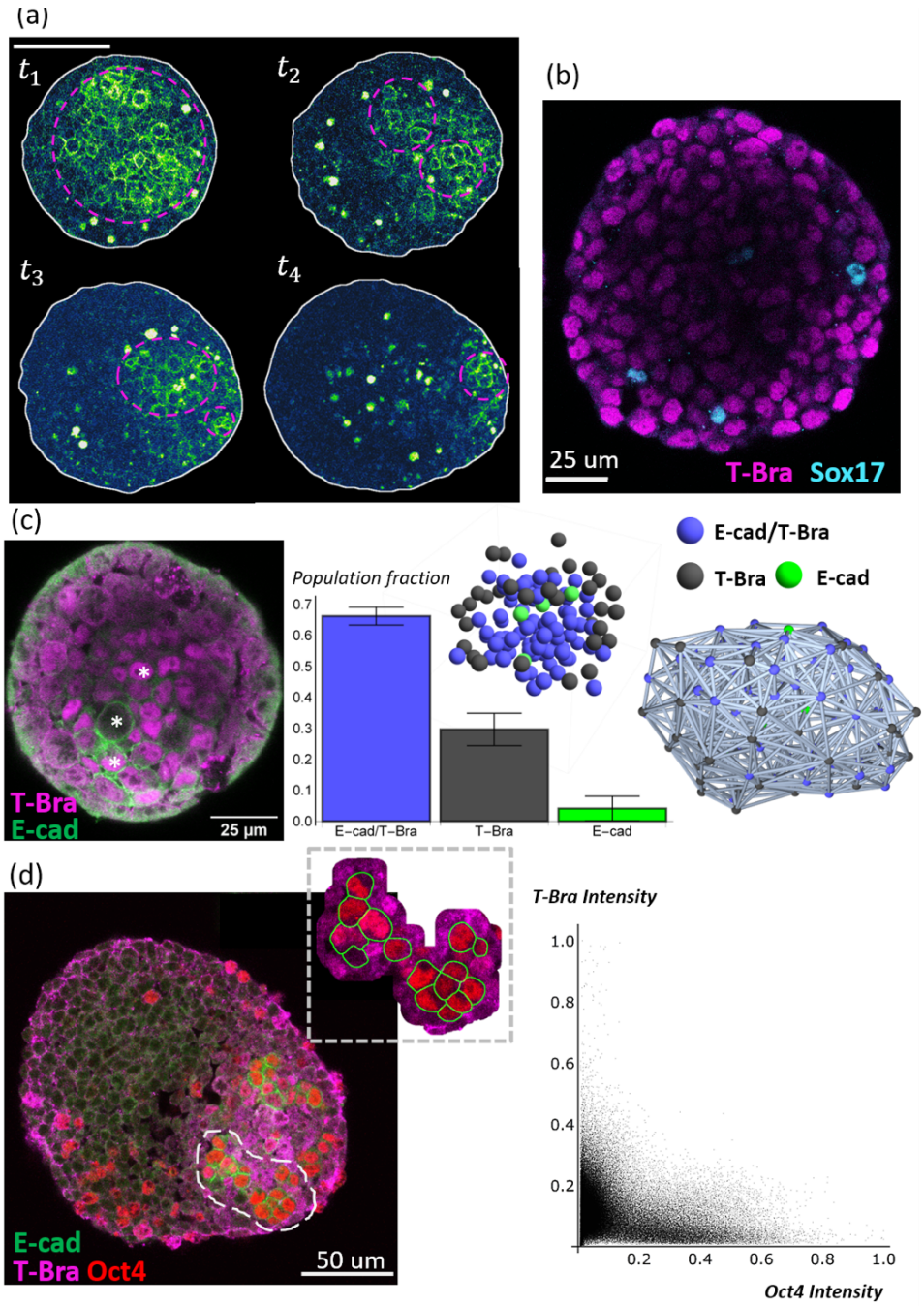

Figure 3: (a) Time lapse images showing a gradual fragmentation and loss of *E-cad* network at t_1 = 72.6 hrs-pp, t_2 = 80.2 hrs-pp, t_3 = 84.7 hrs-pp and t_4 = 91.5 hrs-pp; scale bar represents 100 µm (b) *Sox17* is expressed in a salt and pepper manner in the aggregate at 72 hrs-pp (c) (left) The three 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* (middle) the fraction of the total population of cells for the three classes are provided in the histogram (n = 4 aggregates) with the SD shown by the error bars; inset shows the spatial distribution of the cells in one such aggregate (right) a neighbourhood graph is generated for the aggregate at 72 hrs-pp (d) (left) 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 (right) A pixelwise correlation between *Oct4* and *T-Bra* (for the inset) illustrates an inverse relation

Directed cellular movements and surface tension merge E-cadherin islands together

Our observations raise the possibility that the islands of E-cad expression sort and merge to form a compact group of differentiated cells. To test this possibility, we analyzed cell movements prior to the establishment of the tip. The time averaged-flow field (Figure 4a and Movie 6) clearly shows that there is a directed flow of cells towards the future region that will form the tip of the gastruloid. The islands of E-cad expressing cells move together with the flow and are brought to the region forming the tip, where they subsequently merge. Prior to the tip formation we find no coherent structure in the velocity vector field. 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. The E-cad expressing cells in islands have a significantly smaller contact angle with the surrounding cells than with each other (Figure 4b, inset). Assuming that the cell-cell contacts are determined by interfacial tension (Supplementary material), we compared the angle of contacts between E-cad expressing cells within the islands and with neighbouring cells (Figure 4b). The E-cad cells maintain a mean angle with an SD of 116 ± 6° (n = 20) with respect to each other while they form a mean angle with an SD of 147 ± 12° (n = 13) with the surrounding cells. This corresponds to approximately a two-fold difference in interfacial tension. Such difference could account for the maintenance of E-cad 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.

