REVIEW

The pluripotent state in mouse and human
Kathryn C. Davidson1, Elizabeth A. Mason2,3 and Martin F. Pera3,4,*

ABSTRACT

In the mouse, naïve pluripotent stem cells (PSCs) are thought to represent the cell culture equivalent of the late epiblast in the pre-implantation embryo, with which they share a unique defining set of features. Recent studies have focused on the identification and propagation of a similar cell state in human. Although the capture of an exact human equivalent of the mouse naïve PSC remains an elusive goal, comparative studies spurred on by this quest are lighting the path to a deeper understanding of pluripotent state regulation in early mammalian development.

KEY WORDS: Human, Mouse, Pluripotent state

Introduction

Many of the impressive advances in our ability to manipulate pluripotent stem cells (PSCs) in vitro, and to convert them into differentiated progeny for use in research and therapy, have relied heavily on our knowledge of normal embryogenesis and the intrinsic and extrinsic factors that control cell fate during development. For this reason, it is important to understand where in vitro-cultured PSCs and their lineage-specified progeny fit into the paradigm of the developing embryo in vivo.

Shortly before implantation, the inner cell mass (ICM) of the blastocyst gives rise to the primitive endoderm, which will generate the visceral and parietal yolk sacs, and the epiblast, which represents the precursor of all tissues of the body, including the germ line. As post-implantation development progresses, the late epiblast becomes poised for lineage specification. The pre-implantation epiblast represents a naïve state of pluripotency, prior to any lineage specification, whereas the post-implantation epiblast is said to represent a primed pluripotent state (Fig. 1; Nichols and Smith, 2009). Pioneering work by Richard Gardner showed that, although mouse embryonic stem cells (mESCs) can be derived from the E3.5 implantation embryo, with which they share a unique defining set of features. Recent studies have focused on the identification and propagation of a similar cell state in human. Although the capture of an exact human equivalent of the mouse naïve PSC remains an elusive goal, comparative studies spurred on by this quest are lighting the path to a deeper understanding of pluripotent state regulation in early mammalian development.

argue that mESCs maintained in the naïve state are very similar in phenotype and function to the mouse pre-implantation epiblast (Nichols and Smith, 2009).

The derivation of human ESCs (hESCs) from the pre-implantation blastocyst, and subsequent analysis of their phenotype and growth requirements, revealed that they differed in significant ways from the mouse counterparts (see below). Early on, these findings raised the important question as to which developmental stage the hESCs might correspond to. Some years after the derivation of hESCs, several groups showed that it was possible to isolate PSCs from the mouse post-implantation epiblast (Brons et al., 2007; Tesar et al., 2007). These cells, called epiblast stem cells or EpiSCs, although pluripotent by the criteria of in vitro differentiation, teratoma formation and contribution to chimeras following culture in the post-implantation epiblast in vitro (Huang et al., 2012), were unable to participate in formation of germline chimeras after blastocyst injection, an assay that indicates pluripotency in vivo. The mouse EpiSC (mEpiSC) phenotype was more similar to that of hESCs than to that of mESCs. Many interpreted these data to mean that hESCs were in fact equivalent to mEpiSCs, and therefore representative of a later stage of development than that of mESCs. It followed from this hypothesis that it ought to be possible to isolate hESCs that correspond directly to their naïve mESC equivalents. The alternative hypothesis, that PSCs at equivalent stages of ontogeny in the two species might display different phenotypes, has not been widely explored, but over the past two years a number of attempts to derive naïve human PSCs (hPSCs) have been reported. In this Review, we will consider recent efforts at naïve hPSC derivation in the context of background work in the mouse, and of what is known concerning embryonic development in the human. For a discussion of the historical context to these studies and the potential advantages that might accrue from derivation of naïve hPSCs, see Pera (2014).

Naïve pluripotency in the mouse

 naïve and primed cells differ in vitro

Following the development of defined culture systems that enabled the propagation of naïve mouse PSCs (mPSCs) in a relatively pure form, various studies identified the key features of naïve pluripotency that corresponded to the properties of the late pre-implantation epiblast (Table 1; Hackett and Surani, 2014). These properties include the ability to give rise to all somatic lineages and the germ line in chimeras following either blastocyst injection or tetraploid complementation; expression of a specific set of genes associated with the naïve pluripotent state; a global state of DNA hypomethylation; two active X-chromosomes in female cell lines; a low level of bivalent histone marks; a capacity to generate energy through both oxidative phosphorylation (associated with a high mitochondrial content) and glycolysis; the ability to propagate as single cells with high cloning efficiency; a lack of lineage priming or the co-expression of pluripotency-associated genes and lineage-specific markers.

By contrast, EpiSCs display a different pattern of pluripotency-associated gene expression, enhanced levels of DNA methylation,
X-inactivation, energy generation through glycolysis, poor cloning efficiency as single cells, lineage priming and little or no ability to participate in chimera formation (Hackett and Surani, 2014). In addition to the characteristic changes in expression of a specific subset of genes that marks the transition between the naïve state and the post-implantation epiblast-like EpiSC (Table 1), the expression of all pluripotency-associated genes is more uniform in naïve cells in vivo or in vitro compared with EpiSCs. Low variability of gene expression and the absence of lineage priming suggest that in vitro, the naïve cell might represent a stable, deep attractor state (a stable state toward which a system will evolve). However, the naïve state converts within hours in culture (if conditions are altered) or in the embryo, to intermediates that show wider variability in pluripotent gene expression as well lineage priming. Hence, the naïve attractor state is subject to rapid and facile remodelling in vivo (Kalkan and Smith, 2014).

