

This is the peer reviewed version of the following article:

Menchero, S., Sainz de Aja, J., & Manzanares, M. (2018). Our First Choice: Cellular and Genetic Underpinnings of Trophectoderm Identity and Differentiation in the Mammalian Embryo. *Current Topics in Developmental Biology, 128*, 59-80. doi:10.1016/bs.ctdb.2017.10.009

which has been published in final form at: <u>https://doi.org/10.1016/bs.ctdb.2017.10.009</u>

Our first choice: cellular and genetic underpinnings of trophectoderm identity and differentiation in the mammalian embryo

Sergio Menchero, Julio G. Sainz de Aja and Miguel Manzanares*

Centro Nacional de Investigaciones Cardiovasculares (CNIC), Melchor Fernandez Almagro 3, 28029 Madrid, Spain.

* corresponding author (mmanzanares@cnic.es)

Abstract

The trophectoderm is the first cell population to appear in the mammalian preimplantation embryo, as the result of the differentiation of totipotent blastomeres located on the outer surface of the late morula. Trophectodermal cells arrange in a monolayer covering the expanding blastocyst and acquire an epithelial phenotype with distinct apico-basal polarity and a basal lamina placed toward the blastocyst interior. During later development through the periimplantation and gastrulation stages, the trophectoderm gives rise to extraembryonic membranes and cell types that will eventually form most of the fetal placenta, the specialized organ through which the embryo obtains maternal nourishment necessary for subsequent exponential growth. The specification of the trophectoderm is controlled by the combination of morphological cues arising from cell polarity with differential activity of signaling pathways such as Hippo and Notch, and the restriction to outer cells of lineage-specifiers such as CDX2. This is possibly the first symmetry-breaking decision undertaken by the uncommitted cells produced by a handful of mitosis divisions from the newly fertilized zygote. Understanding how this cell lineage is specified will therefore provide unique information about development, differentiation, and how the interplay between cellular morphology and signaling and regulatory factors results in a correctly 3Dpatterned embryo.

Keywords: trophectoderm, blastocyst, preimplantation development, stem cells, placenta, Cdx2, Hippo, Notch.

During development, a complete organism is built from a unicellular zygote. This single cell must generate not only the embryo itself, but also the extraembryonic structures. The totipotent capacity of the zygote is gradually lost as the cell populations in the embryo differentiate toward committed fates. In mammals, the first symmetry-breaking event occurs during preimplantation development and establishes two cell lineages: the trophectoderm (TE) and the inner cell mass (ICM). The TE gives rise to extraembryonic lineages, such as the extraembryonic ectoderm (ExE) and the ectoplacental cone (EPC), which will later differentiate into the major portion of the fetal placenta. In contrast, the ICM gives rise to the embryo proper, the yolk sac, and the allantois.

Heterogeneity emergence in the mouse embryo

The mouse zygote undergoes a series of divisions that increases the number of cells without altering the size of the embryo. These cells, called blastomeres, are initially equal, and each one can potentially develop into a full embryo (Suwinska et al., 2008; Tarkowski, 1959; Tarkowski and Wroblewska, 1967). Heterogeneities among these blastomeres appear gradually, raising questions about the extent to which these emerging differences bias cells toward a particular lineage contribution or reveal a certain variability within an apparently equivalent population. Thus although the two first distinct cell populations are clearly distinguishable at the blastocyst stage, at embryonic (E) day 3.5, the key to understanding how these cells commit to a specific lineage must be sought earlier.

Recent research into gene expression differences between individual cells of 4cell embryos suggests that molecular heterogeneities between blastomeres

appear some time before the first lineage decision is established, and that differential gene expression at this stage can bias the contribution of descendant cells to one lineage or another (Goolam et al., 2016; Torres-Padilla et al., 2007). However, at these stages cell totipotency is undiminished. Rather, it appears that blastomeres acquire tendencies or preferences to contribute to a particular lineage, while at this stage retaining their capacity to give rise to every population in the embryo (Kelly, 1977; Tabansky et al., 2013; Tarkowski and Wroblewska, 1967). The first morphological evidence of differentiation is seen at the morula stage. After three rounds of division, blastomeres increase their intercellular interactions, and cells on the surface acquire an apical-basal polarity, with the apical domain facing outwards (Ziomek and Johnson, 1980). This process, known as compaction, is driven at the 8-cell stage by periodic cortical waves controlled by the actomyosin cortex in cooperation with CDH1 (E-cadherin) (Maitre et al., 2015). Moreover, cell shape changes during compaction have been found to depend on CDH1-dependent filopodia (Fierro-Gonzalez et al., 2013). The increasing number of blastomeres generated during subsequent divisions forces the cells to rearrange their positions, resulting in the emergence of two distinct cell populations: an outer layer of polarized cells enclosing an inner group of apolar cells. The acquisition of polarity and the subsequent cell divisions have been the focus of the two main hypotheses about how the TE and the ICM are segregated. The inside-outside model proposes that the position of a cell in the morula exposes it to specific conditions that are the key to deciding its fate. Thus, cells on the outside of the morula give rise to the TE and the cells on the inside give rise to ICM (Tarkowski and Wroblewska, 1967). Supporting this model, when blastomeres are placed on the outside of another embryo, those cells contribute

predominantly to the TE (Hillman et al., 1972). In contrast, the polarization model proposes that differentiation is established according to the inheritance of an apical pole in the blastomeres. When a polar cell divides symmetrically, both daughter cells inherit the apical-basal polarity and remain in the outer population. However, if a polarized cell divides asymmetrically, only one daughter cell inherits the apical pole, and the non-polarized daughter cell will form part of the inner population (Johnson and Ziomek, 1981).

