

Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo structures

Berna Sozen^{1,2,10}, Gianluca Amadei^{1,10}, Andy Cox¹, Ran Wang³, Ellen Na⁴, Sylwia Czukiewska¹, Lia Chappell⁵, Thierry Voet^{5,6}, Geert Michel⁷, Naihe Jing^{3,8}, David M. Glover⁹ and Magdalena Zernicka-Goetz^{1*}

¹Mammalian Embryo and Stem Cell Group, University of Cambridge, Department of Physiology, Development and Neuroscience; Downing Street, Cambridge, CB2 3DY, UK

²Department of Histology and Embryology, Faculty of Medicine, Akdeniz University, Antalya, 07070, Turkey

³State Key Laboratory of Cell Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences; University of Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China

⁴Department of Cardiology, Charité, Berlin, Germany

⁵Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK

⁶Laboratory of Reproductive Genomics, Department of Human Genetics, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

⁷Transgene Technologies, Charité, Berlin, Germany

⁸School of Life Science and Technology, ShanghaiTech University, 100 Haike Road, Shanghai 201210, China

⁹Department of Genetics, University of Cambridge, Cambridge, England, UK

¹⁰equal contribution

*corresponding author mz205@cam.ac.uk

Embryonic stem (ES) cells can be incorporated into the developing embryo and its germ line but when cultured alone, their ability to generate embryonic structures is restricted. They can interact with trophoblast stem (TS) cells to generate structures that break symmetry and specify mesoderm, but their development is limited as the epithelial-mesenchymal-transition (EMT) of gastrulation cannot occur. Here we describe a system that, for the first time, allows assembly of mouse ES, TS and extra-embryonic-endoderm (XEN) cells into structures that acquire the embryo's architecture with all distinct embryonic and extra-embryonic compartments. Strikingly, such embryo structures develop to undertake EMT, leading to mesoderm and then definitive endoderm specification. Spatial transcriptomic analyses demonstrate that these morphological transformations are underpinned by gene-expression patterns characteristic of gastrulating embryos. This demonstrates the remarkable ability of three stem cell types to self-assemble *in vitro* into gastrulating embryo structures undertaking spatio-temporal events of the gastrulating mammalian embryo.

Introduction

The mammalian embryo is built from three distinct tissues: the embryonic tissue or epiblast, that generates the body of the embryo, and two extra-embryonic tissues: trophoctoderm that generates the placenta and primitive endoderm that generates the yolk sac. As the embryo implants, it undergoes a morphological transformation during which primitive endoderm-derived visceral endoderm (VE) triggers remodelling of the epiblast into a polarised cellular rosette that evolves to undertake lumenogenesis^{1,2}. This is followed by cavitation of the adjacent trophoctoderm-derived extra-embryonic ectoderm (ExE) and finally the fusion of these cavities to form the single pro-amniotic cavity³. The signalling interactions between these tissues are important for the embryo to break its symmetry to establish anterior and posterior identity and initiate gastrulation^{4,5}.

ES cells can be incorporated into the developing embryo and its germ line^{6,7}. However, when cultured alone, their ability to mimic embryogenesis is restricted⁸⁻¹⁰ - the resulting structures neither acquire embryo architecture nor correctly recapitulate gastrulation. To address this, we previously allowed ES cells to co-operate with TS¹¹ cells *in vitro* in a three-dimensional extra-cellular-matrix scaffold³. This led them to form embryo-like structures that resemble early post-implantation embryos with an ES-derived embryonic compartment and a TS-derived extra-embryonic compartment. These structures develop further and break symmetry

to undertake specification of nascent mesoderm and primordial-germ-cells (PGCs) at the posterior³. However, the establishment of the anterior and the key morphogenetic events of gastrulation, namely the EMT, were limited in this system. We hypothesised that this was because crosstalk between all three tissues would be required to properly recapitulate embryogenesis *in vitro* and these structures were missing the third tissue, the primitive endoderm-derived VE.

Indeed, the VE is known to play important roles during embryogenesis. First, it develops to surround both embryonic and extra-embryonic compartments and then becomes regionalised and develops a signalling centre that will enhance development of the anterior-posterior axis by repressing Nodal and Wnt signalling on the future anterior^{12,13}. Second, it provides a barrier for migration of embryonic cells undertaking the EMT during gastrulation to form mesoderm. Finally, migrating embryonic cells progressively displace VE cells to generate the definitive endoderm^{14,15}.

Here, we wished to test the hypothesis that crosstalk between all three stem cell types found in the natural embryo would enable the generation of whole embryo-like structures able to exhibit the cell rearrangements of gastrulation. To achieve this, we established culture conditions allowing the stem cell lines for all three tissues that comprise mouse embryo - ES, TS and XEN¹⁶ cells - to co-operate *in vitro*. We show that this co-operation results in the spontaneous self-assembly of structures that are remarkably similar to natural mouse embryos and contain all three embryonic and extra-embryonic compartments. We show that these structures develop to undertake anterior-posterior patterning and morphogenesis leading to the key spatio-temporal events of gastrulation, including the EMT and specification of mesoderm and definitive endoderm.

