# Ground rules of the pluripotency gene regulatory network.

# The Ground Rules of Pluripotency Regulatory Networks

**Author(s): Mo Li[1,2,3] and Juan Carlos Izpisua Belmonte[1]**

[1]Salk Institute for Biological Studies, 10010 N. Torrey Pines Rd., La Jolla, CA 92037

[2]Universidad Católica San Antonio de Murcia (UCAM) Campus de los Jerónimos, nº 135 Guadalupe 30107, Murcia, Spain

[3]4700 King Abdullah University of Science and Technology (KAUST) Thuwal 23955-6900, Kingdom of Saudi Arabia

Correspondence: Belmonte@salk.edu

***Abstract:*** Pluripotency is a state that exists transiently in the early embryo and, remarkably, can be recapitulated *in vitro* by deriving embryonic stem cells or by reprogramming somatic cells to induced pluripotent stem cells. This state, which is stabilized by an interconnected network of pluripotency genes, integrates external signals and exerts control over the decision between self-renewal and differentiation at the transcriptional, post-transcriptional and epigenetic levels. Recent evidence of alternative pluripotency states indicates regulatory flexibility of this network. Insights into the underlying principles of the pluripotency network may provide unprecedented opportunities for studying development and for regenerative medicine.

**INTRODUCTION:**

Pluripotent stem cells (PSCs) are defined by their ability to self-renew indefinitely and to differentiate into all cell types (including germ cells) of an adult animal. The term stem cell was first used by scientists in the late 19th century to refer to precursor cells of the germline lineage or the common precursor of the blood system[1-3]. The first association of the terms pluripotency and stem cell in the sense that is commonly understood today was made by Martin Evans in 1972 when he described the *in vitro* culturing of pluripotent cells from teratomas[4]. PSCs had been previously derived from teratomas (benign) and teratocarcinomas (malignant) by others, but they did not use the term PSC[5-7]. The pioneering work involving these rare and bizarre tumours (reviewed in[8]) lead to: first, the development of culture conditions for pluripotent embryonic carcinoma (EC) cells; second, the discovery of embryonal carcinoma cell-specific surface antigens; and third, the establishment of functional assays for testing developmental potency of pluripotent cells. These advancements proved instrumental for isolating pluripotent cells from normal embryos.

Pluripotency is a transient property of cells within the early embryo. This property can be captured *in vitro* as PSCs at different developmental time periods[9]. Pluripotent cells first arise in the inner cell mass (ICM) of the blastocyst stage mouse embryo (cells in earlier stages, including the morula stage and 2-cell and 4-cell embryos are considered totipotent). The pluripotent embryonic stem cells (ESCs) were first derived from the ICM of pre-implantation embryos that were explanted on feeder cells in the presence of serum[10,11]. **[Au: it would be best to state clearly in this sentence or earlier that ESCs are a type of PSC as this may not be clear to general readers]** Pluripotency can also be captured at later time points during germline development by the epigenetic reprogramming of unipotent primordial germ cells (PGCs) *in vitro* [12]. **[Au:OK?]** PSCs obtained in this manner are called embryonic germ cells (EGCs) and they are nearly indistinguishable from ESCs. Pluripotency, either existing *in vivo* or captured *in vitro*, can be defined by several functional assays, including differentiation into the three germ layers *in vitro*, formation of teratomas *in vivo*, chimera formation and germline transmission through blastocyst injection, and tetraploid complementation. The latter is considered the most stringent test for pluripotency. Mouse ESCs, like the early epiblast from which they are derived, satisfy all of these criteria and are considered to be in a naïve state of pluripotency[10,11,13]. By contrast, PSCs derived from post-implantation mouse epiblasts, termed epiblast stem cells (EpiSCs), do not efficiently

contribute to chimera formation at the blastocyst stage, but can readily participate in chimeric embryo formation when introduced into the post-implantation epiblast[14-18]. EpiSCs are thus thought to represent a developmentally more advanced pluripotent state, i.e., a "primed" state of pluripotency[14,15]. Intriguingly, EpiSCs and human ESCs have many commonalities, including dependence on Activin and fibroblast growth factor 2 (FGF2), a flat colony morphology, and poor single cell survival, suggesting that human ESCs are also in the primed state. Two recent studies describe conditions that capture new flavours of primed pluripotency[17,18]. PSCs derived under these conditions take on specific regional properties and show affinity to specific regions of the epiblast when re-introduced into the developing embryo.

The pluripotency state is governed by a highly interconnected pluripotency gene regulatory network (PGRN), which is functionally anchored by a set of core pluripotency transcription factors (TFs). The fundamental importance of these TFs in the acquisition and maintenance of pluripotency is best illustrated by the reprogramming of induced pluripotent stem cells (iPSCs), where certain combinations of pluripotency factors are sufficient to override the epigenetic mechanisms safeguarding the somatic cell fate and to reinstate the pluripotency state de novo[19,20]. In this review, we discuss the mechanisms by which the PGRN governs the acquisition, maintenance, and dissolution of the pluripotency state. We consider interactions between the PGRN and additional layers of regulation that involve chromatin and RNA-mediated mechanisms. We also discuss the emerging wide gamut of alternative states of pluripotency and the transition between different states.

## CORE PLURIPOTENCY GENE NETWORK

A few core pluripotency TFs serve as the hub of the PGRN, instigating a cascade of regulatory events that, together with many other TFs and cofactors, maintain the pluripotent state[21,22]. Octamer-binding transcription factor 4 (OCT4) is uniquely expressed in ESCs and PGCs, and is essential for both *in vivo* and *in vitro* pluripotency[23-25]. Similarly, SRY-box 2 (SOX2) is required for formation of the pluripotent epiblast and is a key regulator of Oct4 expression in ESCs[26,27]. Loss of OCT4 or SOX2 promotes trophectoderm differentiation, whereas overexpression of Oct4 or Sox2 leads to mesendoderm and neural ectoderm, respectively[25,28]. Thus, precisely regulated OCTt4 and SOX2 serve as the foundation of the pluripotency transcriptional circuit. Nanog is also considered a core pluripotency TF because of its important role in the acquisition of pluripotency in the ICM, and its ability to enable leukemia inhibitory factor (LIF)-independent self-renewal when overexpressed in ESCs (although its absence is compatible with the maintenance of ESCs)[29-33].

