Early cell fate decisions in the mouse embryo

Néstor Saiz and Berenika Plusa

Faculty of Life Sciences, The University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK

Correspondence should be addressed to B Plusa; Email: berenika.plusa@manchester.ac.uk

Abstract

During mammalian preimplantation development, the fertilised egg gives rise to a group of pluripotent embryonic cells, the epiblast, and to the extraembryonic lineages that support the development of the foetus during subsequent phases of development. This preimplantation period not only accommodates the first cell fate decisions in a mammal's life but also the transition from a totipotent cell, the zygote, capable of producing any cell type in the animal, to cells with a restricted developmental potential. The cellular and molecular mechanisms governing the balance between developmental potential and lineage specification have intrigued developmental biologists for decades. The preimplantation mouse embryo offers an invaluable system to study cell differentiation as well as the emergence and maintenance of pluripotency in the embryo. Here we review the most recent findings on the mechanisms controlling these early cell fate decisions. The model that emerges from the current evidence indicates that cell differentiation in the preimplantation embryo depends on cellular interaction and intercellular communication. This strategy underlies the plasticity of the early mouse embryo and ensures the correct specification of the first mammalian cell lineages.

Introduction

Animals develop surrounded by extraembryonic membranes, which protect the embryo and mediate the uptake of nutrients and gases as well as the disposal of waste products. Amniotes have the most complex set of membranes, which allow them to develop on land, either within hard-shelled eggs or inside the mother's uterus. In most mammals, some of these membranes have evolved into the placenta, a highly specialised organ that connects directly mother and foetus (Ferner & Mess 2011). The importance of these tissues for development, therefore, cannot be overemphasised. In eutherian mammals the founders of most of these tissues are the trophectoderm (TE) and the primitive endoderm (PrE). The TE is responsible for the implantation of the embryo into the uterus and contributes to the embryonic part of the placenta. The PrE develops into the visceral endoderm (VE) and the parietal endoderm (ParE) after implantation, which form the majority of the yolk sac (Gardner & Johnson 1972, Gardner et al. 1973, Copp 1979, Gardner & Rossant 1979, Papaioannou 1982, Gardner 1983, Nagy et al. 2003). Besides this embryonic nutritive role, the VE also contributes cells to the gut endoderm of the embryo (Kwon et al. 2008). Moreover, it provides the signals that lay down the antero-posterior axis of the foetus during gastrulation (reviewed in Beddington & Robertson (1999), Nagy et al. (2003), and Stern & Downs (2012)).

TE and PrE are established during the preimplantation stages of development, before the embryo attaches to the uterus. Mammalian preimplantation development is entirely dedicated to the formation of these two lineages and their segregation from the epiblast, the tissue that produces the foetus as well as the rest of extraembryonic cell types. In the mouse embryo, the most extensively studied model, the formation of these three lineages occurs in two successive cell differentiation events, which take place over ~3 days. These are the first and, arguably, most critical cell fate decisions during mammalian development. Therefore, they have been the subject of extensive research. Moreover, the progress in embryonic stem (ES) cell research and the generation of in vitro differentiation models have fuelled a renewed interest in understanding cell differentiation in embryo. The current evidence indicates that these lineage decisions are very much the result of cell–cell interactions, which delineate the expression pattern of master regulatory genes. Here we review the latest findings in the field and propose a comprehensive model of the mechanisms behind these early cell fate decisions.

Morphogenesis of the mouse preimplantation embryo

The mouse embryo spends nearly 5 days preparing for implantation (Fig. 1). Over this period, the blastomeres produced by the cleavage of the zygote differentiate and
arrange to form the blastocyst, characterised by the presence of a fluid-filled cavity and an inner cell mass (ICM), both surrounded by the TE (Fig. 1G). In the expanded blastocyst (embryonic days, E4–4.5), the PrE can be distinguished as an epithelium at the surface of the ICM in contact with the cavity. The rest of the ICM represents the epiblast, enclosed between the TE and the PrE (Fig. 1H). Two morphogenetic events during this period, compaction and cavitation, create asymmetries in the embryo that affect lineage formation. Compaction of the eight-cell embryo through the development of intercellular junctions produces the morula, where individual cells are no longer evident (Fig. 1E). Cells located on the inside of the morula after subsequent cell divisions become the ICM of the blastocyst, whereas cells positioned on the surface develop into the TE (Tarkowski & Wróblewska 1967). Through cavitation at around the 30-cell stage (Smith & McLaren 1977), the morula becomes the blastocyst (Fig. 1F). The presence of the cavity eventually determines the position of the PrE in the late blastocyst (Fig. 1H). The consequences of these asymmetries for cell differentiation will be discussed further below.

**Developmental potential vs development**

Cell differentiation results in a concomitant loss of developmental potential. The zygote is totipotent by definition, as it is the mother for all cells in the body and the extraembryonic membranes. However in order to build an organism, totipotency has to be forsaken; this developmental potential needs to be realised into specialised tissues with defined functions. In mammals this high developmental potential is maintained for a remarkable amount of time. In the mouse not only the zygote but all the blastomeres through the morula and even the ICM of cavitating blastocysts are totipotent, as the year area blastogenera is totipotential and extraembryonic lineages (Hillman et al. 1972, Garner & McLaren 1974, Kelly 1977, Handyside 1978, Hogan & Tilly 1978, Spindle 1978, Rossant & Lis 1979, Rossant & Vijh 1980, Suwirńska et al. 2008, Tarkowski et al. 2010, Grabarek et al. 2012). The epiblast, in turn, is pluripotent, as it can produce all tissues except those derived from the TE and PrE (Gardner & Rossant 1979). The persistence of totipotent/pluripotent populations is likely to be a consequence of the particularities of mammalian development, where proper embryogenesis is delayed, sometimes until the post-implantation stages. Such a delay creates the need to maintain a population of pluripotent cells, the epiblast, capable of producing the foetus even when implantation is temporarily postponed – as it is in mammals that undergo diapause (Nichols et al. 2001). From an applied perspective, these extended pluripotent states have allowed the generation of ES cells from the epiblast (Evans & Kaufman 1981,
Martin 1981, Brook & Gardner 1997), therefore allowing us to capture this pluripotency in vitro (Smith 2001). Subsequently, trophoblast stem (TS) cells (Tanaka et al. 1998) and extraembryonic endoderm cells (Kunath et al. 2005) have also been established from the TE and the PrE respectively, thus making stem cells available from all three blastocyst lineages.