Polarization and inner versus outer locations are still considered the basis of the first lineage choice. However, several recent studies have shown that the division angle and initial position does not robustly predict blastomere fate, because substantial cell relocations force blastomeres to internalize prior to the blastocyst configuration (Anani et al., 2014; McDole et al., 2011; Toyooka et al., 2016; Watanabe et al., 2014). Actomyosin cortical networks subject the outer cells to heterogeneous tensile forces, while myosin II accumulates around constricting cells that will become embedded in the inner population (Samarage et al., 2015). Polarized and unpolarized blastomeres have differing contractilities, with unpolarized cells having more cortical myosin and a higher contraction amplitude, triggering their internalization (Maitre et al., 2016).

Since the publication of the first models of TE and ICM segregation, several studies have confirmed the importance of apical-basal polarity in outer cells (Alarcon, 2010; Johnson and Ziomek, 1981; Kondratiuk et al., 2012; Plusa et al., 2005). Interestingly, a recent report showed that the apical domain is sufficient to promote TE fate; the apical domains of polarized 8-cell blastomeres induce asymmetric division when transplanted into apolar 8-cell blastomeres, with the

the daughter cell inheriting the apical domain differentiating toward TE (Korotkevich et al., 2017).

These studies make it clear that the preimplantation embryo is a self-organizing system that is able to interpret morphogenetic cues in order to exit the totipotent state while taking lineage decisions.

The trophectoderm gene regulatory network

The establishment of the first cell populations in the mouse embryo is underpinned by a transcriptional circuitry orchestrating the fate of every cell. The core of these networks is a set of transcription factors (TFs) that drive the commitment of cells to particular lineages. The first lineage choice is controlled by two specific programs: the pluripotency network and the trophectoderm network.

The core genes of the pluripotency network are *Oct4* (*Pou5f1*) (Nichols et al., 1998), *Sox2* (Avilion et al., 2003), and *Nanog* (Chambers et al., 2003; Mitsui et al., 2003). This small set of TFs maintains pluripotency both in vivo and in vitro, and at the same time represses differentiation and lineage commitment. Accordingly, the core pluripotency network actively represses the expression of TE-specific genes (Niwa et al., 2005), and the absence of *Oct4* leads to increased differentiation to the trophoblast lineage both in vivo and in vitro (Nichols et al., 1998; Niwa et al., 2000).

The trophectoderm network is controlled by *Cdx2* (Strumpf et al., 2005), a mammalian homologue of the *Drosophila* homeotic gene *caudal* (Mlodzik et al., 1985) that encodes a homeodomain transcription factor. *Cdx2* was reported to

be expressed in the blastocyst TE and in the ExE at postimplantation stages (Beck et al., 1995). Analysis of $Cdx2^{-/-}$ embryos showed that they were unable to maintain the blastocoel, with suppressed expression of other TE markers and pluripotency factors such as OCT4 and NANOG detected in all cells (Strumpf et al., 2005). Another TE marker, the cytokeratin intermediate filament Krt8 (Chisholm and Houliston, 1987) is also dependent on Cdx2 (Ralston and Rossant, 2008). These results suggest that Cdx2 is necessary for the correct formation of the TE and its differentiation, and place Cdx2 at the core of the regulatory network (Fig 1).

The zinc-finger coding gene Gata3 also regulates trophoblast identity (Home et al., 2009; Ralston et al., 2010) (Fig 1). Gata3 is coexpressed with Cdx2 at preimplantation stages, but at postimplantation stages its expression is higher in the EPC than in the ExE (Ralston et al., 2010). Gata2 was recently found to be coexpressed with Gata3, and deletion of both genes impairs the expression of the TE-related genes Cdx2, Eomes, and Elf5 (Home et al., 2017). The T-box gene *Eomes (Eomesodermin)* is also expressed in the TE at E3.5, in the polar TE at E4.5, and in the ExE at postimplantation stages (Hancock et al., 1999; Ralston and Rossant, 2008; Strumpf et al., 2005). *Eomes* is required later than Cdx2, and while *Eomes* expression is impaired in $Cdx2^{-/-}$ embryos, the reverse is not observed (Ralston and Rossant, 2008; Strumpf et al., 2005). In addition, Eomes regulates some markers of differentiated trophoblast giant cells such as Hand1 and Pl1 (Cross et al., 1995; Faria et al., 1991; Strumpf et al., 2005) (Fig 1). The TF AP-2y (TFAP2C) specifies the TE lineage upstream of Cdx2 and Elf5 (Cao et al., 2015) (Fig 1), and also regulates Pard6b, an important apical domain component required for blastocyst formation (Alarcon, 2010). Fgfr2 is expressed

in the blastocyst TE and could be involved in the maintenance of the TE and its derivatives (Haffner-Krausz et al., 1999; Tanaka et al., 1998). A recent work has shown by means of two reporter mice that Fgfr2 is expressed since the 8-cell stage and *Fafr1* is also expressed in the TE. However the only TE phenotype that Fafr1-/-: Fafr2-/- mutant show is a reduced number of TE cells but no defects in any TE marker has been reported (Kang et al., 2017; Molotkov et al., 2017). Elf5 and *Ets2* are markers of the ExE, and their deletion interferes with proper gastrulation (Donnison et al., 2005; Polydorou and Georgiades, 2013). Intriguingly, Klf5 regulates markers of both the embryonic and the extraembryonic lineages. Klf5-/blastocysts have reduced levels of Cdx2, Eomes, and Krt8 but also of Oct4 and Nanog (Lin et al., 2010) (Fig 1). The advent of next generation sequencing techniques, particularly single cell transcriptomics, has allowed the identification of the earliest molecular differences among cells, and these studies have identified new candidate members of the TE network. For example, Id2 was identified as the earliest marker of the outer population in the embryo (Guo et al., 2010), and new markers such as Dab2 or Lrp2 have been linked to the TE (Posfai et al., 2017).

The TE lineage is thus established and maintained through the progressive activation of TFs that promote the induction of other TE-related markers and their own regulation and at the same time repress the pluripotency network (Fig 1). Direct repression between CDX2 and the pluripotency factors OCT4, NANOG, and SOX2 appears to be important for maintaining lineage segregation (Huang et al., 2017). *In vitro* studies suggest that this repression is reciprocal between both lineages (Chen et al., 2009; Niwa et al., 2005).

