Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

Review

Building principles for constructing a mammalian blastocyst embryo

Peter L. Pfeffer [1,*]

[1] School of Biological Sciences, Victoria University of Wellington, New Zealand; peter.pfeffer@vuw.ac.nz
* Correspondence: peter.pfeffer@vuw.ac.nz; Tel.: +64-4-463-7462

Received: date; Accepted: date; Published: date

Abstract: The self-organisation of a fertilised egg into an embryo of three distinct cell lineages (trophoblast, epiblast and hypoblast) arranged around an off-centre cavity typifies the blastocyst, an embryonic structure that is unique to mammals. While the starting point (the zygote) and endpoint (the blastocyst) are similar in all mammals, the intervening events have diverged. This review examines and compares the descriptive and functional data surrounding embryonic gene activation, symmetry-breaking, first and second lineage establishment and fate commitment in a wide range of mammalian orders. The exquisite detail known from mouse embryogenesis, embryonic stem cell studies and the wealth of recent single cell transcriptomic experiments are used to highlight the building principles underlying early mammalian embryonic development.

Keywords: lineage determination; patterning; blastomere polarization; compaction; cleavage stages; morula; gene regulatory networks

---

Contents

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

1. Introduction

The German word "Bauplan" (building plan/blueprint) was first used by Woodger [1] to describe the archetypical body plan of a particular group of animals. The architectural connotation is useful, as in many ways the formation of an embryo resembles the building of a house: in both, a three-dimensional structure is created from a limited set of building materials, with construction following a set of principles that have been optimised over time for cost, speed and quality. Designs are adapted to local material availability, the environment and the future uses of the building. For an embryo, the building materials are the different types of cells and extracellular secretions, which appear to have been carefully designed to fit together in very limited ways (the building principles) thanks to highly detailed instructions encoded by a temporal series of gene regulatory networks. The truly amazing part though, is that no "outside" help is required in the building of an embryo. How is this possible? In this first review covering only the initial steps of embryogenesis, I will address the remarkable innovations which evolution has selected to automate the generation of a fairly consistent mammalian late-blastocyst Bauplan, while starting out from a myriad of divergent initial environments.

1.1. Mammalian early embryological diversity

Before delving into mechanistic aspects, a brief description of mammalian taxonomy, early embryogenesis and the bewildering nomenclature (Table 1) is warranted. Class Mammalia consists of eutherian (placental) mammals, monotherian (marsupial) mammals and their sister group, the egg-laying monotremes. Within the evolutionary successful eutherians there are 19 extant orders divided into 4 four superordinal groups: Euarchontoglires (e.g. primates, rodents, rabbits), Laurasiatheria (e.g. carnivores, ruminants, pigs, horses, bats, whales), Afrotheria (e.g. elephants, sea cows, tenrecs) and Xenarthra (e.g. armadillos, anteaters), all of which last shared a common ancestor about 100 million years ago (Fig. 1) [2]. Before the molecular era, the embryology of representatives of numerous of these orders were examined [3], however during the last 40 years, investigations increasingly focused on only a couple of species, most prominently the easily kept laboratory animals (mouse, rat, rabbit) with some additional work done on farm animals.

Excitingly, with new molecular technologies, the wealth of mechanistic insight into development that could previously only be obtained from the mouse model, is being applied to re-investigate embryos from other mammalian orders, opening a new phase of comparative molecular embryology that is revealing (somewhat ironically) that early mouse development is atypical of most mammals.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

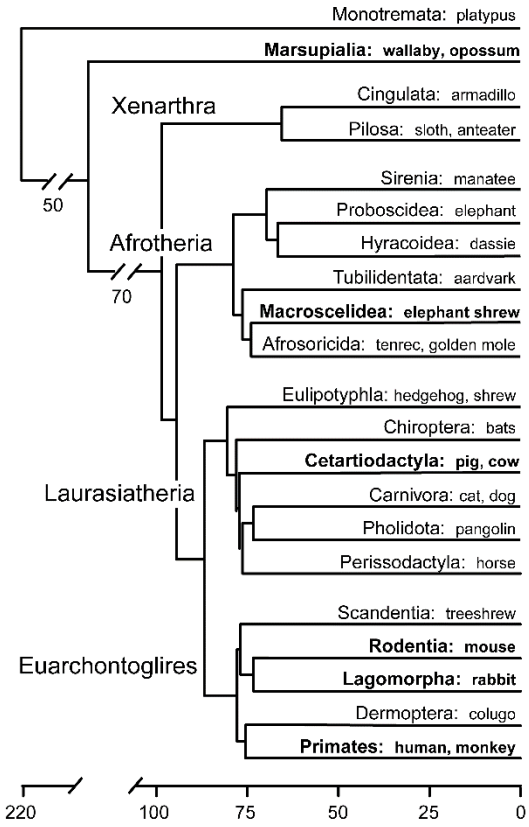


Figure 1. Divergence time (in millions of years) of all extant mammalian orders [2] and their associated superordinal groupings. Groups discussed are shown in bold.

Table 1. Glossary of embryological terms (in bold), alternatives often used (in regular font) in various species and their definition as used in this review.

term and/or abbreviation	equivalent (species)	description
conceptus		Refers to all tissue derived from the zygote (embryonic and extraembryonic).
embryo	embryo proper	Before gastrulation equivalent to conceptus. From gastrulation stages refers to the embryonic parts of a conceptus that will give rise to the foetus as opposed to the extraembryonic membranes and placenta.
Epi, epiblast		Derived from ICM, progenitor population of the three germ layers as well as the amniotic ectoderm and primordial germ cells.
HB, hypoblast	Primitive endoderm (mouse)	Cells differentiated from the ICM not contributing to the epiblast. The hypoblast will give rise to the inner layer of the yolk sac and, in primates, to extraembryonic mesenchyme.
ICM, inner cell mass	Pluriplast (marsupial)	Cells giving rise to epiblast and hypoblast.
primitive endoderm	**HB, Hypoblast**	Primitive endoderm is an alternative name to hypoblast and not to be confused with true (definitive) endoderm
pTB, polar trophoblast	**Rauber's Layer (cow)**	Trophoblast overlying the ICM or epiblast
TB, trophoblast	TE (mouse)	Extraembryonic layer: cells giving rise to the conceptus-derived part of the chorionic membrane and subsequently the fetal part of the placenta
TE, trophectoderm	TB	During blastocyst stages, before overt differentiation, the trophoblast epithelium is often termed trophectoderm

1.2 Morphological events leading to the blastocyst

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

Eutherian mammalian development (Fig. 2) commences after fertilisation, when the zygote undergoes several cleavages without concomitant growth. The resultant blastomeres are contained by a proteinaceous zona pellucida shell. In Laurasiatherians and Euarchontoglirans (which together constitute most mammals), the blastomeres compact together to form a morula with the outermost cells forming an epithelium. The innermost cells are apolar and their number increases by cell division and contributions from the outside layer. Eventually a cavity develops (the blastocoel), marking the onset of the blastocyst stage, characterised by an outside layer of polarised cells termed the trophoblast or trophectoderm and a clump of inner cells called the ICM (inner cell mass). In contrast, in Afrotherians of the orders Macroscelidea (elephant shrew) and Afrosoricida (tenrec), the morula stage is bypassed (Fig. 2), with blastomeres first forming a hollow unilaminar vesicle, from which future ICM cells bud off, as in the elephant shrew [4], or in which future ICM cells proliferate in one region and somehow relocate to the inside of the vesicle as in tenrecs [5]. The net result in all cases is a locally bilayered structure with trophoblast, which faces the maternal environment on the outside, covering the ICM cells located asymmetrically on the inside. Note that in marsupial mammals a trophoblast/ICM bilayer is not formed, instead, a large patch of the unilaminar vesicle becomes the pluriplast with a fate equivalent to the eutherian ICM. The trophoblast will contribute to the fetal part of the placenta but not to the fetus (embryo proper). The next morphological event is the segregation of hypoblast cells, which form a unicellular layer beneath the ICM. From this point in development the ICM is referred to as epiblast. While not discussed in this review, hypoblast cells will give rise only to extraembryonic components of the conceptus, such as the inner layer of the yolk sac and, in primates, extraembryonic mesenchyme which, together with epiblast-derived extraembryonic mesoderm forms most of the umbilical cord and the connective tissue and blood vessels supporting extraembryonic membranes. The epiblast forms all parts of the embryo proper as well as contributing to extraembryonic tissues such as the amnion, allantois and extraembryonic mesoderm.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

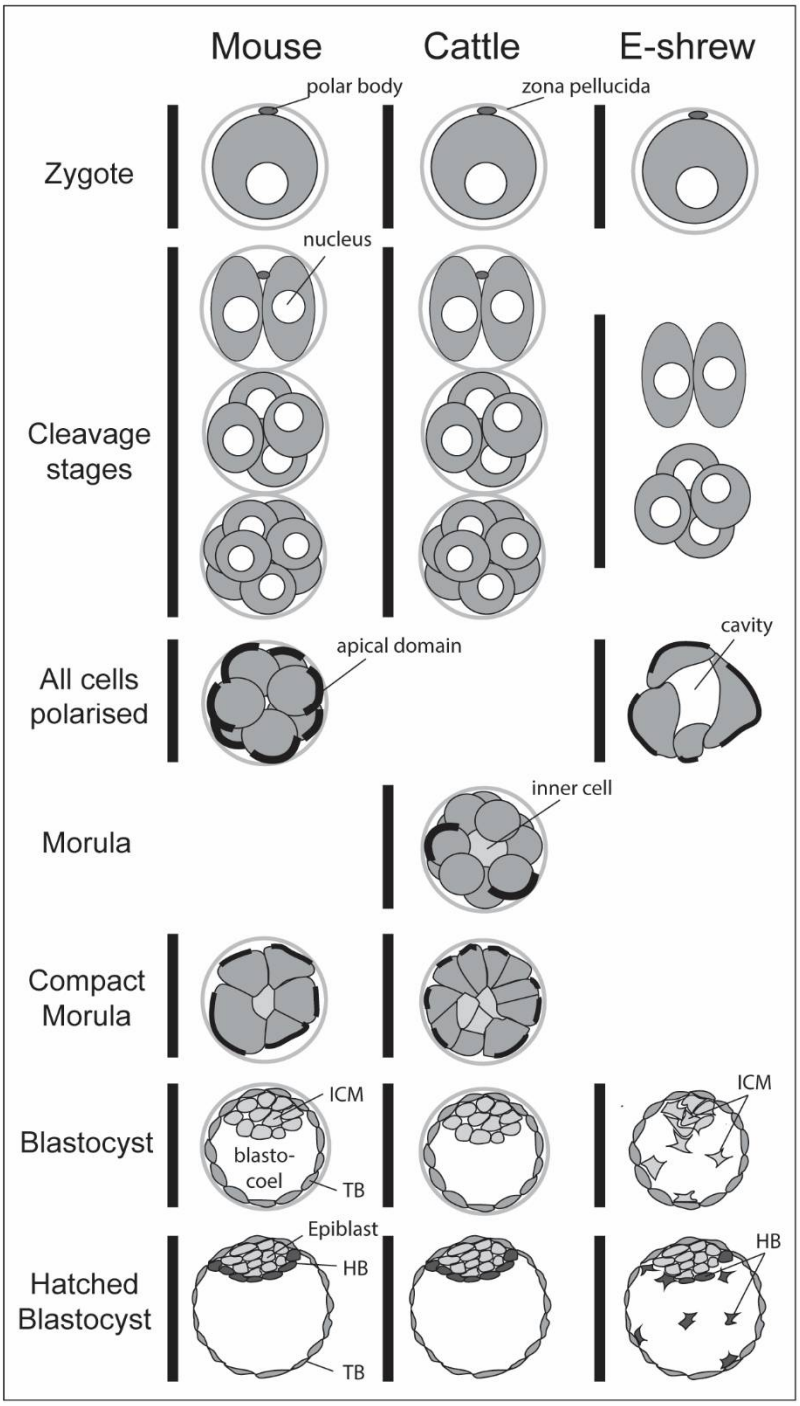


Figure 2. Main developmental stages in an Euarchontoglirian (mouse), Laurasiatherian (cattle) and Afrotherian (elephant shrew). While the sequence of morphological events is largely conserved, some species lack certain stages: mouse embryos compact already at the 8-cell (cleavage) stage, thus skipping the un-compacted morula stage; cattle embryos do not pass through a stage where all blastomeres are polarised and elephant shrews miss out on morula stages altogether and lose their zona pellucida prematurely, immediately after the 1-cell stage. The cattle sequence also applies to the Euarchontoglirian rabbit and human embryo. HB; hypoblast; ICM, inner cell mass; TB, trophoblast.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

2. Gearing up for autonomy (cleavage stages)

One of the most difficult challenges for multicellular animals is the maternal to embryonic transition, which involves the activation of the newly conjoined genetic material so as to drive the embryonic gene program, while at the same time maintaining basic cellular needs using the stores of protein and RNA inherited from the mother via the egg. Embryonic gene activation (EGA, also called zygotic genome activation) occurs in mammals during the cleavage stages in successive "waves" or phases of gene set activations of increasing magnitude. The major phase, involving the switching on of over a thousand genes, happens at the late 2-cell stage in mice, at 4 cells in rabbits, 4-8 cells in humans and pigs and 8-16 cells in cattle [6]. How is this brought about? Two non-exclusive mechanisms are conceivable. 1.) An external trigger, as supplied by the sperm or from the maternal reproductive tract, kick-starts EGA. 2.) An internal time delay mechanism operates, involving the gradual decay/dilution of inhibitors and/or the activation of "pioneer" factors stored in the egg.

2.1. Triggering EGA via an external signal

The fact that eggs (i) can be activated to undergo EGA and further development without sperm, and (ii) can be fertilised and grown outside the maternal environment, argues against the importance of external stimuli. However, very recently the Banerjee lab found an extraordinary link between EGA and pyruvate availability [7]. It has long been known that pyruvate is preferred to glucose as energy substrate for mammalian cleavage-stage embryos [8], as their low energy requirements results in a high ATP:ADP ratio which inhibits phosphofructokinase, the gatekeeper for glucose glycolysis [9]. Nagaraj and co-workers now showed that prior to EGA, pyruvate is required to translocate the mitochondrial enzymes involved in the first half of the tricarboxylic acid (TCA) cycle into the nucleus! Removal of pyruvate blocks the embryo before EGA and prevents histone modifications associated with the opening up of chromatin (H3K4 and H3K27 acetylation; H3K27 trimethylation), due to a reduction of the metabolites required for these modifications [7]. This mechanism appears to be conserved: in humans, pyruvate dehydrogenase, required for Acetyl-CoA production, is found in the nucleus only from the 4 to 8 cell stages, concomitant with EGA in this species [7]. Thus, pyruvate availability in the oviduct regulated by the mother in a time- or location-dependent fashion (the embryo is propelled along the oviduct toward the uterus during cleavage stages), may function as an extrinsic mechanism that, even if not sufficient, is necessary for EGA activation.