Recent work has refined our understanding of the developmental status of stem cell lines corresponding to the post-implantation epiblast. Careful comparison between gene expression in mEpiSCs derived at different stages of post-implantation development and the mouse epiblast from E5.5 to E8.5 shows that mEpiSCs are most similar to the ectoderm of the anterior primitive streak, a relatively late stage of development (Kojima et al., 2014). By modifying the culture conditions for mEpiSCs, Wu et al. were able to derive mEpiSCs that were more similar to late stage post-implantation epiblasts (Wu et al., 2015). Both of these studies indicate that mEpiSCs are poised at a fairly late stage of epiblast development.

The naïve state: similarities and differences in vivo and in vitro

As noted above, the molecular and cellular properties of naïve PSCs reflect the properties of the epiblast (Fig. 1). Prior to epiblast formation, the constituent cells of the ICM are not yet committed to the pluripotent or primitive endoderm lineages (an extra-embryonic tissue that is the precursor of the yolk sac). Thus, prior to epiblast segregation, levels of the pluripotency gene Nanog in the ICM are bimodal, with low levels of expression associated with high levels of expression of the primitive endoderm gene Gata4. After the epiblast forms, Nanog levels are high within the pluripotent compartment and low or undetectable in the primitive endoderm, and gene expression in the epiblast cells reflects a uniform pluripotent state (Boroviak et al., 2014; Guo et al., 2010). The absence of co-expression of lineage and pluripotency markers represents a transient state in mouse development, because very soon after implantation, the post-implantation epiblast begins to express germ layer lineage-specific markers in preparation for gastrulation.

In vitro, it has repeatedly been observed that mESCs are heterogeneous for expression of Nanog (Chambers et al., 2007). Although mPSCs can flux between Nanog+ and Nanog− states in vitro, this is not reflected in the embryo in vivo, where, once epiblast cells become specified, they no longer move backwards in developmental time to give rise to primitive endoderm (Xenopoulos et al., 2015). mPSCs grown under conditions that support naïve

### Table 1. Key characteristics of the naïve and primed pluripotent stem cell states

<table>
<thead>
<tr>
<th>Corresponding in vivo tissue:</th>
<th>Naïve state PSCs</th>
<th>Primed state PSCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colony appearance:</td>
<td>Early epiblast (pre-implantation)</td>
<td>Epiblast (peri-post-implantation)</td>
</tr>
<tr>
<td>Genes expressed:</td>
<td>Oct4, Nanog, Sox2, Klf2, Klf4, Zfp42, Esrb, Dppa3, Tcfcp21, Fgf4, Tbx3</td>
<td>Oct4, Nanog, Sox2, Dmnt3b, Fgf5, Pou3f1, Meis1, Otx2, Sox11, Gdf3</td>
</tr>
<tr>
<td>Oct4 enhancer usage:</td>
<td>Distal</td>
<td>Proximal</td>
</tr>
<tr>
<td>Global DNA methylation:</td>
<td>Hypomethylated</td>
<td>Hypermethylated</td>
</tr>
<tr>
<td>Female X chromosome status:</td>
<td>Two active X</td>
<td>One inactivated X</td>
</tr>
<tr>
<td>Clonogenicity:</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Gene targeting feasibility:</td>
<td>Amenable to gene targeting</td>
<td>Low targeting efficiency</td>
</tr>
<tr>
<td>Chimeric contribution in rodents:</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Level of HERVH* expression:</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Growth factor dependence:</td>
<td>LIF</td>
<td>Activin, FGF2</td>
</tr>
</tbody>
</table>

*HERVH is a primate-specific endogenous retrovirus.

Features are based primarily on data from the mouse (Hackett and Surani, 2014), except for HERVH expression (Wang et al., 2014).
pluripotency show more uniform Nanog expression than those cultured in serum containing medium supplemented with LIF. Nevertheless, a small subset of mESCs co-expresses pluripotency and extra-embryonic endoderm markers even when maintained in naïve culture conditions (i.e. in the presence of LIF plus GSK3 and ERK inhibitors) (Canham et al., 2010; Marks et al., 2012; Morgani et al., 2013). This co-expression pattern is seen at the single-cell and population levels. Cells with this phenotype might represent an earlier developmental state than epiblast cells, as this gene expression phenotype is more equivalent to ICM, and these minority cells can give rise to extra-embryonic endoderm. It has long been observed that mESCs are able to undergo differentiation into primitive endoderm in vitro (Niwa, 2010), a feature that is suggestive of an ICM identity rather than that of a pure epiblast population, and it is in fact possible to derive extra-embryonic endoderm stem cells directly from mouse naïve ESCs (Cho et al., 2012). Thus, even naïve mPSC cultures might contain a subset of cells that resembles the ICM rather than the epiblast.

Alternative pluripotent states in vitro

Two groups have described stem cells that display characteristic corresponding to an interesting intermediate state between the pre-implantation epiblast and the epiblast at 6.5 days (Buecker et al., 2014; Hayashi et al., 2011). These cells have been called epiblast-like cells (EpiLCs). EpiLCs are derived from mESCs under conditions similar to those used for maintenance of mEpiSCs or hESCs, but unlike these cells, they cannot be maintained in culture beyond a few days. As they form, EpiLCs switch off genes expressed in naïve cells and switch on a subset of genes similar to EpiSCs. In contrast to EpiSCs, however, EpiLCs can efficiently give rise to germline cells upon chimera formation. These EpiLCs most likely represent a rather early post-implantation epiblast cell (around E5.5), in contrast to EpiSCs, which resemble more closely cells of the late gastrula or primitive streak (around E6.5). Early post-implantation epiblast cells are poised to receive signals for somatic and germline differentiation, and such cells are thus downstream from naïve pluripotent pre-implantation epiblast cells. This condition has been termed a state of ‘formative’ pluripotency, prior to entry into the primed state (Kalkan and Smith, 2014). Thus, the generation of EpiLCs from naïve PSCs represents a destabilization of the naïve PSC attractor state by exposure to different growth regulatory milieu.

