CHAPTER 1

EARLY EMBRYONIC CELL FATE DECISIONS IN THE MOUSE

Yojiro Yamanaka* and Amy Ralston*

Abstract: During development, initially totipotent cells of the embryo specialize to form discrete tissue lineages. The first lineages to form in the mouse are the extraembryonic tissues. Meanwhile, cells that do not become extraembryonic retain a pluripotent fate since they can give rise to all the germ layers of the fetus. Pluripotent stem cell lines have been derived from the fetal lineage at several stages of development. Interestingly, multipotent stem cell lines have been derived from the extraembryonic lineages around the same time. Examining the regulation of early embryonic cell fate decisions is therefore a rare opportunity to examine establishment of stem cell progenitors. Classical studies have provided considerable insight into specification of the first three lineages and use of modern molecular and imaging techniques has advanced this field further. Here we describe current understanding of the diverse molecular mechanisms that lead to establishment and maintenance of the first three lineages during mouse development.

INTRODUCTION

During the earliest days of mouse development, initially totipotent cells become restricted in their developmental potential to give rise to the first lineages of the mouse. While in nonmammalian species the first lineage decisions might involve specification of the major body axes, mammals have an altogether different first priority: implantation. Thus discrimination between fetal and extraembryonic tissue lineages comprises the first two lineage decisions (Fig. 1) and precedes establishment of the germ layers (ectoderm, mesoderm, endoderm) and the germline by several days. This uniquely mammalian developmental strategy involves unique cell types that can be isolated and expanded
in culture as stable stem cell lines. Understanding the origins of the extraembryonic tissues therefore illuminates our understanding of establishment and differentiation of stem cells. Classical studies provided considerable insight into specification of the first three lineages and use of modern molecular and imaging techniques has advanced this field further.

Three days after fertilization, the mouse embryo, or blastocyst contains three tissue lineages: epiblast (EPI), trophectoderm (TE) and primitive endoderm (PE). Isolation and study of stem cell lines from these lineages has reinforced and extended our understanding of early embryonic cell fate decisions. Three types of stem cell lines have been derived from the blastocyst: embryonic, trophoblast and extraembryonic endoderm stem cells (ES, TS and XEN cells). Each of these exhibits stem cell properties, such as the ability to either self-renew or to differentiate into multiple mature cell types. Yet each stem cell line exhibits features of the lineage from which it derives, including tissue-specific developmental potential, morphology, transcription factor expression and growth factor requirements. These stem cell lines not only provide an expandable source of pure cell populations for studies requiring large amounts of starting material, but they provide an opportunity to understand where stem cells come from.

Studies performed in ES cells have enabled deeper molecular analysis of the role of genes in cell fate selection. Manipulation of levels of certain lineage-regulating genes causes corresponding changes in stem cell fate. For example, the trophoblast transcription factor Cdx2 is sufficient to convert ES cells to TS-like cells. These kinds of observations demonstrate the remarkable plasticity of ES cells, as well as the central role of genes such as Cdx2 as lineage-determining factors. ES cells also provide an opportunity to examine molecular interactions between lineage-determining genes and thus serve as a model for understanding cell fate selection in the embryo. However, examination of the role of lineage-determining genes in the embryo has revealed that lineage-determining genes play a relatively late role in lineage specification, raising the question as to how the first three lineages are initially specified.

A variety of mechanisms are probably involved in specifying the first lineages, including cell position, shape, polarization, signaling and division plane. A new paradigm is emerging, in which an early pre-stem cell program specifies the tissue lineages as the blastocyst forms. Later, around the time of implantation and thereafter, cell fates are maintained by a program that is active in stem cell lines (Fig. 2).

LINEAGE ESTABLISHMENT AND THE PRE-STEM CELL PROGRAM: FORMATION OF THE BLASTOCYST

Here we will consider the first phase of lineage specification: establishment of the TE and inner cell mass (ICM) as the blastocyst forms. The TE will give rise to placenta, while the ICM contains a mixture of fetal and primitive endoderm progenitors. In the blastocyst, the TE surrounds the ICM and hollow blastocoel and lineage-tracing experiments have shown that TE and ICM populations begin as the outside and inside cell populations of the embryo. That is, as cell cleavage partitions the zygote into two, four, eight and sixteen cells, a small number of cells become enclosed by outside cells. Continued cleavages increase numbers of inside and outside cells, the TE epithelializes and the blastocoel expands, forming the blastocyst structure. The mechanism by which topology becomes linked to cell fate has been elusive. Several models have been put
Figure 1. Overview of the first two lineages decisions during mouse development. The initially totipotent zygote develops to the blastocyst, which contains three lineages: EPI (blue), TE (red, crosshatched) and PE (yellow, lined). These lineages will give rise to the fetus, the placenta and a portion of the yolk sac at later stages of development.