2.2. Intrinsic embryonic genome activation

Does an internal delay mechanism operate as well, such as is seen in the fruit fly *Drosophila*, where the maternally encoded zinc-finger protein Zelda has been identified as the main activator of the embryonic genome [10]. It appears mammals may also have a master EGA activator (Fig. 3A), in the form of the double-homeodomain-containing DUX-C protein [11-13]. Both the human homologue, DUX4, and the mouse homologue, Dux, are expressed before the first EGA phase in each species. DUX4/Dux overexpression in embryonic stem cells induces hundreds of genes known to be upregulated at the start of EGA. This gene activation was shown to be associated with the opening up of chromatin specifically around EGA target genes, concomitant with direct Dux binding [13]. Lastly, embryonic Dux depletion leads to defective EGA in mouse embryos [12]. Two features are likely to have contributed to the selection of DUX-C orthologues as the instigators of EGA. First, the increased and potentially synergistic binding provided by two separate DNA binding domains within one protein are likely to make this transcription factor a pioneer factor, and indeed DUX4 was shown to be able to bind at DNaseI inaccessible sites and, via p300/CBP recruitment, to acetylate histone H3 at lysine 27 to open up chromatin [14]. Secondly, the *DUX4* gene is found in tandem repeats within individual units of macrosatellite repeat regions, making it the highest copy-number protein-coding gene within the human genome. Indeed, *DUX-C*-family homologues (*DUX-C, DUX4, Dux*) in species from all four mammalian superordinal groups show a macrosatellite tandem-array organisation with very high copy numbers [15]. Thus even a modest transcriptional activation of each gene copy will result in an overall

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

massive burst of protein, able to drive a crucial event such as the activation of the genome. The *DUX-C*-containing macrosatellites are found either close to the telomere ends (or in mice, a chromosomal fusion point marking a prior telomere end) or close to the centromere (pericentrosomic) [15]. Such regions generally form constitutive heterochromatin which is kept in a transcriptionally tightly repressed state via distinct epigenetic mechanisms [16]. So how are the *DUX-C* genes then activated? It has been shown that upon fertilisation the mouse egg enters a globally transcriptionally permissive state characterised by (i) very extensive low-level transcription even in intergenic and repetitive regions and (ii) an independence of enhancer elements with spurious transcriptional initiation occurring at cryptic promoters [17]. This permissive state is likely driven by a loosening of chromatin, as demonstrated by ATAC-sequencing [18] and the detection of increased histone mobility [19] seen specifically at the mouse zygote stage. Whether such low-level ubiquitous transcription, in combination with the now accessible and extremely high abundance of *DUX-C* genes, is sufficient to generate sufficient Dux-C protein to initiate its downstream effects, remains to be determined.

Comparing human DUX4- and mouse Dux-activated genes in human as well as mouse cells, revealed the existence of two sets of target genes that overlapped extensively with genes activated during EGA [11-13]. One set, recognised by the more conserved (second) DUX4/Dux homeodomain, was recognised by both DUX4 and Dux, as well as by the dog DUX-C homologue. The second set of target genes was specific to each species with many downstream of included distinct retroviral repeat elements: human DUX4 strongly activated the HERVL, whereas mouse Dux activated the MERVL-associated genes [11]. It had previously been discovered that many critical EGA genes are under the control of ERVL long terminal repeat enhancers [20]. Now it became evident that many of these enhancers are driven by DUX-C proteins. The picture that thus emerges (Fig. 3B,C) is one where all DUX-C family members recognise targets that form the core of an ancestral EGA network, whereas EGA genes that have subsequently come under the control of particular retrotransposon classes represent species-specific refinements. The binding and transcriptional activation by DUX-C locally counteracts the increasingly repressive chromatin state that spreads after the brief transcriptionally permissive period, thereby contributing to the establishment of large (median length 40 kbp) stretches of open chromatin during an “early” or “minor” EGA wave [13, 18]. These open regions typically are found downstream of Dux-bound, transcribed ERVL elements which drive the expression of early EGA genes. The early EGA genes presumably include transcription factors initiating the “major” wave of EGA. In the mouse major EGA was shown to involve a novel chromatin signature characterised by open chromatin centred over (i) promoters, (ii) distal stage-specific enhancers and (iii) transcriptional end sites [18]. In humans, one of the early EGA specific genes activated via ERVL-driven transcription is LEUTX, a paired –type homeodomain-containing transcription factor that was shown to activate thousands of genes, including the repressor DPRX, which binds similar targets to LEUTX thereby restricting the extent of EGA [21]. Interestingly no mouse LEUTX homologue exists [21] indicating that these early events have diverged considerably, in line with the use of species-specific DUX-C activated ERVL elements and the different timing of EGA.

2.3. Achieving totipotency

While the EGA gene regulatory networks may show only limited overall resemblance between different mammals, certain critical genes (the ancestral set) may need to be activated by all mammals [22]. One such gene is *Zscan4*, which was shown to be activated in mice and humans by DUX4/Dux [13] and is required to ensure recombination-mediated telomere extension, which is necessary to rejuvenate and preserve the chromosome ends which are eroded during gametogenesis [23, 24]. Most importantly though, is the EGA dependent activation of the master regulatory genes that set up the blastomeres to be able to generate every cell of the conceptus (foetus and extraembryonic/foetal membranes), that is, to be “totipotent”.

Totipotency of peri-EGA embryos was first dramatically shown in rabbits [25] and mice [26], when, after ablation of one blastomere of the 2-cell embryo, the remaining blastomere could result in normal young. In sheep (a Laurasiatherian), four lambs were born by separating the blastomeres from one 4-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

cell embryo and transferring these cells, embedded in agarose blocks into separate recipient ewes [27, 28]. Even in humans related experiments have been executed, showing that individual blastomeres from a 4-cell embryo are able to reform four entire blastocysts [29].

Yet at some point post-EGA, totipotency begins to be eroded as master lineage regulators compete with each other to eventually establish lineage specific gene regulatory networks (GRN) that stably restrict the fate of subpopulations of embryonic cells. Using a variety of transplantation, cell mixing and embryo chimera experiments, it has been shown that the fate of individual cells gradually becomes specified and then stably set ("committed") along fixed developmental trajectories at species-specific times (reviewed previously [30]). An exciting question, dealt with in the next section, is how this specification and commitment is achieved and how well these processes are conserved across different mammalian orders.

In summary, the embryo acquires autonomy post-fertilisation under the permissive influence of externally supplied pyruvate, entering a state of transcriptional promiscuity that leads to the initial activation of pioneering factors such as the multicopy DUX-C family, which instigate a minor wave of embryonic gene activation. Numerous of these early EGA genes are driven by long terminal repeat enhancers derived from species-specific retrotransposon families, thus allowing their co-ordinate activation. Some of the minor EGA genes, such as LEUTX in humans, lead to major EGA, characterised by the totipotency of the blastomeres.

3. Creating two environments in the morula (inside and outside)

3.1. Establishment of protein heterogeneity

How are *different* stable lineage gene regulatory networks first established when starting from embryonic genome activation that occurs in every blastomere? An important consideration is that *EGA is not uniform*. Unequal expression levels of genes in sister blastomeres arise with high statistical likelihood, particularly for genes expressed at relatively low levels. Two often concurrent mechanisms contribute to this: unequal partitioning of cytoplasm (unequal segregation) upon cell division and the inherent "noise" of transcription, arising because transcription occurs in short but intense bursts [31]. Detailed studies in mice have indicated the early presence of biases in mRNA or protein expression from the 2-cell stage onward. These include differences (i) at the 4-cell stage in the epigenetic modifier Prdm14 [32], (ii) in Carm1-mediated histone 3 arginine methylation [33], (iii) in DNA binding kinetics of the ICM-markers Oct4 (Pou5f1) and Sox2 [34, 35] and (iv) in mRNA levels of the trophoblast markers *Cdx2* [36] and *Sox21* [37]. Numerous bimodal expression patterns (high in one blastomere, absent in the other), particularly in genes involved in Wnt signaling, were reported at the mouse 2-cell stage [38]. Mathematical modelling incorporating segregation effects, transcriptional noise and feedback loops, in combination with single cell expression analyses using mouse and human datasets, suggested that such early, stochastically arising biases of potential lineage specifiers are likely to guide, but not fully determine lineage trajectories, as at the 8-cell stage such biases could still be reversed [36]. Thus it appears that *the establishment of heterogeneity is a cell-intrinsic and gradual, progressive process* (Fig. 3D).

3.2. Blastomere polarisation

While these heterogeneities in gene expression arise, cell division without cell growth results in an increasing number of blastomeres of decreasing size confined within the proteinaceous envelope of the zona pellucida. Blastomeres are initially of rounded appearance and loosely apposed to each other but soon acquire polarity, establishing apical domains enriched for microvilli, F-actin and an apical protein-complex containing Pard3,6 and aPKC [39]. In mice, this polarisation process occurs at the 8-cell stage (Fig. 3E). It appears to be a cell-intrinsic autonomous event as it occurs in all blastomeres, even when these are kept separated from the 4-cell stage [40]. However, cell-cell contact, mediated either by the ubiquitously expressed E-Cadherin (Cadh-1) - or potentially by Cadherin-3 [41], synchronises and speeds up the onset, and directs the axis of polarisation such that the apical domain forms in the centre of the contact-free zone [42]. Importantly, the apical domain attracts one end of the spindle axis thereby

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

strongly biasing the plane of the subsequent cleavage [43]. This can lead to completely asymmetric divisions, where one daughter cell does not inherit any part of the apical domain and thus is apolar.

The embryo thus now consists of two different types of cells – polar and apolar (Fig. 3E,F). These cells have measurable differences in key lineage determinants: polar cells express more of the trophoblast lineage determination gene *Cdx2* and less of the ICM lineage marker *Sox2* [43]. *Cdx2* is a target of Tead4, a transcription factor that is upregulated by the Yap protein [44]. To be active, Yap has to escape phosphorylation by Lats [45]. This is achieved only in polarised cells, in which the apical domain sequesters Angiomotin [43], thereby preventing it from activating Lats [45, 46]. The same pathway, involving active Yap and Tead4, simultaneously downregulates *Sox2* in these polar cells [47]. Aside from the apical domain/Yap-mediated pathway in transcriptional partitioning of *Cdx2* and *Sox2*, the apical domain is also able to sequester *Cdx2* messenger RNA via a mechanism that involves the apical domain component aPKC and the terminal 97 nucleotides of the *Cdx2* coding region [48].

3.3. Blastomere positioning within the morula

It follows from this sequence of events that the position of a blastomere within the embryo is important to its eventual fate. The embryo volume is limited by the enclosing zona pellucida, so as the blastomere number increases, some cells will land up on the inside, surrounded on all sides by other cells. These inside cells are apolar. Why? At the 8-cell stage, when mouse blastomeres acquire polarity, the apical domain will face outwards as this is the only contact-free zone. Due to the subsequent bias for a radially orientated spindle axis, cleavage will be predominantly asymmetric [43, 49, 50]. In the original polarity model, it was assumed that this would inherently lead to apolar daughter cells situated on the inside [51]. However, time lapse observations revealed significant numbers of apolar cells on the outside [52, 53]. Some apolar outside cells, characterised by enhanced Yap activity, secondarily acquire polarity and form an apical domain on their contact-free membranes, while others internalise [52]. This internalisation process is closely linked to compaction. In mice compaction is seen nearly concurrently with polarisation and involves the E-Cadherin mediated adhesion process, whereby blastomeres flatten against each other to form a compacted morula [54]. Compaction creates tension between adjacent cells. Between the 8 and 16 cell stage some of this tension is driven by E-Cadherin-expressing filipodia – at the 8-cell stage half the blastomeres extend these filipodia onto the apical surface of two neighbouring cells for 5 hours then retract them during the next hour before the next cell division. Interference with filipodia formation (by knock-down of E-Cadherin, its interaction partners α- and β-Catenin or Myosin-X or by F-actin inhibition) abolished or delayed compaction, while Myosin-X overexpression caused premature compaction at the 6-cell stage [55].

The high tensile forces arising through the compaction process work in concert with constriction occurring at the apical side of those cells that, after division, are positioned closer to the centre. Such apical constriction reduces the apical area and conversely increases the basolateral surface, thereby gradually positioning the cell closer to the embryo centre. Constricting cells, seen predominantly from the 12-cell stage, were shown to be enriched for Mysosin-II [56]. Remarkably, enrichment of cortical myosin coupled to higher contractility could be linked to apolar cells, which thus internalise in preference to the polar cells. Polar cells, by virtue of inheriting the "stiff" apical domain are better in resisting apical constriction and remain on the outside [57]. Whether the extensive changes in tension and cell-shape accompanying internalisation of cells (so-called "mechanosensing") controls YAP activity redundantly to the apical-domain mechanism outlined previously, remains to be seen [57, 58].

So what emerges from these studies in mouse embryos is a self-organising or self-regulative process, whereby distribution of polarity genes in cell-contact free regions leads to asymmetry within each cell. Because the mitotic spindle aligns with the resulting apical domain, asymmetric divisions ensue, generating two types of cells: polar or apolar. Apolar cells have little or no apical domain, which predisposes them to be positioned towards the inside through the tensile forces generated by compaction. The absence of the apical domain concurrently shuts down Yap activity thus biasing inside polar cells to an ICM-like GRN (Fig. 3F). Thus morphological events set up lineage gene regulatory networks. What is still unclear, is whether and how these morphological events relate to the early

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

heterogeneities in gene and protein expression and kinetics seen in 2 to 8-cell mouse embryos. Several questions remain to be answered. Do the early biases affect the decision of an 8-cell blastomere to divide symmetrically (thus generating two polar daughter cells) or asymmetrically (thus contributing a polar cell that is likely to form ICM)? Alternatively, do the early biases affect the timing of cell division? Such altered timing would have downstream lineage determining effects as divisions of blastomeres isolated from 8-cell embryos, after compaction has commenced, are more likely to be asymmetric than divisions of pre-compaction blastomeres [49]. Conversely, do the morphologically-induced biases simply reinforce (or annul) earlier gene expression biases? Additive effects may be necessary to drive expression of particular lineage-determining master genes over a critical threshold that allows, via feedforward loops and negative feedback of alternate lineages, the stabilisation of a particular lineage-GRN [59, 60].