Results

Establishing conditions for spontaneous assembly of ES, TS and XEN cells into embryo-like structures

To address the hypothesis that crosstalk between the three stem cell types *in vitro* would permit generation of embryo-like structures with all three compartments as is the case in natural embryos, we seeded single-cell suspensions of ES, TS and XEN cells in medium supporting the proliferation of all these cell types into inverted-pyramidal micro-well plates

allowing cells to aggregate (Methods). At optimised cell density, the three cell types spontaneously underwent self-organisation into multicellular aggregates within 24h. After 96h, approximately 70% of multi-cellular aggregates contained all three cell types. Of these, 28.6% (n=130/435) had single adjoining ES-derived embryonic and TS-derived extra-embryonic compartments surrounded by a layer of XEN cells and were strikingly similar in morphology to mouse embryos 5-6 days after fertilisation (Fig.1a-c; Supplementary Fig.1a,b; Supplementary Movie 1). In agreement with their morphology, these structures expressed markers characteristic of the three tissues of the embryo (Fig. 1d,e). Cell numbers in the ES, TS and XEN-derived compartments after 96h of culture were similar to those in the embryonic and extra-embryonic compartments of embryos cultured *in vitro* until 5.5 day after fertilisation (Fig.1f). We named these spontaneously organised structures ES+TS+XEN (ETX) embryos or embryo structures.

Mesoderm and PGC specification

We first examined the development of the ETX embryos to assess to which extent it was similar and different from the development of natural embryos and of the previous synthetic embryos (ET embryos) we had generated from only ES and TS cells³. Several developmental events followed similar spatio-temporal patterns. This included asymmetric expression of the mesoderm marker, *T/Brachyury*, in the ES-compartment close to the TS-compartment boundary 4.25 days after plating and extending towards the distal tip by day 5.5, thus resembling *T/Brachyury* expression in early (E6.5) to mid-streak (E7.0) stage embryos (Supplementary Fig.2a, b; 42% ETX embryos, n=101; 90% E6.5 embryos, n=20). The ratio of *T/Brachyury*-expressing to total cells in the ES-compartment of these ETX embryos was significantly higher than in synthetic embryos generated from only ES and TS cells and similar to embryos at equivalent developmental stage (Supplementary Fig.2c, d). In contrast, the majority of structures formed from ES and XEN cells alone did not express *T/Brachyury* in these conditions indicating this is promoted by signalling from the TS-derived compartment (Supplementary Fig.2e-g).

In natural embryogenesis, PGCs are specified at the posterior. To address whether this is also the case for ETX embryos, we next built them using ES cells expressing a T:GFP reporter. We found that expression of PGC markers (assessed on protein and mRNA levels) was induced at day 6 at the boundary between the ES- and TS-derived compartments and proximal to the mesodermal domain (Supplementary Fig. 2h-i). We also found an inverse

gradient of Oct4 and T/Brachyury protein across the ES-derived-embryonic compartment at this stage (Supplementary Fig.2j) indicating that not only posterior, but also aspects of anterior pattern became established as development of ETX embryos progressed, reminiscent of natural embryogenesis. These results indicate that ETX embryos develop to break symmetry close to the ES/TS compartment boundary to induce expression of mesoderm and PGC markers.

XEN cell layer, basal membrane and cavity development

Lumenogenesis of embryos requires signalling from the basement membrane generated by the VE¹. We therefore asked whether this requirement could be met by the XEN layer in ETX embryos. We found the XEN layer formed a laminin-containing basement membrane where the XEN cells met the ES- and TS-derived compartments (Supplementary Fig.3a). The localization of E-Cadherin and Podocalyxin indicated that this basal membrane was functional since it induced polarisation and lumenogenesis in ES- and TS-derived compartments (Supplementary Fig.3b-d) and basement membrane was not detectable between these compartments at day 5 (Supplementary Fig.3e, f). Thus, by providing components of the basal membrane to enable ES- and TS-cell maturation, the XEN layer appears to fulfil one role of the VE.

Regionalisation of the XEN-cell layer

During embryogenesis, VE undergoes functional regionalisation. The VE overlying the embryonic compartment develops as a heterogeneous mix of squamous and cuboidal cells at E5.25, which resolves at E6.75 into defined squamous VE. In contrast, the VE overlying the ExE develops predominantly cuboidal morphology. We found that as ETX embryos developed, XEN cells overlying the ES-derived compartment became squamous by day 5, whereas those overlying the TS-derived compartment remained cuboidal (Fig.2a), thus resembling VE compartmentalisation in natural embryos (Fig.2b). The proportions of cells between the embryonic and extra-embryonic XEN layers in ETX embryos were similar to the distribution in the VE of natural embryos (Fig. 2c). Thus, cellular morphogenesis of the ETX embryo XEN layer resembles natural VE.

Further regionalisation of VE leads to the formation of AVE, which secretes Lefty1, DKK1 and Cer1 to antagonise posteriorizing morphogens and establish an anterior-posterior gradient of Nodal within the epiblast¹⁷. To assess Nodal patterning as the ETX embryos developed,

we built them using ES cells expressing the Nodal “highly bound element” HBE-YFP reporter¹⁸. At 4 days, the ETX embryos showed uniform Nodal expression throughout the ES-derived compartment (Fig.3a, Supplementary Fig.4a-c; 90%, n=18/20), becoming asymmetric and higher on the presumptive posterior side at day 5 (Fig.3b, Supplementary Fig. 4a-c; 80%, n=16/20).

To assess whether this could be a response to antagonism by a developing AVE-like domain, we assessed expression of *Lefty1*¹⁷ at day 4 and 5. We found that some ETX embryos showed asymmetric *Lefty1* expression in the XEN layer (Fig.3c; n=7/17), indicative of AVE-like specification. On day 4 to 5 increased *Lefty1* immuno-fluorescence became evident on the side of the XEN layer opposite *T/Brachyury* or Nodal HBE-YFP expression (Fig.3c-e). Similar regionalisation of gene expression was further assessed by *in situ* hybridisation to localise *Cer1*, *Nodal*, *T/Brachyury*, *Cripto*, *Wnt3* and *Bmp4* in day 4 and day 5 ETX embryos in comparison to natural embryos (Supplementary Fig.4d).