Establishment of trophectoderm fate

Core factors such as OCT4, NANOG, SOX2, and CDX2 play a well-defined role in determining the first lineage decisions in the embryo. However, surprisingly their initial expression is stochastic among blastomeres up to the late-morula stage, after which they gradually become restricted to their definitive domains (Dietrich and Hiiragi, 2007; Guo et al., 2010; Ohnishi et al., 2014). This dynamic raises the question of how the expression of core lineage factors is controlled during these first stages of development.

Cdx2 is initially expressed at the 8-cell stage in a salt-and-pepper pattern, before becoming restricted to the TE at the blastocyst stage (Dietrich and Hiiragi, 2007; Ralston and Rossant, 2008) (Fig 2). *Cdx2* is in fact the first factor to be restricted to a specific lineage in the blastocyst. *Cdx2* lies downstream of the Hippo pathway (Fig1, 2), which was first identified in *Drosophila*, where it is involved in tissue growth control (Pan, 2010). However, in the preimplantation embryo Hippo is not involved in growth regulation but instead provides the link between polarization and cell fate. The differential distribution of pathway components in polarized and unpolarized blastomeres dictates the activity of the pathway (Manzanares and Rodriguez, 2013; Sasaki, 2015).

In apolar inner cells, where the Hippo pathway is active, the junction-associated protein AMOT (angiomotin) localizes at adherent junctions (Hirate et al., 2013; Leung and Zernicka-Goetz, 2013). The kinase LATS1/2 phosphorylates AMOT and stabilizes a complex formed by LATS1/2, AMOT, and NF2 (Hirate et al., 2013). LATS1/2 then phosphorylates the transcriptional coactivator YAP (and the Yap-related protein TAZ), which is retained in the cytoplasm. In polar outer cells,

AMOT is sequestered by the apical complex. The kinases are inactive and, consequently, YAP is not phosphorylated. Therefore, YAP is translocated to the nucleus, where it binds the transcription factor TEAD4 and triggers the expression of its target genes (Nishioka et al., 2009), including the TE lineage factors Cdx2 (Nishioka et al., 2009) and Gata3 (Ralston et al., 2010). If Hippo pathway activation is blocked due to disruption of any of its components, ectopic Cdx2 expression can be detected in inner cells (Cockburn et al., 2013; Hirate et al., 2013; Leung and Zernicka-Goetz, 2013; Nishioka et al., 2009). Most embryos lacking Tead4, the Hippo pathway effector, are unable to cavitate and to implant, and have below-normal expression of Cdx2 (Nishioka et al., 2008; Yagi et al., 2007). Tead4-1- morulae therefore have low levels of CDX2, and in Tead4-1blastocysts CDX2 is not detected (Nishioka et al., 2008; Rayon et al., 2014). Remarkably, this phenotype can be circumvented if *Tead4*^{-/-} embryos are cultured in low oxygen conditions that reduce oxidative stress (Kaneko and DePamphilis, 2013), suggesting a role for the Hippo pathway in the regulation of energy homeostasis in the blastocyst. The role of the Hippo pathway is not only to promote the TE lineage, but also to contribute to the establishment of the ICM by limiting the expression of SOX2, independently of Cdx2, in this population (Wicklow et al., 2014).

There is a good overall correlation between blastomere position in the early embryo and polarization: outer cells tend to be polarized and inner cells unpolarized. Nonetheless, since polarization happens progressively, apolar cells can be found in the outer layer of 16-cell embryos, and interestingly the Hippo pathway is active in these cells, with phosphorylated YAP maintained in the

cytoplasm (Anani et al., 2014; Hirate et al., 2015). Hippo is thus linked to polarity rather than to cell position (Fig 2A).

Hippo was the first signaling pathway identified to play a role in lineage determination in the preimplantation embryo (Nishioka et al., 2009). However, recent evidence shows that activation and maintenance of the TE program is ensured not by one pathway but rather by the combination of signals from different inputs. One of these is the Notch pathway, which cooperates with Hippo to regulate Cdx2 expression (Rayon et al., 2014) (Fig 1, 2). Notch is a cell-to-cell signaling pathway involved in many developmental programs (Koch et al., 2013). Disruption of this pathway does not affect embryo survival until postimplantation stages (Conlon et al., 1995; Shi et al., 2005; Souilhol et al., 2006), and for this reason its putative role in preimplantation stages was ruled out. However, double mutants for Tead4 and Rbpi, the transcriptional effectors of the Hippo and Notch pathways respectively, exacerbate their single phenotypes and lead to embryonic death before the blastocyst stage. Notch is active specifically in the TE, and its blockade diminishes Cdx2 levels. Furthermore, forced expression of the active domain of Notch1 (N1ICD) unbalances the TE/ICM cell ratio (Rayon et al., 2014). Both pathways act in parallel and converge on a regulatory element upstream of Cdx2, called the TE enhancer (TEE) (Rayon et al., 2014), and they do this through interaction with the chromatin remodeler SBNO1 (Strawberry Notch1) (Fig 1). Sbno1^{-/-} embryos fail to cavitate and to express the TE markers Cdx2, Eomes, and Krt8. SBNO1 physically interacts with the YAP/TEAD and NICD/RBPJ transcriptional complexes, and lack of Sbno1 prevents YAP/TEAD and NICD/RBPJ from properly inducing Cdx2 (Watanabe et al., 2017).

TE determination may also involve other signaling inputs (Menchero et al., 2017). For example, Rho/Rock signaling has been shown to mediate cell polarization and Hippo regulation. Blockade of Rho/Rock results in disruption of apical-basal polarity and ectopic activation of Hippo (Kono et al., 2014; Mihajlovic and Bruce, 2016). We can therefore conclude that the first differentiation event is driven not by one pathway but by a multitude of transcriptional and signaling players. The Hippo pathway is the main one orchestrating TE fate, but its role is supported by other players like Notch and Rho. This multiplicity may function to ensure the robustness of embryonic development in the event that any branch of the circuitry is blocked. This interplay between pathways shows the need to study combined inputs in order to unravel unsuspected roles during the establishment of the first lineages.