Numerous studies have used high-throughput systems biology tools to dissect the regulatory networks underlying pluripotency. These tools include, microarray and RNA sequencing (RNA-seq) for profiling gene expression in wild type or genetically modified ESCs, immunoprecipitation coupled with mass spectrometry proteomics for assessing the protein interactome, and high-throughput chromatin immunoprecipitation (ChIP)-based methods for mapping genomic occupancy of core pluripotency TFs. Findings from these studies have been extensively reviewed[21,22,34,35]. Evidence to date can be distilled into our current understanding of the PGRN. OCT4, SOX2, and NANOG function together and co-occupy hundreds of potential regulatory elements in the genome, including their own promoters, thus forming densely interconnected feedback and feedforward regulatory loops. Co-localized core TFs provide a platform for recruiting additional factors that integrate regulatory inputs from signaling pathways, other transcriptional circuits, co-activators/repressors, regulatory RNAs, and epigenetic mechanisms[21,36,37]. In such a system, fluctuations in the concentration of core TFs are either buffered by regulatory redundancy and positive feedback (resulting in stabilization of the pluripotent state), or exploited to initiate differentiation in an appropriate signaling environment (Figure 1). This bi-stability results from protein-DNA and protein-protein interactions that characterize the PGRN[38].

**[Au: unfortunately it's not possible for a main section to contain only one subsection. Could you create a second subsection, either from the text above, or by splitting the contents of the below subsection, which is quite long, into two or more? For example,**

**perhaps you could have a subsection on Oct4-Sox2 complex formation and one on the binding sites or transposable elements]**

***Co-occupancy of core pluripotency TFs***

OCT4 and SOX2 cooperatively bind many genomic sites as heterodimers[39,40]. NANOG binding also shows extensive overlap with OCT4 and SOX2[41]. Sites bound by these three core TFs often exhibit enhancer activity. Importantly, genes adjacent to these potential enhancers are enriched for the gene ontology terms self-renewal and differentiation, again suggesting that this core pluripotency circuit generates a bi-stable state within ESCs[21]. However, there may be functional differences between the OCT4 and SOX2 module and the NANOG module. OCT4 and SOX2 are uniformly expressed at tightly controlled levels across the entire population of ESCs, being highly expressed in both the naïve and primed pluripotency states. NANOG, on the other hand, may be expressed at high or low levels in an ostensibly uniform population of ESCs and is downregulated in primed pluripotency[31,42]. Cells that completely lack NANOG can still self-renew in an undifferentiated state, albeit with a much higher propensity for differentiation[31]. Many putative target genes (including pluripotency genes) bound by NANOG remain expressed in the absence of NANOG[43]. Thus, the role of NANOG seems to be safeguarding naïve pluripotency against pro-differentiation signals, rather than activating self-renewal genes per se.

Recent evidence provides new insights into the molecular basis underlying OCT4 and SOX2 complex formation and their co-binding of DNA targets. Single molecule imaging revealed how OCT4 and SOX2 assemble on their cognate DNA sites. First, SOX2 dynamically searches and engages with the chromatin, priming target sites for OCT4 binding. OCT4 in turn stabilizes binding of the OCT4-SOX2 complex on composite recognition sites[44]. A highly conserved residue of OCT4 (K156) is crucial for the OCT4-SOX2 interaction and for OCT4 protein stability within ESCs. Mutation of the OCT4-K156 residue results in the downregulation of stemness genes, and the upregulation of mesendodermal genes involved in the epithelial-mesenchymal transition (EMT)[45] (Figure 2A). Therefore, post-translational modification of OCT4-K156 may be the mechanism by which the OCT4-SOX2 complex is disassociated, leading to specification of the mesendodermal lineage. Structural studies show that OCT4 and SOX2 can co-bind composite target DNAs in one of two complex configurations, depending on the positioning of their individual DNA-binding motifs[46]. In one configuration, OCT4 and SOX2 bind a no-gap canonical motif where the octamer motif of OCT4 is immediately juxtaposed to the SOX2-binding motif. Examples of this configuration

are seen in the regulatory regions of *POU5F1*, *UTF1*, and *NANOG* by ChIP experiments[47]. **[Au: perhaps you could expand a bit here. Have OCT4 and SOX2 been shown to bind to regulatory regions of these genes or is this merely suggestive?]** In the alternative configuration, OCT4 and SOX2 bind to motifs that are separated by three base pairs, as is seen in the fibroblast growth factor 4 (FGF4) promoter. **[Au: There's an error in the wording here, but I haven't corrected it as I am unsure whether you mean "the OCT4 and SOX2 binding motifs are separated by three base pairs" or "OCT4 and SOX2 bind to motifs that are separated by three base pairs"; The latter option suggest that the binding motifs themselves are different in this configuration. Is that the case or are they just spaced differently. Please clarify ]** Taking advantage of *Sox2* mutations that interfere with specific OCT4-SOX2 heterodimer configurations, a recent study showed that the OCT4-SOX2 configuration that dimerizes on the no-gap canonical motif has a more crucial role in somatic cell reprogramming and in ESC pluripotency than any other configuration[48]. This interesting finding suggests that a subset of target genes regulated by the core pluripotency TFs may help define a minimal pluripotency network (Figure 2B).

***Co-occupancy at transposable elements shapes PGRN***

It is worth mentioning that only a small subset of DNA sites bound by core TFs are *bona fide* regulatory elements of nearby genes[49]. It has been shown that changes in OCT4/SOX2/NANOG (OSN) occupancy do not correlate well with differential gene expression[49,50]. Thus, the precise function of the majority of OSN-occupied sites remains unclear. One possibility is that by binding these sites, OSN may prime the local chromatin for subsequent action by TFs involved in differentiation, thereby promoting exit from pluripotency. Interestingly, up to 25% of the OCT4- and NANOG-bound sites are within retrotransposons. Because of the divergence of transposable elements between mouse and human, this has resulted in a low level of sequence conservation in pluripotency TF binding between the two species[51]. **[Au:OK? Original could be interpreted as intraspecies conservation across sites]** As such, transposable elements have contributed greatly to rewiring the PGRN in different mammalian species, yet the PGRN displays high levels of plasticity. What roles do transposable elements play in the transcriptional regulation of pluripotency? Recent evidence suggests that they participate in species-specific pluripotency gene expression via several mechanisms. First, as discussed above, they may recruit pluripotency TFs and act as enhancers of nearby genes[52,53]. Second, they may be transcribed as non-coding RNAs that perform regulatory functions (discussed below). Third, certain

types of transposable elements are repressed in ESCs by co-repressors via epigenetic mechanisms, which can spread and silence nearby genes. In summary, as a major constituent of the mammalian genome, transposable elements play an important role in shaping the PGRN. Future studies on these elements may help reveal the inner workings of diverse pluripotent states.