In summary, although it is clear that mPSCs maintained under conditions that support the naïve state are distinct from later stages of formative and primed pluripotency, minority populations within these naïve cultures might represent an earlier stage of embryonic development resembling the ICM or even the two cell stage (Macfarlan et al., 2012).

Pluripotent states in the human embryo

Attempts to isolate naïve hPSCs with properties similar to mouse cells are predicated on the assumption that the phenotypes of pluripotent cells in pre- and post-implantation embryonic development in the two species are similar. This is an assumption difficult to validate at present. Data on post-implantation embryonic development in primates are almost entirely confined to high quality morphological and descriptive studies (Enders et al., 1986; Tarara et al., 1987); no molecular characterization of the post-implantation primate epiblast has been reported to date. Through in vitro fertilisation it has been possible to study human development from the zygote through to the blastocyst stage. However, in human embryos grown in vitro, the segregation of the epiblast from primitive endoderm is not really complete until day 6-7 of cultivation (Kuijik et al., 2012; Roode et al., 2012), and then only in high-quality (often hatched) blastocysts. As studies of human embryos in vitro do not often reach this milestone, and because the properties of the three constituent cell types of the 6-7 day embryo are most meaningfully analysed at the single-cell level (which has not yet been carried out in sufficient detail), the properties of the pre-implantation human epiblast remain inadequately defined at present.

Patterns of pluripotency-associated gene expression might differ between human and mouse

Single-cell gene expression studies in pre-implantation embryos provide a more detailed characterization of developmental cell phenotypes, uncovering distinguishing patterns often masked at the population level. Recently, Yan et al. characterised the epiblast, trophectoderm and primitive endoderm lineages in human blastocysts by analysis of single cells using RNAseq (Yan et al., 2013). The RNAseq analysis was also extended to hESCs. We have re-evaluated these data, investigating sample clustering according to a subset of genes associated with mouse naïve and/or primed pluripotent states (Fig. 2). The expression pattern of this gene set is sufficient to identify discrete stages of pre-implantation development and to segregate hPSCs from blastocyst cells. The expression patterns of some genes characteristic of the naïve mouse phenotype do distinguish human blastocyst cells from hPSCs, as predicted if hPSCs are indeed equivalent to primed EpiSCs. These genes include TFCPL1, ESRRB, TRX3 and DPPA3 (all expressed at higher levels in blastocyst cells relative to conventional hESCs, as predicted from the mouse), and OTX2, SOX11, GDF3, NODAL, FST and FGFR1 (all expressed at higher levels in hESCs relative to blastocyst, again as predicted from mouse data). However, many genes did not fit the predicted patterns of naïve and primed states based on data from the mouse. Indeed, we identified patterns of gene expression that challenge the assumption that human and mouse share equivalent developmental states. ZFP42, TEAD4, FBOX15, NROB1, TEAD4 and KLF4 are expressed at similar levels in human blastocyst cells and hESCs, while they are considered characteristic markers for naïve cells in the mouse, and are downregulated in the post-implantation epiblast and mouse EpiSCs. FGF5, POU5F1 and MEIS1, which switch on when mouse cells exit the naïve state, are expressed at similar levels in human blastocyst cells and hESCs, whereas these genes are more strongly expressed in naïve versus primed mouse cells. If the blastocyst cells surveyed in this study represent the pre-implantation epiblast stage, the comparison of gene expression between conventional hESCs and human epiblast does not yield the same demarcation between primed and naïve states observed in the mouse.

There are two important caveats about the study conducted by Yan et al. The first is that the results are based on three blastocysts from a single study, and although they are consistent, further assessment is clearly indicated, particularly as a subset of blastocyst cells in this study maps close to hESCs in the principal component analysis. The second caveat is that, in all cells of the blastocysts studied, NANOG, GATA6 and KRT7 were expressed at high levels. This co-expression of pluripotency and extra-embryonic marker genes raises the possibility that the embryos had not yet actually reached the epiblast stage of development, and that the cells studied equate to the ICM stage in the mouse. If this is the case, comparison of blastocyst gene expression with ESCs in this study is not
The process of X-inactivation varies between human and mouse as demonstrated by the presence of two active X chromosomes in female cell lines. The mouse genome of 6- to 10-week embryos has considerably higher levels of DNA methylation compared with that of the ICM. Demethylation is largely complete by cleavage stages, and ESCs exhibit much higher levels of DNA methylation compared with that of the ICM. Indeed, DNA methylation occurs rapidly during the early phases of ESC derivation from ICM outgrowths in vitro. As expected, the genome of 6- to 10-week embryos has considerably higher levels of methylation.

Another key epigenetic hallmark of the mouse naïve state is the process of X-inactivation, which is determined early in development and occurs in the female during the peri-implantation stages. Studies of X-inactivation in human cells have shown that the process is more complex and involves the activation of the inactive X chromosome (XIST) in a specific way.