Pluripotency is lost as cells differentiate into specialised tissues. In the mouse embryo organising centres do not appear until post-implantation stages, and no maternal determinants have been described to date that can drive cell differentiation at preimplantation stages. In fact, the majority of maternal mRNAs are degraded by the two-cell stage and replaced by products of the zygotic genome (Schultz 2002). Instead, these early lineage decisions are taken based on the position of the cells within the embryo and on intercellular communication. This strategy endows mouse embryos with outstanding regulative abilities. Numerous embryological studies have shown that mouse embryos can withstand rather drastic manipulations, such as rearrangement, addition or elimination of cells (or parts of them), and still develop to term (Tarkowski 1959, 1961, Tarkowski & Wróblewska 1967, Gardner 1968, Kelly 1977, Ziomek & Johnson 1980, Zernicka-Goetz 1998, Suwinski et al. 2008). In the past decades, a vast amount of work using blastocyst-derived stem cells, together with experimental embryology, has helped us gain further insight into the genetic and cellular mechanisms responsible for the maintenance of these pluripotent stages and for taking these early cell fate decisions. In the next sections we will discuss these mechanisms and integrate them into a unified model of our current understanding of early lineage specification.

The specification of TE and ICM

As mentioned earlier, the first differentiation event during mammalian development is the formation of the TE and the ICM. Tarkowski & Wróblewska (1967) proposed that the exposure to different microenvironments of cells located on the periphery or inside of the morula results in differential cell fates, an idea that has come to be known as the ‘inside–outside model’. Later work built onto this concept and showed that blastomeres at the eight-cell stage (E2.5) acquire a polarised apical surface upon compaction, which becomes isolated from the basolateral surfaces with the development of intercellular junctions (Ducibella & Anderson 1975, Handyside 1980, Ziomek & Johnson 1980, Johnson et al. 1981). This ‘polarity model’ postulates that it is the differential inheritance of this polarised surface, after division from 8 to 16 and from 16 to 32 cells, which determines a TE or ICM fate (Johnson & Ziomek 1981). Thus, cell divisions immediately after compaction can produce either two outer cells that inherit the polarised domain and therefore become TE (symmetric or conservative division) or one polar and one apolar, internal cell, which becomes part of the ICM (asymmetric or differentiative division). ICM cells are produced in two successive rounds of asymmetric division, from the 8- to 16-cell stage and from the 16- to 32-cell stage (Pedersen et al. 1986, Fleming 1987; Fig. 1D, E and F). The number of cells internalised in each round varies widely, although they are inversely proportional, resulting in a relatively consistent number of ICM cells in the 32-cell blastocyst (∼E3.25) and suggesting that mechanisms are in place to ensure an ICM of the adequate size is produced (Fleming 1987, Bischoff et al. 2008).

The TE is the first epithelium to arise during mammalian development. As such, its cells are bound through an extensive network of intercellular junctions that ensure selective permeability and epithelial cohesion. These junctions develop sequentially during TE development (Calarco & Epstein 1973, Ducibella et al. 1975). The compaction of the eight-cell embryo is a Ca²⁺-dependent process mediated by changes in the cytoskeleton and by the formation of adherens junctions, composed of E-cadherin (or cadherin 1; Ducibella & Anderson 1975, Hyafil et al. 1980, Vestweber & Kemler 1984). E-cadherin is essential for compaction and for the subsequent formation of an organised TE epithelium (Hyafil et al. 1980, Vestweber & Kemler 1984, Larue et al. 1994). The formation of adherens junctions is followed by the development of tight junctions and desmosomes, which provide the cell cohesion necessary to allow the expansion of the blastocyst cavity (Calarco & Brown 1969, Nadijcka & Hillman 1974, Ducibella et al. 1975, Fleming et al. 1989, 1991, Thomas et al. 2004). These intercellular junctions demarcate an apical domain containing microvilli and the actin-associated protein ezrin, as well as the atypical protein kinase C (aPKC)–partitioning defective 6 (PAR6) complex (Calarco & Epstein 1973, Reeve & Ziomek 1981, Louvet et al. 1996, Pauken & Capco 2000, Plusa et al. 2005, Vinot et al. 2005). Some components of this apical domain seem to be able to affect cell position in the morula, resulting in allocation to either the TE or the ICM. Disruption of aPKC activity in a subpopulation of blastomeres before compaction reduces the size of their apical domain and causes a higher contribution of these cells to the ICM lineage through asymmetrical divisions and internalisation of cells (Plusa et al. 2005, Dard et al. 2009). Therefore, the apico-basal asymmetry established in the compacted morula seems to provide the first cue for the segregation of the TE from the ICM (Johnson & Ziomek 1981, Johnson et al. 1986).

The development of the apical domain in the morula precedes the action of TE-specific transcription factors, such as caudal-type homeobox-2 (CDX2). CDX2, in turn, is not needed for the polarisation of these cells (Ralston & Rossant 2008, Stephenson et al. 2010). CDX2 is one of the first markers for TE cells. It is genetically upstream of other trophodermal genes, such as the
T-box gene eomesodermin (Eomes), integrin α7 (Itga7), or cadherin 3 (Cdh3) and its expression in ES cells is sufficient to induce TE differentiation (Niwa et al. 2005, Strumpf et al. 2005). Cdx2−/− embryos show peri-implantation lethality due to a failure in the formation of a mature TE epithelium and are unable to produce TS cells in culture (Chawengsaksophak et al. 1997, Strumpf et al. 2005, Wu et al. 2010, Blij et al. 2012). CDX2 is not required for compaction or the initiation of cavitation in the morula; however, Cdx2 null embryos fail to expand the cavity (Strumpf et al. 2005, Blij et al. 2012), indicating that epithelialisation initiates independently of CDX2, although CDX2 is necessary for maintaining the epithelial integrity and for maturation of the tissue. Cdx2 is expressed from the eight-cell stage onwards and CDX2 is present at heterogeneous levels across all blastomeres in the morula (≥16 cells), where it overlaps with ICM and epiblast markers, such as octamer-binding protein 4 (OCT4/POU domain, class 5 transcription factor 1 (POU5F1)) and NANOG respectively (Fig. 2B and C). It becomes gradually restricted to TE cells only in the mid and late blastocyst and to the TE-derived extraembryonic ectoderm at post-implantation stages (Beck et al. 1995, Strumpf et al. 2005, Dietrich & Hiragi 2007).