Plasticity during lineage segregation

Much of the research into mammalian preimplantation development has focused on finding key factors and regulators that drive a cell to a specific population. Nevertheless, it is becoming clear that lineage commitment is a gradual process that is not dependent on a single factor, and that cells retain a certain degree of plasticity.

The relocation of cells during the morula to blastocyst transition raises the question of whether, at the time of internalization, cells are uncommitted to a particular population or if their lineage fate gets altered (Anani et al., 2014; Watanabe et al., 2014). The use of a CDX2-GFP reporter line for live imaging during preimplantation stages has proven particularly useful in addressing this issue. These studies revealed that some GFP-positive outer cells internalized,

after which their GFP intensity gradually decreased until it became undetectable (McDole and Zheng, 2012; Toyooka et al., 2016). There is thus patent cell plasticity during this time window, but the precise timing of lineage commitment and its correlation with transcriptional changes is unclear. A recent report used the same CDX2-GFP line to study the potential of single cells during lineage segregation. Transcriptome profiling of single cells from 16-cell to 64-cell embryos revealed a sharp divergence in gene expression between the late 16cell and early 32-cell stages, with genes related to the TE network being the first to be activated when transcriptional differences arise among blastomeres (Posfai et al., 2017). Nevertheless, the triggering of lineage programs does not immediately restrict cell potency. The ability of a cell to contribute to both populations decreases at the 16-cell to 32-cell transition, but interestingly the temporal dynamics of CDX2 expression changes is not the same for the TE and ICM lineages. CDX2-GFP^{high} cells preferentially contribute to the TE, but are still able to contribute to the ICM until the late 32-cell stage. In contrast, CDX2-GFP^{low} cells contribute exclusively to the ICM from the early 32-cell stage. This plasticity is extended in embryos reconstructed exclusively from CDX2-GFP^{high} or CDX2-GFP^{low} cells. Reconstituted embryos composed of late 32-cell stage CDX2-GFP^{low} cells still form both lineages, and it is not until the 64-cell stage that cells lose their ability to form TE. However, embryos made from late 32-cell stage CDX2-GFPhigh cells do not develop because the cells have already lost their ability to form ICM (Posfai et al., 2017).

These differences in plasticity suggest that TE is the first committed lineage in the embryo, whereas ICM cells retain a broader potential when placed in forced situations.

Recapitulating trophectoderm identity in culture

The ability to obtain blastocyst-derived stem cells is a major advance in the analysis of regulatory networks underlying the emergence of the first lineages in the embryo. These populations maintain a high level of plasticity and receptiveness to genetic manipulations that can be used to study molecular mechanisms. Embryonic stem (ES) cells are derived from the ICM (Evans and Kaufman, 1981; Martin, 1981), whereas trophoblast stem (TS) cells are derived from the TE (Tanaka et al., 1998). These stem cell populations mostly preserve their potency and regulatory networks, can be cultured *in vitro* in stemness conditions, and can also be directed to differentiate and contribute to their target lineage in chimera assays (Beddington and Robertson, 1989; Tanaka et al., 1998).

Transdifferentiation assays allow to change the identity of a stem-cell population have underlined the importance of specific factors for the establishment of early lineages. Overexpression of *Cdx2*, *Gata3*, *Eomes*, *Tfap2c*, or *Tead4* in ES cells triggers their transdifferentiation to TS-like cells (Kuckenberg et al., 2010; Nishioka et al., 2009; Niwa et al., 2005; Ralston et al., 2010). The use of TS cells has obvious advantages, such as the possibility to obtain enough material for biochemical or ChIP-seq assays (Home et al., 2012; Latos et al., 2015), which would be technically challenging with embryonic material obtained directly. However, it is important to recognise that TS cells are not completely equivalent to the *in vivo* TE. Differences between the TE and TS cells have been reported in epigenetic status (Senner et al., 2012) and in the usage of *Cdx2* regulatory elements (Rayon et al., 2016). Furthermore, TS cells comprise a population

whose identity is already established, and therefore most of the mechanisms at play are likely related more to maintaining the trophoblast phenotype than to its establishment. For instance, the pivotal role played by cell polarity in establishing TE identity in the embryo cannot be recreated in 2-dimenstional cultures of TS cells.

This could explain why, although the core TFs such as CDX2 are crucial in both the TE and TS cells, there are important differences in signaling pathways acting upstream of these factors. For example, YAP is nuclear in the TE and excluded from nuclei in the ICM (Nishioka et al., 2009), but it is located in the nuclei of ES cells (the counterpart of the ICM), where it has been reported to maintain pluripotency and control differentiation (Chung et al., 2016; Lian et al., 2010). Similarly, heterogeneous Notch activity has been detected in ES cells (Nowotschin et al., 2013), whereas Notch is inactive in the ICM (Rayon et al., 2014).

Trophectoderm derivatives and their role in the postimplantation embryo

After hatching from the zona pellucida, the blastocyst implants in the uterus through the mural TE, formed by the trophoblast population opposite the ICM pole. Implantation requires a complex set of interactions between the embryo and the uterine epithelium, involving several proteases, adhesion molecules, and signaling events (Aplin and Ruane, 2017).