**Epigenetic features define naïve and primed pluripotent states**

Studies of global DNA methylation in the human embryo (Guo et al., 2014; Smith et al., 2014) have found widespread demethylation between fertilisation and the two-cell stage, with demethylation continuing through to the ICM stage. Demethylation is largely complete by cleavage stages, and ESCs exhibit much higher levels of DNA methylation compared with that of the ICM. Indeed, DNA methylation occurs rapidly during the early phases of ESC derivation from ICM outgrowths in vitro. As expected, the genome of 6- to 10-week embryos has considerably higher levels of methylation.

Another key epigenetic hallmark of the mouse naïve state is the presence of two active X chromosomes in female cell lines. The process of X-inactivation varies between human and mouse (Okamoto et al., 2011; van den Berg et al., 2009). X-inactivation is not coupled to the expression of the X inactive specific transcript (XIST) in the human, as it is in the mouse, and much more of the...
human X chromosome escapes X-inactivation. In the human, a long non-coding RNA called X active specific transcript (XACT, which does not exist in the mouse) coats the active X chromosome and might be responsible for protecting the active X from inactivation (Vallot et al., 2013). The timing and degree of synchrony of X-inactivation during human development have not been clearly elucidated, but the available data indicate that there are two active X chromosomes in most cells of the ICM which must undergo inactivation shortly after this stage of development.

**Signalling pathways involved in early lineage specification differ between mouse and human**

The extrinsic signalling pathways governing early lineage specification are not equivalent between mouse and human embryos. For example, FGF signalling activation in pre-implantation development has opposing effects in the ICM of the two species. FGF signalling in the mouse ICM is required for differentiation of some of its constituent cells into primitive endoderm. However, blockade of FGF signalling has no effect on the segregation of the primitive endoderm lineage from the epiblast in human embryos (Kuijk et al., 2012; Roode et al., 2012). The role of FGF signalling is important because this pathway is also required for the self-renewal of conventional hESCs and, indeed, most naïve populations of hESCs in vitro (see below).

In summary, the embryological roadmap for understanding the true identity of hPSCs in the context of pre- and post-implantation development in human is still fragmentary, but it is improving steadily. It is already apparent that, whereas many features of the development of the pluripotent cells in the embryo are conserved between mouse and human, there are some important and significant differences.

**Towards human naïve PSCs?**

Human or non-human primate PSCs with some features of naïve pluripotency have been derived by the conversion of conventional hESCs or induced PSCs (iPSCs) in vitro, through reprogramming of somatic cells to a pluripotent state or by direct derivation from the embryo (Tables 2 and 3) (Buecker et al., 2010; Chan et al., 2013; Fang et al., 2014; Gafni et al., 2013; Hanna et al., 2010; Li et al., 2009; Takashima et al., 2014; Theunissen et al., 2014; Valamehr et al., 2014; Ware et al., 2014; Wang et al., 2014). Early studies used exogenous transgenes to trigger the conversion from primed to naïve pluripotency and, in some cases, to stabilise and maintain the naïve state. Dependence on continued transgene expression is not a desirable feature for many applications of hPSCs; thus, subsequent studies focused on conditions that could maintain cells in the naïve state without the need for expression of exogenous transgenes. Several protocols have been established for the maintenance of naïve hPSCs in vitro in the absence of transgene expression, but each produces a different transcriptional profile depending on the signalling pathways targeted. Below, we describe the methods used to generate human naïve cells and the results of the characterization of the cells phenotype in vitro and their developmental potential in vivo.

**Modifying culture protocols to generate human naïve cells**

The culture systems devised to support naïve pluripotent human cells vary considerably. Most are based around the original concept of using inhibitors of MEK/ERK and GSK3B along with LIF in the absence of a feeder layer, as described for naïve mESC propagation (Ying et al., 2008). However, this system, in its original form, has universally failed to maintain hPSCs (see references in Tables 2 and 3). Therefore, a range of small molecule inhibitors that support some aspects of naïve pluripotency in human cells has been identified using various assays, mostly comprising empirical screening approaches. Although these systems are rather diverse, many employ FGF2 and some form of TGFR-β superfamily agonist, and most employ mouse embryonic fibroblast feeder cell support and complex protein additives such as knockout serum replacement (KSR) (Tables 2 and 3). These systems are thus much less defined than the elegantly minimalist 2i conditions described by Smith and co-workers (Ying et al., 2008). It thus appears that if human cells analogous to naïve mESCs exist, their growth requirements are rather different. In particular, the activity of FGF in mouse versus human PSCs appears to represent a fundamental difference between the species, reflected in the activity of this factor in the embryos of the two species, as mentioned above. The action of FGF2 in hPSC maintenance is probably mediated through PI3K rather than the MEK/ERK pathway (Singh et al., 2012), so inhibition of the latter pathway, probably mediated through PI3K rather than the MEK/ERK, has universally failed to maintain hPSCs (see references in Tables 2 and 3). Therefore, a range of small molecule inhibitors that support some aspects of naïve pluripotency in human cells has been identified using various assays, mostly comprising empirical screening approaches. Although these systems are rather diverse, many employ FGF2 and some form of TGFR-β superfamily agonist, and most employ mouse embryonic fibroblast feeder cell support and complex protein additives such as knockout serum replacement (KSR) (Tables 2 and 3). These systems are thus much less defined than the elegantly minimalist 2i conditions described by Smith and co-workers (Ying et al., 2008). It thus appears that if human cells analogous to naïve mESCs exist, their growth requirements are rather different. In particular, the activity of FGF in mouse versus human PSCs appears to represent a fundamental difference between the species, reflected in the activity of this factor in the embryos of the two species, as mentioned above. The action of FGF2 in hPSC maintenance is probably mediated through PI3K rather than the MEK/ERK pathway (Singh et al., 2012), so inhibition of the latter pathway, as originally described in the mouse and implemented in most naïve hPSC protocols, will not disrupt the positive action of the FGF2 in promoting hPSC self-renewal. It is most important to note, in undertaking comparisons of putative naïve cells with conventional hPSCs, that culture systems for the latter differ from laboratory to laboratory (Tables 2 and 3).
Most human naïve systems described so far support higher cloning efficiency than that attained with conventional hPSCs (Table 4). However, it is well known that cells can undergo adaptation to propagation as single cells even in conventional culture systems (Chan et al., 2008; Hasegawa et al., 2006). Adaptation to single-cell survival is often associated with both epigenetic and characteristic genetic changes (Amps et al., 2011; Gafni et al., 2013; Takashima et al., 2014; Valamehr et al., 2014). The studies to date have examined a number of parameters to determine whether the culture systems they describe in fact support the maintenance of naïve pluripotency (Table 4). As noted above, there is a characteristic set of genes that distinguish naïve from primed cells in the mouse. In many of these naïve culture systems, these genes are expressed at higher levels and more consistently compared with conventional cultures of hESCs (Table 4). However, it is important to recognise some limitations of these comparisons. Conventional hPSC cultures are heterogeneous, and a careful examination of gene expression in these cultures shows that subsets of cells grown under conventional conditions have distinct patterns of gene expression (see below). In fact, the data