Cdx2 expression is induced by the transcriptional enhancer activator (TEA) domain transcription factor (TEAD4; Yagi et al. 2007, Nishioka et al. 2008), in cooperation with either of the transcriptional co-activators Yes-associated protein 1 (YAP1 or YAP) or WW domain containing transcriptional regulator 1 (WWTR1), both repressed by the Hippo signalling pathway (Kanai et al. 2000, Zhao et al. 2007, Nishioka et al. 2009; Fig. 2C). TEAD4 is found in all cells at the early blastocyst stage (Nishioka et al. 2008). TEAD4 has also been placed upstream of another TE transcription factor, the zinc-finger, GATA-binding protein 3 (GATA3; Ralston et al. 2010). GATA3 co-localises with CDX2 from the eight-cell stage and can promote Cdx2 expression (Home et al. 2009, Ralston et al. 2010). Their regulation and transcriptional targets suggest that CDX2 and GATA3 act in parallel to promote TE fate downstream of TEAD4 (Ralston et al. 2010; Fig. 2C). Intriguingly, it has also been recently reported that TEAD4 is excluded from the nuclei of ICM cells in later blastocysts and suggested that
its localisation plays a role in the specification of TE and ICM cells (Home et al. 2012). However, the functional relevance of this observation is still unclear.

On the other hand OCT4 is necessary for the maintenance of both the ICM and the ES cells (Schöler et al. 1990b, Nichols et al. 1998, Niwa et al. 2000). Zygotic Pou5f1 is expressed in all blastomeres from the two-cell stage through the mid blastocyst. It subsequently becomes restricted to the ICM (late blastocysts), the epiblast (peri-implantation blastocysts (E4.5)) and the primitive ectoderm (post-implantation), being only expressed in primordial germ cells from mid gestation (Schöler et al. 1989, 1990a, Rosner et al. 1990). OCT4 displays a similar pattern, except that its levels have been reported to peak in the PrE before becoming completely restricted to the epiblast (Palmieri et al. 1994). In line with this observation, high levels of POUSF1 promote differentiation of ES cells (Niwa et al. 2000). In contrast, more recent reports show no such increase in OCT4 levels (Nichols et al. 2009, Frankenberg et al. 2011), or even a reduction in PrE cells, which correlates with lineage commitment (Grabarek et al. 2012). OCT4 is, nevertheless, necessary for maintaining pluripotency. Its absence in either the ICM or the ES cells results in them acquiring TE features (Nichols et al. 1998, Niwa et al. 2000). Its expression pattern, essentially restricted to pluripotent cells, and its requirement for the maintenance of these cells have placed OCT4 at the centre of the gene network that preserves pluripotency. Interestingly OCT4 has been observed to display different kinetics in outer and inner cells of the morula, indicating that the intracellular dynamics of transcription factors are different in these two-cell populations (Plachta et al. 2011). Outer cells of the morula were found to have a faster rate of OCT4 import and export to and from the nucleus (fast kinetics), whereas inner cells presented slower rates of nuclear import and export (slow kinetics). These two types of OCT4 kinetics were also found in blastomeres at the 8-cell stage and correlate with the type of division blastomeres undertake at the 8- to 16-cell transition (fast kinetics, symmetric division; slow kinetics, asymmetric division; Plachta et al. 2011). However, how much of a causal role these kinetics play during lineage specification and what is the relationship between the dynamics of OCT4 and the interaction of OCT4 with other transcription factors (see below) still need to be determined.

CDX2 and OCT4 eventually display complementary expression domains in the late blastocyst, where CDX2 is found only in TE cells and OCT4 only in the ICM. In fact, CDX2 and OCT4 can repress each other's activity (Niwa et al. 2005) and is the interplay between them and TEAD4–YAP that results in the establishment of the TE and ICM populations (Fig. 2C). In outer cells of the morula, where the Hippo pathway is inactive, TEAD4–YAP promote Cdx2 expression, which is reinforced by CDX2 itself and GATA3 (Niwa et al. 2005, Home et al. 2009, Nishioka et al. 2009; Fig. 2C). CDX2 in turn activates the expression of later TE genes, alongside GATA3, and represses OCT4 activity in outer cells, which eventually leads to the restriction of OCT4 to the ICM by the late blastocyst (Niwa et al. 2005, Strumpf et al. 2005; Fig. 2C). In the absence of either Cdx2 or Tead4, a mature TE fails to develop and OCT4 remains in both ICM and TE cells (Strumpf et al. 2005, Yagi et al. 2007, Nishioka et al. 2008). Conversely, in inner cells of the morula, activation of the Hippo pathway causes inactivation and cytoplasmic localisation of YAP via phosphorylation by the large tumour suppressor (LATS) kinase (Nishioka et al. 2009). In the absence of nuclear YAP, TEAD4 cannot stimulate Cdx2 expression and CDX2 is repressed by OCT4, preventing the self-activation of Cdx2 and TE genes (Niwa et al. 2005, Nishioka et al. 2009; Fig. 2C). Consistent with this, lack of OCT4 results in TE differentiation of the ICM (see above; Nichols et al. 1998), indicating that OCT4 actively suppresses TE fate in ICM cells. Moreover, OCT4 is necessary for the expression of fibroblast growth factor 4 (Fgf4) in the ICM (Nichols et al. 1998). Paracrine FGF4 is necessary for the proliferation of TE and extraembryonic ectoderm cells at late blastocyst and post-implantation stages respectively, a mechanism that ensures the co-ordinated growth of TE and ICM (Gardner & Johnson 1972, Nichols et al. 1998, Tanaka et al. 1998, Goldin & Papaioannou 2003, Saba-El-Leil et al. 2003).