## UPSTREAM SIGNALING PATHWAYS AND THE PGRN

**[Au: this section is quite a large chunk of uninterrupted text. Would you be able to break it down into more easily digestible subsections to improve readability?]**

### *Signaling pathways of conventional ESC*

The derivation and maintenance of ESCs require activation or inhibition of multiple signaling pathways. Mouse ESCs were originally derived in the presence of irradiated fibroblasts and serum, which together supply LIF, WNT, bone morphogenetic protein 4 (BMP4), and other factors that support self-renewal[10,11,21]. Part of the reason ESC cultures had been difficult to establish is that self-renewal is not an inherent feature of *in vivo* pluripotency. As alluded to above, the core pluripotency factors OCT4 and SOX2 activate FGF4, which is an autocrine factor that feeds back through the FGF–ERK signaling pathway to promote differentiation[54,55]. ESCs lacking FGF4 or ERK have severely limited propensity for neural and mesendoderm differentiation. ERK1 and ERK2 (ERK1/2) activation modulates chromatin occupancy of polycomb recessive complex 2 (PRC2) and RNAPII at developmental genes, presumably promoting a poised status for their transcription once a differentiation cue is received[56]. The addition of two cytokines, LIF and BMP4, can stabilize *in vitro* the pluripotent epiblast in a so-called “conventional” ESC state (Box1).

LIF signals through the transcription factor signal transducer and activator of transcription 3 (STAT3) to activate the self-renewal gene program[57]. Among LIF/STAT3 targets are the pluripotency genes *Klf4* and *Tfcp2l1*. Overexpression of *Klf4* or *Tfcp2l1* enables ESC self-renewal without LIF[58-60]. Transcription factor CP2-Like 1 (TFCP2L1) stimulates a parallel circuit of self-renewal by upregulating NANOG, which itself supports LIF-independent self-renewal when overexpressed[60]. B cell leukemia/lymphoma 3 (BCL3) was recently shown to act downstream of LIF–STAT3 signaling to positively regulate pluripotency genes, possibly by associating with OCT4 and beta-catenin[61].

BMP acts via the SMAD TFs to induce the inhibitor of differentiation (Id) genes to prevent differentiation[62]. The effect of BMP signalling seems to be to suppress neuroectoderm fate, whereas LIF represses non-neural differentiation. Therefore, conventional ESCs integrate competing signals from the FGF–ERK, LIF–STAT3, and BMP–SMAD pathways to exist in a meta-stable state.

***2i and ground state ESC***

It is also possible to maintain ESCs without extrinsic signals (that is, without LIF and BMP). This is achieved by dual inhibition (2i) of the pro-differentiation MEK–ERK signaling pathway and glycogen synthase kinase 3 (GSK3) with the small molecule inhibitors PD0325901 and CHIR99021, respectively[63]. ESCs maintained in 2i media efficiently contribute to chimera formation and germline transmission, display homogenous expression of *Nanog*, display significantly reduced expression of differentiation-associated genes, and resemble *in vivo* naïve epiblast cells at the transcriptome level[63-65]. Because of these features, ESCs cultured in 2i media are said to be in the ground pluripotency state. **[Au:OK?or would you prefer "under 2i conditions"?]**

Blockade of FGF4–ERK signaling stabilizes *Klf2*, a gene essential for ground state pluripotency[66]. Similarly, WNT–beta-catenin signaling activates *klf2* expression[67]. Thus, dual inhibition of ERK and GSK3 acts in concert to promote the pluripotency circuit controlled by KLF2. ERK also phosphorylates NANOG (another TF critical for the ground state), thereby reducing its transactivation activity and stability[68]. These data engendered a view that ERK signaling was dispensable for naïve pluripotency and that FGF2–ERK signaling was instead required for primed EpiSC and human ESCs. This view has been challenged by recent genetic data. Surprisingly, ERK1–2 double knockout ESCs cannot be maintained. Acute loss of both ERKs quickly leads to telomere shortening, altered expression of pluripotency genes, reduced cell proliferation, G1 cell-cycle arrest, and increased apoptosis[69]. These data also suggest that MEK inhibition in 2i ESCs may function through both ERK-dependent and ERK-independent mechanisms. Understanding the ERK-independent function of MEK will probably provide significant insights into the ground state pluripotency gene network.

Effectors of the LIF, BMP4, and WNT pathways (STAT3, SMAD1, and beta-catenin–transcription factor (TCF) 3, respectively) directly modulate the OSN core circuit by co-occupying enhancers bound by these core TFs[21,36]. In the absence of nuclear beta-catenin,

TCF3 functions as a transcriptional repressor, antagonizing the action of OCT4 and SOX2, which colocalize with TCF3 at pluripotency genes[70]. WNT stimulation relieves TCF3 repression on its target genes in the PRGN, thereby reinforcing ESC self-renewal. For example, derepression of *Esrrb* is both necessary and sufficient for supporting self-renewal downstream of GSK3 inhibition[71]. On the other hand, GSK3 inhibition is not necessary for ESC self-renewal, as LIF–STAT3 can act alone to support naïve pluripotency in the absence of Esrrb. These data suggest that ground state pluripotency is supported by parallel pathways. A systematic characterization of how upstream signals alter the OSN global-binding landscape has recently been provided. During the 2i-induced transition from conventional ESC to the ground state, all three core pluripotency TFs show rapid and widespread binding rearrangements. Differentially bound sites are enriched at distal enhancers and tend to contain binding motifs for TFs associated with the canonical WNT and ERK signaling pathways (Figure 3A). Specifically, sites with higher OSN binding in 2i are enriched for binding motifs for LEF1, TCF7, and HOX proteins, which belong to the WNT pathway, whereas sites with decreased OSN binding are enriched for binding motifs for co-factors of the ERK pathway, such as early growth response 1 (EGR1), specificity protein 1 (SP1) and KLF7[50]. While more work is needed to understand the functional relevance of these changes in OSN binding, this study provides mechanistic insights into the 2i state and the dynamics of the PGRN.

**COFACTORS: THE CHROMATIN CONNECTION [Au: unfortunately main headings can only be a maximum of 39 characters (including white space). This is so that they will fit above a single column of text in the final layout. Please can you shorten?]**

Transcriptional co-activators and co-repressors are protein complexes that do not bind DNA on their own, but regulate the action of sequence-specific TFs via chromatin-mediated mechanisms. Because they serve as general factors of transcriptional regulation, a pluripotency-specific role of these factors was not suspected. Thus it came as a surprise when several RNA interference (RNAi) screens revealed that ESCs are extremely sensitive to deficiency in some of these cofactors (e.g., the mediator and cohesin complexes[72], the Tip60-p400 chromatin remodelling complex[73], the RNA polymerase associated factor (PAF1) complex[74], and the corepressors CCR4-NOT transcription complex subunit 3 (CNOT3) and tripartite motif containing 28 TRIM28 (KAP1))[75].