### Table 3. Protocols for transgene dependent derivation or conversion of primate naïve PSCs

<table>
<thead>
<tr>
<th>Transgenes expressed</th>
<th>Basal medium</th>
<th>Growth factors/inhibitors</th>
<th>Feeders/matrix</th>
<th>O2</th>
<th>Conventional primed conditions for comparison</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>Albumax+N2 or 20% KSR</td>
<td>2IL, FGF2, TGF-β, ROCK1,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>JNK1, p38i, PKC, insulin</td>
<td>Gelatine/VN or</td>
<td></td>
<td>15% KSR+FGF2, TGF-β on gelatine/VN or MEFs</td>
<td>Gafni et al., 2013</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>N2B27</td>
<td>2IL, FGF2, Activin, SRCi,</td>
<td>MEFs</td>
<td>5%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Theunissen et al., 2014</td>
</tr>
<tr>
<td>NANOG, KLF2*</td>
<td>N2B27</td>
<td>2IL, PKCi</td>
<td>MEFs, matrigel</td>
<td></td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Takashima et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4, NANOG, c-MYC†</td>
<td>20% KSR</td>
<td>hLIF</td>
<td>MEFs</td>
<td>20%</td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Buecker et al., 2010</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4 or OCT4, KLF2†</td>
<td>N2B27</td>
<td>2IL</td>
<td>MEFs</td>
<td>20%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Hanna et al., 2010</td>
</tr>
<tr>
<td>OCT4, KLF4 or KLF4, KLF2†</td>
<td>N2B27</td>
<td>2IL, Forskolin</td>
<td>MEFs</td>
<td>20%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Fang et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>15% KSR+N2</td>
<td>2IL, FGF2, JNK1, p38i</td>
<td>MEFs</td>
<td>20%</td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Valamehr et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, SV40Lts+NANOG*</td>
<td>20% KSR (+N2, B27 to reprogram)</td>
<td>2IL, FGF2, ROCK1, (+TGF-β to reprogram)</td>
<td>MEFs, Matrigel or VN</td>
<td>20%</td>
<td>2% KSR+FGF2 on MEFs</td>
<td>Wang et al., 2011</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4, c-MYC, RARG, NRS5A*</td>
<td>20% KSR</td>
<td>2IL±FGF2i</td>
<td>STO feeders</td>
<td>20%</td>
<td>20% KSR+FGF2 on STO feeders</td>
<td></td>
</tr>
</tbody>
</table>

*Temporary expression or induction of transgenes (constitutive expression not required).
†Constitutive expression of transgenes required to maintain the naïve state over multiple passages.
‡Unless specified otherwise, assume atmospheric O2 (20%).