We next examined the localisation of *Otx2*, which also regulates AVE and the anterior-posterior axis specification¹⁹. We found *Otx2* was expressed only in XEN cells overlying the embryonic, and not the extra-embryonic compartment (Fig.3f, 42% n=20) and became progressively restricted to one side of ETX embryos from day 4 to 5, resembling *Otx2* expression in E5.75-E6.75 embryos (Fig.3f). Together these results indicate that the XEN layer in the ETX embryo undergoes regionalisation and, at least in some structures, can develop to express AVE markers on the side opposite posterior *T/Brachyury* and Nodal expression, thus mimicking AVE specification and localisation.

Cell migratory events characteristic of the EMT at gastrulation

During embryogenesis, epiblast cells expressing *T/Brachyury* undertake the EMT and ingress into the space between the epiblast and VE during gastrulation²⁰. When we analysed ETX embryos developing to day 5 we could identify bottle-shaped *T/Brachyury*-expressing cells with nuclei oriented parallel rather than perpendicular to the basal membrane within the ES-derived embryonic compartment and between the ES and XEN layers (Fig. 4a-b, Supplementary Fig. 5a). To examine if these cells indeed had characteristics of migratory mesenchyme, we micro-surgically isolated T:GFP negative and T:GFP positive cells and carried out RT-qPCR. This revealed upregulation of EMT markers *Snai1*, *Vimentin*, *Mmp9*

and *N-Cadherin* in T:GFP positive cells (Fig.4c), a transcriptional program reflecting the EMT.

To further confirm that such cells behaved as migrating mesenchyme, we next built ETX embryos using CAG:GFP ES cells and filmed their development. At the onset of T/*Brachyury* induction at day 4.5, CAG:GFP expressing ES cells on one side maintained intact, organised epithelium whereas on the opposite side, they changed their orientation at the ES/TS boundary (arrowheads, Fig.4d, Supplementary Fig.5b). We could identify cells undergoing apical constriction within the ES cell layer and transitioning through a bottle- to a rounded-mesenchymal shape (asterisks, Fig.4e, Supplementary Movie 2), supporting observations on fixed synthetic embryos and mirroring natural development²⁰. We confirmed cell polarity changes in the posterior ES-derived embryonic compartment through loss of the apical positioning of Golgi marker, GM130, as cells adopted their bottle-shaped, migratory morphology (asterisk, Fig.4f,g). Finally, newly formed mesenchymal cells also changed polarity from an apico-basal to a front-rear orientation, in a manner not due to the rounding of mitotic cells (Supplementary Fig.5c-d). Re-oriented cells in this newly formed mesoderm layer expressed N-cadherin but did not form a continuous epithelium (Supplementary Fig.5e). These observations suggest that *T/Brachyury*-expressing cells in ETX embryos can undertake migratory events typical of the EMT.

To achieve ingression during the EMT, cells have to breach the basement membrane^{20,21}. To determine if this occurs in ETX embryos, we analysed the major component of basement membrane, laminin. We found laminin was disrupted in the region of *T/Brachyury* expression (Fig. 4h, Supplementary Fig. 5f,g). To determine whether this disruption depended upon Wnt signalling, as is the case during natural embryogenesis, we subjected ETX embryos to the Wnt antagonist, DKK1. This prevented *T/Brachyury* expression and basement membrane disruption (Fig. 4i. Supplementary Fig. 5h, 70% versus 25%, $P < 0.05$ Fisher's exact test, n=20 ETX embryos per group). Thus, the ETX embryos can undertake the migratory events of the EMT and therefore gastrulate.

Axial mesoderm and definitive endoderm development

Since in natural development, the EMT and mesoderm specification are followed by mesoderm patterning, we next determined whether mesoderm in gastrulating embryo structures could develop to express axial mesodermal markers. We found that *Foxa2*, a

transcription factor also expressed later in definitive endoderm^{22,23}, was expressed in XEN cells and a cluster of intermediate posterior ES cells at day 5, resembling early axial mesoderm induction 6.5 days after fertilization (Fig.5a, middle). Cells expressing both *Foxa2* and *T/Brachyury* were at the most distal part of the posterior domain at day 6 (Fig.5a-d, Supplementary Fig.6a, b). This expression pattern is typical of cells fated to become axial mesoderm, confined to the anterior-most primitive streak in E7.0 embryos²⁴ (Fig.5a-d, Supplementary Fig.6a, b) and which will then express genes identifying the notochord and node²⁵. Thus, progression of mesoderm development in gastrulating embryo structures appears similar to natural embryos.

In natural embryos, definitive endoderm cells emerging from the anterior primitive streak become incorporated into the VE, which is progressively displaced anteriorly and proximally¹⁵. We found that similar events can be achieved in gastrulating ETX embryos as we observed a subset of *Foxa2* and *Sox17*-expressing cells between the ES and XEN compartments and contacting the XEN layer at day 5 (Fig.6a). To substantiate these observations, we built ETX embryos with EGFP-reporter XEN-cells. Strikingly, as development progressed, GFP-negative areas appeared in the continuous layer of GFP-positive XEN cells at day 6, as egressing cells from the ES-derived compartment were accommodated there (Fig.6b, second row). Accordingly, cells in these GFP-negative areas expressed the definitive endoderm marker *Foxa2* (Fig.6b, second and third row). By the end of day 6, there were few remaining GFP-positive cells suggesting ES-derived definitive endoderm cells had largely replaced the XEN-derived VE-like layer (Fig.6b, fourth row). To confirm this further, we built ETX embryos using XEN:GFP cells and membrane-marked CAG:GFP or mTmG ES cells (Fig.6c,d). Again, we found egressing *Sox17* positive mTmG cells that contacted the XEN layer at day 5 and intercalated into the XEN layer at day 6 (Fig.6d). These results demonstrate that gastrulating ETX embryos develop not only mesoderm but also definitive endoderm.