The current evidence therefore indicates that the differentiation of TE and ICM is initiated by the allocation of cells to the periphery or the inside of the morula. Cell position and polarisation of outer cells precede lineage-specific gene expression and seem to play an instructive role in the specification of these lineages. Further support for this model comes from the ability of inner cells to re-polarise and develop a new TE when isolated from early blastocysts (Handyside 1978, Hogan & Tilly 1978, Spindle 1978, Rossant & Lis 1979, Stephenson et al. 2010). Thus, inheritance of apical components by outer cells and differential contacts between cells result in the patterning of lineage-specific genes, which are initially ubiquitously expressed (Fig. 2B). A recent report also argues that differential contacts are necessary for the establishment of the TE- and ICM-specific gene expression patterns (Lorthongphanich et al. 2012). Restriction of these regulatory transcription factors to each TE and ICM cells in turn allows the activation of the genetic programmes necessary for the maturation and further development of each lineage. The subsequent addition of epigenetic marks allows the long-term stabilisation of gene expression in each cell type (for a review on the role of epigenetic modifications in early lineage decisions, see Albert & Peters (2009) and Saitou et al. (2012)). In other words, the gradual loss of plasticity observed in these cells (Handyside 1978, Hogan & Tilly 1978, Spindle 1978, Rossant & Lis 1979, 2005, Home et al. 2009, Nishioka et al. 2009; Fig. 2C).
Suwińska et al. 2008) seems to be a reflection of the gradual commitment to each lineage as a result of the stabilisation of gene expression, induced and likely reinforced, by the position of the cells within the embryo.

In this context, aPKC is a particularly interesting molecule, as it has been shown to direct the allocation of cells to TE or ICM (Plusa et al. 2005, Dard et al. 2009). Interestingly, it has been observed that Cdx2 binding sites are present in the promoter of the gene(s) for aPKC (Jedrusik et al. 2008), indicating the possibility of a feedback loop. Jedrusik et al. (2008) found that the premature expression of Cdx2 (from the two-cell to four-cell) phenocopies aPKC overexpression, allocating cells preferentially to the TE, and causes an increase in the extent of the apical aPKC domain at the 4/8-cell stage. Jedrusik et al. (2008) thus propose that high expression of Cdx2 is the factor driving the allocation of cells to the TE lineage. However, from their results, elevated levels of Cdx2 are not obvious. It seems more likely that forcing Cdx2 expression prematurely, rather than increasing the levels of Cdx2, results in an increase in aPKC in the injected blastomeres. These blastomeres in turn would develop, prematurely, a more extensive aPKC apical domain, which would increase their chances of becoming TE. This interpretation would agree with the reports indicating that Cdx2 lies downstream of cell polarity in outer cells (Ralston & Rossant 2008, Stephenson et al. 2010). In normal conditions, polarisation of aPKC would indirectly promote Cdx2 expression, which would in turn reinforce cell polarity via Prkc expression in outer cells. Moreover, aPKC has been shown to phosphorylate KIBRA/WWC1 (Büther et al. 2004), an upstream member of the Hippo pathway (Baumgartner et al. 2010; reviewed in Genevet et al. (2010), Yu et al. (2010), and Genevet & Tapon (2011)), suggesting that aPKC could provide the link between polarity and the regulation of Cdx2 and TE specification by the Hippo pathway. This possibility certainly deserves future investigation.

The segregation of PrE and epiblast

The second differentiation event during preimplantation development results in the formation of the PrE and the epiblast. These two lineages arise in the ICM through mechanisms similar to but also remarkably different from those driving the differentiation of TE and ICM cells. The precise moment when these two lineages first emerge (if such a moment exists) is still unclear, although we now know that ICM cells begin to show differences in gene expression and cell plasticity already at the early blastocyst stage (E3.25), which becomes more obvious as the blastocyst expands further (E3.5) (Rossant et al. 2003, Chazaud et al. 2006, Kurimoto et al. 2006, Plusa et al. 2008, Guo et al. 2010, Morris et al. 2010, Niakan et al. 2010, Grabarek et al. 2012). However, because of the dynamic nature of this process, it is an issue that still remains controversial. Early work proposed that the cells located on the surface of the ICM in contact with the cavity become the PrE by virtue of their position, in a process analogous to the induction of the TE (Rossant 1975). However, the ICM has been suggested to be a heterogeneous population before the appearance of a PrE layer on the surface of the ICM (Koutsourakis et al. 1999). Indeed, more recent studies have shown that PrE and epiblast precursors are found intermingled in the ICM of the E3.5 blastocyst, in a ‘salt and pepper’ pattern (Rossant et al. 2003, Chazaud et al. 2006, Kurimoto et al. 2006) and that these two populations become subsequently resolved into two defined layers by a number of mechanisms (Plusa et al. 2008, Meilhac et al. 2009). Thus, in the morula, cell positioning precedes the establishment of the TE- and ICM-specific gene expression patterns, whereas PrE and epiblast precursors arise intermingled within the ICM, with their spatial organisation being a rather late event during their differentiation.