Figure 2. Overview of molecular interactions leading to cell fate specification and maintenance during early mouse development. The Tead4/Yap complex selects TE fates (red, crosshatched) from initially totipotent cells (grey). Cells that do not become TE, then adopt a mixture of EPI (blue) and PE (yellow, lined) fates. Signaling within this lineage facilitates the sorting out of EPI and PE fates. Lineage-specific transcription factors participate in maturation of each lineage.
forward. For example, cell fate could be a consequence of cell position (Fig. 3A). Alternatively, predetermined cell fates could drive cells into appropriate topological positions (Fig. 3B). This latter mechanism predicts that pre-inside and pre-outside cells would be detectable prior to formation of overt inside and outside cell populations. In spite of extensive effort in the field, however, there is currently no support for this predetermination mechanism.

Two main strategies have been used to look for evidence of predetermination among cells prior to the blastocyst stage: lineage tracing and molecular analysis. In terms of lineage tracing, reports of biased developmental potential among cells at the two-cell stage are not relevant to the TE/ICM lineage decision since these studies demonstrate contribution of both cells to the TE and ICM. Likewise, all cells of four and eight-cell embryos can also contribute to both TE and ICM lineages. Although one group reported restricted lineage potential from the four-cell stage, extraembryonic lineages were incompletely scored. Thus there is no evidence from lineage tracing experiments to suggest that cells are predetermined to make TE or ICM prior to formation of inside and outside groups. In terms of molecular analyses, no protein has been detected within a subset of cells prior to the 16-cell stage that instructs the TE/ICM lineage decision. The level of one type of histone methylation is reported to exhibit uneven distribution among blastomeres at the 4-cell stage and correlates with reduced potential to contribute to viable mice in chimeras. The functional importance of these observations in TE/ICM lineage specification needs to be clarified. Therefore, no molecular evidence supports the existence of pre-TE or pre-ICM cells prior to formation of inside and outside cell populations. Rather, inside and outside cells could acquire fates once they have acquired their positions within the embryo.

If cell position acts upstream of cell fate, mechanisms must exist for cells to sense their position within the embryo. Longstanding evidence that cells polarize around the 8-cell stage supports the claim that there are differences along the inside/outside axis at the cellular level. Polarization by conserved polarity proteins such as atypical PKC (aPKC),
Par3 and Par6 is required for maintaining cell position and cell contact has been shown to be required for cell polarization. However the link between position, polarization and cell fate has not been examined at the molecular level. This area is challenging to study using conventional knockout techniques. Many of the proteins involved in cell position and cell contact, such as aPKC, are members of large gene families, suggesting that genetic redundancy may mask their requirements in single gene knockout studies. In addition, this early developmental stage may be regulated in part by maternally supplied protein, requiring germline gene deletion to detect a phenotype. Finally, many of these proteins are involved in basic cellular processes, such as cell division, making it difficult to study their effects during development. On the other hand, overexpression of dominant-negative or siRNA constructs leads to only short-term or partial loss of function, which can also impede phenotype resolution.

Ultimately, to convert inside/outside differences into changes in gene expression, a differentially localized transcription factor is needed. Several strategies have led to the identification of transcription factors involved in early lineage development. Candidates have been identified by microarray analysis of transcripts expressed in pre-implantation development, followed by in situ hybridization to screen for those with restricted expression in the blastocyst. Alternatively, candidates have been identified by microarray comparison of blastocyst-derived stem cell lines. Advances have also come from fortuitous discovery of an unexpectedly early lethal phenotype in knockouts, which led to identification of Cdx2 and Tead4.