Gastrulating embryo structures resemble E7.0 mid-gastrula embryos in gene expression

Finally, we sought to compare gene expression patterns between gastrulating ETX embryos and natural embryos. We therefore built ETX embryos from *T/Brachyury-GFP*-expressing ES cells and dissected at day 5 their individual non-fluorescent prospective anterior parts (side 1) and fluorescent prospective posterior parts (side 2) to isolate mRNA for deep RNA-seq (Methods). Principal Component Analysis indicated close clustering of side 2 and more

loose clustering, or higher variability, in side 1 (Fig.7a, Supplementary Fig.7a-c), indicating distinct patterning between the two sides.

We identified 499 genes upregulated on side 2 compared to side 1 and 239 upregulated on side 1 compared to side 2 with good agreement between biological replicates (Fig.7b-c; Supplementary Fig.7a-c). GO analysis of genes upregulated on side 1 identified terms associated with anterior embryonic development, including neural precursor and nervous system development, whereas side 2 analysis identified terms relevant to posterior development including gastrulation, node formation, cilium assembly and left-right symmetry (Fig.7c, Supplementary Fig. 7c-e). Moreover, genes differentially expressed on side 2 included posterior fate determinants such as Nodal, Wnt, Eomes, Mixl1 and Nanog (Fig. 7b, c). Thus side 1 of gastrulating ETX embryos had an anterior, and side 2, a posterior gene-expression signature.

To further assess developmental relationships of gastrulating ETX embryos and the anterior/posterior parts of natural embryos, we performed a Neighbour-joining tree analysis²⁶ to generate a gene signature for different sample domains using differentially expressed genes to compute topological distances and a distance matrix. These signatures allowed us to group ETX embryos and epiblast of natural embryos at E6.5, E7.0 and E7.5 into a phylogenetic tree whose branching indicates samples most similar in gene expression (Fig. 7d). All side 2 samples of ETX embryos clustered closely with the posterior part of the E7.0 embryo and side 1 samples with the anterior part of E7.5 embryo, supporting the idea that axis patterning of gastrulating ETX embryos occurs as in natural embryos (Fig.7d).

Having identified such an anterior signature for side 1 and posterior signature for side 2, we next wished to compare the side 1 and 2 transcriptomes to spatial domains of gene expression in epiblast of embryos. To this end, we calculated the correlation between gene expression levels from side 1 or side 2 samples from three gastrulating ETX embryos (A, B, C; Fig. 7e) and particular epiblast regions at E6.5, E7.0 and E7.5, visualising the outcome on 'corn plots'²⁷. Side 2 samples of gastrulating ETX embryos showed few regionalised differences with E6.5 embryos but strongly resembled the posterior of gastrulating embryos at E7.0 (Fig. 7e). This correlation of posterior gene expression between side 2 of ETX embryos and E7.5 embryos was not evident. Side 1 samples of ETX embryos showed very weak correlation in gene expression pattern with the anterior epiblast of E6.5 embryos and an increasingly

striking correlation with the anterior subdomains of the epiblast at E7.0 and E7.5 (Fig. 7e). Thus, the transcriptomes of sides 1 and 2 of ETX embryos correlate with the respective anterior and posterior epiblast of gastrulating embryos, supporting the conclusion that ETX embryos develop to specify regionalised domains characteristic of natural gastrulation.

DISCUSSION

Here we describe the ability of the stem cell types derived from the three component tissues of the mouse embryo to collaborate to form structures that morphologically and transcriptionally resemble embryos developing to exhibit key landmarks of gastrulation, including mesoderm and definitive endoderm formation. In this model system, we have made no attempt to direct stem cells to co-operate but allowed them to self-sort and self-organise. Thus, it is remarkable that although not all, many of the resulting multicellular aggregates develop to resemble gastrulating natural embryos in both morphology and gene expression. We conclude that the ability for such a large proportion of ETX embryos to form spontaneously reflects the robust self-organising properties of mammalian development.

Some features of the gastrulating ETX embryo model we describe here were also seen in our previous ET embryo model comprising only ES and TS cells³. As with ET embryos, ETX embryos form distinct, ES-derived and TS-derived compartments that undertake lumenogenesis and then form a common cavity resembling the pro-amniotic cavity of natural pre-gastrulating embryos. They then take the next step and specify mesoderm as shown by T/Brachyury expression in the embryonic-like compartment abutting the extra-embryonic boundary (Fig.7f). Moreover, both synthetic ET and ETX embryos initiate PGC formation in the proximal-most part of this T/Brachyury domain. However, whereas development of ET embryos requires exogenous extra-cellular matrix, as we previously showed³, the XEN-derived VE-like layer in the ETX embryos overcomes this requirement permitting them to develop further (Fig.7f). As synthetic embryogenesis progresses, this enveloping VE-like layer becomes regionalised, as in natural development and, within the squamous cells covering the embryonic-like compartment of some ETX embryos, we could observe an emerging group of cells expressing anterior signalling centre markers. The VE-like layer also permits the orderly morphogenetic events of gastrulation evident from the onset of EMT-associated gene expression as well as the shape changes and migration of T/Brachyury expressing cells to create a cell layer between the ES-derived embryonic compartment and VE-like cells. These T/Brachyury-expressing cells develop to express axial mesoderm

markers and to migrate outwards to intercalate into the VE-like layer to form cells resembling definitive endoderm. Our results also show that at this stage, cells from the anterior and posterior parts of the gastrulating ETX embryos exhibit gene expression characteristic of natural gastrulating embryos at E7.0/E7.5 of development.