Cell fate regulators in the ICM

A number of lineage-specific transcription factors have been identified in PrE and epiblast cells. Epiblast cells are marked by the homeodomain transcription factor NANOG (Chambers et al. 2003, Mitsui et al. 2003) and in late stages by OCT4 and the SRY-related HMG box-containing transcription factor SOX2 (Rosner et al. 1990, Schöler et al. 1990a, Avilion et al. 2003). PrE cells, on the other hand, are marked by a number of endodermal proteins, such as GATA6, platelet-derived growth factor receptor alpha (PDGFRα), SOX17, GATA4, SOX7, hepatic nuclear factor 4 (HNF4), HNF1β, or hematopoietically expressed homeobox (HEX), which appear sequentially during PrE development (Arcesi et al. 1993, Chen et al. 1994, Duncan et al. 1994, Thomas et al. 1998, Barbacci et al. 1999, Koutsourakis et al. 1999, Rossant et al. 2003, Kurimoto et al. 2006, Plusa et al. 2008, Artus et al. 2010, 2011, Morris et al. 2010, Niakan et al. 2010). These regulatory transcription factors sit at the top of the hierarchy regulating the genetic programmes responsible for the identity of these cells. Nanog and Sox2 are necessary for maintaining the epiblast (Avilion et al. 2003, Mitsui et al. 2003), and have been thus placed, alongside OCT4, at the core of the pluripotency network. On the other hand, Gata6 mutants show defects in the formation of the VE, the immediate derivative of the PrE (Morrisey et al. 1998, Koutsourakis et al. 1999). In Gata4 mutants, the VE and ParE develop normally in most cases (Molkentin et al. 1997), but yolk sac closure is affected (Kuo et al. 1997, Molkentin et al. 1997). Moreover, expression of either Gata6 or Gata4 in ES cells is sufficient to induce differentiation towards extraembryonic endoderm (Fujikura et al. 2002). However, some of these genes...
display an overlapping pattern at early stages. *Gata6* is the earliest PrE gene to be expressed and GATA6 is present in all blastomeres from the eight-cell stage, alongside NANOG, OCT4 and CDX2 (Fig. 2B; Dietrich & Hiiiragi 2007, Plusa et al. 2008, Guo et al. 2010). *Gata6* is followed by Pdgfra expression in the morula (8–16 cells) and Sox17 from the late morula/early blastocyst (16–32 cells). On the other hand Sox2 is exclusively expressed in ICM cells from the early blastocyst stage (Guo et al. 2010). Consequently, GATA6, PDGFRA and SOX17 co-localise with NANOG, SOX2 and OCT4 in ICM cells of early blastocysts (∼32 cells; Dietrich & Hiiiragi 2007, Plusa et al. 2008, Guo et al. 2010, Niakan et al. 2010). Interestingly, during the transition from early to mid blastocysts (∼64 cells), the expression of these genes becomes restricted, so that they define two different populations: *Nanog* is now expressed only in epiblast precursors, whereas *Gata6* and *Sox17* are only expressed in PrE precursors (Rossant et al. 2003, Chazaud et al. 2006, Kurimoto et al. 2006, Plusa et al. 2008, Guo et al. 2010, Niakan et al. 2010). These two cell types, which can now be identified by their gene expression profile, appear scattered throughout the ICM, in a ‘salt and pepper’ fashion (Fig. 3A; Rossant et al. 2003, Chazaud et al. 2006). At this stage, PrE precursors up-regulate Gata4 expression, thus making GATA4 the first exclusive marker of PrE cells (Kurimoto et al. 2006, Plusa et al. 2008). This shift in gene expression, from an overlapping to an exclusive pattern, raises some fundamental questions: i) what are the mechanisms responsible for this change? and ii) is this exclusive gene expression pattern a reflection of lineage commitment by PrE and epiblast precursors?

This transition from a homogeneous to a heterogeneous, ‘salt and pepper’ population suggests that some non-cell autonomous mechanism could act as a

![Figure 3](image_url)

**Figure 3** Molecular mechanism for the specification of PrE and epiblast precursors. Heterogeneous expression of the ligand/receptor pair fibroblast growth factor 4 (FGF4)/FGF receptor 2 (FGFR2) leads to the random specification of PrE and epiblast precursors, which are scattered throughout the ICM in a ‘salt and pepper’ manner. Cell fate specification is followed by cell sorting (green arrows) so that PrE cells become arranged in a monolayer at the surface of the ICM in the late blastocyst. Stimulation of FGFR2 by FGF4 leads to inhibition of NANOG via ERK1/2, resulting in Gata6 expression and promotion of PrE identity. Low expression of Fgfr2 and active repression of Gata6 by NANOG prevents endodermal differentiation and maintains epiblast precursors in a naïve state. PrE precursors maintain OCT4 and SOX2 until very late stages, although these cells down-regulate the expression of Fgf4. FGF4 produced by the epiblast also stimulates cells in the polar TE (D) to promote their proliferation.
symmetry breaker in the ICM. Fibroblast growth factor (FGF) signalling has long been known to be involved in PrE formation. Fgf4 is expressed in epiblast cells (Niswander & Martin 1992, Rappolee et al. 1994, Frankenberg et al. 2011). Its expression is activated by OCT4 and SOX2 throughout the ICM (Yuan et al. 1995, Nichols et al. 1998) and is up-regulated only in epiblast precursors during the early to mid blastocyst transition (Fig. 3B and C), whereas its receptor, Fgfr2, is down-regulated in these cells (Guo et al. 2010). PrE precursors, on the other hand, maintain Fgfr2 expression and down-regulate Fgf4 (Guo et al. 2010; Fig. 3C). Mutations in either Fgf4, Fgfr2, or members of their intracellular pathway (such as the adaptor molecule growthfactor receptor-bound protein 2 (Grb2)) prevent PrE differentiation and the expression of endodermal genes (Feldman et al. 1995, Arman et al. 1998, Cheng et al. 1998, Chazaud et al. 2006), as does pharmacological inhibition of FGF signalling (Nichols et al. 2009, Yamanaka et al. 2010). Ablation of Fgfr2 produces a somewhat earlier phenotype than elimination of mutants shows defects already at the peri-implantation stage and display defective implantation (Arman et al. 1998), whereas Fgf4 or Gata6 null embryos look normal at blastocyst stage, but also die at early post-implantation stages due to the absence of a VE (Feldman et al. 1995, Morrisey et al. 1995, Koutsourakis et al. 1999). Importantly, it has been recently reported that Fgf4 mutant blastocysts fail to maintain GATA6 past the early blastocyst (Kang et al. 2013). Moreover, expression of Gata6 is able to rescue endoderm formation in Grb2 mutants, which have defects in the FGF/ERK transduction pathway (Wang et al. 2011). This body of evidence suggests that, although the expression of some lineage regulators precedes cell differentiation, FGF signalling lies upstream of these genes during lineage specification.