3.4. Polarisation, compaction and cell positioning in other mammals

Is the mouse situation characteristic for other mammals? Similar to embryonic genome activation, polarisation lags behind the mouse in cattle and rabbits, occurring at the 9-15-cell stage in cattle and from the 32-cell stage in rabbits [61]. Interestingly though, only around 40% of cattle blastomeres are polarised between the 9 and 74-cell stage (maximally 52%) - similarly, only 46% of rabbit blastomeres are polarised [61]. Thus unlike mice, in these mammals the *acquisition of polarity*, as opposed to the *loss* of previously acquired polarity (via asymmetric divisions), leads to two morphologically distinct cells (Fig. 3E,F). Secondly, compaction in cattle does not occur at the same stage as polarisation, starting only at the 32-cell stage, more than a full cell division later than polarisation [62]. Allocation of cattle blastomeres to the inside was seen to occur both well before compaction (1-2 inner cells in 20% of 16-cell embryos) and after compaction [62]. Thus blastomere positioning is temporally more loosely coupled to any tensile compaction forces that may actively drive ICM-lineage biased apolar cells to the interior of the embryo. Remarkably, in members of two mammalian orders belonging to Afrotheria (Figure 1), there are no "inside" cells at all initially and no morula stage exists. Instead, blastomeres of the streaked tenrecs of Madagascar (*Hemicentetes semispinosus*; Order Afrosoricida) line up after the second division (4- to 8-cell) along the zona pellucida to form a one-layered (unilaminar) embryo with a central cavity. By the 16-cell stage inner cells have accumulated at one pole of this cavity [5]. Elephant shrews (*Elephantulus myurus*; Order Macroscelidea; Fig. 2) lose their zona pellucida already by the 2-cell stage and from the 4-cell stage must have formed tight junctions as they have developed a cavity in their centre [4]. As the blastomeres divide they remain on the outside of this unilaminar blastula-like embryo until the 120-cell stage when cell divisions originating in all areas of the blastula wall lead to rounded daughter cells that protrude towards the inside and develop long pseudopodia. These cells detach and eventually lose their pseudopodia and coalesce to form the inner cell mass [4, 63]. The formation of a unilaminar vesicle/blastula is reminiscent of marsupials, the sister group of the eutherians, though in these animals the cells equivalent in fate to the eutherian ICM do not migrate to the inside. These comparisons suggest that the early differential loss of polarity (or differential absence of a gain in polarity) is the critical eutherian adaptation that inherently led to two distinguishable cell types (polar/epithelial versus apolar) that would consequentially be positioned in two territories (outside versus inside; Fig. 3G). The mouse studies clearly demonstrate how changes in polarity can be coupled to lineage cues. From an evolutionary perspective, the early partitioning into inside and outside lineages enables early outer lineage-specific specialisations required for subsequent implantation. Secondly, the epithelialisation of the outer layer provides a protected environment (in the absence of an egg shell) for the further specialisation of the inner lineage.

3.5. Epithelialisation of the outer cells

The polarisation of outer cells combined with the upregulation of proteins required to form various junctions (tight junctions, adherent junctions, desmosomes and gap junctions [64]), allows the formation of an epithelial layer. This layer becomes impermeable from the end of the morula stage [65].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

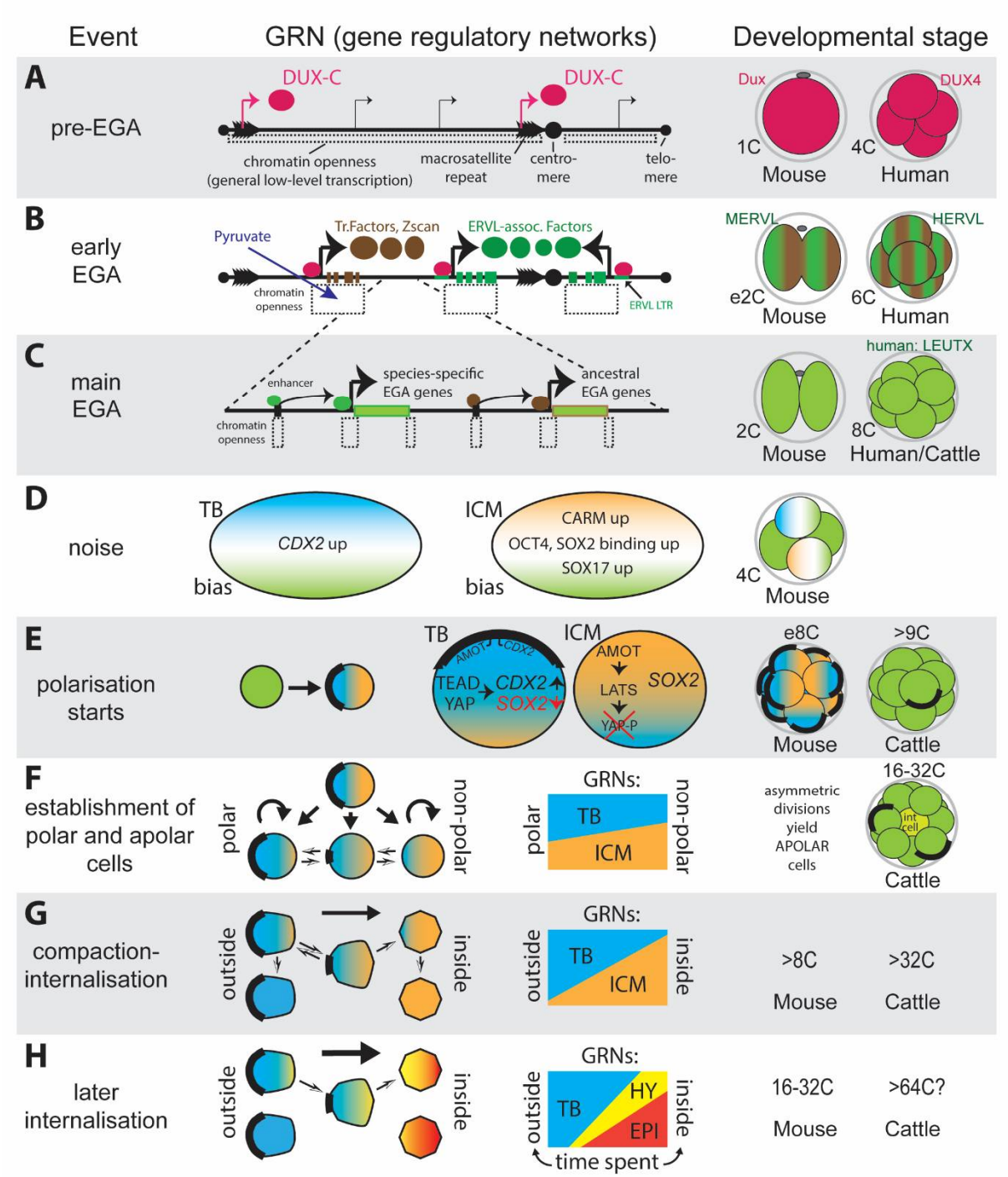
The sealing of the epithelial layer is accomplished by F-Actin rings that form around microtubule-enriched apical domains to then expand toward cell-cell junctions, where they couple with neighbouring actin rings. The coupling involves tight junction molecules such as ZO1 and adherent junction components including E-Cadherin and β-Catenin [66]. Once coupled, Myosin-II is attracted and provides a mechanical tensional force that is necessary for the "zipping up" of the Actin rings, which seals the embryo [66]. Once sealed, the activity of various outer-cell specific Na^+/K^+-ATPase dependent pumps in combination with Aquaporins leads to the directional flow of solutes and water into the interior of the embryo [67]. This results in the formation of the blastocoel, a fluid filled cavity that marks the transition from the morula to the blastocyst stage.

Specifically in outer cells, ATP requirements rise as the energy hungry Na^+/K^+-ATPases begin pumping sodium ions into the emerging blastocoel cavity [68]. Secondly, from the morula stage all blastomeres require more energy (ATP) than at the early cleavage stages as cell divisions are accompanied with growth and therefore increased protein synthesis [69, 70]. The higher energy demands are met by the consumption of glucose in addition to pyruvate. Interestingly, in mice inner cells convert the majority of the glucose substrate to lactate via the oxygen-independent, but energy-inefficient metabolic pathway of glycolysis [71]. The cytosolic NAD^+ used during glycolysis is replenished (i) partly by reducing pyruvate to lactate which is then secreted, and (ii) partly by blastomeres taking up aspartate and converting it to malate (the Aspartate-Malate shuttle, or MAS), which is subsequently oxidised in the mitochondrion [9]. For outer cells, glycolysis does not provide sufficient energy for their higher demands and thus a significant fraction of the consumed glucose is, after conversion to pyruvate, transported into the mitochondria to undergo oxidative phosphorylation. Overall, oxygen consumption increases [72, 73], particularly in the outer cells, and this creates free oxygen radicals. Some of the radicals are mopped up by blastomeres diverting glucose to the pentose phosphate pathway which results in the production of the antioxidant glutathione [9]. This is critical as excess free radicals created by increased oxidative phosphorylation (oxidative stress) leads to an arrest or impairment of development [74]. Intriguingly, Tead4, which is one of the Tead proteins interacting with Yap1 and thus involved in outer cell lineage specification, was shown to be essential for mediating the embryos balancing act between obtaining sufficient energy while not succumbing to oxidative stress damage [75]. While *Tead4* deficient embryos do not develop to the blastocyst stage *in vivo*, they are able to develop normally and activate lineage specific genes *in vitro*, but only under conditions in which oxidative stress is minimised. It is unclear how Tead4 mediates this effect, but it is likely to be linked to its unique (among Tead proteins) translocation to mitochondria [75]. Notably, expression of TEAD4 predominantly in outer cells has been reported in rat, cattle, rhesus monkey and human blastocysts [76]. While knock-down of *TEAD4* in cattle embryos via siRNA did not result in increased susceptibility to oxidative damage [77], this evidence is inconclusive as residual TEAD4 message and protein may have sufficed to prevent developmental arrest.

In view of the observation that TEAD4 is also detected in marsupial cleavage stage embryos [78], it is clear that this protein mediates an ancestral function during mammalian early development. It will be interesting to see whether the prime (ancestral) function lies in protecting the future trophoblast cells from oxidative damage, incurred through the energy demands to generate a blastocoel, or in specifying these very same cells via the TEAD family's interactions with YAP1 to regulate lineage determination genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

416



417

418 **Figure 3. A summary of pre-blastocyst mammalian development with a focus on gene expression. A.** Post-
419 fertilisation but prior to embryonic gene activation (EGA) a loosening of chromatin leads to transcription of the
420 highly repetitive *DUX-C* genes found in macrosatellite regions. **B.** DUX-C protein binding leads to further
421 chromatin opening and transcription of early-EGA target genes. This may require exogenous pyruvate dependent
422 tanslocation of mitochondrial TCA enzymes into the nucleus. Binding of the more divergent first homeodomain
423 of DUX-C occurs particularly in the LTR of ERVL-like retrotransposons and leads to species specific transcription
424 of EGA factors. Binding via the second DUX-C homeodomain in turn activates sets of evolutionarily conserved
425 EGA genes. **C.** During main EGA (shown in green), transcription factors previously activated by DUX-C, such as
426 LEUTX in humans, lead to EGA in an increasingly restrictive chromatin environment. **D.** Levels or binding kinetics
427 of lineage specific transcription factor protein or transcripts, activated during EGA, show regional heterogeneity
428 due to stochastic (noise) events and leads to a potential bias in lineage, shown by blue for trophoblast and orange

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

for inner cell mass. **E.** Formation of apical domains (AD) leads to asymmetry within blastomeres due to tethering of *Cdx2* RNA and the YAP-inactivator AMOT. Cells with an AD thus are biased towards a trophoblast fate. **F.** In mice, the default state is polarisation: non-polar cells are generated via asymmetrical division. Cells with less AD are biased toward the ICM lineage. In other mammals the default state is non-polar and blastomeres gradually acquire polarity. A relationship between presence of an AD and lineage bias has not yet been examined. **G.** Compaction aids the internalisation of non-polar cells to the inside of the embryo. A strong lineage bias is seen. **H.** Cells internalised earlier appear to be biased toward the epiblast lineage, those later to the hypoblast lineage. At this stage numerous cells are committed to either TB or ICM-derived lineages while the hypoblast-epiblast lineages within ICM progenitor cells are not yet resolved.

4. Establishing the first lineages (the blastocyst)

As the blastocoel cavity expands, the inner cells are pushed to one side of the embryo. Morphologically, two different cell types corresponding to the first two lineages are now visible. These are, firstly, the mesenchymal-like inner cells, now termed the inner cell mass/ICM, and secondly, the surrounding unicellular layer of epithelialized prospective trophoblast cells. The ICM will give rise to the epiblast and hypoblast (primitive endoderm) forming the embryo proper as well as numerous extraembryonic tissues. The outer epithelial layer will contribute cells only to the placenta [79, 80]. It is trophoblast cells that establish the initial contacts with the maternal uterine epithelium, which leads to attachment and subsequently implantation of the conceptus. Interestingly, embryos of different mammalian species attach at widely disparate stages of their embryonic development ranging from pre-blastocyst (elephant shrew) to late blastocyst stages in which the epiblast has just begun to form (rodents, primates), to embryos that are already in the midst of neurulation and early somite formation (sheep, cattle, pigs) or have even reached limb bud stages (horses) [3]. Such alternate timing may impose different requirements on the relative speed at which the trophoblast has to commit to its fate [81].

4.1. The timing of the first lineage commitments

Commitment is defined experimentally by the ability of cells to maintain their fate when placed in a different ectopic environment. Waddington famously pictured the analogy of a ball running down a slope containing forked valleys – once in a valley it is committed to that trajectory or fate as it can no longer reach parallel valleys due to the intermediate ridges [82].

In mice, detailed lineage tracking in unperturbed embryos revealed that the 16-cell embryo is the earliest stage exhibiting cell fate bias, but only blastomeres of the 32-cell embryo are lineage restricted, giving rise to either trophoblast or ICM *but not both* [50]. Competency-probing aggregation experiments showed that the outer, prospective trophoblast cells of 16-cell (compacted morula stage, see Fig. 2) embryos were not yet committed to that fate [83, 84], but once embryos had started to cavitate after the 32-cell stage (early blastocysts), commitment had occurred [85]. This is about two cell cycles or one and a quarter days before implantation at the 160-cell stage on Day 4.5 [86]. The interval between commitment and implantation provides time for differentiation of trophoblast to occur. Interestingly, mouse blastomeres from the inside of the embryo only commit to an *ICM* fate later, somewhere between mid (64-cell) and late blastocyst stages [85, 87-89]. In *human* mid-blastocyst stage embryos (up to 45-cell), the outer cells, while fated to become trophoblast, are not committed, as shown by being able to form NANOG-expressing ICM when aggregated to each other. By late Day 5 (late/expanded blastocysts or >180 cells) they are committed [90]. Embryos implant at E7, > 260-cell stage [86], which provides an interval of one and a half days between commitment and implantation, which corresponds to only one cell division. Commitment of *cattle* trophoblast cells was tested by sandwiching outer prospective trophoblast cells between younger uncommitted blastomeres and tracking their fate. This revealed no commitment even at the late blastocyst stage (Day 7), but trophoblast cells were committed by Day 14, a time point just prior to gastrulation [81]. Attachment to the uterus occurs only on Day 21 [91]. In this mammal, though commitment is much delayed in relation to mice, it nevertheless precedes implantation, in this case by at least a week. In sum we can conclude that trophoblast lineage

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

commitment is a prerequisite for implantation and that trophoblast commitment occurs at species-specific developmental time points that are related to the time point of implantation.