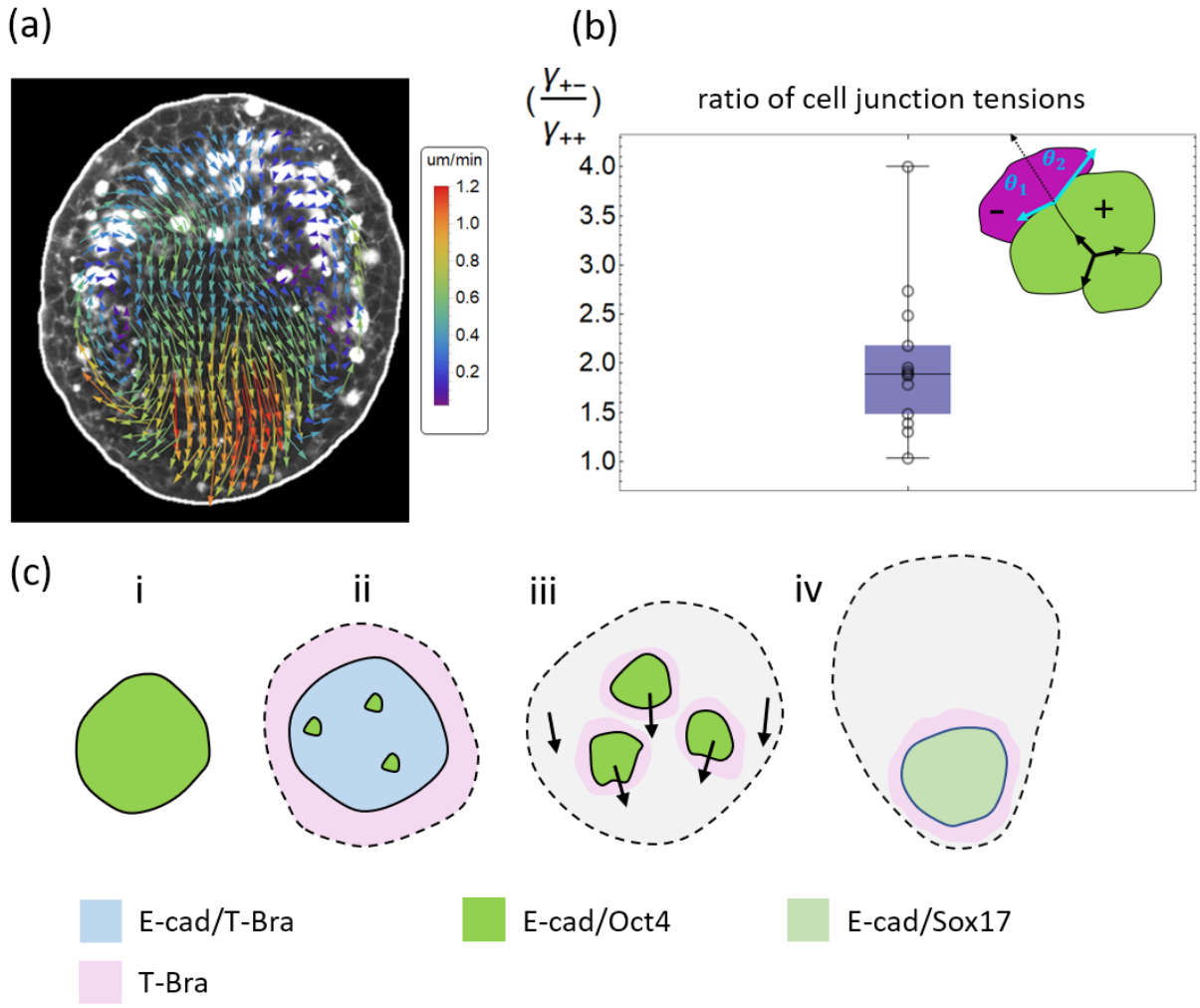

Figure 4: (a) The velocity field within the aggregate shows a directed flow of cells towards the tip (representative aggregate shown) (b) distribution of the ratio of the cell-junction tension heterogeneity between *E-cad* expressing cells (+), and between *E-cad* expressing cells and their neighbours (-) (n = 13 junctions); 25th and 75th quartiles together with the minimum and maximum values are shown (c) the cartoon depicts the underlying model that explains the formation of the endoderm-like region within the gastruloid. (i) upon addition of *Chi*, majority of the *E-cad* expressing cells begin to express *T-Bra* (ii) The transient expression of *T-Bra* represses *E-cad* expression in *T-Bra* +/*E-cad* + (iii) cell divisions cause the *E-cad* expressing cells (green) to grow in size forming islands of *E-cad* expressing cells surrounded by *T-Bra* + cells; these islands are transported to one end of the aggregate via a coordinated cellular movement (iv) islands of *E-cad* expressing cells fuse to give rise to the endoderm.

Discussion

In contrast to the previously reported assays on 3D gastruloids, aggregates were prepared in Activin and FGF to transit naïve cells to an epiblast-like state that is characterized by epithelial and pluripotent traits (E-cad/Oct4). Exposure of developing gastruloids to Chi on the third day results in a polarized expression of T-Bra and a coordinated multicellular movement that mimics some aspects of mammalian gastrulation, in particular the epithelial organization of the gastruloid at the time of Chiron application. This protocol results in the emergence of a distinct population of cells with features of the endoderm, a situation that mimics that of the embryo, Our results suggest a three-tier mechanism for the emergence of endoderm in mESC aggregates: (1) a release of E-cad+/T-Bra- cells that generate compact islands; (2) a directed global flow that transports such islands at a pole; (3) their segregation to form an endoderm-like region. This mirrors what has been described in the embryo[29] but also raises some possibilities for the way the endoderm arises.