<table>
<thead>
<tr>
<th>Transgenes expressed</th>
<th>Basal medium</th>
<th>Growth factors/inhibitors</th>
<th>Feeders/matrix</th>
<th>O2</th>
<th>Conventional primed conditions for comparison</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>Albumax+N2 or 20% KSR</td>
<td>2IL, FGF2, TGF-β, ROCK1,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>JNK1, p38i, PKC, insulin</td>
<td>Gelatine/VN or</td>
<td></td>
<td>15% KSR+FGF2, TGF-β on gelatine/VN or MEFs</td>
<td>Gafni et al., 2013</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>N2B27</td>
<td>2IL, FGF2, Activin, SRCi,</td>
<td>MEFs</td>
<td>5%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Theunissen et al., 2014</td>
</tr>
<tr>
<td>NANOG, KLF2*</td>
<td>N2B27</td>
<td>2IL, PKCi</td>
<td>MEFs, matrigel</td>
<td></td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Takashima et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4, NANOG, c-MYC†</td>
<td>20% KSR</td>
<td>hLIF</td>
<td>MEFs</td>
<td>20%</td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Buecker et al., 2010</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4 or OCT4, KLF2†</td>
<td>N2B27</td>
<td>2IL</td>
<td>MEFs</td>
<td>20%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Hanna et al., 2010</td>
</tr>
<tr>
<td>OCT4, KLF4 or KLF4, KLF2†</td>
<td>N2B27</td>
<td>2IL, Forskolin</td>
<td>MEFs</td>
<td>20%</td>
<td>15% FBS, 5% KSR+FGF2 on MEFs</td>
<td>Fang et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4*</td>
<td>15% KSR+N2</td>
<td>2IL, FGF2, JNK1, p38i</td>
<td>MEFs</td>
<td>20%</td>
<td>20% KSR+FGF2 on MEFs</td>
<td>Valamehr et al., 2014</td>
</tr>
<tr>
<td>OCT4, SOX2, SV40Lts+NANOG*</td>
<td>20% KSR (+N2, B27 to reprogram)</td>
<td>2IL, FGF2, ROCK1, (+TGF-β to reprogram)</td>
<td>MEFs, Matrigel or VN</td>
<td>20%</td>
<td>2% KSR+FGF2 on MEFs</td>
<td>Wang et al., 2011</td>
</tr>
<tr>
<td>OCT4, SOX2, KLF4, c-MYC, RARG, NRS5A*</td>
<td>20% KSR</td>
<td>2IL±FGF2i</td>
<td>STO feeders</td>
<td>20%</td>
<td>20% KSR+FGF2 on STO feeders</td>
<td></td>
</tr>
</tbody>
</table>

Assaying the pluripotent state in vitro

The studies to date have examined a number of parameters to determine whether the culture systems they describe in fact support the maintenance of naïve pluripotency (Table 4). As noted above, there is a characteristic set of genes that distinguish naïve from primed cells in the mouse. In many of these naïve culture systems, these genes are expressed at higher levels and more consistently compared with conventional cultures of hESCs (Table 4). However, it is important to recognise some limitations of these comparisons. Conventional hPSC cultures are heterogeneous, and a careful examination of gene expression in these cultures shows that subsets of cells grown under conventional conditions have distinct patterns of gene expression (see below). In fact, the data

### Table 4. Features used to assess candidate naïve primate PSCs

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compact dome colony morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gene expression similar to mouse naïve ESC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Two active X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Low levels of DNA methylation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Histone marks</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oct4 distal enhancer usage</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Improved clonogenicity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chimera formation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oxidative metabolism</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Diphid</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Derived from blastocyst</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

A, Gafni et al., 2013; B, Ware et al., 2014; C, Takashima et al., 2014; D, Theunissen et al., 2014; E, Chan et al., 2013; F, Fang et al., 2014; G, Hanna et al., 2010; H, Wang et al., 2014; I, Buecker et al., 2010; J, Valamehr et al., 2014.

✓, positive finding; no, characteristic was not observed; –, test was not performed.

The various cellular and molecular features used to assess naïve cells are listed. The table is not intended to provide a rating of the various studies but rather to indicate the range of tests performed on the cells by the various groups. Interpretation of these findings is complicated, owing to gaps in our knowledge of primate development (e.g. gene expression, X-chromosome inactivation) and limitations of the assays themselves (e.g. clonogenicity, chimera formation), as discussed in the text.
on single-cell gene expression in the human embryo discussed above would question whether expression of the mouse naïve gene set should serve as a gold standard test to distinguish naïve and primed pluripotent states. It is probably fair to say that an appropriate baseline for comparison of cultured cells to the human pre-implantation epiblast has yet to be established.

Nonetheless, much can be learned from the analysis of the results to date. Huang et al. performed a meta-analysis in which they compared gene expression patterns in mouse and human naïve and primed cells, and pre-implantation embryos of both species (Huang et al., 2014). Using a systems biology approach, these authors analysed the gene networks in the datasets to identify co-expressed modules and their conclusions are informative. First, as expected, mouse naïve cells mapped to the late mouse blastocyst. Second, the various types of human naïve cells generated with different protocols appeared to be very different. Interestingly, the only gene expression module conserved across several studies comprised genes involved in RNA processing and ribosomal biogenesis, as well as mitochondrial genes, suggesting that a common cellular metabolic profile is a key component of these various naïve states. Despite the variation between the studies, a general consensus was that the human naïve cells mapped closer to the human blastocyst stage than to the mouse blastocyst, and more nearly approximated the late blastocyst than primed human cells did. The protocols of Theunissen et al. (2014) and Takashima et al. (2014) yielded cells that bore the greatest similarity to the human blastocyst. Surprisingly, the primed cells with which the naïve cells were being compared varied considerably between studies, even more so than the naïve cells grown under a diverse range of conditions. This finding suggests that variations between cell lines, culture methodology and laboratories contribute to significant heterogeneity in the primed (and most likely naïve) cell phenotype. Given that a standardised methodology for the propagation of conventional hPSCs has been in place for a decade or so, this level of variation is surprising and should prompt a reassessment of how we define a baseline for primed human pluripotency. Although this meta-analysis confronts the usual challenges of cross-platform comparisons of transcriptome data, it does raise some important questions concerning variability between both primed and naïve states in these studies.

As mentioned above, the epigenetic state of mouse naïve cells is characterised by global DNA demethylation, a low level of bivalent (activator and repressor) histone marks on chromatin and the absence of X-inactivation in female lines. Most of the studies reported to date describe cells that fit all of these criteria (although the methods for the experimental definition of X-chromosome status are variable) (Table 4). However, it is important to remember that DNA methylation and X-inactivation status are quite variable across conventional hPSC lines, so, again, any absolute statement must take this heterogeneity into account. It is also worth noting that, whereas mESCs can retain two active X-chromosomes, it has been well known for many years that a majority of the XX mESC lines converts to an aneuploid XO status after propagation in vitro. Therefore, the status of X-chromosome activation in PSC lines might be impacted by unknown factors influencing their tolerance to a lack of dosage compensation on the X chromosome during long-term propagation in vitro.