4.2. Setting up stable lineage gene regulatory networks (GRNs)

Preceding fate commitment is lineage specification which is dictated via a cell's gene regulatory network. With the advent of technologies allowing dozens of genes or even the entire transcriptome to be measured in single cells, in combination with dimensionality reduction methods such as Principle Component Analysis (PCA) or Diffusion Maps (DM) it has been possible to graphically compare the GRNs among blastomeres of an embryo and from different embryos. Such analyses in human, monkey, mouse and cattle blastomeres [92-96] revealed that blastomeres between egg and morula stages separate according to developmental age, forming a distinct stage-dependent progression in a PCA or DM representation. Blastomeres from the same stage clustered together. However, from morula to blastocyst stages, the GRNs of individual blastomeres of the same stage segregated either into two populations, correlating to trophoblast and ICM cells, or into three populations mapping to trophoblast, epiblast and hypoblast cells.

How does lineage specification relate to the experimentally determined lineage commitment time points? In mice, specification and commitment follow each other closely. Trophoblast cells became committed as well as transcriptionally distinct from ICM cells and uncommitted 16-cell blastomeres at the early blastocyst stage (32-cell). In contrast, uncommitted 32-cell ICM cells were already transcriptionally distinct, though they did show a closer resemblance to the GRN of uncommitted 16-cell blastomeres than did 32-cell trophoblasts [93]. The commitment of human trophoblast progenitors at late blastocyst stages [90] tracks a day after the emergence of their distinct GRN in *early* blastocysts [92, 94, 97]. In cattle, trophoblast cells begin to acquire a distinct GRN beyond the *late* blastocyst stage and even at this stage (the latest examined) some 7% of blastomeres could not be assigned by their gene expression profile to either lineage [95]. Commitment had not occurred at this stage [81]. We can conclude that GRN lineage specification generally occurs before commitment, but can also occur simultaneously as seen for mouse trophoblast. Secondly, GRN distinction between trophoblast and ICM cells occurs at distinct stages in different mammals - for mouse and human in early blastocysts, in cynomolgus macaque monkeys at early to mid-blastocyst stages (E6-7,75 – 200 cells) [96] and in cattle beyond the late blastocyst stage. Yet the order in which the first three lineages become transcriptionally distinct is conserved. In mice the segregation of the first two lineages (trophoblast and ICM) occurs one cell cycle before ICM-like cells resolve into epiblast and hypoblast lineages [93]. In humans initial reports suggested that all three lineages appear to arise simultaneously [94], however subsequent reanalysis of the data and a wealth of novel data indicated that in both human and monkeys a cell cycle separates these events [92, 96, 97].

In view of the different timings of implantation, lineage specification and commitment as well as the prior events of embryonic genome activation driven by enhancers from species-specific transposable elements, how (dis)similar are the first lineage GRNs and the mechanisms of their establishment and stabilisation? Up to morula stages distinct gene-co-expression modules identified in humans were largely preserved in mice, though the timing differed [22]. Thereafter, upon lineage specification, both conserved and species-specific (non-conserved) networks could be detected when comparing lineage-GRNs of different mammals [96, 98, 99]. While global expression studies are ideal for identifying new lineage-specific marker genes, additional loss and gain of function experiments are required to determine the functional importance of candidate genes. Overwhelmingly, such in vivo experiments have been performed only in mice. However, our understanding of lineage networks has also benefited enormously from in vitro studies using embryo-derived stem cell populations representing cells of various lineages that have adapted to proliferate stably in the presence of particular culture media and components. It should though be born in mind that such cell populations (2-cell-like, naïve and primed mouse and human embryonic stem cells (ESC), induced pluripotent stem cells (iPSC), epiblast stem cells, mouse trophoblast (TS) and hypoblast (XEN) stem cells) only

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

approximate their source cells as the embryonic environment differs from culture media and rapidly changes over time (an exception being diapause).

4.3. The trophoblast lineage

Focussing only on transcription factors, as these are fundamental in setting up transcriptional programs, the earliest trophoblast-specific markers in mouse embryos are *Cdx2* and *Id2,* followed by *Eomes, Elf5, Gata2, Gata3, Tfap2a* and *Tfap2c* [93, 99-101]. Of these eight factors, *ID2, ELF5* and *EOMES* are not expressed at all in human preimplantation embryos [98]. Similarly in cattle and pigs, *ELF5* and *EOMES* are not expressed in trophoblast before the blastocyst has hatched [102-104]. While *ID2* is initially ubiquitously expressed in cattle embryos, at late blastocyst stage it is found in epiblast and hypoblast rather than trophoblast [95].

4.3.1. CDX2

Notably though, unique CDX2 expression in preimplantation trophoblast is highly conserved, having been recorded in humans [105], monkeys [96], cattle [106], pigs [107], rabbits [108] and even in the opossum, a marsupial mammal [78]. Yet the onset of Cdx2 expression differs among species, suggesting that its ancestral function lies not in the specification, but rather the stabilisation/survival of the trophoblast lineage. This is borne out by loss of function experiments. In the mouse, loss of Cdx2 prevented the upregulation of trophoblast-specific genes and affected blastocyst hatching and trophoblast cell number yet trophoblast was formed [109-112]. Similarly, in rhesus monkeys, cattle and pigs, Cdx2 knock-down still allowed trophoblast formation, but led to proliferation and hatching defects and downregulation of trophoblast-specific genes [81, 107, 113-115].

In mice *Cdx2* is activated at the 8-cell stage primarily via the Yap-TEAD4 pathway, with Yap1 activated via a mechanism involving the apical domain of polarised blastomeres or potentially via mechanosensing (see previous section). However redundant mechanisms appear to exist. Firstly, mouse embryos lacking Tead4 can still activate Cdx2 under conditions where Tead4's unique metabolic functions are bypassed [75]. Potentially Tead1 and/or Tead2, which (i) are also expressed at the 8-cell stage [116], (ii) have identical DNA binding properties to Tead4 [117] and (iii) are able to interact with Yap [118], can compensate for Tead4 under these conditions. Secondly, *Cdx2* activation in the mouse involves Tfap2c protein, which can directly regulate *Cdx2* transcription via an enhancer located in its first intron [119]. Tfap2c also affects components of the apical domain, thereby indirectly (via YAP) activating *Cdx2* [119]. Thirdly, the Gata3 transcription factor regulates *Cdx2* transcription via the same intronic *Cdx2* enhancer [120]. Fourthly, Notch signalling can co-operate with Yap in activating a second trophoblast enhancer (termed TEE) located upstream of the *Cdx2* gene [121].

Yet unlike the early, pre-morula activation of Cdx2 in mice, CDX2 protein is seen in outer cells only from blastocyst stages in rabbit, pig, cattle and human. This is well after polarisation and differential Yap signalling, suggesting that the key role of YAP activation in the onset of *CDX2* transcription may not be conserved. The gene networks activated by CDX2 similarly may be quite distinct in different mammals. For example, the mouse Cdx2 targets *Eomes* and *Elf5*, critical for mouse trophoblast maintenance, are not expressed early on in numerous other mammals. Secondly, the downregulation of *OCT4* (which drives the ICM/pluripotency GRN) by CDX2 is seen in the mouse [110, 111, 122], but not in cattle [81]. On the other hand, activation of *BMP4* may be a more common function of this transcription factor. Cdx2 binds to a trophoblast-enhancer of *Bmp4*, activating transcription in mouse trophoblast stem cells [123]. Furthermore, in pig blastocyst embryos, modulation of CDX2 levels affected *BMP4* transcription [107]. Thus while CDX2 is likely to be a pan-mammalian master regulator of the trophoblast lineage and required for its maintenance, the mechanism of its activation as well as of its actions appear to have diverged among mammals.

4.3.2. GATA2, GATA3

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

In mice Gata2 and 3 are both expressed from cleavage stages in all cells but become restricted to the outer, trophoblast cells in late blastocysts [100]. Similar to *Cdx2*, *Gata3* requires Tead4 for its activation [124]. Gata2 and 3 are at least partially redundant, being required for activation of trophoblast specific genes such as *Cdx2* and the efficient formation of blastocysts and, at later stages, for the correct differentiation of trophoblast cells [100]. In cattle there is in vitro evidence using a trophoblast cell line, that GATA2 and 3 may affect the expression of trophoblast genes including that coding for the ruminant pregnancy recognition signal, Interferon-tau [125]. Cattle and human GATA2 and 3 are expressed in the trophoblast of late blastocysts though GATA3 protein in both species could also be detected at lower levels in the presumptive hypoblast [94, 95, 98, 126].

4.3.3 TFAP2a, TFAP2c

The Tfap2 family of transcription factors have a key role in the specification and maintenance of the trophoblast lineage, based on their expression and knock-out phenotype in mouse embryos and trophoblast stem cells [127, 128]. In mice Tfap2c becomes restricted to the trophoblast in early blastocysts and appears to be involved in lineage specification via its effects on *Cdx2* expression, polarisation of outer cells and modulation of the Hippo (Yap) signalling pathway [119]. Notably though, TFAP2c is not trophoblast-specific in human, Cynomolgus monkey, cattle and pig embryos, where it is also highly expressed in prospective epiblast cells [95, 96, 98, 129] and has a later role in mesoderm and primordial germ cell specification [130]. In contrast, TFAP2a is specifically expressed in trophoblast cells in both mouse and cattle blastocyst embryos [93, 95]. TFAP2a appears to have overlapping functions with TFAP2c [127], raising the possibility that different mammals may have substituted TFAP2a for TFAP2c as the key gene for trophoblast specification.

We can conclude firstly that current results point to a conserved trophoblast-lineage GRN dominated by CDX2, GATA2/3 and TFAP2a/c. Secondly, the initiation of this core network in mammals other than mice is still unclear. Thirdly, downstream targets of the trophoblast-GRN appear to have diverged, which is in line with the different implantation strategies and timings among mammals.

4.4. The pluripotent inner cell mass lineage

The ICM gives rise to all cells of the conceptus with the exception of trophoblast [131] and thus the term "*pluripotent*" was coined to reflect this difference from the *totipotent* cleavage stage blastomeres. Morphologically and functionally, the ICM is a very transitory state in that its constituent cells rapidly assume either an epiblast or hypoblast identity [132]. In mice [93], humans [92, 97] and Cynomolgus monkeys [96], the trophoblast:ICM split in GRN-identity precedes that between the ICM derivative lineages by about one division (Fig. 3G,H). The transitory existence of the ICM may explain why the ICM state has not been captured in stem cell lineages. Embryonic stem cells (ESC), isolated from mouse ICM, were shown to be pluripotent [133], however their contribution to hypoblast was minimal [134], indicating that functionally, mouse ESC represent epiblast as opposed to bipotential ICM. This was verified recently in detailed analyses of the derivation and GRNs of such ESC cells [99, 135]. This potential of a cell to contribute to all three embryonic germ layers and the germline, but not to hypoblast, is now termed "*naïve*" pluripotency or "*ground-state*" [136]. Primate ESC cultured under the original "standard-ESC" conditions (in the presence of Activin and FGF) differ from mouse ESC in that they correspond to a *late* "*primed*" stage of epiblast development [96], resembling epiblast undergoing epithelialisation. From 2013, human "naïve" ESC of an early epiblast character have been derived [137, 138], reviewed in [139].

4.5. The mouse ICM-epiblast gene regulatory network

The regulatory network stably sustaining mouse naïve ESC has been refined to a core set of 12 factors including the transcription factors Oct4 (Pou5f1), Sox2, Nanog, Sall4, Klf2, Klf4, Esrrb, Gbx2 and Tfcp2l1 [140]. Apart from Gbx2, all of these transcription factors are expressed in compacted morulas,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

mid-blastocyst ICM and late-blastocyst epiblast [99], suggesting a key role in ICM and epiblast lineage establishment and maintenance. This is supported by a wealth of data:

1) *In vivo knock-out experiments*: Oct4-deficient embryos develop to the early blastocyst stage but the inner cells stop expressing some epiblast (however Nanog is upregulated) and hypoblast markers, and instead start expressing trophoblast markers. Subsequently, all ICM–derived tissue is lost [141, 142]. Sox2, while not required for the initial specification of ICM and epiblast, is critical for maintaining epiblast identity, including continued *Oct4* and *Nanog* expression [47, 143]. Loss of Nanog led to normal early E3.5 blastocysts but subsequent loss of epiblast, with blastocyst ICM-culture outgrowths forming only hypoblast [144, 145]. Sall4 is required for both ICM and hypoblast derivation [146]. Some of the factors though did not appear to be involved in lineage decisions in this in vivo functional assay: double knock-outs of the closely related Klf2 and 4 genes [147] or of Tfcp2l1 [148], Esrrb [149] or Gbx2 [150] led to no impairment of early development in mice.

2) *An early differential expression in inner cells*: *Sox2, Nanog, Klf2* and *Essrb* are among the first genes seen to be uniquely expressed in inner cells of 16-24-cell morulas [93].

3) *Downregulation of the trophoblast GRN*: for example, Oct4 [122] and Nanog [151] directly repress the key trophoblast gene *Cdx2*.

4) *Pluripotent reprogramming ability*: Overexpression of a cocktail of genes has been shown to be able to reprogram somatic cells to a naïve pluripotent state (so called "induced pluripotent stem cells" or iPSC). The initial cocktail contained three of the core pluripotency factors - Oct4, Sox2, Klf4 – as well as c-Myc [152], Subsequently c-Myc was shown to be dispensable, and Nanog and Sall4 to aid, in the derivation of iPSC [153, 154].

4.6. Conservation of the ICM-epiblast pluripotency GRN

Is this mouse pluripotency/ICM/epiblast core transcription factor GRN conserved across mammals? Studies in primates, using naïve ESC and examining embryonic gene and protein expression, have highlighted differences, but also a large degree of conservation with mice. In marmoset and Cynomolgus monkeys, all mouse homologous naïve ESC core factors apart from KLF2 (and GBX2) are expressed in the epiblast lineage [96, 99]. In humans, OCT4, SOX2, NANOG, SALL4, KLF4 and TFCP2L1 are expressed, but ESRRB and KLF2 (and GBX2) are not [98, 137]. Instead of KLF2, humans and marmosets express KLF17 [98, 99], which potentially may substitute for KLF2 in an analogous fashion to zebrafish, where KLF17 and KLF2 have partially redundant functions during embryogenesis [155]. Primates and rodents are Euarchontogliran species (Fig. 1). The more distantly related cattle, which belong to the sister clade (Laurasiotheria), closely follow the primate pattern with *OCT4, SOX2, NANOG, SALL4* and *KLF4* expressed predominantly in the epiblast lineage at late blastocyst stages and *KLF2* absent and *ESRRB* more abundant in trophoblast cells [95]. In pigs, ESSRB is downregulated at the blastocyst stage though expressed in, and important for, the maintenance of pig induced pluripotent cells [156].