In ES cells blocking differentiation signals mediated by the FGFR–ERK pathway is necessary to maintain them in a pluripotent, ‘ground’ state (Silva & Smith 2008, Ying et al. 2008). This is achieved either by inhibition of FGF signalling or through transcriptional inhibition of endodermal genes by NANOG (Chambers et al. 2003, Mitsui et al. 2003, Singh et al. 2007, Ying et al. 2008). In the embryo epiblast precursors expressing Nanog, Pou5f1 and Sox2 but not Gata6 would represent this ground state, maintained by low FGF4 stimulation and active NANOG (Nichols et al. 2009, Silva et al. 2009, Messerschmidt & Kemler 2010; Fig. 3C). In the complete absence of maternal and zygotic FGF4, embryos express Gata6 at the eight-cell and morula stages but fail to maintain it past the early blastocyst, and all ICM cells contain NANOG thereafter (Kang et al. 2013). Conversely stimulation of the FGFR–ERK pathway maintains Gata6 expression throughout the ICM via the inhibition of NANOG, thus promoting PrE fate (Yamanaka et al. 2010, Frankenberg et al. 2011, Kang et al. 2013; Fig. 3C and D).

Incidentally, Nanog also seems to be essential for the transition from a totipotent, presumably unstable, state, represented by the ICM of the early blastocyst (containing NANOG, OCT4 and GATA6; Fig. 3B), to an ICM comprising epiblast and PrE precursors (NANOG and GATA6 mutually exclusive, in a ‘salt and pepper’ fashion: Fig. 3C). Intriguingly, during this transition, Nanog expression switches from being monoallelic to biallelic (Miyanari & Torres-Padilla 2012). Moreover, Nanog is essential for the reactivation of the paternal X chromosome (Xp) in epiblast cells at this stage (Mak et al. 2004, Okamoto et al. 2004, Silva et al. 2009). Nanog null embryos maintain the Xp inactive throughout the ICM and fail to develop an epiblast. Furthermore, Nanog−/− embryos display a non-cell autonomous defect in PrE formation – although all ICM cells in these embryos have GATA6, they show increased apoptosis and no FGF4 production (Silva et al. 2009, Messerschmidt & Kemler 2010, Frankenberg et al. 2011). Therefore, in the absence of Nanog, OCT4 and SOX2 would be unable to repress Gata6 and promote the epigenetic changes necessary to produce epiblast cells. The presence of GATA6 in these cells and the lack of FGF4 prevent the formation of neither PrE nor pluripotent epiblast cells and result in ICM degeneration (Frankenberg et al. 2011). Therefore, paracrine FGF4 and NANOG provide a mechanism for the restriction of gene expression and for promoting the divergence of the early ICM into the two distinct populations (epiblast and PrE precursors) that make up the ICM of the mid blastocyst.

From a functional standpoint, could this transition from an overlapping to a restricted gene expression pattern in ICM cells reflect the transition from an uncommitted, undifferentiated state to a lineage-restricted state? This hypothesis has been recently tested using chimera assays (Grabarek et al. 2012). Single presumptive PrE and epiblast precursors were isolated from blastocysts at different stages and transplanted to recipient 8- or 8- to 16-cell stage embryos, which were allowed to develop to late blastocysts. Unexpectedly, it was found that both PrE and epiblast precursors could contribute to several lineages in the chimaera, even when isolated from late blastocysts, after they display restricted gene expression (Grabarek et al. 2012). Therefore a restricted transcription factor profile does not preclude cell fate switch when challenged by the new environment. Strikingly epiblast precursors from early blastocysts very rarely contributed cells to the PrE of the chimaeric blastocysts, whereas PrE precursors were much more plastic and were normally able to contribute to all three lineages (Grabarek et al. 2012). These findings indicate that epiblast cells are somewhat lineage restricted from very early stages. The lower amount of FGF2 in epiblast cells (Guo et al. 2010) should result in a lower sensitivity to FGF4, which could account for their inability to produce PrE in chimaeras. In agreement with this hypothesis, addition of FGF4
restored the ability of epiblast cells to form PrE, whereas inhibition of FGF signalling restricted the plasticity of PrE precursors, which mainly produced epiblast cells (Grabarek et al. 2012). This lower sensitivity of epiblast precursors (or a subpopulation of them) to differentiation signals could be the means to ensure the presence of the foetal lineage in the embryo (Grabarek et al. 2012).

On the origin of the PrE and epiblast

Despite the evidence presented earlier, the timing and mechanism for the origin of these two populations have been a matter of debate. Early work suggested that epiblast and PrE could originate from ICM cells produced in the first and second wave of asymmetric division respectively during the morula to blastocyst transition, due to differential inheritance of cellular components, such as cytokeratin (Chisholm & Houlston 1987). This is an interesting hypothesis that has been addressed recently with opposing results (Morris et al. 2010, Yamanaka et al. 2010). Morris et al. (2010) found that ICM cells produced in the first round of asymmetrical division had a strong tendency to contribute to the epiblast, whereas cells produced in the second round mainly formed the PrE. On the other hand, Yamanaka et al. (2010) found no such correlation. These discrepancies have been attributed to the different mouse strains used and the methodological differences between the two studies (Morris 2011, Yamanaka 2011). Embryos from different strains produce different numbers of ICM cells in the first round of asymmetric division (Fleming 1987), which is also the case here (2.8 in Morris’s experiments vs 4.8 in Yamanaka’s experiments). When these data are analysed in the light of more recent work (Guo et al. 2010, Frankenberg et al. 2011, Grabarek et al. 2012), these differences in cell numbers become relevant. When few cells become internalised in the first round the overall amount of FGF4 they produce would be insufficient to stimulate their few FGF4 receptors and thus inhibit NANOG. As a consequence they would acquire epiblast identity, down-regulating Fgr2 and thus becoming refractory to FGF4. Interestingly these cells could represent the epiblast cells that we have found are unable to form PrE in chimaeras (Grabarek et al. 2012). These cells in turn would produce FGF4 that would induce PrE identity in the cells internalised in the second round. Conversely when a larger number of cells become internalised in the first round, the net amount of FGF4 they produce could then reach the threshold necessary to stimulate each other and therefore induce PrE identity in some of them.