The first definitive cells to differentiate from the TE upon implantation are the trophoblast giant cells (TGC). TGCs cover the entire embryo except for the region covered by the polar TE, which directly overlies the ICM. The polar TE gives rise

to two different cone shaped structures: the ExE, in direct contact with the epiblast, and the EPC, which is located proximal to the ExE and mediates early interactions with the maternal endometrium. By E6.5, both the ExE and EPC are present, and the chorion has formed. The chorion is a bilayered tissue derived from embryo-derived mesoderm that has migrated to the extraembryonic region and from the trophoblast progenitor cells of the ExE. Another structure, the allantois, is also derived from embryonic mesoderm that migrates to the extraembryonic compartment at the posterior end of the embryo; when the allantois reaches the chorion, chorio-allantoic fusion occurs. The resulting chorioallantoic attachment will form the labyrinth zone of the placenta, the main route for nutrient exchange between mother and embryo (Cross et al., 2006). The labyrinth is formed by three trophoblast cell types derived from distinct precursors in the chorion: spongiotrophoblasts, glycogen cells, and secondary TGCs. TGCs give rise to several subtypes, including maternal arterial canal-associated TGCs, parietal TGCs, and spiral artery-associated TGCs (Latos and Hemberger, 2016; Rossant and Cross, 2001). These cells, together with glycogen cells, invade the decidua and associate with maternal tissue (Mould et al., 2012; Simmons et al., 2008). Cells of another TGC subtype, the sinusoidal TGCs, establish themselves in the maternal blood sinusoids. Invasion of maternal tissue occurs with the help of TGC-expressed proteases such as CTS7 and CTS8 (Screen et al., 2008).

One of the most interesting phases in trophoblast differentiation is ExE differentiation during gastrulation. The ExE establishes a complex crosstalk with the epiblast that is necessary not only for the maintenance and differentiation of ExE trophoblasts and for establishing the anterior-posterior axis, but also for

germ-cell formation and gastrulation in the epiblast (Beck et al., 2002; Feldman et al., 1995; Yoshimizu et al., 2001) (Fig 3).

The ExE is formed after implantation at around E5.5 and is the reservoir for TS cells that will proliferate and contribute to different cell types of the placenta during its differentiation and growth. TS cell proliferation depends on diffusible signals from neighboring cells in the epiblast, including FGF4, which acts through FGFR2 expressed on TS cells in the ExE. FGF4 is induced by NODAL in the epiblast and has a proliferative effect on ExE TS cells, activating the expression of genes encoding critical transcription factors necessary for their maintenance, such as *Cdx2*, *Eomes*, and *Esrrb*, and at the same time blocking specification genes such as *Ascl2*, expressed in the EPC (Guzman-Ayala et al., 2004) (Fig 3). ExE development is also determined by the levels of *Elf5* and *Ets2*. *Elf5* is important for TS maintenance and its deletion causes TS cells to differentiate to EPC cells. Knockout mice for both genes show loss of ExE upon gastrulation, although a compensatory upregulation of *Ets2* partially rescues the defects in *Elf5* mutants (Donnison et al., 2005; Donnison et al., 2015; Yamamoto et al., 1998).

One of the major players in the crosstalk between the ExE with the embryonic epiblast is NODAL, a secreted protein of the TGFβ family. *Nodal* is expressed in the epiblast, producing the uncleaved precursor pro-NODAL. Secreted pro-NODAL activates the expression of BMP4 and the convertases PACE4 and FURIN in the ExE through Activin receptor complexes (Fig 3). PACE4- and FURIN- mediated pro-NODAL cleavage and maturation sets up a positive loop in which FOXH1 and SMAD2 activate further *Nodal* expression (Brennan et al., 2001; Norris et al., 2002). BMP4 is able to activate another feedback loop by inducing WNT3, which induces *Nodal* expression. In the posterior epiblast, BMP4

and WNT3 also induce *Brachyury* expression, necessary for correct gastrulation and mesoderm fomation (Ben-Haim et al., 2006). BMP4 and its downstream effector SMAD1 are also important for allantois formation and are ultimately responsible for correct placenta development (Tremblay et al., 2001).

The importance of the ExE in primordial germ cell (PGC) differentiation was revealed by the observation that the proximal region of the epiblast gives rise to PGCs while the distal part never does so unless transplanted to a more proximal location (Tam and Zhou, 1996). Later experiments confirmed that the ExE secretes BMP4 and BMP8b, which are essential for PGC generation (Lawson et al., 1999).

These observations demonstrate the critical role of the crosstalk at the embryonic-extraembryonic interface during early mouse development. The possible evolutionary conservation of these mechanisms has not been resolved because of the different geometries of other mammalian postimplantation embryos, such as human or rabbit, which are not cup-shaped like the mouse but flat disk-shaped (Sheng, 2015), which would require crosstalk without regionalized apposition of embryonic and extraembryonic territories.

The instructive role of extraembryonic cells in patterning non-committed pluripotent embryonic cells was recently addressed by constructing ES-TS aggregates in three-dimensional culture (Harrison et al., 2017). These structures are capable of expressing *Brachyury* and PGC markers such as *Dppa3* in the ES cell component in a spatially defined fashion and dependent on NODAL signaling. This level of complexity regarding NODAL pathway and PGC formation is not observed in similar aggregates generated exclusively from ES or TS cells. These experiments demonstrate that trophoblast-derived cells can drive pluripotent

embryonic cells along spatial and temporal developmental paths similar to those occurring in the embryo. On the other hand, it has been shown that embryonic organoids derived exclusively from ES cells self-organize spatially and acquire axial organization, independently of extraembryonic tissue and, interestingly, of BMP signaling (van den Brink et al., 2014). It can therefore be concluded that the embryo has an intrinsic ability to break its symmetry, and extraembryonic tissues may simply provide the cues needed to bias this intrinsic ability (Turner et al., 2017). This could partly explain how the A-P axis can be established in embryos of species such as humans, in which ExE and epiblast contact is not polarized but takes place over the entire embryo surface.

Conclusions

The emergence of the trophectoderm is the first cell-type differentiation event in mammalian embryos. The acquisition of an epithelial phenotype is tightly linked to the differential expression of core transcription factors, which leads to lineage commitment in the blastocyst. Genetic and morphological properties of the cell, such as apical-basal polarity, are both required for this initial symmetry-breaking event. The study of trophectoderm specification provides important clues about how cellular and transcriptional events combine to generate lineage commitment from the naïve state.