Hence the consensus at this stage is that, of the core pluripotency factors, GBX2 and KLF2 are mouse specific, while ESSRB expression is variable even between closely related species. The greater similarity in the expression of core naïve pluripotency transcription factors between primates and cattle compared to mice may explain the recent finding that chimera formation can be obtained by injecting naïve human ESC into cattle and pig, but not mouse blastocysts [157]. Notably, the OCT4-SOX2-NANOG triumvirate, the members of which have been shown to activate each other and co-regulate pluripotency targets as well as repress the trophoblast GRN in human primed ESC [158], have been detected specifically in the epiblast of all mammalian species examined, including in addition to the aforementioned species, rabbits [159], pigs [160] and two species of marsupials [78, 161]. The key role of these three factors is underlined by the observation that homologs and/or paralogs are involved in the generation of the initial pluripotent embryonic ground state not only in mammals, but also non-mammalian amniotes [162, 163] and even non-amniotic vertebrates such as fish and amphibians [164, 165]. The dual requirement for OCT4 and SOX2 is likely to stem from the fact that OCT4 heterodimerises well with SOX2, with the complex binding a unique set of targets that forms the most

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

common motif associated with pluripotency target genes [166]. Secondly, OCT4 and SOX2 (as well as KLF4) appear to be "pioneering" factors, able to bind target sites even when these are embedded in epigenetically silenced chromatin, thus acting as the transcriptional pioneers of the ICM/epiblast (pluripotent) lineage [167, 168]. The widespread use of NANOG for establishment of the pluripotency network may seem surprising considering its low sequence conservation outside its DNA-binding homeodomain (for example, mouse and human NANOG are only 54% identical). Remarkably, a 70 amino acid fragment comprising only the homeodomain (a quarter of the protein) was sufficient to induce naïve pluripotency in mouse cells [165]. This fragment, similar to non-mammalian NANOG proteins, does not contain the WD domain responsible for cooperative interactions with core pluripotency factors OCT4 and SALL4 [169], suggesting that a critical ancestral subset of NANOG interactions encased in its unique homeodomain has been co-opted to achieve mammalian pluripotency.

While non-mouse mammalian germline knockout models for genes have been rare, CRISPR technology is beginning to overcome this limitation and has recently allowed the introduction of deletions in the *OCT4* gene in all cells of about half the human embryos injected [170]. This study revealed an earlier role for OCT4 in human embryos compared to mouse embryos. Similar to mice, expression of OCT4 target genes, including epiblast and hypoblast markers was lost, but unlike mouse *Oct4* knockout embryos, *NANOG* expression was lost as well and trophoblast genes (including *CDX2* and *GATA2*) were downregulated instead of upregulated, with a concomitant failure to form and/or maintain blastocyst stage embryos. The human OCT4 loss-of-function phenotype resembles the CRISPR-induced cattle OCT4 knockout, achieved using somatic cell nuclear transfer embryos. These embryos, in which maternal OCT4 persisted to morula stages, initially switched on NANOG, but at blastocyst stages lost NANOG expression [171]. Other lineage markers for all three lineages were downregulated as in the human embryos. siRNA Oct4-depleted cattle embryos also exhibited reduced expression of the trophoblast gene *CDX2*, as well as of the OCT4 target gene *FGF4*, while *NANOG* expression was reduced, albeit not significantly [115].

The accessibility of (non-pioneering) transcription factors is modulated by epigenetic modifiers which affect the methylation state and chromatin state of DNA. Conversely, many transcription factors can affect the localised epigenetic state of DNA via their ability to tether chromatin or DNA modifiers. Recently TET1 (a factor involved in active DNA demethylation) and THAP11/RONIN, a DNA binding factor recruiting epigenetic modifiers, were shown to be exclusively expressed in human ICM and, together with MCRS1 (a factor involved in chromatin remodelling) to be sufficient to reprogram fibroblasts into naïve ESC [92]. Tet1 and Thap11 are also essential in mice for ICM development as proven by in vivo functional studies [172, 173]. TET1 is also expressed in cattle blastocysts, but ubiquitously [95]. Other pluripotency-associated DNA binding factors that function predominantly in attracting epigenetic modifiers such as PRDM14, DPPA2 and DPPA4 are expressed specifically in the ICM and epiblast of mice, humans and marmosets [98, 99]. In cattle PRDM14 too is specific to the ICM/epiblast though DPPA2 is not [95]. However, most epigenetic modifiers (TET1 being an exception) do not show a lineage specific distribution [32], suggesting that epigenetic modifiers may have more of an indirect function in lineage network establishment.

It can be concluded that the ICM lineage leading to epiblast (naïve pluripotency) is highly conserved with mice showing some differences that may be linked to this species' requirement to more rapidly fully separate the trophoblast from the ICM lineages to accommodate precocious implantation [81].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

5. The third lineage (hypoblast)

The third lineage to arise in the embryo is the hypoblast (termed primitive endoderm in mice). In most eutherian mammals, hypoblast progenitors arise within the inner cell mass and subsequently cover the blastocoel surface of the epiblast to eventually line the entire surface of the blastocoel. ICM cells in mouse [174], rabbit [159], human [105], marmoset [99] and cattle [175] embryos initially co-express epiblast and hypoblast markers, but at later blastocyst stages this expression resolves into a "salt and pepper" pattern, where hypoblast and epiblast progenitor cells, now identified via exclusive expression of one or the other lineage marker, are interspersed in a seemingly random fashion.

5.1. The mouse hypoblast gene regulatory network (GRN)

In the mouse there is a definite sequence of hypoblast lineage marker activation (and/or concurrent shut-down in the alternate epiblast lineage) of Gata6 (8-cell), followed by Pdgfra (16-cell), Sox17 (32-cell), Gata4 (58-cell) and Sox7 (>64-cell) [176, 177]. Earlier markers are expressed more ubiquitously than later ones and, by the > 64-cell late-blastocyst stage, all become confined to ICM cells that have downregulated the initially ubiquitously expressed epiblast marker Nanog. Gata6 lies at the top of the hypoblast GRN network, in that loss of this gene prevents the activation of all subsequent hypoblast markers. Even only a mild reduction in Gata6 levels (in *Gata6*+/- heterozygous embryos) is sufficient to reduce the number of hypoblast progenitor cells, causing a delay in hypoblast specification [178, 179]. Furthermore *Gata6* overexpression is able to reprogram ES cells into hypoblast stem ("XEN") cells [180]. Gata6 is initially found in all blastomeres, then becomes progressively restricted to the subset of ICM-cells that will form the hypoblast [176]. Hence the pertinent question regarding hypoblast lineage determination is how Gata6 expression is *maintained* in prospective hypoblast cells, while concomitantly shut down in prospective epiblast cells. This transition is asynchronous, occurring heterogeneously in individual cells between the early (32-cell) and late (120-cell) blastocyst stages, such that over time more and more cells have transited from a Gata6-Nanog double positive state to expressing one or the other marker [181]. Subjecting single positive cells to an alternate signalling environment did not change their fate, indicating that lineage commitment is achieved once an ICM cell expresses Nanog or Gata6 in a mutually exclusive fashion [181, 182].

5.2. FGF signalling in the mouse hypoblast/epiblast lineage decision

The mutually exclusive Gata6/Nanog pattern appears to be established predominantly via cell-cell signalling mediated by FGF. Fgf4 is necessary for the specification of the hypoblast, as shown in vivo by loss of function experiments of Fgf4 [183, 184], both receptors Fgfr1 and Fgfr2 [185, 186], as well as the downstream component Grab2, which is required for FGF-mediated Ras-Mapk signalling [174]. In all these mutant embryos, Gata6 expression is lost and all inner cell mass cells express Nanog to adopt an epiblast (naïve pluripotency) fate. Similarly, chemical inhibition of FGF receptors (with PD173074) or of the downstream kinases MAP2K1/2 (=ERK1/2; inhibited with PD0325901) could, in a reversible fashion, direct ICM cells to an epiblast-only fate if applied from the 32-cell early blastocyst stage [187], that is, before any cells show a reciprocal expression of Gata6 and Nanog. Conversely, exposure of embryos to exogenous FGF could direct Nanog/Gata6 double positive ICM cells to a hypoblast fate in a dosage dependent fashion [181, 183, 187]. This instructive role of FGF in directing ICM cells to a hypoblast fate is mediated predominantly by Fgfr1 [185, 186] and acts via Gata6, as FGF treatment could not rescue the hypoblast defect seen in *Gata6*-/- embryos [178, 179].

Fgf4 expression is under dual control by Oct4 and Sox2 [47, 141, 143, 188, 189]. While Oct4 is expressed more widely, Sox2 is the limiting factor, being restricted first to the ICM then to the epiblast. Thus at the "salt and pepper" late-blastocyst stage, the Sox2/Oct4/(Nanog)-expressing epiblast progenitors are the predominant source of Fgf4. These cells secrete Fgf4 which mediates its effects in a paracrine fashion on the surrounding hypoblast progenitors via both Fgfr1 and the hypoblast-specific Fgfr2 to activate/maintain the later hypoblast markers Sox17, Pdgfra and Gata4 [145, 185, 186]. Yet at the 34-50 cell early blastocyst stage, when Sox2 (and Oct4) show no lineage-specific restriction *within*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

the ICM, differential *Fgf4* mRNA expression is already seen [93, 190] and may be selectively inducing hypoblast differentiation in the most responsive surrounding cells. It is still unclear how this early differential expression in *Fgf4* expression arises. On the one hand, intrinsic stochastic fluctuation in expression levels could generate cells that by chance express either more FGF signal or a better response to FGF signalling, leading to respective biases toward the epiblast or hypoblast state. Such biases could subsequently be stabilised by signal reinforcement [190]. Alternatively, the bias in FGF signalling could be imparted by the history of the blastomeres. The reasoning is as follows. Inner apolar cells are not only generated from polar outside cells at the 8 cell stage, but also during the subsequent one or two rounds of outer cell divisions. It has been suggested that cells internalising at these later time points may be biased to form hypoblast [39]. Initial studies using lineage tracing of microinjected outer blastomeres did not support the hypothesis [187], but other studies using non-invasive tracing did find such a bias [191, 192]. It could be argued that the later internalising cells had been subject to more Yap signalling which prevents *Sox2* induction (see previous section; [47]), thus resulting in a delay in Fgf4 synthesis. Secondly, cells remaining on the outside may have accumulated more Fgfr2 protein as this gene is progressively upregulated specifically in outer cells between the 16 and 32-cell morula stages [186, 192]. Temporary retention of Fgfr2 after internalisation may have sensitised these cells to FGF signals. As internalised cells are able to move within the ICM [187, 191], such hypoblast-biased cells would disperse, creating a subsequent random salt and pepper pattern equivalent to the stochastic model. Interestingly, while in mice extended residency in the outer trophoblast fated environment may bias inner cells toward the hypoblast lineage, such a mechanism is less disputable in the distantly related (Afrotherian) elephant shrew. In these mammals, which do not transit through a morula stage, hypoblast progenitors delaminate from the outer cells only *after* the ICM has formed and assemble on the ICM surface to directly form a distinct hypoblast layer (see panels 18-20 in [4]). It thus appears that their extended residency in the outer, future trophoblast, layer strongly biased them toward the hypoblast lineage allowing them to bypass a fate-refinement period within the ICM. In marsupials, hypoblast cells may also require extended exposure to a trophoblast-like environment in that the majority of hypoblast cells are seen to arise on the margin of the pluripotent (ICM-equivalent) disc, where they are in close proximity to the abutting trophoblast [161].

How the initial stochastic or ontogenic FGF-signalling induced bias in fate becomes stabilised in mice is not quite clear either. One mechanism may involve reciprocal negative feedback of Gata6 and Nanog on each other's transcription to amplify initially subtle differences in expression. In vivo evidence is that in Nanog-deficient embryos, Gata6 is upregulated [145] and in Gata6-deficient embryos, Nanog is upregulated [178, 193]. In vitro results suggest that these effects may be direct, as Nanog can bind the *Gata6* enhancer in ESC cells to downregulate expression [194], whereas Gata6 overexpression in ESC cells downregulates *Nanog* within 12 hours of induction and Gata6 binding was enriched upstream of the *Nanog* gene [180]. A second mechanism may relate to different downstream effectors mediating the FGF response in epiblast and hypoblast progenitors, although the details are yet to be worked out [185, 186].

5.3. A common hypoblast gene regulatory network (GRN)

Thus for the mouse two key elements in hypoblast formation are FGF signalling and the central role of Gata6 at the top of the hypoblast GRN. How well are these elements conserved in mammals? Comparing the hypoblast GRNs, it appears that the sequentially activated and progressively refined mouse hypoblast markers GATA6, PDGFRA, SOX17 and GATA4 not only are all expressed in the hypoblast, but also appear in a similar temporal order in rabbit, human, old and new world monkeys, pig and cattle embryos (references as in Fig. 4 as well as [98, 99, 195, 196]). The progressive reciprocal restriction of expression of the epiblast markers NANOG and SOX2 relative to the hypoblast markers GATA6, SOX17 and GATA4 in mouse late blastocysts is also seen in the other eutherians at the late blastocyst stage, just prior to hatching (Table 2, Fig. 4). GATA6 and NANOG have been examined in six species and invariably are detected ubiquitously in most cells from the earliest blastocyst stage, to then progressively resolve into hypoblast and epiblast domains respectively, with little to no overlap

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

in expression upon blastocyst hatching (Fig. 4). The eventual segregation of NANOG and GATA6 to epi and hypoblast lineages is even detected in marsupials [161]. There are though subtle differences in the regulation of GATA6 and NANOG. In the rabbit, the downregulation of NANOG in prospective hypoblast cells is delayed by one cell cycle, resulting in a transitory period where NANOG is still expressed in all ICM cells, whereas GATA6 downregulation in prospective epiblast cells has already commenced [159]. This suggests that if a direct inhibitory action of GATA6 and NANOG on each other's transcription exists, it is likely not important for rabbits. Secondly, GATA6 expression is rapidly shut off in trophoblast cells in mice, but maintained to mid-blastocyst stages in humans and to late-blastocyst stages in rabbits, Cynomolgus monkeys and cattle (Fig. 4). The extended expression of the early-hypoblast marker GATA6 in the outer prospective trophoblast layer may well be indicative of an extended potential of these cells to give rise to hypoblast, in analogy to the elephant shrew mode of hypoblast formation discussed previously. Such extended fate plasticity has indeed been demonstrated for cattle [81].