The requirement of mediators and cohesins in maintaining pluripotency highlights the importance of 3D genome organization in the PGRN. Mediators and cohesins are large protein complexes that facilitate physical interaction between TF-bound enhancers and promoters[76,77]. The function of cohesin in forming chromatin loops in interphase nuclei is analogous to but independent from its role in sister-chromatid cohesion during mitosis. Transcription of pluripotent genes depends on interactions between distant regulatory elements. For instance, the expression of *Oct4* requires its upstream enhancer (when bound by OCT4, SOX2, KLF4, mediator, and cohesion) to come into contact with its promoter[78,79]. KLF4 and cohesins are important organizers of these chromosomal interactions, as knockdown of either factor abolishes this *Oct4* enhancer-promoter interaction, leading to reduced *Oct4* expression and dissolution of pluripotency[78,79]. A high-throughput survey of such 3D chromosomal interactions in ESCs and neural progenitor cells using the Chromosome Conformation Capture Carbon Copy (5C) technology (Box2) revealed that the vast majority of these interactions involve distinct combinations of Mediator, cohesion, and CTCF[80]. Interestingly, these interactions show hierarchical organization at different length scales. Cohesin and CTCF anchor long-range interactions (>1 Mb) that are invariant between cell types, whereas Mediator and cohesin (together with OSN) organize short-range (<100 kb) enhancer-promoter interactions in an ESC-specific fashion[80].

The co-repressors CNOT3 and TRIM28 were identified in a genome-wide RNAi screen in mouse ESCs for factors essential for self-renewal. They co-occupy gene promoters with c-MYC and ZFX, rather than OSN, and therefore form a distinct module of the PGRN that likely regulates proliferation, cell death, and cancer[75]. TRIM28 is critical for suppressing the transposition of endogenous retroviruses in ESCs and for silencing enhancers harboured in these retro-elements (i.e., preventing nearby genes from inappropriate activation)[81]. TRIM28 accomplishes these functions by interacting with SET domain bifurcated 1 (SETDB1) to induce heterochromatin formation[82]. TRIM28 has recently been shown to interact with polycomb repressive complex 1 (PRC1) and binds cooperatively with PRC1 at promoters of differentiation-inducible genes to repress transcription. In contrast, TRIM28 binds to pluripotency-associated genes without recruiting PRC1, thereby de-repressing transcription[83]. Thus, the function of TRIM28 in ESCs is context-dependent and may involve different epigenetic partners.

Several co-factors that are important for the PGRN have been identified by proteomic analysis of protein interaction networks involving pluripotency genes. The orphan nuclear hormone receptor, NR0B1 (also known as DAX1) was identified as a novel OCT4-interacting protein in a yeast two-hybrid screen[84]. Although NR0B1 is not essential for pluripotency, it is required to repress the transcription of *Zscan4c* (a 2-cell stage embryo-specific gene) in mouse ESCs[85]. Overexpression of *Zsan4c* in the absence of NR0B1 disrupts normal self-renewal due to G2 cell cycle arrest and cell death. Overexpression of NR0B1 supports LIF-independent self-renewal and does so without affecting STAT3 activation or ERK phosphorylation[86]. NR0B1 can inhibit extra-embryonic endoderm differentiation by binding to the GATA6 promoter and inhibiting its transcription. NR0B1 also inhibits trophectoderm differentiation either independently or cooperatively with OCT4. It is proposed that NR0B1 and NANOG act in parallel to maintain an optimal pluripotent state[86]. It is worth noting that these investigations concerning the role of Nr0b1 in PGRN were performed using the conventional LIF/serum condition. It will be interesting to determine whether Nr0b1 carries out similar functions in the ground pluripotency state.

Mass spectrometry analysis of immunoprecipitated SOX2 protein complexes in ESCs revealed TEX10 as a novel transcriptional cofactor in the PGRN. Functionally, TEX10 is required for ESC maintenance, early embryo development, and efficient reprogramming of both mouse and human somatic cells[87]. The distribution of TEX10 binding across the genome strongly correlates with that of OSN and Mediators. **[Au:OK?]** In particular, TEX10 is enriched at ESC-specific super-enhancers (SEs), which are large clusters of enhancers that are bound by OSN and Mediators, which positively regulate their activity. Mechanistically, TEX10 regulates SE activity and transcription of enhancer RNAs (eRNAs) by modulating histone acetylation and DNA demethylation through interactions with p300 and Tet1[87]. The domain structure of the TEX10 protein is suggestive of interfaces for protein, DNA, and RNA binding, which may enable TEX10 to regulate gene expression at multiple levels. It will be of great interest to further dissect the function of TEX10 in the PGRN.

CBFA2/RUNX1 translocation partner 2 (CBFA2T2) is another co-repressor important for the regulation of pluripotency. CBFA2T2 was identified via the proteomic analysis of proteins that interact with PR domain containing 4 (PRDM14), which itself is a pluripotency factor that regulates DNA methylation and germ cell specification[88,89]. Similar to PRDM14 knockout ESCs, CBFA2T2 knockout ESCs cannot be maintained under conventional serum

conditions[89,90]. Like PRDM14, CBFA2T2 represses lineage commitment genes in ESCs. Overexpression of CBFA2T2 also enhances iPSC reprogramming efficiency, as reported for PRDM14. CBFA2T2 knockout mice show severe defects in PGC maturation and epigenetic reprogramming, suggesting that CBFA2T2 and PRDM14 belong to a regulatory network shared by the embryonic pluripotency and the latent pluripotency of the germline. CBFA2T2 and PRDM14 co-localize extensively throughout the genome and share many sites with OSN. The current mechanistic understanding is that CBFA2T2 oligomerizes to form a scaffold, which helps to stabilize OCT4 and PRDM14 at their binding sites[90].

## RNA-BASED REGULATORY INPUTS INTO PGRN

It is now well established that many RNA-based regulatory mechanisms are crucial for the regulation of self-renewal and pluripotency. The diverse mechanisms of gene regulation by versatile RNA molecules not only expand and fine-tune the regulatory capacity of the PGRN, but also increase its proteomic diversity.

### *microRNAs*

A class of small, non-coding RNAs called microRNAs (miRNAs) are important for pluripotency[91]. This was first demonstrated by knocking out genes essential for their biogenesis, namely *Dicer* and *DGCR8*. Data from these studies suggested that the main function of miRNAs in ESCs is to promote cell cycle progression[92]. ESC-specific miRNA clusters, miR-290 and -302, promote self-renewal and inhibit somatic differentiation[93]. miR-290 and -302 oppose effects of the let-7 family of miRNAs, which are broadly expressed in differentiated tissues and are required for maintaining the differentiated state[94]. A number of ESC-specific miRNAs facilitate somatic cell reprogramming by targeting genes involved in multiple aspects of the reprogramming process[95]. Paradoxically, a recent study showed that the miR-290 and -302 clusters promote dismantlement of the naïve pluripotency program. These miRNAs effectively silence naïve pluripotency in *Dgcr8* knockout ESCs, which cannot exit naïve pluripotency or establish primed pluripotency. Several mechanistic targets (e.g., the MEK pathway, *Akt1*, and *Klf2*) have been proposed to explain the seemingly contradictory and context-dependent functions of these miRNAs[96]. However, the precise mechanism remains unclear.