While required for TE development, Cdx2 probably does not play an instructive role in TE formation. Nevertheless, Cdx2 mRNA, but not protein, has been reported to localize to the outside surface of cells at the 8-cell stage. Since Cdx2 is not required for specification of TE at either morphological or molecular levels, evidenced by the continued expression of the TE marker Gata3 in Cdx2 null embryos, it is difficult to imagine that localized Cdx2 mRNA plays an instructive role in lineage establishment. Recently, a new pathway, involving Tead4 and cofactors, has been shown to play an instructive role in the first lineage decision. The transcriptional coactivator Yap and related protein Taz, exhibit cell position-sensitive changes in activation of Cdx2 expression. Prior to the blastocyst stage, Yap/Taz localize to nuclei of outside cells and cytoplasm of inside cells. This localization is regulated by phosphorylation by the Hippo signaling pathway members Lats1/2. In addition, manipulation of cell position led to corresponding changes in Yap localization: outside cells embedded inside an aggregate of cells lost nuclear Yap, while inside cells stripped of surrounding outer cells acquired nuclear Yap. Yap/Taz interact directly with Tead4 a DNA binding protein required for expression of Cdx2 and other trophectoderm markers. The identity or nature of Yap/Taz-regulating signals that can sense cell position are unknown, but probably involve the Hippo signaling pathway and possibly proteins involved in cell contact such as cadherins. This will undoubtedly be an exciting area of research to follow in the future.

Besides what is working upstream of Yap/Tead4, it is not entirely clear what is working downstream. Tead4 is required for Cdx2 expression, but Tead4 null embryos die prior to blastocyst formation, while Cdx2 null embryos die after blastocyst formation. Tead4 is not required in the ICM, so additional genes must operate in parallel to Cdx2 in the TE. Some of these, such as Gata3 are beginning to be identified. It will be important to identify Tead4 targets that participate in promoting outside cell proliferation and construction of the blastocyst.
LINEAGE MAINTENANCE AND THE STEM CELL PROGRAM: BEYOND THE BLASTOCYST

In the blastocyst, interactions between lineage-determining transcription factors reinforce TE and ICM fates established at earlier stages. Central players at this stage are Oct4 (Pou5f1) and Cdx2. Oct4 is required for maturation of the ICM, while Cdx2 is required for maturation of the TE. Mutual antagonism between these two factors was initially speculated to cause the first lineage decision. Cdx2 is required for repression of Oct4 and other ICM genes in the TE of the blastocyst. But the TE still forms in Cdx2 null embryos and other TE markers are still expressed. Similarly, Oct4 represses Cdx2 in the ICM, but not until implantation, a full day after blastocyst formation. Thus lineage specification is initially normal in the absence of either Oct4 or Cdx2, but embryos fail to maintain correct expression of lineage genes. Nevertheless, in spite of adoption of ICM gene expression, Cdx2 null TE does not fully adopt ICM fate. The TE marker Gata3 is still expressed in the TE of Cdx2 null embryos and Cdx2 null embryos exhibit higher levels of apoptosis in the TE than do wild type embryos. Cdx2 must therefore enable survival and/or proliferation of cells that are already committed to being TE. This is consistent with its continued expression in the proliferative region of the trophoblast at later stages. The reason for the lethality of Oct4 null embryos is currently unclear.

The antagonistic relationship between Oct4 and Cdx2 is borne out by stem cells from the blastocyst. ES cells cannot be derived from Oct4 null embryos and TS cells cannot be derived from Cdx2 null embryos. Loss of Oct4 from existing ES cell lines leads to upregulation of Cdx2 and formation of TS-like cells in the presence of TS cell culture medium. Similarly, overexpression of Cdx2 in ES cells leads to repression of Oct4 and formation of TS-like cells. Other trophoblast factors, such as Eomes and Gata3 can also induce trophoblast gene expression in ES cells and these also play relatively late roles in trophoblast maturation rather than allocation. Maintenance of the TE/ICM lineage restriction in stem cells therefore appears to use genetic programs that become active once the blastocyst has formed. This makes sense given that stem cell derivation requires culture beyond the blastocyst stage. Understanding the further development of the ICM, however, requires a look at the second lineage decision in development, discussed next.

THE SECOND LINEAGE DECISION: SUBDIVIDING THE ICM

Three days after fertilization, the ICM of the blastocyst contains two cell types: the epiblast (EPI) and the primitive endoderm (PE). Only the EPI gives rise to the fetus, whereas the PE is an extraembryonic lineage, which contributes to the yolk sac (Fig. 1). The PE lineage plays two important roles just after implantation. The first is that it provides nutrients to the embryo and the second is that it serves as a signaling center that helps confer anterior-posterior polarity upon the gastrulating embryo. As for the TE lineage, a special stem cell line can be derived from the PE lineage. Multipotent stem cell lines, called XEN cells, have been derived from the PE lineage (Fig. 2). In addition, PE-like cells can be induced from ES cells by overexpression of PE transcription factors, such as Gata4 and Gata6. Yet Gata4/6 act relatively late in PE development, suggesting that, as for the TE lineage, the PE is specified by a mechanism acting upstream of the stem cell genes. Insight into specification of the PE lineage has revealed a unique cell signaling-based strategy.
Heterogeneity and Progenitor Sorting