Acknowledgments

We thank Teresa Rayon and members of the Manzanares lab for fruitful discussions and Simon Bartlett (CNIC) for English editing. Work in our lab is

supported by the Spanish government (FPI-SO Fellowship to SM, FPI Fellowship to JGSA, and grant BFU2014-54608-P to MM). The CNIC is supported by the Ministry of Economy, Industry and Competitiveness (MEIC) and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

References

Alarcon, V. B. (2010). *Biol Reprod* 83, 347-58.

- Anani, S., Bhat, S., Honma-Yamanaka, N., Krawchuk, D., and Yamanaka, Y. (2014). *Development* **141**, 2813-24.
- Aplin, J. D., and Ruane, P. T. (2017). J Cell Sci 130, 15-22.
- Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). *Genes Dev* **17**, 126-40.
- Beck, F., Erler, T., Russell, A., and James, R. (1995). Dev Dyn 204, 219-27.
- Beck, S., Le Good, J. A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F., and Constam, D. B. (2002). *Nat Cell Biol* **4**, 981-5.
- Beddington, R. S., and Robertson, E. J. (1989). Development 105, 733-7.
- Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E. J., and Constam, D. B. (2006). *Dev Cell* **11**, 313-23.
- Brennan, J., Lu, C. C., Norris, D. P., Rodriguez, T. A., Beddington, R. S., and Robertson, E. J. (2001). *Nature* **411**, 965-9.
- Cao, Z., Carey, T. S., Ganguly, A., Wilson, C. A., Paul, S., and Knott, J. G. (2015). *Development* **142**, 1606-15.
- Cockburn, K., Biechele, S., Garner, J., and Rossant, J. (2013). *Curr Biol* 23, 1195-201.
- Conlon, R. A., Reaume, A. G., and Rossant, J. (1995). *Development* **121**, 1533-45.
- Cross, J. C., Flannery, M. L., Blanar, M. A., Steingrimsson, E., Jenkins, N. A., Copeland, N. G., Rutter, W. J., and Werb, Z. (1995). *Development* **121**, 2513-23.
- Cross, J. C., Nakano, H., Natale, D. R., Simmons, D. G., and Watson, E. D. (2006). *Differentiation* **74**, 393-401.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). *Cell* **113**, 643-55.

- Chen, L., Yabuuchi, A., Eminli, S., Takeuchi, A., Lu, C. W., Hochedlinger, K., and Daley, G. Q. (2009). *Cell Res* **19**, 1052-61.
- Chisholm, J. C., and Houliston, E. (1987). *Development* **101**, 565-82.
- Chung, H., Lee, B. K., Uprety, N., Shen, W., Lee, J., and Kim, J. (2016). *EMBO Rep* **17**, 519-29.
- Dietrich, J. E., and Hiiragi, T. (2007). *Development* **134**, 4219-31.
- Donnison, M., Beaton, A., Davey, H. W., Broadhurst, R., L'Huillier, P., and Pfeffer, P. L. (2005). *Development* **132**, 2299-308.
- Donnison, M., Broadhurst, R., and Pfeffer, P. L. (2015). Dev Biol 397, 77-88.
- Evans, M. J., and Kaufman, M. H. (1981). Nature 292, 154-6.
- Faria, T. N., Ogren, L., Talamantes, F., Linzer, D. I., and Soares, M. J. (1991). *Biol Reprod* **44**, 327-31.
- Feldman, B., Poueymirou, W., Papaioannou, V. E., DeChiara, T. M., and Goldfarb, M. (1995). *Science* **267**, 246-9.
- Fierro-Gonzalez, J. C., White, M. D., Silva, J. C., and Plachta, N. (2013). *Nat Cell Biol* **15**, 1424-33.
- Goolam, M., Scialdone, A., Graham, S. J., Macaulay, I. C., Jedrusik, A., Hupalowska, A., Voet, T., Marioni, J. C., and Zernicka-Goetz, M. (2016). *Cell* 165, 61-74.
- Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D., and Robson,P. (2010). *Dev Cell* 18, 675-85.
- Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D. B. (2004). *Proc Natl Acad Sci U S A* **101**, 15656-60.
- Haffner-Krausz, R., Gorivodsky, M., Chen, Y., and Lonai, P. (1999). *Mech Dev* **85,** 167-72.
- Hancock, S. N., Agulnik, S. I., Silver, L. M., and Papaioannou, V. E. (1999). *Mech Dev* **81**, 205-8.
- Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C., and Zernicka-Goetz,M. (2017). *Science* 356.

- Hillman, N., Sherman, M. I., and Graham, C. (1972). *J Embryol Exp Morphol* **28**, 263-78.
- Hirate, Y., Hirahara, S., Inoue, K., Kiyonari, H., Niwa, H., and Sasaki, H. (2015). *Dev Growth Differ* **57**, 544-56.
- Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V. B., Akimoto, K., Hirai,
 T., Hara, T., Adachi, M., Chida, K., Ohno, S., Marikawa, Y., Nakao, K.,
 Shimono, A., and Sasaki, H. (2013). *Curr Biol* 23, 1181-94.
- Home, P., Kumar, R. P., Ganguly, A., Saha, B., Milano-Foster, J., Bhattacharya,
 B., Ray, S., Gunewardena, S., Paul, A., Camper, S. A., Fields, P. E., and Paul,
 S. (2017). *Development* 144, 876-888.
- Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M., and Paul, S. (2009). *J Biol Chem* **284**, 28729-37.
- Home, P., Saha, B., Ray, S., Dutta, D., Gunewardena, S., Yoo, B., Pal, A., Vivian, J. L., Larson, M., Petroff, M., Gallagher, P. G., Schulz, V. P., White, K. L., Golos, T. G., Behr, B., and Paul, S. (2012). *Proc Natl Acad Sci U S A* **109**, 7362-7.
- Huang, D., Han, X., Yuan, P., Ralston, A., Sun, L., Huss, M., Mistri, T., Pinello,L., Ng, H. H., Yuan, G., Ji, J., Rossant, J., Robson, P., and Guo, G. (2017).*bioRxiv*.
- Johnson, M. H., and Ziomek, C. A. (1981). *Cell* 24, 71-80.
- Kaneko, K. J., and DePamphilis, M. L. (2013). Development 140, 3680-90.
- Kang, M., Garg, V., and Hadjantonakis, A. K. (2017). Dev Cell 41, 496-510.e5.
- Kelly, S. J. (1977). J Exp Zool 200, 365-76.
- Koch, U., Lehal, R., and Radtke, F. (2013). Development 140, 689-704.
- Kondratiuk, I., Bazydlo, K., Maleszewski, M., and Szczepanska, K. (2012). *Dev Biol* **368**, 54-62.
- Kono, K., Tamashiro, D. A., and Alarcon, V. B. (2014). Dev Biol 394, 142-55.
- Korotkevich, E., Niwayama, R., Courtois, A., Friese, S., Berger, N., Buchholz, F., and Hiiragi, T. (2017). *Dev Cell* **40**, 235-247.e7.