Figure 4. NANOG and GATA6 protein distribution in early (E), mid (M), late (L) and hatched blastocysts of various mammals. Lineages are colour coded: blue being trophoblast, red epiblast and yellow hypoblast. Blastocyst stages E, early, M, mid, L, late and H, hatched defined as per Table 2. Note extended expression of GATA6 in rabbit, old world monkey and cattle trophoblast at late blastocyst stages and extended maintenance of NANOG in rabbit prospective hypoblast at the late blastocyst stage. Based on data from the following references: mouse [47, 176, 177], rabbit [159], human [105, 197, 198], cynomolgus old world monkey [96] and cattle [175, 199].

Table 2. Characteristics of early to hatched blastocyst stages in mammals [1]

Species:	mouse	rabbit	human	cynomolgus	cattle

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

Early (cavity	From 32 cells	From 64 cells	From 35 cells	From 50 cells	From 64 cells
visible, <30% vol)	E3.25	E3 ("VI")	E4 - early E5	E5 - 6	E6
Mid (ca 30-70%)	<64 cells	>128 cells	64 – 100 cells		100-130 cells
	E3.5	E3.25 ("VII")	late E5		E7
Late (max cavity	>64 cells	>256 cells	128-256 cells	200-300 cells	140-200 cells
zona enclosed)	E3.75	E3.5 ("VIII")	early E6	E7-8	E7
Hatched (Hypo	>100 cells	>512 cells	> 256 cells	300-600 cells	>250 cells
forming layer)	E4.25	E3.75 ("IX")	late E6	E8-9	E8

Table footnote: [1]Based on descriptions from references: mouse [176], rabbit [159], human [105], cynomolgus [96] and cattle [62, 200].

5.4. FGF signalling in other mammals

It appears thus that overall the hypoblast regulatory network is very similar among all mammals examined to date. Is FGF signalling though universally involved in the establishment of this network? In rabbits, which are closely related to rodents, (i) exogenous FGF4 treatment transformed nearly all ICM cells to SOX17-expressing hypoblast progenitors and (ii), inhibition of FGF signalling via a MAP2k1/2 inhibitor prevented the formation of SOX17 positive hypoblast [159]. However, subtle differences existed: FGF pathway inhibition in rabbit embryos did not lead to an increase in the number of NANOG positive cells as in the mouse, but rather an increase in apoptosis. Secondly, GATA6 expression was not lost upon FGF pathway inhibition [159]. This difference to mice is more extreme in human embryos which were shown to be refractory to FGF pathway inhibition (via FGFR or MAP2K1/2 inhibitors) in regards to GATA6 and NANOG marker expression [175, 197]. Marmoset monkey embryos responded uniquely to FGF pathway inhibition by maintaining expression of GATA6 as well as the trophoblast marker CDX2 in most cells, while NANOG expression was normal [99]. In the Laurasiatherian cattle and pig embryos, exogenous FGF could drive ICM conversion to GATA4/6-positive hypoblast as in mice and rabbits, but MAP2k1/2 inhibition resulted only in a partial conversion of hypoblast to epiblast progenitors as monitored via NANOG and GATA4/6 expression [175, 196, 201]. Interestingly, use of inhibitors targeting the FGF receptors as opposed to the downstream MAP2k1/2 had no effect on GATA4/6 expression in either species [175, 196], but did affect the total number of ICM cells specifically in pig blastocysts [196]. This indicates a role for FGF signalling in pig ICM proliferation, but not GATA6 maintenance.

When considering these gain and loss of function experiments, it must be borne in mind that activation of hypoblast markers was achieved using extremely high concentrations of FGF4, usually 1000 ng/ml. This should be contrasted to the more physiological 25 ng/ml that is required to sustain FGF-dependent mouse trophoblast cells in culture [202]. Thus hypoblast marker activation via exogenous FGF cannot be taken as proof that FGF4 is the endogenous lineage-determining ligand. This places more emphasis on the FGF pathway inhibitor results. In mice, use of such inhibitors led to a loss of Gata6 and downstream hypoblast gene expression, with all cells expressing epiblast markers such as Nanog and Sox2. As listed, this conversion to epiblast was at best only partially evident in the other mammalian species. The conclusion is that while FGF signalling may be sufficient for hypoblast specification, it is not necessary for this purpose in numerous mammalian species.

5.5. Alternative signalling

The different phenotypes seen in various mammals when inhibiting FGF receptors (no effect) as opposed to the main downstream FGF signalling pathway (MAP2K1/2: some effects) could be explained by alternate signalling ligands and receptor tyrosine kinases that mediate their effects through MAP2K1/2. PDGFA is also able to signal through MAP2K1/2, however while PDGFRA is one of the earliest hypoblast markers, PDGF signalling is not involved in hypoblast specification in mice, but rather is required for prevention of apoptosis after establishment of the hypoblast lineage [203, 204].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

Furthermore, treatment of mouse embryos with exogenous PDGF-A (500ng/ml) is unable to mimic the hypoblast-inducing effects seen with FGF [203]. Similarly, treatment of cultured mouse [205], human [206], cattle [207, 208] or rabbit [209] embryos with other ligands able to activate MAP2K1/2 signalling, such as EGF and IGF1, have shown improvements in survival rates and increases in ICM cell number. However these studies did not examine the hypoblast to epiblast ratio. Inhibiting most receptor tyrosine kinases (VEGFR, EGFR, PDGFR, FGFR) with the broad-spectrum inhibitor BI-BF1120 resulted in minimal changes in *NANOG* and *SOX17* and no effect on *SOX2* and *PDGFRA* mRNA expression in cattle embryos [210]. Hence in mammals other than mice and rabbits, the signals leading to MAP2K1/2 activation and the in vivo importance of such signalling in hypoblast specification are still unclear.

What about other signalling pathways? In cynomolgus monkeys *NODAL* is expressed in prospective epiblast cells [96]. In human embryos, genes of the NODAL/ACTIVIN/TGFβ pathway (the ligands NODAL and GDF3, the receptor ALK5/TGFBR1, NODAL-coreceptor TDGF1/CRIPTO, intracellular mediators SMAD2 and 4 and the target LEFTY1) are enriched in prospective epiblast cells at the late blastocyst stage [98]. Inhibition of this pathway using the ALK4/5/7 inhibitor SB431542 abolished NANOG expression indicative of a failure in epiblast establishment. However, expression of the hypoblast marker SOX17 was also lost in most embryos and levels of the ICM marker OCT4 reduced [98]. This argues for a role of NODAL-pathway signalling in human epiblast maintenance as opposed to hypoblast/epiblast lineage decisions. The lack of NODAL involvement in second lineage determination was also seen in marmoset monkey [99], cattle [175] and pig [196] embryos, where NODAL pathway inhibition via either SB431542, or the even more potent A83-01 inhibitor, had no effect on the epiblast to hypoblast cell number ratio. In mice, which differ from the other mammals by expressing Nodal only later, namely in mature epiblast (E5.5, corresponding to "primed" pluripotency) [96], SB431542 treatment had no effect on *Nodal*, *Sox17* or *Oct4* expression at late blastocyst stages [98]. Thus in mammals, the role of NODAL signalling appears not to be involved in lineage decision, but may affect subsequent epiblast maintenance.

The third signalling pathway of potential influence is that of WNT. Canonical WNT signalling prevents β-Catenin degradation, allowing it to translocate to the nucleus to displace repressors from TCF/LEF sites so as to activate WNT target genes [211]. Among these target genes in naïve ESC are *Nanog*, *Klf2* and *Essrb* [212]. Activation of Wnt signalling by preventing β-Catenin degradation through inhibition of GSK3β with the chemical CHIR99021 is required for the maintenance of mouse and human pluripotent stem cells in vitro [213, 214]. However, in mice it has been conclusively shown that in embryos, WNT signalling is only required from gastrulation stages onward and is dispensable during preimplantation and lineage determination [215]. In cattle, treatment with 3μM CHIR99021 from zygote, but not from morula, stages resulted in slightly more NANOG cells with no change in the number of GATA6 positive hypoblast cells [175], indicating at most a minor effect of this pathway. Treatment of cattle embryos with DKK1 protein, which is an endogenous canonical WNT pathway inhibitor, resulted in a reduction in the number of blastocyst ICM cells, with a decrease in the epiblast (NANOG) to hypoblast (GATA6) cell ratio [199], suggesting that WNT signalling may be required specifically for prospective epiblast cell survival or proliferation. In marmoset monkey early blastocysts, the WNT signalling pathway may be specifically inhibited as low expression of *β-Catenin* in combination with high expression of the WNT inhibitor *DKK1* and the β-Catenin destabiliser *GSK3β* genes was noted (Boroviak15). This inhibition appeared to be relieved during ICM lineage decision stages as DKK1 expression was diminished at the late blastocyst stage. Maintaining WNT signal inhibition via CHIR99021 from the early blastocyst led to a strong upregulation of the blastomere protein levels of NANOG, GATA6 as well as the trophectoderm marker CDX2 in late blastocysts. This pattern was similar to the coexpression of these markers seen before lineage segregation, at early blastocyst stages, leading the authors to infer a role for WNT signalling in marmoset ICM lineage determination (Boroviak15). In human embryos, functional data is not available, however *WNT3* mRNA increases sharply from early to late blastocyst stages, where it is restricted to the epiblast [97].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

5.6. The third lineage - conclusion

The picture that emerges from the available data is that the GRN driving the hypoblast lineage is well conserved across eutherian mammals and may extend to marsupials as well. This GRN is driven by GATA6 with successive deployment of PDGFRA, SOX17 and GATA4 with concomitant exclusion of the key epiblast genes NANOG and SOX2 while OCT4 expression is maintained. However the deployment of the hypoblast GRN has diverged in that FGF signalling is strongly implicated in the mouse and rabbit, but less so in primates and Laurasiatherians. No clear candidates substituting for FGF's role have emerged and it may be that in these species the timing of internalisation and thus the length of exposure in the trophoblast precursor environment, is important for generating the initial differences in ICM cells that will bias cells to either the hypoblast or epiblast lineage. I have discussed some data indicating that such a bias may also exist in mice though its influence may be subjugated in this species by FGF dependent mechanisms.

6. Concluding remarks

In summary, early mammalian development up to the hatched blastocyst stage is driven by a very limited set of building principles. The arguably most important ones can be listed as follows:

1. An intrinsic trigger to switch on the embryonic gene expression program. This trigger (DUX-C) is nearly fail-proof thanks to being present in the genome in high copy numbers.
2. The use of inherent random fluctuations (noise) in the gene expression machinery to generate asymmetries between blastomeres, which is likely to play a part in biasing cells during the first and second lineage decisions.
3. The adaptation of basic cellular processes (polarisation, compaction - as seen during mesenchymal to epithelial transitions) to asymmetrically segregate lineage specifiers during subsequent cell divisions.
4. Amplification of small differences in GRN-biases via reciprocal inhibition between alternative GRN programs. For the first lineage decision such inhibition is achieved predominantly through a small set of master transcriptional regulators, for the second decision additional control is achieved through the use of diffusible signalling molecules.

As already anticipated over 2000 years ago in Aristotle's epigenesis idea, each step during the building of the blastocyst is dependent on the prior one and as new lineages and distinct cell populations form, new avenues for autonomous interactions between these cell populations arise, leading to the ever-increasing complexity of the mammalian embryo.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

7. References

1. Woodger, J., On biological transformations: Pp. 95–120 in E. Le Gross Clark and PB Medawar, eds. Essays on growth and form presented to D'Arcy Wentworth Thompson. **1945**, Clarendon Press, Oxford.
2. Foley, N.M., M.S. Springer, and E.C. Teeling, Mammal madness: is the mammal tree of life not yet resolved? *Philosophical Transactions of the Royal Society B: Biological Sciences,* **2016**. *371*(1699).
3. Wimsatt, W.A., Some comparative aspects of implantation. *Biol Reprod,* **1975**. *12*(1): p. 1-40.
4. Van der Horst, C.J., Early stages in the embryonic development of *Elephantulus. S. Afr. J. Med. Sci.,* **1942**. *7*(Biol. Suppl.): p. 55-67.
5. Goetz, R.H., Early development of the Tenrecoidea. *BioMorphosis,* **1938**. *1*: p. 67-79.
6. Graf, A., et al., Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Anim Reprod Sci,* **2014**. *149*(1-2): p. 46-58.
7. Nagaraj, R., et al., Nuclear Localization of Mitochondrial TCA Cycle Enzymes as a Critical Step in Mammalian Zygotic Genome Activation. *Cell,* **2017**. *168*(1): p. 210-223.e11.
8. Khurana, N.K. and H. Niemann, Energy metabolism in preimplantation bovine embryos derived in vitro or in vivo. *Biol Reprod,* **2000**. *62*(4): p. 847-56.
9. Gardner, D.K. and A.J. Harvey, Blastocyst metabolism. *Reprod Fertil Dev,* **2015**.
10. Liang, H.L., et al., The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature,* **2008**. *456*(7220): p. 400-3.
11. Whiddon, J.L., et al., Conservation and innovation in the DUX4-family gene network. *Nat Genet,* **2017**. *49*(6): p. 935-940.
12. De Iaco, A., et al., DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet,* **2017**. *49*(6): p. 941-945.
13. Hendrickson, P.G., et al., Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Genet,* **2017**. *49*(6): p. 925-934.
14. Choi, S.H., et al., DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res,* **2016**. *44*(11): p. 5161-73.
15. Leidenroth, A., et al., Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma,* **2012**. *121*(5): p. 489-497.
16. Perrod, S. and S.M. Gasser, Long-range silencing and position effects at telomeres and centromeres: parallels and differences. *Cell Mol Life Sci,* **2003**. *60*(11): p. 2303-18.
17. Abe, K., et al., The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3' processing. *Embo j,* **2015**. *34*(11): p. 1523-37.
18. Wu, J., et al., The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature,* **2016**. *534*(7609): p. 652-657.
19. Ooga, M., et al., Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics,* **2016**. *11*(1): p. 85-94.
20. Peaston, A.E., et al., Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell,* **2004**. *7*(4): p. 597-606.
21. Jouhilahti, E.-M., et al., The human PRD-like homeobox gene LEUTX has a central role in embryo genome activation. *Development,* **2016**. *143*(19): p. 3459-3469.
22. Xue, Z., et al., Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature,* **2013**. *500*(7464): p. 593-7.
23. Liu, L., et al., Telomere lengthening early in development. *Nat Cell Biol,* **2007**. *9*(12): p. 1436-41.
24. Zalzman, M., et al., Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature,* **2010**. *464*(7290): p. 858-63.
25. Seidel, F., Die Entwicklungspotenzen einer isolierten Blastomere des Zweizellenstadiums im Säugetierei. *Naturwissenschaften,* **1952**. *39*(15): p. 355-356.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