The gastrulation of ETX embryos is not perfect. We do not, for example, observe formation of mesodermal wings as during natural gastrulation. Cell ingress into the mesodermal layer is also not uniform along the T/Brachyury-positive domain. Thus, although the ETX embryos are gastrulating, the final structures are incomplete. However, it is important to note that even natural embryos do not gastrulate *in vitro* as perfectly as embryos developing *in vivo*. Nevertheless, despite this current limitation, our results show that the principal morphogenetic events of the early/mid stages of gastrulation *in vivo* can occur in the correct spatio-temporal sequence in ETX embryos established from three stem cell types *in vitro*.

In conclusion, the results we present here demonstrate the remarkable self-organising ability of the three distinct stem cell types for the three founding tissues to build embryo structures that recapitulate not only embryo morphogenesis and gene expression but also, uniquely, the gastrulation cell movements leading to axial mesoderm and definitive endoderm specification.

Author contributions: B.S, G.A, A.C with the help of S.C carried out experiments and data analysis. R.W and N.J analysed the sequencing data. E.N and G.M contributed to the stem cell derivation and the experimental design. L.C. prepared cDNA libraries. T.V. supervised the cDNA library preparation. D.M.G co-supervised parts of the study. M.Z.G conceived, supervised the study and wrote the paper with the help of B.S, G.A and D.M.G.

Acknowledgments

We are grateful to our colleagues in M.Z.G lab for their insightful comments. The M.Z.G lab is supported by grants from the European Research Council (669198) and the Wellcome Trust (098287/Z/12/Z). B.S. is also supported by Akdeniz University, Turkey. T.V and L.C are funded by Wellcome, T.V is also funded by the University of Leuven, Belgium (PFV/10/016). We thank A. Weberling, M. Mole, N. Christodoulou, C. Kyprianou and J. Guo

for their help; A. Hupalowska for inspiration for a model in Fig. 7f; L. Wittler, I. Urban, A. Landsberger, C. Schick and H. Schlenger for technical support.

References

- 1 Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032-1044, doi:10.1016/j.cell.2014.01.023 (2014).
- 2 Shahbazi, M. N. *et al.* Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239-243, doi:10.1038/nature24675 (2017).
- 3 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* **356**, doi:10.1126/science.aal1810 (2017).
- 4 Rivera-Perez, J. A. & Magnuson, T. Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. *Dev Biol* **288**, 363-371, doi:10.1016/j.ydbio.2005.09.012 (2005).
- 5 Robertson, E. J. Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Semin Cell Dev Biol* **32**, 73-79, doi:10.1016/j.semdb.2014.03.028 (2014).
- 6 Wood, S. A., Allen, N. D., Rossant, J., Auerbach, A. & Nagy, A. Non-injection methods for the production of embryonic stem cell-embryo chimaeras. *Nature* **365**, 87-89, doi:10.1038/365087a0 (1993).
- 7 Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-256 (1984).
- 8 ten Berge, D. *et al.* Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* **3**, 508-518, doi:10.1016/j.stem.2008.09.013 (2008).
- 9 van den Brink, S. C. *et al.* Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231-4242, doi:10.1242/dev.113001 (2014).
- 10 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* **11**, 847-854, doi:10.1038/nmeth.3016 (2014).
- 11 Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A. & Rossant, J. Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072-2075 (1998).
- 12 Kimura-Yoshida, C. *et al.* Canonical Wnt signaling and its antagonist regulate anterior-posterior axis polarization by guiding cell migration in mouse visceral endoderm. *Dev Cell* **9**, 639-650, doi:10.1016/j.devcel.2005.09.011 (2005).
- 13 Yamamoto, M. *et al.* Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. *Nature* **428**, 387-392, doi:10.1038/nature02418 (2004).
- 14 Tam, P. P. & Beddington, R. S. Establishment and organization of germ layers in the gastrulating mouse embryo. *Ciba Found Symp* **165**, 27-41; discussion 42-29 (1992).
- 15 Viotti, M., Nowotschin, S. & Hadjantonakis, A. K. Afp::mCherry, a red fluorescent transgenic reporter of the mouse visceral endoderm. *Genesis* **49**, 124-133, doi:10.1002/dvg.20695 (2011).
- 16 Kunath, T. *et al.* Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649-1661, doi:10.1242/dev.01715 (2005).

- 17 Takaoka, K., Yamamoto, M. & Hamada, H. Origin and role of distal visceral endoderm, a group of cells that determines anterior-posterior polarity of the mouse embryo. *Nat Cell Biol* **13**, 743-752, doi:10.1038/ncb2251 (2011).
- 18 Papanayotou, C. *et al.* A novel nodal enhancer dependent on pluripotency factors and smad2/3 signaling conditions a regulatory switch during epiblast maturation. *PLoS Biol* **12**, e1001890, doi:10.1371/journal.pbio.1001890 (2014).
- 19 Hoshino, H., Shioi, G. & Aizawa, S. AVE protein expression and visceral endoderm cell behavior during anterior-posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. *Dev Biol* **402**, 175-191, doi:10.1016/j.ydbio.2015.03.023 (2015).
- 20 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* **241**, 270-283, doi:10.1002/dvdy.23711 (2012).
- 21 Laurie, G. W., Leblond, C. P. & Martin, G. R. Localization of type IV collagen, laminin, heparan sulfate proteoglycan, and fibronectin to the basal lamina of basement membranes. *J Cell Biol* **95**, 340-344 (1982).
- 22 Sasaki, H. & Hogan, B. L. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development* **118**, 47-59 (1993).
- 23 Lewis, S. L. & Tam, P. P. Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Dev Dyn* **235**, 2315-2329, doi:10.1002/dvdy.20846 (2006).
- 24 Burtscher, I. & Lickert, H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029-1038, doi:10.1242/dev.028415 (2009).
- 25 Balmer, S., Nowotschin, S. & Hadjantonakis, A. K. Notochord morphogenesis in mice: Current understanding & open questions. *Dev Dyn* **245**, 547-557, doi:10.1002/dvdy.24392 (2016).
- 26 Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-425, doi:10.1093/oxfordjournals.molbev.a040454 (1987).
- 27 Peng, G. *et al.* Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo. *Dev Cell* **36**, 681-697, doi:10.1016/j.devcel.2016.02.020 (2016).