### *Long non-coding RNAs*

Long non-coding RNAs (lncRNAs) have emerged as another important class of molecules that regulate pluripotency. Broadly defined as non-coding transcripts longer than 200 nucleotides, lncRNAs have diverse modes of action, including acting as molecular decoys, interfering with the transcription machinery, providing scaffolds for TFs, and targeting epigenetic enzymes. Several conserved lncRNAs were identified as direct targets of OCT4 and NANOG and directly modulate pluripotency gene expression and differentiation[97]. One such large intergenic non-coding RNA (LincRNA-ROR, for regulator of reprogramming) modulates reprogramming of human iPSCs[98]. A later study demonstrated that LincRNA-ROR functions as a "miRNA sponge" to trap other miRNAs, thereby regulating levels of core pluripotency TFs. LincRNA-ROR and core TF mRNAs share seed sequences of miR-145, which binds core TF mRNAs and represses their translation[99]. In self-renewing ESCs, levels of LincRNA-ROR exceed miR-145 levels, allowing LincRNA-ROR to protect core TFs from miR-145 mediated suppression. During differentiation, the situation reverses, allowing miR-145 to coordinate the exit from pluripotency[100] (Figure 4). This fascinating tug of war between three RNA components (miRNA, LincRNA, and mRNA) represents a new paradigm of posttranscriptional regulation in the PGRN.

Another abundant source of non-coding RNAs in human ESCs and iPSCs are endogenous retroviruses[101]. Depleting human endogenous retrovirus subfamily H (HERV-H) transcripts (including lincRNA-ROR) results in a differentiation-like phenotype in human ESCs[102]. HERV-H lincRNA-ROR is transiently reactivated during human iPSC reprogramming[98]. Importantly, a recently identified naïve-like cell population in human ESC and iPSC cultures exhibit elevated HERV-H transcription[103]. Together, these data indicate that transcription of HERV-H is essential for human ESC. Studying HERV-H RNAs may facilitate the derivation of naïve human PSCs. The question of how these retrotransposon RNAs regulate gene expression is still unanswered. Recent work has shown that HERV-H lincRNAs can act as scaffolds to recruit OCT4 or transcriptional coactivators (e.g., CBP, p300, MED6, and MED12) to enhancers to regulate their activity[102].

***Alternative splicing***

Alternative splicing expands proteome diversity by allowing multiple protein isoforms (with potentially distinct functions) to be produced from a single gene. Recently, alternative splicing has emerged as a new regulatory layer that is woven into the PGRN to safeguard pluripotency and self-renewal. Several ESC-specific isoforms of pluripotency factors,

including TCF3[104], FOXP1[105], NANOG[106], and MBD2[107], have been shown to regulate pluripotency. **[Au:OK?]** In addition, regulators of the alternative splicing programs, such as the muscleblind-like family of RNA binding proteins and the splicing factor SRSF2, also play important regulatory roles in ESCs[107,108]. For example, the pluripotency TF (OCT4), the splicing regulator (SRSF2), and alternatively spliced protein isoforms (MBD2a and MBD2c) participate in a positive feedback loop, which is further regulated by miRNAs, nicely illustrating how different regulatory mechanisms work in concert to maintain pluripotency[107] (Figure 4).

***Messenger RNA modification and ESC fate determination***

$N^6$-methyladenosine ($m^6A$) is the most abundant form of reversible chemical modification on messenger RNAs in eukaryotes[109]. Recent reports have shown widespread $m^6A$ modification in both human and mouse ESCs[110,111]. The data revealed significant conservation of $m^6A$ modified genes between the two species, suggesting functional significance of this RNA modification in PSC biology[110]. Notably, mRNAs of core pluripotency TFs SOX2 and NANOG and many developmental regulators are modified with $m^6A$, whereas OCT4 mRNA lacks this modification. The $m^6A$-bearing transcripts showed shorter half-life and increased rate of mRNA decay, suggesting that $m^6A$ is a mark for mRNA turnover[110,111]. To understand the functional significance of $m^6A$ in PSCs, one study knocked down (KD) Mettl3, a component of the $m^6A$ methylase complex, which lead to reduced $m^6A$ levels in a majority of the modified genes and to compromised self-renewal[111]. A later study relying on genetic knockout (KO) of METTL3 confirmed the global reduction of $m^6A$, but challenged the previous conclusion that METTL3 is required for self-renewal. It showed that *Mettl3* KO mouse ESCs and *METTL3* KD human ESCs have no defect in self-renewal or viability. Instead, METTL3-depleted cells displayed persistent *Nanog* expression and impaired lineage commitment upon differentiation *in vitro* and *in vivo*[110]. Together, these studies show that $m^6A$ is an important regulator of transcriptome flexibility of PGRN and is required for lineage differentiation of ESCs.

**THE WIDE GAMUT OF PLURIPOTENT STATES [Au: please shorten to <39 characters. Perhaps just remove the word 'alternative' as I think this will be inferred anyway. I can't 100% guarantee that the word Gamut ]**

Since ESCs were first successfully isolated from mouse embryos 35 years ago, pioneering research has greatly expanded the spectrum of pluripotent states that can be either

stabilized using defined culture parameters, or artificially induced using genetic factors. From the original conventional naïve ESCs, the gamut of pluripotent states now extends from the 2i ground state and transient totipotent-like cells (with both embryonic and extra-embryonic potentials)[112,113] to primed EpiSCs and human ESCs (which are more developmentally advanced). Between these two ends of the spectrum, there exist a range of intermediate states with distinct functional characteristics and practical advantages (reviewed in[114]). For example, the spatial property of pluripotency was recently explored in the derivation of a novel class of primed PSCs. These so-called region-selective EpiSCs (rsEpiSCs) selectively integrate into the posterior, proximal region of the post-implantation epiblast. Analogous human rsPSCs can robustly contribute to interspecific chimeras when injected into the posterior epiblast of post-implantation mouse embryos[18]. Another unexpected finding revealed that elevated expression of reprogramming factors in somatic cells can lead to an artificially stabilized NANOG-positive alternative pluripotent state, thus extending the concept of pluripotency outside of its normal developmental context[115].