Four days after fertilization, the blastocyst implants. At this stage, the PE appears as a distinct monolayer on the blastocoel surface of the ICM. For this reason, the PE was originally assumed to arise from ICM cells directly facing the blastocoel around the time of implantation. Microenvironmental differences between blastocoel-facing and deeper cells were postulated to participate in lineage specification at this stage. However, recent studies have shown that EPI and PE progenitors can be detected in the blastocyst one full day before implantation. At this stage, the ICM appears as a mixed population of EPI and PE progenitors, expressing lineage-specific transcription factors. Prior to this stage, Nanog and Gata6 are coexpressed in all cells of the ICM and expression gradually becomes mutually exclusive to specify the two progenitors in a position-independent manner during blastocyst expansion. Notably, there is no stereotyped pattern of distribution of the two progenitors within the ICM. Rather, they are sprinkled randomly throughout the ICM like salt and pepper.

These results suggest that the two randomly distributed lineage progenitors sort out to form two morphologically distinct layers by implantation. Indeed, support for this model has been provided by live imaging of blastocyst expansion in transgenic mice expressing fluorescent lineage markers. In the Pdgfra\(^{H2B\cdotGFP}\) mouse line histone H2B-GFP is expressed in the PE and revealed that separation of the two lineages involves both apoptosis and cell migration. Cells within the growing ICM appear to rearrange constantly, but once PE progenitors come to the ICM surface they stay there. Consistent with this, the maturation of the PE takes place progressively and this is correlated with position within the ICM. One outstanding question is whether PE cells sort out by directional cell movement or a combination of random movement and position recognition.

Several mutants exhibit a defect in formation of a cohesive PE layer. In these mutants, Gata4-expressing, presumptive PE cells, are found clustered within the middle of the ICM, suggesting that PE progenitors are specified but fail to form a morphologically distinct surface layer. This contrasts with the TE, in which lineage allocation (position) precedes lineage specification. For the PE, lineage specification precedes allocation. Understanding how PE fates are selected from within the ICM is therefore key to understanding PE/EPI lineage choice.

CELL SIGNALING REGULATES PE/EPI SPECIFICATION

Early heterogeneity in the ICM suggests that position-independent mechanisms regulate specification of PE and EPI lineages. FGF signaling has been shown to be necessary for PE formation in vivo and in vitro. How extracellular signaling pathways, such as the FGF signaling pathway, could participate in the generation of a salt and pepper distribution of PE and EPI within the ICM is not clear. For example, certain pre-PE cells within the embryo could be predisposed to respond to signals, or cells could randomly receive signals and thereby become PE progenitors.

These possibilities are summarized in two models: the origin-dependent model and the signaling-dependent model (Fig. 4A,B). The origin-dependent model relies on understanding the process of inner cell generation during the cleavage stages. Inner cells of the morula, which will become the ICM of the blastocyst, are generated from two rounds of asymmetric divisions at 8-16 and 16-32 cell stages. According to the
origin-dependent model, the developmental origins of individual ICM cells determine their fate. That is, inner cells generated in the first round of divisions (primary inner cells) would preferentially adopt the EPI fate, whereas cells generated in the second round (secondary inner cells) would preferentially become PE (Fig. 4A). Secondary inner cells would be predisposed to become extraembryonic due to their prolonged external position since TE cells are also external. To test the origin-dependent model, generation of inner cells was first directly observed in live embryos and then the contribution of their progeny to EPI and PE lineages was analyzed at later stages. No difference in lineage potential was detected between primary and secondary inner cells since both primary and secondary inner cell progeny contributed to EPI and PE lineages without an obvious bias. These observations therefore suggest that the origin-dependent model is unlikely.

The second model is a signaling-dependent model, in which individual ICM cells stochastically respond to certain levels of FGF signaling to choose EPI or PE fates (Fig. 4B). As described above, FGF signaling is necessary for PE formation in the embryo. When FGF signaling is blocked, using either chemical inhibitors or by gene knockouts, all ICM cells adopt EPI fates. Interestingly, high doses of exogenous FGF4 can induce the converse phenotype: all ICM cells adopt PE fates. This suggests that all early ICM cells have the potential to respond to FGF signaling and become PE. During normal development, however, limited amounts of endogenous FGFs would restrict the proportion of FGF-responding ICM cells (Fig. 5). Whether or not individual ICM cells...
respond to the limited amount of FGFs could be stochastically determined by cell-to-cell variation in sensitivity determined by cell-autonomous or non-autonomous mechanisms. Endogenous levels of FGFs, governed by developmental genetic programs, would thereby generate roughly equal proportions of EPI/PE lineages reproducibly, without need for deterministic developmental mechanisms.