FIGURE LEGENDS

Figure 1. Self-assembly of embryo structures. **a.** Schematic of an AggreWell with ES, TS and XEN cells on Day-0 and representative examples of ETX embryos stained for lineage markers after progressive times in culture. 10 experiments. Bar=20 μ m. **b.** Representative examples of ETX embryos generated from a typical experiment after 4 days demonstrate 3 distinct compartments. Lineage markers; Oct4, red (embryonic); Tfap2c, blue and Gata6, grey (extra-embryonic). 10 experiments. Bar=20 μ m. **c.** Frequency of structures comprising 3 stem cell types (ES, TS, and XEN cells) compared to other combinations of stem cell types. A total of 435 structures was scored from 3 separate experiments. Pie chart: after 4-days, 69.2% of structures comprised 3 lineages; 24.7%, 2 lineages; and rarely a single cell type (1.87%, ES; 2.9%, TS; and 1.1%, XEN cells). Bar chart: Of structures with 3 lineages, 28.6% were cylindrical with single adjoining ES and TS cell compartments surrounded by a XEN layer corresponding to ETX embryos (left-most image); the remaining twin structures had either 2 TS (middle image) or ES (right-most image) compartments at polar positions. ESC, Embryonic Stem Cell; TSC, Trophoblast Stem Cell; XEN, Extra-embryonic Endoderm stem cells. **d, e.** Immunostaining of ETX embryos after 4- (d, top, 15 embryos) or 5-days (e, top, 10 embryos) of culture compared to natural embryo cultured for 48h from blastocyst stage equivalent to E5.5 (d, bottom, 10 embryos) or E6.0 embryo (e, bottom, 8 embryos) collected in 3 experiments. Yellow dashed line outlines embryonic/extra-embryonic border; white dashed line demarcates VE/XEN layer. Non-nuclear, unspecific anti-Oct4 VE fluorescence was removed from E6.0 embryo to enable visualisation of the Gata4 positive VE cells. Bar=20 μ m. **f.** Quantification of mean embryonic and extra-embryonic cells in day-4 ETX embryos and cultured embryos equivalent to E5.5. Two-sided Student's *t*-test, $n=10$ ETX embryos per group, 3 experiments, not significant. Columns are means \pm SEM.

Figure 2. XEN cell layer in embryo structures. **a.** ETX embryos after 3 (representative of 9 embryos), 4 (8 embryos), 5 (17 embryos) days of culture stained for F-actin to reveal shape of XEN cells. Middle: Surface rendering of ETX embryo showing embryonic (em) compartment, red; and extra-embryonic (ex) compartment, blue. Top-right: Boundary (dashed lines) of em and ex XEN layer based on boundary of Oct4 immunofluorescence. Bottom-right: Tracing of em and ex XEN layer of ETX embryos based on F-actin immunofluorescence. Bottom: magnified fields showing em (left) or ex (right) XEN cells to highlight the cell shape. Scatter plots of cell aspect ratios in XEN layer. Mean cell aspect ratio is significantly different between ex and emXEN compartments of ETX embryos at days -4 and -5. Each dot represents a cell. Value less than 1 indicates cells with predominant cuboidal morphology; value greater than 1, cells with prevalent squamous morphology. Two-sided Student's *t*-test with Welch's correction. Day-3, $n=102$ emXEN cells, $n=95$ exXEN cells; day-4, $n=267$ emXEN cells, $n=181$ exXEN cells; day-5, $n=318$ emXEN cells, $n=191$ exXEN cells. Plots show means \pm SEM (black lines). Bar=20 μ m. **b.** Natural embryos at E5.25 (14 embryos), E5.5 (16 embryos), E6.75 (5 embryos) stained for F-actin to reveal cell shape in VE layer. Right: Surface rendering of embryo showing embryonic (em) compartment, red; extra-embryonic (exe) compartment, blue. Bottom: magnified fields showing em (left) or ex (right) VE cells to highlight the cell shape. Scatter plots of cell aspect ratios in VE overlying ex and em compartments with time. Each dot represents a cell. Value less than 1, cuboidal cells; value greater than 1, squamous cells. Two-sided Student's *t*-test with Welch's correction. E5.25, $n=509$ emVE cells, $n=500$ exVE cells; E5.5, $n=508$ emVE cells, $n=538$ exVE cells; E6.75, $n=505$ emVE cells, $n=502$ exVE cells. Plots show means \pm SEM (black lines). Bar=20 μ m. **c.** Comparison of proportions of cells in em and ex XEN layer

in ETX embryos ($n=7$ day-4, $n=12$ day-5 ETX embryos) versus VE layer of natural embryos ($n=3$ E5.5, $n=3$ E6.75 embryos). Two-sided Student's t-test. Columns are means \pm SEM.