26. Tarkowski, A.K., Experiments on the development of isolated blastomers of mouse eggs. *Nature,* **1959**. *184*: p. 1286-7.
27. Willadsen, S.M., The development capacity of blastomeres from 4- and 8-cell sheep embryos. *J Embryol Exp Morphol,* **1981**. *65*: p. 165-72.
28. Fehilly, C.B. and S.M. Willadsen, Embryo manipulation in farm animals, in *Oxford Reviews of Reproductive Biology*, J.R. Clarke, Editor. **1986**, Clarendon Press: Oxford. p. 379-413.
29. Van de Velde, H., et al., The four blastomeres of a 4-cell stage human embryo are able to develop individually into blastocysts with inner cell mass and trophectoderm. *Hum Reprod,* **2008**. *23*(8): p. 1742-7.
30. Pfeffer, P.L., ed. Lineage commitment in the mammalian preimplantation embryo. Reproduction in Domestic Ruminants VIII, ed. J. Juengel, A. Miyamoto, and R. Webb. Vol. 8. **2014**, Context: Obihiro, Japan. 89-103.
31. Huh, D. and J. Paulsson, Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat Genet,* **2011**. *43*(2): p. 95-100.
32. Burton, A., et al., Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep,* **2013**. *5*(3): p. 687-701.
33. Torres-Padilla, M.-E., et al., Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature,* **2007**. *445*: p. 214.
34. Plachta, N., et al., Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat Cell Biol,* **2011**. *13*(2): p. 117-23.
35. White, Melanie D., et al., Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. *Cell,* **2016**. *165*(1): p. 75-87.
36. Shi, J., et al., Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell RNA-seq. *Development,* **2015**. *142*(20): p. 3468-77.
37. Goolam, M., et al., Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell,* **2017**. *165*(1): p. 61-74.
38. Biase, F.H., X. Cao, and S. Zhong, Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. *Genome Research,* **2014**. *24*(11): p. 1787-1796.
39. Yamanaka, Y., et al., Cell and molecular regulation of the mouse blastocyst. *Dev Dyn,* **2006**. *235*(9): p. 2301-14.
40. Ziomek, C.A. and M.H. Johnson, Cell surface interaction induces polarization of mouse 8-cell blastomeres at compaction. *Cell,* **1980**. *21*(3): p. 935-42.
41. Stephenson, R.O., Y. Yamanaka, and J. Rossant, Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. *Development,* **2010**. *137*(20): p. 3383-91.
42. Johnson, M.H., B. Maro, and M. Takeichi, The role of cell adhesion in the synchronization and orientation of polarization in 8-cell mouse blastomeres. *J Embryol Exp Morphol,* **1986**. *93*: p. 239-55.
43. Korotkevich, E., et al., The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse Embryo. *Developmental Cell,* **2017**. *40*(3): p. 235-247.e7.
44. Nishioka, N., et al., The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell,* **2009**. *16*(3): p. 398-410.
45. Hirate, Y., et al., Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. *Curr Biol,* **2013**. *23*(13): p. 1181-94.
46. Leung, C.Y. and M. Zernicka-Goetz, Angiomotin prevents pluripotent lineage differentiation in mouse embryos via Hippo pathway-dependent and -independent mechanisms. *Nature Communications,* **2013**. *4*: p. 2251.
47. Wicklow, E., et al., HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. *PLoS Genet,* **2014**. *10*(10): p. e1004618.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

48. Skamagki, M., et al., Asymmetric localization of Cdx2 mRNA during the first cell-fate decision in early mouse development. *Cell Rep,* **2013**. *3*(2): p. 442-57.
49. Humięcka, M., et al., Mouse blastomeres acquire ability to divide asymmetrically before compaction. *PLOS ONE,* **2017**. *12*(3): p. e0175032.
50. Strnad, P., et al., Inverted light-sheet microscope for imaging mouse pre-implantation development. *Nat Methods,* **2016**. *13*(2): p. 139-42.
51. Johnson, M.H. and C.A. Ziomek, The foundation of two distinct cell lineages within the mouse morula. *Cell,* **1981**. *24*(1): p. 71-80.
52. Anani, S., et al., Initiation of Hippo signaling is linked to polarity rather than to cell position in the pre-implantation mouse embryo. *Development,* **2014**. *141*(14): p. 2813-2824.
53. Watanabe, T., et al., Limited predictive value of blastomere angle of division in trophectoderm and inner cell mass specification. *Development,* **2014**. *141*(11): p. 2279-2288.
54. de Vries, W.N., et al., Maternal β-catenin and E-cadherin in mouse development. *Development,* **2004**. *131*(18): p. 4435-4445.
55. Fierro-González, J.C., et al., Cadherin-dependent filopodia control preimplantation embryo compaction. *Nature Cell Biology,* **2013**. *15*: p. 1424.
56. Samarage, Chaminda R., et al., Cortical Tension Allocates the First Inner Cells of the Mammalian Embryo. *Developmental Cell,* **2015**. *34*(4): p. 435-447.
57. Maître, J.-L., et al., Asymmetric division of contractile domains couples cell positioning and fate specification. *Nature,* **2016**. *536*(7616): p. 344-348.
58. Dupont, S., et al., Role of YAP/TAZ in mechanotransduction. *Nature,* **2011**. *474*: p. 179.
59. Davidson, E.H. and M.S. Levine, Properties of developmental gene regulatory networks. *Proc Natl Acad Sci U S A,* **2008**. *105*(51): p. 20063-6.
60. MacArthur, B.D., et al., Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. *Nat Cell Biol,* **2012**. *14*(11): p. 1139-1147.
61. Koyama, H., et al., Analysis of polarity of bovine and rabbit embryos by scanning electron microscopy. *Biol Reprod,* **1994**. *50*(1): p. 163-70.
62. Van Soom, A., et al., Timing of compaction and inner cell allocation in bovine embryos produced in vivo after superovulation. *Biol Reprod,* **1997**. *57*(5): p. 1041-9.
63. van der Horst, C.J., The placentation of Elephantulus. *Transactions of the Royal Society of South Africa,* **1949**. *32*(5): p. 435-629.
64. Barcroft, L.C., et al., Trophectoderm differentiation in the bovine embryo: characterization of a polarized epithelium. *J Reprod Fertil,* **1998**. *114*(2): p. 327-39.
65. Moriwaki, K., S. Tsukita, and M. Furuse, Tight junctions containing claudin 4 and 6 are essential for blastocyst formation in preimplantation mouse embryos. *Developmental Biology,* **2007**. *312*(2): p. 509-522.
66. Zenker, J., et al., Expanding Actin Rings Zipper the Mouse Embryo for Blastocyst Formation. *Cell,* **2018**. *173*(3): p. 776-791.e17.
67. Watson, A.J. and L.C. Barcroft, Regulation of blastocyst formation. *Front Biosci,* **2001**. *6*: p. D708-30.
68. Benos, D.J. and R.S. Balaban, Energy requirements of the developing mammalian blastocyst for active ion transport. *Biol Reprod,* **1980**. *23*(5): p. 941-7.
69. Leese, H.J., et al., Metabolism of the viable mammalian embryo: quietness revisited. *Mol Hum Reprod,* **2008**. *14*(12): p. 667-72.
70. Du, Z.F. and R.G. Wales, Glycolysis and glucose oxidation by the sheep conceptus at different oxygen concentrations. *Reprod Fertil Dev,* **1993**. *5*(4): p. 383-93.
71. Hewitson, L.C. and H.J. Leese, Energy metabolism of the trophectoderm and inner cell mass of the mouse blastocyst. *Journal of Experimental Zoology,* **1993**. *267*(3): p. 337-343.
72. Houghton, F.D., et al., Oxygen consumption and energy metabolism of the early mouse embryo. *Molecular Reproduction and Development,* **1996**. *44*(4): p. 476-485.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

73. Trimarchi, J.R., et al., Oxidative Phosphorylation-Dependent and -Independent Oxygen Consumption by Individual Preimplantation Mouse Embryos1. *Biology of Reproduction,* **2000**. *62*(6): p. 1866-1874.
74. Dumollard, R., et al., Mitochondrial function and redox state in mammalian embryos. *Semin Cell Dev Biol,* **2009**. *20*(3): p. 346-53.
75. Kaneko, K.J. and M.L. DePamphilis, TEAD4 establishes the energy homeostasis essential for blastocoel formation. *Development,* **2013**. *140*(17): p. 3680-3690.
76. Home, P., et al., Altered subcellular localization of transcription factor TEAD4 regulates first mammalian cell lineage commitment. *Proc Natl Acad Sci U S A,* **2012**. *109*(19): p. 7362-7.
77. Sakurai, N., et al., Effects of downregulating TEAD4 transcripts by RNA interference on early development of bovine embryos. *J Reprod Dev,* **2017**. *63*(2): p. 135-142.
78. Morrison, J.T., et al., Expression patterns of Oct4, Cdx2, Tead4, and Yap1 proteins during blastocyst formation in embryos of the marsupial, Monodelphis domestica Wagner. *Evolution & Development,* **2013**. *15*(3): p. 171-185.
79. Dyce, J., et al., Do trophectoderm and inner cell mass cells in the mouse blastocyst maintain discrete lineages? *Development,* **1987**. *100*(4): p. 685-98.
80. Copp, A.J., Interaction between inner cell mass and trophectoderm of the mouse blastocyst. II. The fate of the polar trophectoderm. *J Embryol Exp Morphol,* **1979**. *51*: p. 109-20.
81. Berg, D.K., et al., Trophectoderm lineage determination in cattle. *Dev Cell,* **2011**. *20*(2): p. 244-55.
82. Waddington, C.H., The strategy of the genes. A discussion of some aspects of theoretical biology. **1957**: London: George Allen & Unwin, Ltd. ix +-262 pp.
83. Rossant, J. and K.M. Vijh, Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. *Dev Biol,* **1980**. *76*(2): p. 475-82.
84. Tarkowski, A.K., et al., Individual blastomeres of 16- and 32-cell mouse embryos are able to develop into foetuses and mice. *Dev Biol,* **2010**. *348*(2): p. 190-8.
85. Suwinska, A., et al., Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos. *Dev Biol,* **2008**. *322*(1): p. 133-44.
86. Niakan, K.K., et al., Human pre-implantation embryo development. *Development,* **2012**. *139*(5): p. 829-841.
87. Rossant, J. and W.T. Lis, Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. *Dev Biol,* **1979**. *70*(1): p. 255-61.
88. Szczepanska, K., L. Stanczuk, and M. Maleszewski, Isolated mouse inner cell mass is unable to reconstruct trophectoderm. *Differentiation,* **2011**. *82*(1): p. 1-8.
89. Grabarek, J.B., et al., Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. *Development,* **2012**. *139*(1): p. 129-39.
90. De Paepe, C., et al., Human trophectoderm cells are not yet committed. *Hum Reprod,* **2013**. *28*(3): p. 740-9.
91. Guillomot, M. and P. Guay, Ultrastructural features of the cell surfaces of uterine and trophoblastic epithelia during embryo attachment in the cow. *Anat Rec,* **1982**. *204*(4): p. 315-22.
92. Durruthy-Durruthy, J., et al., Spatiotemporal Reconstruction of the Human Blastocyst by Single-Cell Gene-Expression Analysis Informs Induction of Naive Pluripotency. *Developmental Cell,* **2016**. *38*(1): p. 100-115.
93. Guo, G., et al., Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell,* **2010**. *18*(4): p. 675-85.
94. Petropoulos, S., et al., Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell,* **2016**. *165*(4): p. 1012-1026.
95. Wei, Q., et al., Bovine lineage specification revealed by single-cell gene expression analysis from zygote to blastocyst. *Biol Reprod,* **2017**. *97*(1): p. 5-17.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

96. Nakamura, T., et al., A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature,* **2016**. *537*(7618): p. 57-62.
97. Stirparo, G.G., et al., Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development,* **2018**. *145*(3).
98. Blakeley, P., et al., Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development,* **2015**. *142*(18): p. 3151-65.
99. Boroviak, T., et al., Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Developmental Cell,* **2015**. *35*(3): p. 366-382.
100. Home, P., et al., Genetic redundancy of GATA factors in the extraembryonic trophoblast lineage ensures the progression of preimplantation and postimplantation mammalian development. *Development,* **2017**. *144*(5): p. 876-888.
101. Deng, Q., et al., Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science,* **2014**. *343*(6167): p. 193-196.
102. Degrelle, S.A., et al., Molecular evidence for a critical period in mural trophoblast development in bovine blastocysts. *Dev Biol,* **2005**. *288*(2): p. 448-60.
103. Pearton, D.J., et al., Elf5 regulation in the Trophectoderm. *Dev Biol,* **2011**. *360*: p. 343-350.
104. Valdez Magana, G., et al., Paracrine effects of embryo-derived FGF4 and BMP4 during pig trophoblast elongation. *Dev Biol,* **2014**.
105. Niakan, K.K. and K. Eggan, Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Developmental Biology,* **2013**. *375*(1): p. 54-64.
106. Madeja, Z.E., et al., Changes in sub-cellular localisation of trophoblast and inner cell mass specific transcription factors during bovine preimplantation development. *BMC Dev Biol,* **2013**. *13*: p. 32.
107. Bou, G., et al., CDX2 is essential for cell proliferation and polarity in porcine blastocysts. *Development,* **2017**. *144*(7): p. 1296-1306.
108. Piliszek, A., et al., Differentiation of trophectoderm in rabbit embryos is initiated in the absence of Gata3 and Cdx2. *Mechanisms of Development,* **2017**. *145*: p. S79.
109. Ralston, A. and J. Rossant, Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Dev Biol,* **2008**. *313*(2): p. 614-29.
110. Wu, G., et al., Initiation of trophectoderm lineage specification in mouse embryos is independent of Cdx2. *Development,* **2010**. *137*(24): p. 4159-4169.
111. Strumpf, D., et al., Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development,* **2005**. *132*(9): p. 2093-102.
112. Jedrusik, A., et al., Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition. *Dev Biol,* **2014**.
113. Sritanaudomchai, H., et al., CDX2 in the formation of the trophectoderm lineage in primate embryos. *Dev Biol,* **2009**. *335*(1): p. 179-87.
114. Goissis, M.D. and J.B. Cibelli, Functional characterization of CDX2 during bovine preimplantation development in vitro. *Molecular Reproduction and Development,* **2014**. *81*(10): p. 962-970.
115. Sakurai, N., et al., The Necessity of OCT-4 and CDX2 for Early Development and Gene Expression Involved in Differentiation of Inner Cell Mass and Trophectoderm Lineages in Bovine Embryos. *Cell Reprogram,* **2016**. *18*(5): p. 309-318.
116. Nishioka, N., et al., Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech Dev,* **2008**. *125*(3-4): p. 270-83.
117. Kaneko, K.J. and M.L. DePamphilis, Regulation of gene expression at the beginning of mammalian development and the TEAD family of transcription factors. *Dev Genet,* **1998**. *22*(1): p. 43-55.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