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 4c, i & ii). The cells near the periphery specifically exhibit high level of T-Bra and low levels of E-cad (supplementary Fig. 9). A similar observation has been made on 2D micropatterns in which a T-Bra wave emanates and travels inwards, from the boundary where the E-cad is absent[30]. The inner core of cells expressing both markers envelope the cells that are expressing E-cad/Sox2 (E-cad/Oct4). Nevertheless, we observe that these E-cad/Sox2 expressing cells increase in numbers and give rise to individual islands of E-cad expressing cells. The T-Bra expressing cells situated in the vicinity of the islands are produced via loss of E-cad from the E-cad/T-Bra cells in the aggregate's core (figure 4c, iii). 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 merge at the tip and are found to express markers characteristic of the endoderm (Figure 4c, iv). In contrast, in the absence of Chi pulse the mass of cells maintains high levels of E-cad at the cell-cell contacts; these aggregates generally are also devoid of any pole.

Interestingly, in contrast to the prevailing view[29], we observe that the entire process of establishing an endoderm-like region in the aggregates does not require the endodermal progenitors to undergo an Epithelial-Mesenchymal transition (EMT) and vice versa. The very cells that give rise to the endoderm retain E-cad and remain in compact islands. These cells maintain their epithelial nature as well as integrity and are only spatially relocated to a position where they will define the endoderm. Moreover, these endodermal-like progenitor islands in the aggregate emerge from a different pool of cells (E-cad + /T-Bra -) than the mesodermal-like precursor cells (T-Bra +) which is consistent with an earlier observation[31].

There has been a longstanding debate as to how the different germ layers are established during gastrulation. Previous studies suggest that the endoderm is established from the progenitor cells at the primitive streak that undergo an active EMT and MET[25]. Our study indicates another possible route: while EMT occurs at the vicinity of the endodermal progenitors, the former cells remain epithelial-like and become endoderm without undergoing EMT/MET. Our data indicates that the cells that undergo EMT and become mesoderm-like wrap islands of endodermal progenitors. They move collectively to form a pole. The segregation of endodermal-like cells into a distinct zone is thus a consequence of a wave of cell state-transitions, a global flow and surface tension heterogeneity. Whether this mechanism

might exist at the primitive streak in the mouse embryo remains to be explored. Our work examines how signaling, flows, forces and patterning, might be coupled during endoderm formation, and perhaps more generally during the morphogenesis of germ layers.

MATERIAL AND METHODS

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

- **Cell Culture and Gastruloids**
- **Immunostaining**
- **Live- and fixed-cell imaging**
- **Cryosection**
- **Relative surface tension**
- **Image Processing, Segmentation and Data Analysis**

SUPPLEMENTAL INFORMATION

Material and methods, supplementary figures, movies, analysis scripts and statistics

ACKNOWLEDGMENTS

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-04-01). 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.

AUTHOR CONTRIBUTIONS: PFL, AMA and AH conceived the project; AH, ST, PP and PFL performed the experiments; AH and PFL analyzed the data with some inputs from ST; AH and PFL wrote the paper and all authors commented on it; and PFL supervised the research.

DECLARATION OF INTERESTS

The authors declare no competing interests.

References:

1. Wolpert, L., Tickle, C. & Arias, A. M. *Principles of development*. (Oxford University Press, 2015).

2. Solnica-Krezel, L. & Sepich, D. S. Gastrulation: Making and Shaping Germ Layers. *Annu. Rev. Cell Dev. Biol.* (2012). doi:10.1146/annurev-cellbio-092910-154043

3. Williams, M. L. & Solnica-Krezel, L. Regulation of gastrulation movements by emergent cell and tissue interactions. *Current Opinion in Cell Biology* (2017). doi:10.1016/j.ceb.2017.04.006

4. Snow, M. H. L. Gastrulation in the mouse: growth and regionalization of the epiblast. *J. Embryol. Exp. Morphol.* (1977).