Figure 3. Regionalisation within XEN cell layer in ETX embryos. a,b. Representative ETX embryos built from Nodal HBE-YFP ES cells cultured for 4 or 5 days. Orthogonal XZ and YZ views are shown. 20 ETX embryos/group, 3 experiments. Bar= $20\mu\text{m}$. **c.** Lefty1 immunostaining of E5.75 (top-left) and E6.75 (top-right) natural embryos, and day-4 (bottom-left) and day-5 (bottom-right) ETX embryos. Yellow dashed line demarcates embryonic VE/XEN layer; white boxes, magnified region; yellow arrowheads, Lefty1 positive VE/XEN cells; white dashed lines outline border of VE or XEN layer. Epiblast or ES-compartment occasionally displayed low Lefty1 expression. Non-nuclear anti-Oct4 staining of VE represents non-specific binding. 7/7 natural embryos; 7/17 ETX embryos with asymmetric Lefty1, 3 experiments. Bar= $20\mu\text{m}$. Bars in zoomed images= $10\mu\text{m}$. **d.** Representative ETX embryos generated from Nodal HBE-YFP ES cells with uniform Nodal expression in ES-compartment and no Lefty1 expression (top) or induced Lefty1 expression in XEN layer (arrowhead, middle) and Nodal gradient in ES-compartment with Lefty1 expression on opposite side of XEN layer (bottom). Nodal HBE-YFP and Lefty1 signal pseudocoloured with 'fire' lookup table in Fiji to highlight intensity gradients. White dashed line demarcates embryonic VE/XEN layer. Boxed insets show embryonic compartment as a gradient map. L, low; H, high. 15 ETX embryos, 3 experiments. **e.** Quantification of asymmetric Lefty1 immunofluorescence intensity in XEN or VE layer in representative embryos shown in c. Representative image of an ETX embryo (right) indicates line of intensity profiles along XEN layer. White line indicates midline of the structure, dashed lines outline the ES-derived compartment. Box plots represent $n=9$ Z-planes for day-4, $n=7$ Z-planes for day-5 ETX embryos; $n=6$ Z-planes for E5.5, $n=10$ Z-planes for E6.75 embryos (see methods). Two-sided Student's t-test. Box plots means \pm SEM. Bar= $20\mu\text{m}$. **f.** 42.5% (12/28) of ETX embryos displayed Otx2 induction in emXEN region. 'Fire' Pseudocolour lookup table in Fiji used to highlight intensity gradients. L, low; H, high. Yellow dashed lines outline the emXEN/VE region. White lines indicate the plane used to plot intensity profiles (on right). Purple dashed lines outline TS cell-derived extra-embryonic compartment or ExE for clarity. Bar= $20\mu\text{m}$; 20 ETX embryos, 5 experiments; 10 natural embryos, 2 experiments.

Figure 4. Epithelial-to-mesenchymal transition events in gastrulating ETX embryos. a. *T/Brachyury* expression in nascent mesoderm. Dashed line outlines the ROI, ES-compartment. Magnified fields: (1) bottle-shaped *T/Brachyury* positive cell (dashed lines), (2) re-oriented *T/Brachyury* positive mesenchymal cells between ES/XEN layers (dashed lines). 10 ETX embryos, 4 experiments. Bar= $20\mu\text{m}$. **b.** Boxed area (3) from panel (a), posterior side of ES-compartment magnified and rotated 90° . **Top:** Asterisks: re-oriented mesenchymal cells between ES/XEN layers. Bar= $20\mu\text{m}$. **Second-row:** white dashed lines: border of ES/XEN layers; purple dashed lines outline cells between layers. **Third-row:** Superimposed schematic: bottle-shaped ES cells (yellow) interspersed between epithelial cells (white). **Bottom:** Different focal plane shows repositioned T-positive cells (purple), between ES/XEN layers. **c.** RT-qPCR of candidate EMT genes performed on cell extracts of day-5 ETX embryos (see methods). Two-sided Student's t-test. Columns are means \pm SEM. $n=4$ biological replicates. **d.** Stills from time-lapse movie of live ETX embryo built from CAG:GFP ES cells. Cells on one side of ES-compartment (labelled-1, prospective anterior) maintains epithelial integrity. Cells on opposite (labelled-2, prospective posterior) re-orient at ES-TS boundary, arrowheads. Boxes, magnified fields at each time-point. 3 separate movies, 7 ETX embryos. Bar= $20\mu\text{m}$. **e. Top:** magnified stills of ES-compartment revealing a cell (asterisks) undergoing apical constriction. **Second-row:** 1-4, tracings of stills shown above.

Third/fourth rows: tracking of same cell illustrating epithelial-bottle-mesenchymal shape change between 15^h00^m (2), 16^h30^m (3) and 17^h00^m (4) time-points. Bar=5 μ m. **f.** Polarised Golgi in anterior indicating epithelial integrity (arrowheads); asymmetrically distributed Golgi within mesenchymal cells on posterior. Dashed lines: embryonic compartment. 12 ETX embryos, 3 E6.75 natural embryos; 4 experiments. Bar in zoomed images=10 μ m. **g.** Magnified images from posterior shown in f. Migratory bottle-cells losing epithelial polarity (asterisks, left and middle; arrowheads, right) Bar=5 μ m. **h.** Laminin break-down on *T/Brachyury*-expressing posterior. White boxes, magnified regions. Purple dashed lines: TS-derived extra-embryonic compartment. Rightmost panels, laminin staining inverted for better contrast of magnified regions. Purple arrows: laminin breakdown. Non-nuclear anti-*T/Brachyury* VE fluorescence represents non-specific staining. Bar=20 μ m; 15 ETX embryos, 4 experiments; 6 E6.75 embryos, 2 experiments. **i.** Day-5 ETX embryo with or without DKK1-treatment (200ng/ml) for 24h. Right-most: maximum projections of ES-compartments, pseudocoloured to highlight intensity gradients. Bar=20 μ m. 20 ETX embryos per group, 2 experiments.