Studying the relationships and potential interconversions between different pluripotency states can provide interesting insights into the regulation of pluripotency. In addition to differences in embryonic origin, morphology, signalling requirement and chimera competency discussed in the previous section, naïve and primed PSCs have additional molecular differences (recently reviewed in[114,116]). For instance, global DNA hypomethylation is a hallmark of the naïve epigenome. Interestingly, the reorganization of core TF binding induced by the 2i ground conditions in conventional ESCs happens rapidly with minimal changes to global DNA methylation[50]. Therefore, the genome-wide DNA demethylation that is observed in conventional ESCs following 2i induction[117] is likely driven by the reorganization of core TF binding. Mechanistically, global loss of DNA methylation has been recently attributed to the simultaneous reduction of UHFR1 (an essential recruiting factor of DNMT1) and its cognate histone mark H3K9me2, which synergistically impair the DNA methylation maintenance machinery[118]. By contrast, the transition from naïve ESCs to primed epiblast-like cells (EpiLCs) is associated with genome-wide changes in NANOG binding, which are facilitated by epigenetic resetting of regulatory elements. In this reconfigured epigenetic landscape, NANOG alone can induce germ cell fate by activating enhancers of key germline TFs[119]. Recent transcriptomic analysis coupled with genome-wide mapping of epigenetic markers and core TF binding showed that the transition from naïve to primed pluripotency involves global genomic retargeting of OSN and remodelling of the

enhancer landscape[120,121]. These dynamic changes do not result from changes in core TF expression, but are rather caused by redirection of core TFs by their binding partners (e.g., OTX2) in different cellular contexts (Figure 3B).

Recently, the generation of naïve-like human ESCs has garnered much attention from the field. Many methods have been devised to achieve this state, including expression of naïve TFs or application of specific combinations of cytokines and inhibitors (reviewed in[114]). It was recently shown that a small molecule inhibitor of MLL1, a histone H3K4 methyltransferase, could efficiently reprogram mouse EpiSCs to naïve pluripotency[122]. It will be interestingly to determine if similar methods are effective in human ESCs. Despite interest and effort, ethical issues have prevented the development of definitive tests for naïve pluripotency in human cells (e.g., germline transmission and tetraploid complementation assays). Thus, the equivalency of these human naïve conditions to the gold standard mouse naïve ESCs has not been established.

Overall, research into alternative pluripotent states has greatly enriched our understanding of the PGRN. In turn, this knowledge has broadened the range of potential applications in regenerative medicine, including high-quality iPSCs, the generation of 3D organoids, efficient genome editing, and precise modelling of diseases.

**PERSPECTIVES**

The PGRN is one of the most studied gene regulatory networks. Combined efforts of multiple laboratories have helped elucidate fundamental rules of TF action, epigenetic modes of gene regulation, and the 3D organization of the genome. From a developmental biology perspective, studying the PGRN provides unprecedented access to early development, especially in human development, for which an experimental model was lacking. ESCs not only paved the way for developing genetic engineering by homologous recombination, they are currently serving as the anvil against which new genome editing technologies (e.g., TALENS and CRISPR-CAS9) are being hammered out[123-125]. Knowledge gained concerning the PGRN has guided progress toward safe and efficient iPSC technologies, which is an area of explosive development, holding great promise for regenerative medicine[126]. The development of human naïve pluripotent cells with the ability to contribute to interspecific chimeras raises the captivating possibility of generating xenogenic organs[114]. It is clear that a foundational framework for understanding

pluripotency has been established, but there are important discoveries yet to be made. We believe an important task will be to further our understanding of the heterogeneity and plasticity of the PGRN, either during early embryonic development[127], or induced by *in vitro* culture. To this end, single-cell analysis of the transcriptome and epigenome of PSCs has emerged as a powerful new method to study the heterogeneity and dynamics of PGRN[128-130]. To date, study of the PGRN in different pluripotent states using multi-omics approaches has provided a series of snapshots of this important biological process. Likephotographs of a ballet, these still shots capture the inherent beauty of this biological system, but the fluidity and elegance are left to the imagination. Armed with new technologies to precisely manipulate the genome and to quantitatively assess the transcriptome, epigenome, proteome, and metabolome at the single cell level, we are hopeful that the fascinating choreography governing the acquisition and dissolution of pluripotency in the embryo will be fully appreciated in the near future.

**Box 1**: **States of pluripotent stem cells** **[Au: please add a title for this box]** Mouse ESCs were first isolated on feeder cells in the presence of serum. Leukemia inhibitory factor (LIF) was later identified as a critical factor for self-renewal, and can eliminate the need for feeders. ESCs cultured in serum supplemented with LIF are referred to as **conventional ESCs**. Conventional ESCs display considerable heterogeneity and contain subpopulations that perform poorly in pluripotency assays. The conventional condition is not permissive for the derivation of ESC from NOD or FVB stains of mice. Replacing serum with inhibitors of MEF and GSK3 (2i) allows conventional ESCs to enter the so-called "**ground state**", which is thought to closely resemble unrestricted naïve pluripotency *in vivo*. ESCs in 2i plus LIF culture are characterized by uniform expression of key pluripotency TFs, reduced expression of differentiation genes, global hypomethylation, and the ability to contribute to chimeras. The 2i/LIF condition also enables derivation of ESCs from all stains of mice tested to date. Developmentally more advance or **primed** PSCs require FGF2 and Activin A signaling to self-renew. Primed PSCs are characterized by lower expression of certain pluripotent factors (e.g., NANOG, KLF4, and REX1), bivalent epigenetic marks at lineage genes, female X chromosome inactivation, and poor contribution to blastocyst chimeras[14,15]. **[Au: Please reference this statement.]**

**Box 2: 3D genome organization of ESCs**. Like differentiated cells, the genomes of ESCs are hierarchically organized. Traditional techniques, including fluorescence in situ hybridization (FISH) and immunostaining, revealed that each chromosome occupies its own nuclear space, called a **chromosome territory (CT**). The invention of chromosome conformation capture (3C) and 3C-derived technologies (e.g., 5C and Hi-C) greatly facilitated our understanding of the spatial organization of the genome at much finer scales[77]. This has produced the concepts of: A and B compartments, which are large active and inactive networks of interactions, respectively; Topologically-associated domains (TADs), which are self contained regions of frequent local contacts; and short-range cis-regulatory interactions between enhancer and promoters[77]. **[Au: Please reference this statement.]** TADs are relative static structural units, whereas cis-regulatory contacts are highly dynamic and cell-specific. In ESCs, high OSN occupancy is correlated with regions involved in long-range interactions, suggesting that OSN directly participates in organizing the genome in 3D[131,132]. **[Au: Please reference this statement.]**

**Acknowledgements:**

We apologize to the colleagues whose works are not covered due to space constraint. We would like to thank David O'Keefe and May Schwarz for critical reading and generous help during the preparation of the manuscript. Work in the laboratory of J.C.I.B. was supported by the G. Harold and Leila Y. Mathers Charitable Foundation, The Leona M. and Harry B. Helmsley Charitable Trust (2012-PG-MED002), the Moxie Foundation, Fundacion Dr. Pedro Guillen and the Universidad Católica San Antonio de Murcia (UCAM).