**ESTABLISHMENT AND MODULATION OF PLURIPOTENCY IN THE EPI LINEAGE**

After two rounds of lineage specification, first the TE and then the PE, the EPI is established as a pluripotent lineage. While pluripotency is generically defined as the ability to form tissues from all three embryonic germ layers, recent identification of
multiple pluripotent stem cell lines makes apparent that pluripotency is not a single state. Rather, pluripotency may comprise a range of states with developmental equivalence in the embryo. There are at least two states of pluripotency in the mouse embryo, represented by two types of pluripotent stem cells: ES cells and epiblast-derived stem cells (EpiSCs). These cell lines are derived from the EPI lineage, but represent two distinct embryonic stages: ES cells are equivalent to the EPI cells of the implanting embryo, while EpiSCs are equivalent to EPI cells of the embryo just after implantation and prior to gastrulation. Although EpiSCs cannot contribute to embryos when they are injected into blastocysts, probably due to failure to integrate into host ICMs, they can generate all three germ layers in teratomas. Pluripotency genes such as Oct4 and Sox2 are both expressed during early and late stages, but several features differ between EPI cells over the course of implantation. For example, cell morphology changes from an unorganized cell mass to an epithelial monolayer. In addition, expression of some genes change dramatically during the transition, such as Rex1, which is downregulated and Fgf5, which is upregulated. After the transition, late EPI cells are competent to receive inductive signals to generate three germ layers. Understanding how the pluripotent state is safeguarded during establishment of first, second and subsequent lineages is an active area of research.

One gene potentially involved in safeguarding the pluripotent state is Nanog. Nanog was originally identified as a transcription factor essential for maintaining pluripotency in ES cells, but subsequent studies have revealed that it acts rather as a gate-keeper instead. That is, Nanog levels fluctuate in ES cells in an FGF-signaling dependent manner and ES cells are more prone to differentiate when Nanog levels are low. Downregulation of Nanog does not initiate differentiation but permits it. This is consistent with the endogenous Nanog expression which is transiently downregulated during implantation. Nanog null blastocysts are morphologically normal but ICM cells degenerate soon after the blastocyst stage. Nanog null ICM cells appear to be trapped in a prepluripotent state, specified as neither EPI nor PE, but as a nonviable indeterminate state.

Not much is known regarding mechanisms regulating the transition from early to late EPI tissues in the embryo. In vitro stem cell studies have provided some insight. ES cells have been found to readily become EpiSCs, when cultured in EpiSC culture conditions, including FGF2 and activin A. Interestingly, Fgf4 is required for ES cell differentiation. These observations suggest that the quality or amount of FGF signaling may participate in the transition from ES cells to EpiSCs and possibly ICM to EPI fates.

Interestingly, it is also possible to reverse the transition between the two pluripotent states. That is, EpiSCs can become ES-like following overexpression of KLF4, one of the original reprogramming factors, although reversal occurs with very low frequency. However, when EpiSCs or epiblast cells from gastrula embryos are cultured in conventional ES cell culture conditions, reprogrammed ES-cell-like cells (rES cells) emerge after 10-20 days. Although the reversion takes more time than progression from ES cells to EpiSCs, rES cells have fully reestablished the early pluripotent state.

CONCLUSION

Here, we have described lineage specification in the mouse, currently the most extensively analyzed mammalian embryo. One of the most interesting lessons from the mouse lays in the observation that multiple pluripotent states exist, evidenced by the
existence of ES cells and EpiSCs. Thus far, rat ES cells are the only other mammalian ES cell line similar to mouse ES cells. In contrast, human ES (hES) cells are more similar to mouse EpiSCs than they are to mouse ES cells in morphology, gene expression, and growth factor dependency.\textsuperscript{61,62,72} At this point, it is not known whether the human embryo also has multiple pluripotent states. Interestingly, adult human and mouse cells reprogrammed by identical factors resemble ES cells of their respective species. That is, human iPS cells resemble hES cells,\textsuperscript{73,74} while mouse iPS cells resemble mouse ES cells\textsuperscript{70} and not mouse EpiSCs. Perhaps there are species-specific differences in the stability of pluripotent states. Even though all mammals develop using a blastocyst, the developmental timing of implantation and morphology of early postimplantation embryos are highly varied. Further analysis of lineage specification and stem cells from other mammalian species should provide exciting insight into these issues.

REFERENCES