Sorting the PrE and epiblast populations

Whatever their origin PrE and epiblast precursors become specified independently of their position. These two populations have a ‘salt and pepper’ distribution in the ICM of the mid blastocyst. However, their spatial arrangement by the time of implantation is very different: PrE cells form an absorbent epithelium at the interface between the ICM and the blastocyst cavity and thus leave the epiblast enclosed between the TE and PrE (Fig. 3 A). The PrE in peri-implantation blastocysts (E4.5) is an epithelium with a polarised apical surface facing the cavity and a basement membrane at the interface with the epiblast. The apical surface has microvilli, endocytic receptors such as cubulin (CUB) and LDL receptor-related protein 2 (LRP2), the endocytic adaptor disabled 2 (DAB2) and aPKC (Morisey et al. 1998, 2000, Yang et al. 2002, Gerbe et al. 2008, Saiz & Plusa (unpublished observations). The basement membrane contains laminin and collagen IV (Adamson & Aiers 1979, Dziadek & Timpl 1985, Smyth et al. 1999). This change in cell distribution implies that there must be a mechanism for PrE and epiblast precursors to become sorted into two separate compartments. Unless these cells are to change their gene expression depending on their position, one would expect these precursors to become rearranged into their final position. It has been described that cells within the ICM can indeed migrate in a non-directional, actin-dependent manner (Meilhac et al. 2009). Interestingly live imaging of embryos where PrE precursors express a reporter for platelet-derived growth factor receptor alpha (Pdgfra) expression (PdgfraH2B-GFP; Hamilton et al. 2003) revealed that PrE cells migrate until they reach the surface of the ICM exposed to the cavity (Plusa et al. 2008). Cells that reach this surface either by active migration or by passive intercalation retain their new position and up-regulate Pdgfra. Conversely, PrE cells that remain deeper within the ICM eventually lose Pdgfra expression or die, indicating that cell position is a critical factor during PrE formation (Plusa et al. 2008). These cell behaviours constitute an elegant mechanism to resolve a mixed cell population with no need for directional cell migration. It has been proposed that the segregation of these two populations is due to differential adhesion properties between epiblast and PrE cells (Chazaud et al. 2006, Gerbe et al. 2008, Plusa et al. 2008). However, whether differential adhesion is sufficient to drive cell sorting is unclear (reviewed in Krens & Heisenberg 2011) and work using embryoid bodies (a model for PrE differentiation) suggests endogenous properties of PrE cells, such as their ability to polarise, are the driving force of this sorting process (Rula et al. 2007, Moore et al. 2009). Further support for the importance of positional cues in the segregation of PrE and epiblast comes from the observation that PrE precursors produce surface proteins, such as LRP2 or DAB2, which localise to the apical surface of these cells when they come into contact with the blastocyst cavity (Gerbe et al. 2008). The polarisation of PrE cells at the surface of the ICM would be the first step in their epithelialisation and may be accompanied...
by the development of cell–cell contacts with other PrE precursors that would help them maintain their position.

In summary, the evidence available indicates that PrE and epiblast cells arise in the ICM as a result of the interplay between regulatory transcription factors and FGF signalling (Fig. 3). FGF4 produced during the morula to blastocyst transition maintains Gata6 expression in some ICM cells, therefore inducing PrE fate, whereas NANOG counteracts this to promote epiblast formation. This interaction leads to the formation of two populations, epiblast and PrE precursors, with distinct gene expression profiles and epigenetic features, which are scattered through the ICM in a ‘salt and pepper’ manner (Chazaud et al. 2006). This restricted gene expression, however, does not result in loss of cell plasticity (Yamanaka et al. 2010, Grabarek et al. 2012), which is largely dependent on the environment surrounding the cell. Continued presence of FGF4 reinforces PrE fate and consequently SOX17 and GATA4 (Frankenberg et al. 2011). PrE and epiblast precursors become sorted into two separate compartments by a number of mechanisms, including cell migration and anchorage at the ICM surface to form a coherent epithelium (Gerbe et al. 2008, Plusa et al. 2008, Meilhac et al. 2009). Their localisation at the surface induces later endodermal markers, such as SOX7 (Artus et al. 2011), suggesting that correct positioning means a step forward in the maturation of the PrE. Finally, the complete separation of PrE and epiblast and the restriction of OCT4 to the epiblast seem to indicate irreversible lineage commitment at the time of implantation (Grabarek et al. 2012).

Variations in the lineage specification mechanisms across mammals

Preimplantation development has been best studied in mice, for evident technical and ethical reasons. Many of the molecular players involved in the specification of the blastocyst lineages in mice are known (see above and Table 1) and the regulatory networks they constitute are starting to be unveiled. However, as our knowledge advances, differences between mouse and other mammals’ preimplantation development are beginning to emerge. For instance, implantation in rodents (mice and rats) occurs on the fifth day of development. By contrast, in humans, implantation takes place around the seventh day and in cattle and sheep over two weeks into