**Figure legends:**

**Figure 1 Bistability of the pluripotency gene regulatory network.** An optimal PGRN is the result of balanced expression of the core pluripotency TFs OCT4, SOX2 and NANOG. ESCs with an optimal PGRN can be imagined as existing in a stable energy state. Random fluctuation in levels of core pluripotency TF could disturb this state and make ESCs prone to differentiation. On the one hand, these destabilizing effects may be buffered by the expression of other pluripotency genes, continued presence of self-renewal signals, or other mechanisms (e.g., miRNAs that oppose differentiation) that restore levels of the core TFs and return the network to an optimal self-renewal state. On the other hand, in the presence of differentiation cues, expression of lineage TFs, epigenetic remodeling, and pro-differentiation miRNAs act in concert to dismantle the pluripotency network and to initiate differentiation of the three germ layers. **[Au: Please can you explain the figure more thoroughly in the legend i.e. please explain what is happening at each 'energy' state and mention all of the processes and features that are depicted. Figures need to be able to stand alone from the text. ]**

**Figure 2 Cooperative binding of pluripotency factors. A.** Through the process of 3D diffusion and the 1D sliding along chromosomes, SOX2 first finds and binds to its target sites. This is followed by recruitment of OCT4. OCT4 in turn stabilizes binding of the OCT4-SOX2 complex, ensuring optimal expression of the self-renewal program. The K156 residue of OCT4 is essential for stabilizing the salt bridge between D107 of SOX2 and K151 of OCT4. Posttranslational modification (PTM) or mutation of K156 impairs SOX2-OCT4 interaction and target OCT4 for degradation. This results in the upregulation of mesendodermal genes involved in the epithelial-to-mesenchymal transition (EMT). **B.** The SOX2 and OCT4 heterodimer can assume two different configurations. It was recently revealed that the SOX2-OCT4 configuration that binds to the no-gap canonical motif plays a more important role in the maintenance of pluripotency. The gene networks represent a conceptual model

for how target genes of the canonical SOX2-OCT4 heterodimer could form a network that has more relevance to pluripotency. Structural models of OCT4-SOX2 binding on DNA are reproduced with permission from REF. 44. **[Au: please state what the dashed lines and the lines that cross between the two networks represent]**

**Figure 3 Global retargeting of OSN binding and remodelling of the enhancer landscape during the transition between pluripotency states. A.** Switching from the conventional serum culture to the 2i condition induces rapid and widespread changes in ESCs. The 2i condition increases WNT signals in ESCs, while supressing the ERK signaling pathway. Within 24 hours, hundreds of genomic sites display differential OSN occupancy. Sites with increased OSN binding in 2i are enriched in binding motifs for the WNT pathway, whereas those with reduced OSN binding are associated with binding motifs for the ERK signaling pathway. **B.** In naïve ESCs, core TFs bind naïve-specific enhancers that are marked with H3k4me1 and H3k27Ac. OCT4 and SOX2 bind to these naïve active enhancers together with other pluripotency factors expressed in the naïve state, such as ESRRB. In the primed pluripotent state, naïve-specific enhancers become decommissioned, which is correlated with downregulation of *Esrrb* and upregulation of *Otx2*. OTX2 retargets OCT4 to primed active enhancers, which gain enhancer-specific histone marks after transition from the naïve to the primed state.

**Figure 4 An RNA-mediated "tug of war" in the regulation of the PGRN.** Many RNA-mediated mechanisms are critical for the maintenance of pluripotency, including miRNAs, lincRNAs, and alternative splicing. For example, OSN (OCT4, SOX2 and NANOG) binds to the transposable element HERV-H and activates transcription of lincRNA-ROR, which can relieve miR-145-mediated translational suppression of OSN mRNA (whose transcription also depends on binding of OSN) by acting as a miRNA sponge. OSN also regulates expression of the splicing factor SRSF2. SRSF2 in turn regulates alternative splicing of MBD2, resulting in two MBD2 isoforms–MBD2a and MBD2c. Although both MDB2a and MBD2c can bind the OCT4 and NANOG promoters, they have opposing functions in pluripotency, possibly through recruitment of different cofactors. The miR-301 and miR-302 families participate in the regulation of SRSF2 and MBD2 through positive feedback loops. **[Au: please ensure that all parts of the figure are explained in the legend. For example, the left part of the figure that shows OSN binding upstream of the *OCT4*, *SOX2* and *NANOG* genes.]**

1
2
3
4
5

**[Au: For references that are particularly worth reading (5-10% of the total), please provide a single bold sentence that indicates the significance of the work. Please could you also put the references into the following format. ]**

**Highlighted references:**


10 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).

11 Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7634-7638 (1981).

**These seminal papers described the successful derivaiton and stable culture of ESCs from mouse blastocysts, thereby set the stage for the study of pluripotency gene network.**

14 Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195, doi:10.1038/nature05950 (2007).

15 Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199,

**These papers described the derivaiton of ESCs from mouse blastocysts, thereby set the stage for the study of pluripotency gene network.**

18 Wu, J. *et al.* An alternative pluripotent state confers interspecies chimaeric competency. *Nature* **521**, 316-321, doi:10.1038/nature14413 (2015).

**This paper described a new type of pluripotent stem cells that selectively integrate into the posterior, proximal region of the post-implantation epiblast and can contribute to interspecific chimeras when injected into the posterior epiblast of post-implantation mouse embryos.**

19 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).

**The semimal discovery of induced pluripotent stem cells by reprogramming somatic cells with pluripotency-associated transcription factors.**

38 Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956, doi:10.1016/j.cell.2005.08.020 (2005).

**This semial paper discovered the genome-wide co-occupancy of OCT4, SOX2, and NANOG and suggested that PGRN consists of autoregulatory and feedforward loops.**

50 Galonska, C., Ziller, M. J., Karnik, R. & Meissner, A. Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. *Cell stem cell* **17**, 462-470, doi:10.1016/j.stem.2015.07.005 (2015).

**This paper demonstrated global retargeting of OSN binding and remodelling of the enhancer landscape during the transition between pluripotency states.**

63 Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523, doi:10.1038/nature06968 (2008).

**This is the original report of the 2i condition and ground state pluripotency.**

69 Chen, H. *et al.* Erk signaling is indispensable for genomic stability and self-renewal of mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E5936-5943, doi:10.1073/pnas.1516319112 (2015).

**This study challenged the idea that ERK signaling was dispensable for naïve pluripotency by showing that ERK1/2 double knockout ESCs cannot be maintained.**

75 Hu, G. *et al.* A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes & development* **23**, 837-848, doi:10.1101/gad.1769609 (2009).

**This paper described a screen that has discovered a large number of new genes in PGRN and a new PGRN module involving Cnot3, Trim28 and c-Myc.**

78 Wei, Z. *et al.* Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell stem cell* **13**, 36-47, doi:10.1016/j.stem.2013.05.010 (2013).

79 Zhang, H. *et al.* Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell stem cell* **13**, 30-35, doi:10.1016/j.stem.2013.05.012 (2013).

80 Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295, doi:10.1016/j.cell.2013.04.053 (2013).