5. Anderson, K. V. & Ingham, P. W. The transformation of the model organism: A decade of

developmental genetics. *Nat. Genet.* (2003). doi:10.1038/ng1105

6. Artzt, K. Mammalian developmental genetics in the twentieth century. *Genetics* (2012). doi:10.1534/genetics.112.146191

7. McDole, K. *et al.* In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell* (2018). doi:10.1016/j.cell.2018.09.031

8. Vianello, S. & Lutolf, M. P. Understanding the Mechanobiology of Early Mammalian Development through Bioengineered Models. *Developmental Cell* (2019). doi:10.1016/j.devcel.2019.02.024

9. Shahbazi, M. N., Siggia, E. D. & Zernicka-Goetz, M. Self-organization of stem cells into embryos: A window on early mammalian development. *Science (80-.).* (2019). doi:10.1126/science.aax0164

10. Heemskerk, I. *et al.* Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells. *Elife* (2019). doi:10.7554/eLife.40526

11. Van Den Brink, S. C. *et al.* Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Dev.* (2014). doi:10.1242/dev.113001

12. Turner, D. A. *et al.* Anteroposterior polarity and elongation in the absence of extraembryonic tissues and of spatially localised signalling in gastruloids: Mammalian embryonic organoids. *Dev.* (2017). doi:10.1242/dev.150391

13. Beccari, L. *et al.* Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* (2018). doi:10.1038/s41586-018-0578-0

14. van den Brink, S. C. *et al.* Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* (2020). doi:10.1038/s41586-020-2024-3

15. Veenvliet, J. V *et al.* Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *bioRxiv* 2020.03.04.974949 (2020). doi:10.1101/2020.03.04.974949

16. Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. & Zernicka-Goetz, M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science (80-.).* (2017). doi:10.1126/science.aal1810

17. Sozen, B. *et al.* Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat. Cell Biol.* (2018). doi:10.1038/s41556-018-0147-7

18. Sagy, N. *et al.* Prediction and control of symmetry breaking in embryoid bodies by environment and signal integration. *Dev.* (2019). doi:10.1242/dev.181917

19. Blin, G. *et al.* Geometrical confinement controls the asymmetric patterning of brachyury in cultures of pluripotent cells. *Dev.* (2018). doi:10.1242/dev.166025

20. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* (2014). doi:10.1038/nMeth.3016

21. Etoc, F. *et al.* A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Dev. Cell* (2016). doi:10.1016/j.devcel.2016.09.016

22. Morgani, S. M., Metzger, J. J., Nichols, J., Siggia, E. D. & Hadjantonakis, A. K. Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate

patterning. *Elife* (2018). doi:10.7554/eLife.32839

23. Xue, X. *et al.* Mechanics-guided embryonic patterning of neuroectoderm tissue from human pluripotent stem cells. *Nat. Mater.* (2018). doi:10.1038/s41563-018-0082-9

24. Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M. & Sutherland, A. Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev. Dyn.* (2012). doi:10.1002/dvdy.23711

25. Nowotschin, S., Hadjantonakis, A. K. & Campbell, K. The endoderm: A divergent cell lineage with many commonalities. *Development (Cambridge)* (2019). doi:10.1242/dev.150920

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

27. Engert, S. *et al.* Wnt/β-catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. *Dev.* (2013). doi:10.1242/dev.088765

28. Fernando, R. I. *et al.* The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *J. Clin. Invest.* (2010). doi:10.1172/JCI38379

29. Viotti, M., Nowotschin, S. & Hadjantonakis, A. K. SOX17 links gut endoderm morphogenesis and germ layer segregation. *Nat. Cell Biol.* (2014). doi:10.1038/ncb3070

30. Martyn, I., Brivanlou, A. H. & Siggia, E. D. A wave of WNT signaling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. *Dev.* (2019). doi:10.1242/dev.172791

31. Burtscher, I. & Lickert, H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* (2009). doi:10.1242/dev.028415