A key feature of the naïve state in the mouse is an enhanced capacity for oxidative phosphorylation, along with more active mitochondria, in contrast to EpiSCs, which rely more on glycolysis and have less-developed mitochondria (Teslaa and Teitell, 2015). Thus, mESCs can achieve both oxidative phosphorylation and glycolysis. These features are important because it is now strongly appreciated that metabolism in PSCs is tightly integrated with the epigenetic status. Indeed, cellular pools of α-ketoglutarate (αKG, or 2-oxoglutarate) are maintained by oxidative phosphorylation in order to keep adequate levels of this co-factor available for use by enzymes such as TET methylcytosine dioxygenases and Junonji domain-containing histone demethylases, both of which require Vitamin C and αKG as co-factors, and which maintain DNA and histones in an hypomethylated state in mESCs (Carey et al., 2015; Blaschke et al., 2013; Wang et al., 2011). Therefore, changes in the culture milieu, including alterations in low molecular weight nutrients and co-factors that alter metabolism or modulate key enzymatic pathways can profoundly impact on the cellular epigenome. It remains to be elucidated to which extent the phenotype of the naïve states described in these studies may be driven by metabolic response to the nutrient and physical environment, rather than by fundamental differences in developmental status.

Assaying the pluripotent state in vivo

The developmental potential of the candidate naïve cells has been assessed in different ways. Several studies describe the participation of the human cells in chimera formation following transfer into mouse blastocysts (for references, see Table 4). The extent of chimerism observed was relatively low, and it might well be that species differences in the pace of embryonic development preclude extensive human cellular contribution to chimeras in these experiments. Other studies used both teratoma formation and in vitro differentiation as hallmarks of pluripotency. These are certainly useful benchmarks but might lack the resolution to identify significant differences in developmental potential. Future studies of chimera formation in non-human primates might provide more meaningful data on development potential of these cells. In this regard, it is important to note that one report demonstrating the failure of primate ESCs to contribute to conspecific embryos in a chimeric assay revealed, unexpectedly, that ICM cells also failed to participate in chimera formation (Tachibana et al., 2012). Only four-cell-stage blastomeres were found to be capable of chimera formation in this study.

A recent study using an in vitro interspecies chimera assay reported that human naïve cells failed to incorporate into mouse embryos (Masaki et al., 2015). By contrast, using a modification of the naïve culture system of Gafni et al. (2013) to convert cynomolgus monkey primed ESCs to a naïve state, Chen et al. (2015) were able to demonstrate contributions of the resulting naïve cells to multiple tissues representative of all three embryonic germ layers, including germline progenitors, in two foetuses following injection of the stem cells into host morulae and transfer back to a foster mother. Levels of chimerism were once again low, however, and the experiment, which required use of specific combinations of culture media to ensure stem cell survival during embryo culture, was not attempted with primed cells.

Modulation of the primed state of pluripotency through manipulation of the Wnt signalling pathway can enhance the ability of human cells to engraft mouse post-implantation embryos in vitro (Wu et al., 2015). By contrast, naïve mESCs do not engraft in post-implantation chimeras. In summary, chimera formation, although a powerful assay, is one in which readout is very dependent upon the precise state of the host cells and of the recipient embryo, and is influenced by other factors apart from whether cells are in a naïve or primed state.

It is known from reprogramming studies that synthetic cell states with no analogue in normal development, such as the pluripotent F-
class cells described by Tonge et al. (2014); see also Wu and Izpisua Belmonte (2014), can be isolated and maintained in vitro. Many of the studies of naïve cells involve cultivation of cells in what might well be highly selective media, and it is known that adaptation to culture can involve irreversible genetic and epigenetic changes. For these reasons, it is important to demonstrate that any culture system purporting to support the equivalent of a normal blastocyst should support de novo derivation of cell lines directly from the embryo, and that the naïve cells can convert to the primed state when cultured under appropriate conditions, just as embryonic cells would progress in vivo. These conditions are not met by many of the studies to date (Table 4).

**A re-examination of hESCs**

Current thinking concerning conventional hESCs considers them to be most similar to primed state EpiSCs in the mouse. This conclusion is based primarily on population studies that have examined the properties of hESCs in bulk cultures. Moreover, much of the data leading to this conclusion have been based on culture systems that, similar to LIF plus serum in mouse, present cells with a complex mixture of extrinsic signalling factors present in serum, feeder cells or serum replacements, some of which drive self-renewal and others that promote differentiation. Defined culture systems that supply self-renewal factors in the absence of undefined components established in the early 2000s are now coming into wide use (Hasegawa et al., 2010; Villa-Diaz et al., 2013) and provide a better platform for comparing human and mouse naïve states (many of the studies comparing naïve and conventional hESCs have not used defined culture systems for either).