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein name</th>
<th>Relevant references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epiblast</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nanog</td>
<td>NANOG</td>
<td>Chambers et al. (2003), Mitsui et al. (2003), Silva et al. (2009), and Messerschmidt &amp; Kemler (2010)</td>
</tr>
<tr>
<td>Pou5f1</td>
<td>Octamer-binding protein 4 (OCT4)/POU domain, class 5 transcription factor 1 (POU5F1)</td>
<td>Rosner et al. (1990), Schöler et al. (1990a, 1990b), Palmieri et al. (1994), and Nichols et al. (1998)</td>
</tr>
<tr>
<td>Sox2</td>
<td>SRY-related HMG box-containing transcription factor 2 (SOX2)</td>
<td>Avilion et al. (2003)</td>
</tr>
<tr>
<td><strong>Trophectoderm</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdx2</td>
<td>Caudal-type homeobox transcription factor 2 (CDX2)</td>
<td>Chawengsaksophak et al. (1997), Niwa et al. (2005), Strumpf et al. (2005), and Wu et al. (2010)</td>
</tr>
<tr>
<td>Gata3</td>
<td>GATA-binding protein 3 (GATA3)</td>
<td>Home et al. (2009) and Ralston et al. (2010)</td>
</tr>
<tr>
<td>Eomes</td>
<td>Eomesodermin</td>
<td>Ciruna &amp; Rossant (1999) and Hancock et al. (1999)</td>
</tr>
<tr>
<td>Tead4</td>
<td>TEA domain protein family member 4 (TEAD4)</td>
<td>Yagi et al. (2007) and Nishioka et al. (2008)</td>
</tr>
<tr>
<td><strong>Primitive endoderm</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gata6</td>
<td>GATA6</td>
<td>Morrisey et al. (1996, 1998), Koutsourakis et al. (1999), and Fujikura et al. (2002)</td>
</tr>
<tr>
<td>Gata4</td>
<td>GATA4</td>
<td>Arceci et al. (1993), Soudais et al. (1995), Morrisey et al. (1996), Kuo et al. (1997), and Molkentin et al. (1997)</td>
</tr>
<tr>
<td>Sox17</td>
<td>SOX17</td>
<td>Kanai-Azuma et al. (2002), Shimoda et al. (2007), Morris et al. (2010), Niakan et al. (2010), and Artus et al. (2011)</td>
</tr>
<tr>
<td>Sox7</td>
<td>SOX7</td>
<td>Shimoda et al. (2007) and Artus et al. (2011)</td>
</tr>
<tr>
<td>Pdgfra</td>
<td>Platelet-derived growth factor receptor alpha (PDGFRα)</td>
<td>Plusa et al. (2008) and Artus et al. (2010)</td>
</tr>
<tr>
<td>Fgf2</td>
<td>FGF receptor 2 (FGFR2)</td>
<td>Arman et al. (1998) and Haffner-Krausz et al. (1999)</td>
</tr>
<tr>
<td>Grb2</td>
<td>Growth factor receptor bound protein 2 (GRB2)</td>
<td>Cheng et al. (1998) and Chazaud et al. (2006)</td>
</tr>
<tr>
<td>Dab2</td>
<td>Disabled 2 (DAB2)</td>
<td>Morrissey et al. (2000), Morris et al. (2002), Yang et al. (2002), and Gerbe et al. (2008)</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdh1</td>
<td>Cadherin 1/E-cadherin</td>
<td>Hyafil et al. (1980), Vestweber &amp; Kemler (1984), and Larue et al. (1994)</td>
</tr>
<tr>
<td>Prkci</td>
<td>Protein kinase C, iota (aPKC)</td>
<td>Leitges et al. (2001), Soloff et al. (2004), Plusa et al. (2005), and Dard et al. (2009)</td>
</tr>
<tr>
<td>Prkcz</td>
<td>Protein kinase C, zeta (aPKC)</td>
<td></td>
</tr>
<tr>
<td>Yap1</td>
<td>Yes-associated protein 1 (YAP1/YAP)</td>
<td>Zhao et al. (2007) and Nishioka et al. (2009)</td>
</tr>
</tbody>
</table>
Early cell fate decisions in the mouse embryo

Conclusions

Work in the last decade has allowed the identification of some of the molecular players and mechanisms orchestrating the earliest cell fate decisions in mammalian development. Perhaps not surprisingly, it has also become clear that cell interaction and cell communication are just as important as the expression of master regulatory genes for lineage differentiation. The embryonic nuclei take command of their own fate early in development and start producing some of the transcription factors that regulate the first cell differentiation events. However, their co-expression in all the blastomeres of the morula and early blastocyst calls into question how much of an inductive role these transcription factors play at the earliest stages. Instead they rather seem to prime cells for differentiation upon the reception of the appropriate stimuli, these being cell position during the formation of the TE or paracrine FGF4 and cell position during the formation of the PrE. Cell differentiation then proceeds in a gradual manner and plasticity is retained until TE and PrE have become stable epithelia where the gene regulatory networks have become stabilised and de-differentiation is no longer possible. These mechanisms result in remarkable regulatory abilities, which allow the accommodation into the particular context of the embryo of any biases early gene expression might impose on cells. This robust mode of development, therefore, ensures the correct formation of these first lineages, critical for any further progression of mammalian development.

Despite the progress in our understanding of the mechanisms governing early cell fate decisions, many important questions still remain unanswered. We do not yet fully comprehend how the initial heterogeneities among cells arise and how are they translated into divergent patterns of gene expression (Chambers et al. 2007, Singh et al. 2007, Plusa et al. 2008, Kalmar et al. 2009, Canham et al. 2010). What is the relevance of these heterogeneities in the cell differentiation process and what is their importance for the maintenance of pluripotency are currently exciting research questions (Loh & Lim 2011, Martinez Arias & Brickman 2011). Finally, although we know a good deal about the gene networks that regulate lineage commitment and pluripotency, the mechanisms responsible for the remarkable plasticity of mammalian cells remain largely unknown. The challenge now remains to establish precisely how these gene networks maintain particular cellular states or make cells progress (or revert) from one state to another. Only through the deep understanding of these processes will we be able to control them to our advantage.


Brook FA & Gardner RL 1997 The origin and efficient derivation of embryonic stem cells in the mouse. PNAS 94 5709–5712. (doi:10.1073/pnas.94.11.5709)


Copp AJ 1979 Interaction between inner cell mass and trophoblast of the mouse blastocyst. II. The fate of the polar trophoblast. Journal of Embryology and Experimental Morphology 51 109–120.


Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF & Darnell JE 1994 Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. PNAS 91 7598–7602. (doi:10.1073/pnas.91.16.7598)


Hogan B & Tilly R 1978 In vitro development of inner cell masses isolated immunosurgically from mouse blastocysts. II. Inner cell masses from 3.5- to 4.0-day p.c. blastocysts. Journal of Embryology and Experimental Morphology 45 107–121.


Rossant J 1975 Investigation of the determinative state of the mouse inner cell mass. II. The fate of isolated inner cell masses transferred to the ooviduct. Journal of Embryology and Experimental Morphology 33 991–1000.


Tarkowski AK 1961 Mouse chimaeras developed from fused eggs. Nature 190 857–860. (doi:10.1038/190857a0)


Yamanaka Y 2011 Response: cell fate in the early mouse embryo — sorting out the influence of developmental history on lineage choice. Reproductive Biomedicine Online 22 525–527 (discussion 528).


Zernicka-Goetz M 1998 Fertile offspring derived from mammalian eggs lacking either animal or vegetal poles. Development 125 4803–4808.