- Kuckenberg, P., Buhl, S., Woynecki, T., van Furden, B., Tolkunova, E., Seiffe, F., Moser, M., Tomilin, A., Winterhager, E., and Schorle, H. (2010). *Mol Cell Biol* 30, 3310-20.
- Latos, P. A., and Hemberger, M. (2016). Development 143, 3650-3660.
- Latos, P. A., Sienerth, A. R., Murray, A., Senner, C. E., Muto, M., Ikawa, M., Oxley, D., Burge, S., Cox, B. J., and Hemberger, M. (2015). *Genes Dev* 29, 2435-48.
- Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P., and Hogan, B. L. (1999). *Genes Dev* **13**, 424-36.
- Leung, C. Y., and Zernicka-Goetz, M. (2013). Nat Commun 4, 2251.
- Lian, I., Kim, J., Okazawa, H., Zhao, J., Zhao, B., Yu, J., Chinnaiyan, A., Israel, M. A., Goldstein, L. S., Abujarour, R., Ding, S., and Guan, K. L. (2010). *Genes Dev* 24, 1106-18.
- Lin, S. C., Wani, M. A., Whitsett, J. A., and Wells, J. M. (2010). *Development* **137**, 3953-63.
- Maitre, J. L., Niwayama, R., Turlier, H., Nedelec, F., and Hiiragi, T. (2015). *Nat Cell Biol* **17**, 849-55.
- Maitre, J. L., Turlier, H., Illukkumbura, R., Eismann, B., Niwayama, R., Nedelec, F., and Hiiragi, T. (2016). *Nature* **536**, 344-8.
- Manzanares, M., and Rodriguez, T. A. (2013). Curr Biol 23, R559-61.
- Martin, G. R. (1981). Proc Natl Acad Sci U S A 78, 7634-8.
- McDole, K., Xiong, Y., Iglesias, P. A., and Zheng, Y. (2011). *Dev Biol* **355**, 239-49.
- McDole, K., and Zheng, Y. (2012). *Genesis* **50**, 775-82.
- Menchero, S., Rayon, T., Andreu, M. J., and Manzanares, M. (2017). *Dev Dyn* **246**, 245-261.
- Mihajlovic, A. I., and Bruce, A. W. (2016). Reprod Biomed Online.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). *Cell* **113**, 631-42.

Mlodzik, M., Fjose, A., and Gehring, W. J. (1985). *Embo j* 4, 2961-9.

- Molotkov, A., Mazot, P., Brewer, J. R., Cinalli, R. M., and Soriano, P. (2017). *Dev Cell* **41**, 511-526.e4.
- Mould, A., Morgan, M. A., Li, L., Bikoff, E. K., and Robertson, E. J. (2012). *Genes Dev* **26**, 2063-74.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). *Cell* **95**, 379-91.
- Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R. O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E. M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., and Sasaki, H. (2009). *Dev Cell* **16**, 398-410.
- Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). *Mech Dev* **125**, 270-83.
- Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Nat Genet 24, 372-6.
- Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). *Cell* **123**, 917-29.
- Norris, D. P., Brennan, J., Bikoff, E. K., and Robertson, E. J. (2002). *Development* **129**, 3455-68.
- Nowotschin, S., Xenopoulos, P., Schrode, N., and Hadjantonakis, A. K. (2013). *BMC Dev Biol* **13**, 15.
- Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oles, A. K., Arauzo-Bravo, M. J., Saitou, M., Hadjantonakis, A. K., and Hiiragi, T. (2014). *Nat Cell Biol* 16, 27-37.
- Pan, D. (2010). Dev Cell 19, 491-505.
- Plusa, B., Frankenberg, S., Chalmers, A., Hadjantonakis, A. K., Moore, C. A., Papalopulu, N., Papaioannou, V. E., Glover, D. M., and Zernicka-Goetz, M. (2005). *J Cell Sci* **118**, 505-15.
- Polydorou, C., and Georgiades, P. (2013). Nat Commun 4, 1658.
- Posfai, E., Petropoulos, S., de Barros, F. R., Schell, J. P., Jurisica, I., Sandberg, R., Lanner, F., and Rossant, J. (2017). **6**.