In all culture systems studied to date, hPSCs exhibit a range of biological and molecular properties that more accurately reflect a continuous spectrum of cell states rather than a single state. Subpopulations of cells along this continuum can be isolated on the basis of their levels of expression of commonly used stem cell surface markers, such as GCTM-2, CD9 and EPCAM (Hough et al., 2009, 2014; Laslett et al., 2007). In cultures grown in serum using feeder-cell support, isolation of cells negative for these surface markers shows that these cells have lost the expression of some, but not all, pluripotency-associated genes and are primed for neural differentiation. However, in cultures grown in defined conditions with TGF-β and FGF supplementation, very few cells are negative for the stem cell surface markers, and very few display lineage priming. Under such conditions, it is possible to isolate a subpopulation with the highest levels of expression of stem cell surface markers that expresses uniformly high levels of pluripotency-associated genes, no lineage-specific genes and shows high levels of self-renewal (Hough et al., 2014). Preliminary results obtained in our lab suggest that these cells are located towards the periphery of growing stem cell colonies in vitro and show high levels of staining with dyes that detect mitochondrial membrane potential (unpublished data). Previous studies have shown that the clonogenic fraction of hESCs can be identified by mitochondrial dye staining (Ramirez et al., 2013). These cells located at the periphery of the stem cell colony also show high levels of staining with SMAD3 antibodies, indicating high levels of Activin/Nodal activity (Hough et al., 2014). The patterning of hPSC colonies, in which cells with high levels of expression of pluripotency markers and high levels of SMAD2/3/4 signalling are located at the periphery, was also seen in another recent study (Warmflash et al., 2014).

Although features such as X-inactivation and DNA methylation have not been examined in these hPSC subpopulations yet, it appears that the cells at the colony periphery with high self-renewal capacity do show some features of naïve pluripotency and represent a minority population in metastable equilibrium with other cell types. These cells are capable of robust differentiation into derivatives expressing genes characteristic of all three germ layers in embryoid body assays, but in cultures maintained under defined conditions and in the presence of TGF-β and FGF, some cells expressing high levels of stem cell surface antigens and pluripotency-associated genes co-express markers of extra-embryonic endoderm (Hough et al., 2014). This suggests that, as in naïve mESCs, a subpopulation of hESCs might represent an ICM-like state, capable of producing extra-embryonic endoderm. The laboratory of Dr Brivanlou (Warmflash et al., 2014) suggested that cells on the outer edge of the colonies could be primed to form extra-embryonic ectoderm, again indicative of a more primitive developmental capacity than that of mouse epiblast.

Hence, it is possible that the interpretation that hESCs correspond to primed mEpiSCs is misleading, and is, in part, the result of examining the properties of bulk populations grown under conditions that sub-optimally support maintenance of the naïve stem cell population. These findings are reminiscent of the comparisons of mESCs maintained in serum compared with cells maintained in 2i. In this light, another recent study identified a small subpopulation of hESCs (~4%) grown under conventional culture conditions that express the human retrowiral element HERV, a primate-specific regulator of pluripotency that is, in turn, regulated by TFCP2L1 (a marker gene for the mouse naïve state) (Wang et al., 2014). The minority population displayed some properties of naïve cells including gene expression and X-inactivation. The HERV element targets a human-specific pluripotency gene network that is active specifically in cells expressing the retrovirus. These data, along with previous work showing that retrotransposon integration alters regulation of the pluripotency network in the human (Kunarov et al., 2010), indicates that any consideration of the status of hPSCs should make reference to these primate-specific networks on the developmental roadmap. The existence of sub-states of pluripotency within conventional hESC cultures shows clearly that conclusions concerning the status of hESCs must take into account the cellular hierarchies that exist within the cultures. This is particularly important when comparing population data collected from conventional hESCs grown in the presence of serum, serum replacements or feeder cell support, with data obtained in naïve cells. The question of heterogeneity within the naïve populations of human cells has not yet been addressed at the single-cell level.

**Conclusions**

The pioneering studies of human naïve cells conducted over the past several years and described above indicate that it is possible to derive and maintain cultured hESC lines that display some properties similar to that of the human blastocyst. Whether these naïve hESCs are actually close analogues of the human epiblast remains to be determined. We need to expand our knowledge of early primate embryonic development if we are to answer this question definitively, because mounting evidence suggests that there are significant differences in the molecular and cellular biology of pluripotency between species. The intriguing possibility that conventional hPSC cultures harbour a subset of cells with identity to some earlier pluripotent cell type in pre-implantation development likewise awaits further clarification. In this respect, it is possible that a naïve epiblast state is simply much more difficult to stabilise in primate versus mouse. The embryological identity of the stable attractor state epitomised by human teratocarcinoma in vitro
and in vivo, and by conventional hESCs and iPSCs, remains to be defined, even after many years of study (Pera, 2014). Only through further in-depth analysis of primate development will we be able to ascertain which, if any, normal stage of human development the various types of hPSCs equate to.

It could be argued that the search for human naïve cells is predominantly an academic exercise and that many applications of hESCs and iPSCs are, and will be, addressed using conventional cell lines. However, if we are to use hPSCs as a high-fidelity model for human development, if we wish to shed light on poorly understood differences in differentiation potential and genetic and epigenetic stability between PSC lines, if we wish to improve the outcomes of hPSC differentiation and if we are to improve the fidelity of cell reprogramming to pluripotency, then it becomes important to understand how hPSCs fit into the paradigm of embryonic development. The pursuit of naïve pluripotency in human has already made important contributions towards this end.

Competing interests
The authors declare no competing or financial interests.

Funding
The authors are supported by the Australian Research Council Special Research Initiative in Stem Cell Sciences (Stem Cells Australia).

References


and in vivo, and by conventional hESCs and iPSCs, remains to be defined, even after many years of study (Pera, 2014). Only through further in-depth analysis of primate development will we be able to ascertain which, if any, normal stage of human development the various types of hPSCs equate to.