118. Vassilev, A., et al., TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm. *Genes Dev,* **2001**. *15*(10): p. 1229-41.
119. Cao, Z., et al., Transcription factor AP-2gamma induces early Cdx2 expression and represses HIPPO signaling to specify the trophectoderm lineage. *Development,* **2015**. *142*(9): p. 1606-15.
120. Home, P., et al., GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. *J Biol Chem,* **2009**. *284*(42): p. 28729-37.
121. Rayon, T., et al., Notch and Hippo Converge on Cdx2 to Specify the Trophectoderm Lineage in the Mouse Blastocyst. *Developmental Cell,* **2014**. *30*(4): p. 410-422.
122. Niwa, H., et al., Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell,* **2005**. *123*(5): p. 917-29.
123. Murohashi, M., et al., An FGF4-FRS2alpha-Cdx2 axis in trophoblast stem cells induces Bmp4 to regulate proper growth of early mouse embryos. *Stem Cells,* **2010**. *28*(1): p. 113-21.
124. Ralston, A., et al., Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development,* **2010**. *137*(3): p. 395-403.
125. Bai, H., et al., Regulation of Trophoblast-Specific Factors by GATA2 and GATA3 in Bovine Trophoblast CT-1 Cells. *Journal of Reproduction and Development,* **2011**. *57*(4): p. 518-525.
126. Negron-Perez, V.M., Y. Zhang, and P.J. Hansen, Single-cell gene expression of the bovine blastocyst. *Reproduction,* **2017**. *154*(5): p. 627-644.
127. Winger, Q., et al., Analysis of transcription factor AP-2 expression and function during mouse preimplantation development. *Biol Reprod,* **2006**. *75*(3): p. 324-33.
128. Kuckenberg, P., et al., The transcription factor TCFAP2C/AP-2gamma cooperates with CDX2 to maintain trophectoderm formation. *Mol Cell Biol,* **2010**. *30*(13): p. 3310-20.
129. Lee, S.H., et al., Expression and function of transcription factor AP-2? in early embryonic development of porcine parthenotes. *Reprod Fertil Dev,* **2015**.
130. Kobayashi, T., et al., Principles of early human development and germ cell program from conserved model systems. *Nature,* **2017**. *546*(7658): p. 416-420.
131. Gardner, R.L., Clonal analysis of early mammalian development. *Philos Trans R Soc Lond B Biol Sci,* **1985**. *312*(1153): p. 163-78.
132. Gardner, R.L. and J. Rossant, Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J Embryol Exp Morphol,* **1979**. *52*: p. 141-52.
133. Evans, M.J. and M.H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos. *Nature,* **1981**. *292*: p. 154.
134. Beddington, R.S. and E.J. Robertson, An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development,* **1989**. *105*(4): p. 733-7.
135. Boroviak, T., et al., The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nature Cell Biology,* **2014**. *16*: p. 513.
136. Nichols, J. and A. Smith, Naive and Primed Pluripotent States. *Cell Stem Cell,* **2009**. *4*(6): p. 487-492.
137. Takashima, Y., et al., Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell,* **2014**. *158*(6): p. 1254-1269.
138. Theunissen, Thorold W., et al., Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell,* **2014**. *15*(4): p. 471-487.
139. Ávila-González, D., et al., Capturing the ephemeral human pluripotent state. *Developmental Dynamics,* **2016**. *245*(7): p. 762-773.
140. Dunn, S.J., et al., Defining an essential transcription factor program for naive pluripotency. *Science,* **2014**. *344*(6188): p. 1156-60.
141. Nichols, J., et al., Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell,* **1998**. *95*(3): p. 379-391.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

142. Frum, T., et al., Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev Cell,* **2013**. *25*(6): p. 610-22.
143. Avilion, A.A., et al., Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev,* **2003**. *17*(1): p. 126-40.
144. Mitsui, K., et al., The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* **2003**. *113*: p. 631-642.
145. Frankenberg, S., et al., Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. *Dev Cell,* **2011**. *21*(6): p. 1005-13.
146. Elling, U., et al., Murine inner cell mass-derived lineages depend on Sall4 function. *Proc Natl Acad Sci U S A,* **2006**. *103*(44): p. 16319-24.
147. Chiplunkar, A.R., et al., The Kruppel-like factor 2 and Kruppel-like factor 4 genes interact to maintain endothelial integrity in mouse embryonic vasculogenesis. *BMC Dev Biol,* **2013**. *13*: p. 40.
148. Yamaguchi, Y., S. Yonemura, and S. Takada, Grainyhead-related transcription factor is required for duct maturation in the salivary gland and the kidney of the mouse. *Development,* **2006**. *133*(23): p. 4737-4748.
149. Luo, J., et al., Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. *Nature,* **1997**. *388*(6644): p. 778-82.
150. Wassarman, K.M., et al., Specification of the anterior hindbrain and establishment of a normal mid/hindbrain organizer is dependent on Gbx2 gene function. *Development,* **1997**. *124*(15): p. 2923-34.
151. Chen, L., et al., Cross-regulation of the Nanog and Cdx2 promoters. *Cell Res,* **2009**. *19*(9): p. 1052-61.
152. Takahashi, K. and S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* **2006**. *126*(4): p. 663-76.
153. Yu, J., et al., Induced pluripotent stem cell lines derived from human somatic cells. *Science,* **2007**. *318*(5858): p. 1917-20.
154. Tsubooka, N., et al., Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells,* **2009**. *14*(6): p. 683-94.
155. Kotkamp, K., et al., A Pou5f1/Oct4 dependent Klf2a, Klf2b, and Klf17 regulatory sub-network contributes to EVL and ectoderm development during zebrafish embryogenesis. *Dev Biol,* **2014**. *385*(2): p. 433-47.
156. Yang, F., et al., ESRRB plays a crucial role in the promotion of porcine cell reprograming. *J Cell Physiol,* **2018**. *233*(2): p. 1601-1611.
157. Wu, J., et al., Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell,* **2017**. *168*(3): p. 473-486.e15.
158. Boyer, L.A., et al., Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell,* **2005**. *122*(6): p. 947-56.
159. Piliszek, A., Z.E. Madeja, and B. Plusa, Suppression of ERK signalling abolishes primitive endoderm formation but does not promote pluripotency in rabbit embryo. *Development,* **2017**. *144*(20): p. 3719-3730.
160. Sun, R., et al., Morphological changes and germ layer formation in the porcine embryos from days 7-13 of development. *Zygote,* **2013**: p. 1-11.
161. Frankenberg, S., et al., Early cell lineage specification in a marsupial: a case for diverse mechanisms among mammals. *Development,* **2013**. *140*(5): p. 965-975.
162. Lavial, F., et al., The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development,* **2007**. *134*(19): p. 3549-63.
163. Yoshida, M., et al., Conserved and divergent expression patterns of markers of axial development in reptilian embryos: Chinese soft-shell turtle and Madagascar ground gecko. *Developmental Biology,* **2016**. *415*(1): p. 122-142.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

164. Tapia, N., et al., Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins. *Nat Commun,* **2012**. *3*: p. 1279.
165. Theunissen, T.W., et al., Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. *Development,* **2011**. *138*(22): p. 4853-65.
166. Loh, Y.H., et al., The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet,* **2006**. *38*(4): p. 431-40.
167. Soufi, A., et al., Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell,* **2015**. *161*(3): p. 555-568.
168. Malik, V., D. Zimmer, and R. Jauch, Diversity among POU transcription factors in chromatin recognition and cell fate reprogramming. *Cell Mol Life Sci,* **2018**.
169. Wang, J., D.N. Levasseur, and S.H. Orkin, Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A,* **2008**. *105*(17): p. 6326-31.
170. Fogarty, N.M.E., et al., Genome editing reveals a role for OCT4 in human embryogenesis. *Nature,* **2017**. *550*(7674): p. 67-73.
171. Simmet, K., et al., OCT4/POU5F1 is required for NANOG expression in bovine blastocysts. *Proc Natl Acad Sci U S A,* **2018**.
172. Ito, S., et al., Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* **2010**. *466*(7310): p. 1129-33.
173. Dejosez, M., et al., Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell,* **2008**. *133*(7): p. 1162-74.
174. Chazaud, C., et al., Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell,* **2006**. *10*(5): p. 615-24.
175. Kuijk, E.W., et al., The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. *Development,* **2012**. *139*(5): p. 871-82.
176. Plusa, B., et al., Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development,* **2008**. *135*(18): p. 3081-91.
177. Artus, J., A. Piliszek, and A.K. Hadjantonakis, The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. *Dev Biol,* **2011**. *350*(2): p. 393-404.
178. Bessonnard, S., et al., Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development,* **2014**. *141*(19): p. 3637-3648.
179. Schrode, N., et al., GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. *Developmental Cell,* **2014**. *29*(4): p. 454-467.
180. Wamaitha, S.E., et al., Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes & Development,* **2015**. *29*(12): p. 1239-1255.
181. Saiz, N., et al., Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. *Nat Commun,* **2016**. *7*: p. 13463.
182. Xenopoulos, P., et al., Heterogeneities in Nanog Expression Drive Stable Commitment to Pluripotency in the Mouse Blastocyst. *Cell Rep,* **2015**.
183. Krawchuk, D., et al., FGF4 is a limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of the mouse blastocyst. *Developmental Biology,* **2013**. *384*(1): p. 65-71.
184. Kang, M., et al., FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development,* **2013**. *140*(2): p. 267-79.
185. Kang, M., V. Garg, and A.-K. Hadjantonakis, Lineage Establishment and Progression within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. *Developmental Cell,* **2017**. *41*(5): p. 496-510.e5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

186. Molotkov, A., et al., Distinct Requirements for FGFR1 and FGFR2 in Primitive Endoderm Development and Exit from Pluripotency. *Developmental Cell,* **2017**. *41*(5): p. 511-526.e4.
187. Yamanaka, Y., F. Lanner, and J. Rossant, FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development,* **2010**. *137*(5): p. 715-24.
188. Le Bin, G.C., et al., Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development,* **2014**. *141*(5): p. 1001-10.
189. Ambrosetti, D.-C., et al., Modulation of the Activity of Multiple Transcriptional Activation Domains by the DNA Binding Domains Mediates the Synergistic Action of Sox2 and Oct-3 on the Fibroblast Growth Factor-4Enhancer. *Journal of Biological Chemistry,* **2000**. *275*(30): p. 23387-23397.
190. Ohnishi, Y., et al., Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nature Cell Biology,* **2013**. *16*: p. 27.
191. Morris, S.A., et al., Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. *Proc Natl Acad Sci U S A,* **2010**. *107*(14): p. 6364-9.
192. Morris, S.A., et al., The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. *Open Biol,* **2013**. *3*(11): p. 130104.
193. Schröter, C., et al., FGF/MAPK signaling sets the switching threshold of a bistable circuit controlling cell fate decisions in embryonic stem cells. *Development,* **2015**. *142*(24): p. 4205-4216.
194. Singh, A.M., et al., A heterogeneous expression pattern for Nanog in embryonic stem cells. *Stem Cells,* **2007**. *25*(10): p. 2534-42.
195. Cauffman, G., et al., Markers that define stemness in ESC are unable to identify the totipotent cells in human preimplantation embryos. *Hum Reprod,* **2009**. *24*(1): p. 63-70.
196. Rodríguez, A., C. Allegrucci, and R. Alberio, Modulation of Pluripotency in the Porcine Embryo and iPS Cells. *PLOS ONE,* **2012**. *7*(11): p. e49079.
197. Roode, M., et al., Human hypoblast formation is not dependent on FGF signalling. *Dev Biol,* **2012**. *361*(2): p. 358-63.
198. Deglincerti, A., et al., Self-organization of the in vitro attached human embryo. *Nature,* **2016**. *advance online publication*.
199. Denicol, A.C., et al., The WNT signaling antagonist Dickkopf-1 directs lineage commitment and promotes survival of the preimplantation embryo. *The FASEB Journal,* **2014**. *28*(9): p. 3975-3986.
200. Smith, C., et al., Simultaneous gene quantitation of multiple genes in individual bovine nuclear transfer blastocysts. *Reproduction,* **2007**. *133*(1): p. 231-42.
201. Harris, D., B. Huang, and B. Oback, Inhibition of MAP2K and GSK3 signaling promotes bovine blastocyst development and epiblast-associated expression of pluripotency factors. *Biol Reprod,* **2013**. *88*(3): p. 74.
202. Tanaka, S., et al., Promotion of trophoblast stem cell proliferation by FGF4. *Science,* **1998**. *282*(5396): p. 2072-5.
203. Artus, J., et al., PDGF signaling is required for primitive endoderm cell survival in the inner cell mass of the mouse blastocyst, in *Stem Cells*. **2013**. p. 1932-41.
204. Artus, J., J.J. Panthier, and A.K. Hadjantonakis, A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. *Development,* **2010**. *137*(20): p. 3361-72.
205. Fabian, D., et al., Inhibitory effect of IGF-I on induced apoptosis in mouse preimplantation embryos cultured in vitro. *Theriogenology,* **2004**. *61*(4): p. 745-55.
206. Spanos, S., et al., Anti-apoptotic action of insulin-like growth factor-I during human preimplantation embryo development. *Biol Reprod,* **2000**. *63*(5): p. 1413-20.
207. Byrne, A.T., et al., Regulation of apoptosis in the bovine blastocyst by insulin and the insulin-like growth factor (IGF) superfamily. *Mol Reprod Dev,* **2002**. *62*(4): p. 489-95.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2018 doi:10.20944/preprints201806.0206.v1

208. Ahumada, C.J., et al., Effect of supplementation of different growth factors in embryo culture medium with a small number of bovine embryos on in vitro embryo development and quality. *Animal,* **2013**. *7*(3): p. 455-62.
209. Navarrete Santos, A., et al., Cell lineage-specific signaling of insulin and insulin-like growth factor I in rabbit blastocysts. *Endocrinology,* **2008**. *149*(2): p. 515-24.
210. Meng, F., et al., Signal Inhibition Reveals JAK/STAT3 Pathway as Critical for Bovine Inner Cell Mass Development. *Biol Reprod,* **2015**. *93*(6): p. 132.
211. Daniels, D.L. and W.I. Weis, Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol,* **2005**. *12*(4): p. 364-71.
212. Martello, G., et al., Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell,* **2012**. *11*(4): p. 491-504.
213. Sato, N., et al., Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med,* **2004**. *10*(1): p. 55-63.
214. ten Berge, D., et al., Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat Cell Biol,* **2011**. *13*(9): p. 1070-5.
215. Biechele, S., et al., Porcn-dependent Wnt signaling is not required prior to mouse gastrulation. *Development,* **2013**. *140*(14): p. 2961-71.