- Ralston, A., Cox, B. J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J. S., and Rossant, J. (2010). *Development* 137, 395-403.
- Ralston, A., and Rossant, J. (2008). Dev Biol 313, 614-29.
- Rayon, T., Menchero, S., Nieto, A., Xenopoulos, P., Crespo, M., Cockburn, K., Canon, S., Sasaki, H., Hadjantonakis, A. K., de la Pompa, J. L., Rossant, J., and Manzanares, M. (2014). *Dev Cell* **30**, 410-22.
- Rayon, T., Menchero, S., Rollan, I., Ors, I., Helness, A., Crespo, M., Nieto, A., Azuara, V., Rossant, J., and Manzanares, M. (2016). *Sci Rep* **6**, 27139.
- Rossant, J., and Cross, J. C. (2001). Nat Rev Genet 2, 538-48.
- Samarage, C. R., White, M. D., Alvarez, Y. D., Fierro-Gonzalez, J. C., Henon, Y., Jesudason, E. C., Bissiere, S., Fouras, A., and Plachta, N. (2015). *Dev Cell* 34, 435-47.
- Sasaki, H. (2015). Semin Cell Dev Biol 47-48, 80-7.
- Screen, M., Dean, W., Cross, J. C., and Hemberger, M. (2008). *Development* 135, 3311-20.
- Senner, C. E., Krueger, F., Oxley, D., Andrews, S., and Hemberger, M. (2012). Stem Cells **30**, 2732-45.
- Sheng, G. (2015). Dev Biol 401, 17-24.
- Shi, S., Stahl, M., Lu, L., and Stanley, P. (2005). Mol Cell Biol 25, 9503-8.
- Simmons, D. G., Natale, D. R., Begay, V., Hughes, M., Leutz, A., and Cross, J.C. (2008). *Development* 135, 2083-91.
- Souilhol, C., Cormier, S., Tanigaki, K., Babinet, C., and Cohen-Tannoudji, M. (2006). *Mol Cell Biol* **26**, 4769-74.
- Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). *Development* **132**, 2093-102.
- Suwinska, A., Czolowska, R., Ozdzenski, W., and Tarkowski, A. K. (2008). *Dev Biol* **322**, 133-44.

- Tabansky, I., Lenarcic, A., Draft, R. W., Loulier, K., Keskin, D. B., Rosains, J., Rivera-Feliciano, J., Lichtman, J. W., Livet, J., Stern, J. N., Sanes, J. R., and Eggan, K. (2013). *Curr Biol* 23, 21-31.
- Tam, P. P., and Zhou, S. X. (1996). *Dev Biol* **178**, 124-32.
- Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A., and Rossant, J. (1998). *Science* **282**, 2072-5.
- Tarkowski, A. K. (1959). Nature 184, 1286-7.
- Tarkowski, A. K., and Wroblewska, J. (1967). *J Embryol Exp Morphol* **18**, 155-80.
- Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T., and Zernicka-Goetz, M. (2007). *Nature* **445**, 214-8.
- Toyooka, Y., Oka, S., and Fujimori, T. (2016). *Dev Biol* **411**, 50-60.
- Tremblay, K. D., Dunn, N. R., and Robertson, E. J. (2001). *Development* **128**, 3609-21.
- Turner, D., Alonso-Crisostomo, L., Girgin, M., Baillie-Johnson, P., Glodowski, C.
 R., Hayward, P. C., Collignon, J., Gustavsen, C., Serup, P., Steventon, B.,
 Lutolf, M., and Martinez Arias, A. (2017). *bioRxiv*.
- van den Brink, S. C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A. K., Nowotschin, S., Turner, D. A., and Martinez Arias, A. (2014). *Development* 141, 4231-42.
- Watanabe, T., Biggins, J. S., Tannan, N. B., and Srinivas, S. (2014). *Development* **141**, 2279-88.
- Watanabe, Y., Miyasaka, K. Y., Kubo, A., Kida, Y. S., Nakagawa, O., Hirate, Y., Sasaki, H., and Ogura, T. (2017). *Sci Rep* **7**, 46135.
- Wicklow, E., Blij, S., Frum, T., Hirate, Y., Lang, R. A., Sasaki, H., and Ralston, A. (2014). *PLoS Genet* **10**, e1004618.
- Yagi, R., Kohn, M. J., Karavanova, I., Kaneko, K. J., Vullhorst, D., DePamphilis,M. L., and Buonanno, A. (2007). *Development* 134, 3827-36.

Yamamoto, H., Flannery, M. L., Kupriyanov, S., Pearce, J., McKercher, S. R., Henkel, G. W., Maki, R. A., Werb, Z., and Oshima, R. G. (1998). *Genes Dev* 12, 1315-26.

Yoshimizu, T., Obinata, M., and Matsui, Y. (2001). Development 128, 481-90.

Ziomek, C. A., and Johnson, M. H. (1980). Cell 21, 935-42.

Figure legends

Figure 1. The Gene Regulatory Network driving trophectoderm identity. The establishment and differentiation of the trophectoderm (TE) in the embryo requires the sequential activation of transcription factors and other TE-related genes (middle and right panels) as well as the repression of the pluripotency network (left panel). The transcriptional effectors of the Hippo and Notch pathways are indicated in light red (DNA binding factors) and in purple (coactivators). SBNO1 (dark blue) is a chromatin remodeler. Green arrows indicate activation; red lines indicate repression.

Figure 2. Hippo and Notch regulate trophectoderm fate. The stochastic expression of CDX2 (green) in the early morula cannot be fully explained by the different levels of nuclear YAP (purple) and Notch intracellular domain (NICD, red) in each blastomere (A). Polarized blastomeres (apical domain in yellow) correlate with nuclear YAP. Cytoplasmic YAP is highlighted in unpolarized blastomeres (A). At the blastocyst stage, there is complete overlap of nuclear YAP, NICD and CDX2, which are all restricted to the trophectoderm (B).

Figure 3. Role of the Extra-Embryonic Ectoderm in patterning the embryonic Epiblast. Diagram showing the signaling interactions taking place between the embryonic and extraembryonic portions of the E6.5 embryo. Localized expression of genes (indicated by italics) and secreted factors (upper case) is indicated in the different regions (Epiblast; Proximal Extraembryonic Ectoderm, PrExE;, Distal Extraembryonic Ectoderm, DiExE; Ectoplacental Cone,

EPC). Direct interactions are indicated by solid lines; dashed lines indicate suggested interactions. The gradient of Nodal through the epiblast is indicated by a black triangle.



Figure 1