Figure 5. Spatio-temporal specification of axial mesoderm in gastrulating ETX embryos. **a.** Top: mid-sagittal confocal sections of pre-streak (PS) E6.0, early streak (ES) E6.5, and mid-streak (MS) E7.0 stage embryo. Bottom: Mid-sagittal confocal sections of ETX embryos at days-4, -5 and -6. Immunostaining of Foxa2 reveals axial mesoderm specification. White boxes indicate region of magnified inset. Purple dashed lines outline ExE or TS cell-derived extra-embryonic compartment for clarity. Bar=20 μ m; 20 ETX embryos, 3 experiments; 15 embryos, 2 experiments. **b.** E7.0 embryo and day-6 gastrulating ETX embryo presented in a. Left: maximum projection images show *T*:GFP positive posterior domain expanding from the embryonic/extra-embryonic boundary (yellow dashed lines) to distal tip of embryonic compartment. White dashed lines outline VE or XEN layer. Right: YZ orthogonal views provided to visualise posterior domain. White arrowheads indicate *T*:GFP positive cells on the boundary; yellow arrows, *T*:GFP-Foxa2 double positive mesoderm cells marking axial mesoderm induction on intermediate posterior domain. 10 natural embryos, 3 experiments; 10 ETX embryos, 3 experiments. **c.** Single plane images: white boxes indicate region of magnified insets showing Foxa2-*T* double positive cells. White lines indicate two separate planes (marked 1 and 2) used to plot intensity profiles of *T*:GFP and Foxa2 fluorescence within epiblast/ES-compartment shown in **d.** Pr: Proximal, D: Distal, A: Anterior, P: Posterior. 10 natural embryos, 3 experiments; 10 ETX embryos, 3 experiments. Bar=20 μ m.

Figure 6. Spatio-temporal formation of definitive endoderm in gastrulating ETX embryos. **a.** Magnified fields (white boxes) indicate subset of Foxa2-Sox17 double positive definitive endoderm cells (yellow arrows) in the process of egression within epiblast/ES-cell layer and making contact with embryonic emVE/emXEN layer (white dotted lines). Purple dashed lines outline ExE- or TS-derived extra-embryonic compartment for clarity. Non-nuclear anti-Oct4 VE fluorescence represents non-specific binding. Bar=20 μ m; 4 synthetic embryos; 3 embryos, 2 experiments. **b.** XEN:EGFP reporter ETX embryos at day-5 (10 embryos), day-6 (10) and mid/late day-6 (6) showing progressive XEN cell dispersal. White boxes indicate zoomed inset. Magnified insets on second row show rotated X axis for better visualization. White dashed lines mark embryonic/extra-embryonic boundary. Lower panel: white dashed lines outline cells egressed into EGFP-positive XEN layer. Bar=20 μ m; 4 experiments. **c.** Images from live ETX embryos generated from CAG:GFP ES, EGFP XEN and wild type TS cells at indicated stages. Purple arrows, intercalated ES cells in XEN layer.

GFP-labelled embryonic compartments are inverted for better contrast of the zoomed inset (bottom). Purple dashed lines outline ExE or TS cell-derived extra-embryonic compartment for clarity. Bar=20 μ m; 10 embryos/time point. **d.** EGFP XEN and mTmG ES cell reporter ETX embryos at day-5 and -6 showing progression of ES cell intercalation in XEN layer. White boxes indicate the region of the magnified insets. Yellow arrows show Sox17 upregulation in presumptive definitive endoderm cells within mTmG labelled ES cell compartment in process of egression (upper) or having egressed (lower) into XEN:EGFP layer. 3 embryos/time point, 2 experiments. Bar for all zoomed images=10 μ m.

Figure 7. Transcriptional profiling of gastrulating ETX embryos reveals global similarity of anterior-posterior patterning to gastrulating natural embryos. **a.** Principal Component Analysis (PCA) of the transcriptome from each side (Side1-prospective anterior vs Side2-prospective posterior) of each gastrulating ETX embryo (GE-A, GE-B and GE-C). $n=3$ biological replicates, divided in Side1 and Side2. **b.** Volcano plot showing differentially expressed genes (DEGs) in Side1 and Side2 of gastrulating ETX embryos. Plotted genes have a p -value <0.05 and a fold change >1.5 . Red dots, selected genes required for posterior development. Green dots, genes required for anterior development. $n=3$ biological replicates, divided in side1 and side2. **c.** Heatmap showing expression of 45 selected genes in each side of each ETX embryos. Highest expression, dark red; lowest, dark green. **d.** Neighbour-joining (NJ) tree inferring phylogenetic trajectories of side 1 and 2 transcriptomes of gastrulating ETX embryos and anterior and posterior parts of natural embryos (see methods). Green dashed lines outline clustered samples showing anterior identity, red dashed lines outline clustered samples showing posterior identity. **e.** Corn plots representing the Pearson Correlation Coefficient (PCC) between transcriptomes from each side of gastrulating ETX embryos ($n=3$ biological replicates, divided in side1 and side2) and spatial domains of natural embryo epiblast at different developmental timepoints ($n=13$ E6.5, $n=42$ E7.0, $n=46$ E7.5). Stage of natural embryo development indicated above corn plots. Numbers on the left side of a corn plot indicate each LCM sample collected from along the proximal-distal axis of the mouse epiblast (see methods) (i.e. 1 indicates LCM sample from the distal tip of the epiblast, whereas the largest number in the series for a given corn plot indicates LCM sample from the embryonic-extraembryonic boundary). Right hand colour scale indicates magnitude of PCC. Red indicates highest correlation; grey, no correlation. Transverse sections of natural embryos were collected along the proximal-distal axis and divided into subdomains. A=anterior, P=posterior, L=left, R=right. **f.** Schematic representation of our findings comparing development of a gastrulating ETX embryo (bottom) with a natural embryo (top). EMT: epithelial-to-mesenchymal transition; DE: definitive endoderm; Ax: axial (mesoderm); PAC: pro-amniotic cavity.