**These studies showed that pluripotency genes regulates 3D genome organization, which in turn influences pluripotency gene expression and cell fate determination.**

97 Sheik Mohamed, J., Gaughwin, P. M., Lim, B., Robson, P. & Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *Rna* **16**, 324-337, doi:10.1261/rna.1441510 (2010).

98 Loewer, S. *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature genetics* **42**, 1113-1117, doi:10.1038/ng.710 (2010).

**These papers highlighted the roles of long noncoding RNAs in modulating PGRN.**

107 Lu, Y. *et al.* Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. *Cell stem cell* **15**, 92-101, doi:10.1016/j.stem.2014.04.002 (2014).

108 Han, H. *et al.* MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498**, 241-245, doi:10.1038/nature12270 (2013).

**These reports showed the roles of alternative splicing in regulating pluripotency and self-renewal.**

120 Factor, D. C. *et al.* Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. *Cell stem cell* **14**, 854-863, doi:10.1016/j.stem.2014.05.005 (2014).

121 Buecker, C. *et al.* Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell stem cell* **14**, 838-853,

**These studies demonstrated global retargeting of OSN binding and remodelling of the enhancer landscape during the transition between pluripotency states.**

1. Author, A. B. & Author, B. C. Title of the article. Nature Cell Biol. 6, 123-131 (2001).
2. Author, A. B. Title of the article. Nature Struct. Mol. Biol. 7, 101-109 (2003).
3. Author, A. B., Author, Z. X. & Author, B. C. Title of the article. EMBO J. 25, 3454-3461 (2006).
4. Author, A. B. in Title of Book (ed. Surname, I. N.) 75-98 (Publisher, City, 2000).

**[Au: If there are 6 or more authors, only the first should be listed, followed by et al.']**

**References:**


1 Ramalho-Santos, M. & Willenbring, H. On the origin of the term "stem cell". *Cell stem cell* **1**, 35-38, doi:10.1016/j.stem.2007.05.013 (2007).

2 Pappenheim, A. Ueber Entwickelung und Ausbildung der Erythroblasten. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin* **145**, 587-643, doi:10.1007/bf01969901 (1896).

3 Haeckel, E. Anthropogenie. *(3rd edn) Wilhelm Engelmann, Leipzig* (1877).

4 Evans, M. J. The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. *Journal of embryology and experimental morphology* **28**, 163-176 (1972).

5 Rosenthal, M. D., Wishnow, R. M. & Sato, G. H. In vitro growth and differetiation of clonal populations of multipotential mouse clls derived from a transplantable testicular teratocarcinoma. *Journal of the National Cancer Institute* **44**, 1001-1014 (1970).

6 Kahan, B. W. & Ephrussi, B. Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *Journal of the National Cancer Institute* **44**, 1015-1036 (1970).

7 Finch, B. W. & Ephrussi, B. Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture in vitro and their extinction upon hybridization with cells of permanent lines. *Proceedings of the National Academy of Sciences of the United States of America* **57**, 615-621 (1967).

8 Solter, D. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. *Nature reviews* **7**, 319-327, doi:10.1038/nrg1827 (2006).

9 Wu, J., Yamauchi, T. & Izpisua Belmonte, J. C. An overview of mammalian pluripotency. *Development (Cambridge, England)* **143**, 1644-1648, doi:10.1242/dev.132928 (2016).

10 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).

11 Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7634-7638 (1981).

12 Matsui, Y., Zsebo, K. & Hogan, B. L. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841-847 (1992).
13 Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8424-8428 (1993).
14 Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195, doi:10.1038/nature05950 (2007).
15 Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199, doi:10.1038/nature05972 (2007).
16 Huang, Y., Osorno, R., Tsakiridis, A. & Wilson, V. In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell reports* **2**, 1571-1578, doi:10.1016/j.celrep.2012.10.022 (2012).
17 Kojima, Y. *et al.* The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell stem cell* **14**, 107-120, doi:10.1016/j.stem.2013.09.014 (2014).
18 Wu, J. *et al.* An alternative pluripotent state confers interspecies chimaeric competency. *Nature* **521**, 316-321, doi:10.1038/nature14413 (2015).
19 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).
20 Gonzalez, F., Boue, S. & Izpisua Belmonte, J. C. Methods for making induced pluripotent stem cells: reprogramming a la carte. *Nature reviews* **12**, 231-242, doi:10.1038/nrg2937 (2011).
21 Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940-954, doi:10.1016/j.cell.2011.01.032 (2011).
22 Ng, H. H. & Surani, M. A. The transcriptional and signalling networks of pluripotency. *Nature cell biology* **13**, 490-496, doi:10.1038/ncb0511-490 (2011).
23 Scholer, H. R., Hatzopoulos, A. K., Balling, R., Suzuki, N. & Gruss, P. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. *The EMBO journal* **8**, 2543-2550 (1989).
24 Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391 (1998).
25 Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics* **24**, 372-376, doi:10.1038/74199 (2000).
26 Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & development* **17**, 126-140, doi:10.1101/gad.224503 (2003).
27 Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nature cell biology* **9**, 625-635, doi:10.1038/ncb1589 (2007).
28 Thomson, M. *et al.* Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**, 875-889, doi:10.1016/j.cell.2011.05.017 (2011).

29 Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642 (2003).
30 Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655 (2003).
31 Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-1234, doi:10.1038/nature06403 (2007).
32 Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-737, doi:10.1016/j.cell.2009.07.039 (2009).
33 Suzuki, A. *et al.* Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10294-10299, doi:10.1073/pnas.0506945103 (2006).
34 Macarthur, B. D., Ma'ayan, A. & Lemischka, I. R. Systems biology of stem cell fate and cellular reprogramming. *Nature reviews. Molecular cell biology* **10**, 672-681, doi:10.1038/nrm2766 (2009).
35 Orkin, S. H. *et al.* The transcriptional network controlling pluripotency in ES cells. *Cold Spring Harbor symposia on quantitative biology* **73**, 195-202, doi:10.1101/sqb.2008.72.001 (2008).
36 Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell stem cell* **15**, 416-430, doi:10.1016/j.stem.2014.09.015 (2014).
37 Adamo, A. *et al.* LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nature cell biology* **13**, 652-659, doi:10.1038/ncb2246 (2011).
38 Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956, doi:10.1016/j.cell.2005.08.020 (2005).
39 Chew, J. L. *et al.* Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Molecular and cellular biology* **25**, 6031-6046, doi:10.1128/MCB.25.14.6031-6046.2005 (2005).
40 Rodda, D. J. *et al.* Transcriptional regulation of nanog by OCT4 and SOX2. *The Journal of biological chemistry* **280**, 24731-24737, doi:10.1074/jbc.M502573200 (2005).
41 Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106-1117, doi:10.1016/j.cell.2008.04.043 (2008).
42 Festuccia, N. *et al.* Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell stem cell* **11**, 477-490, doi:10.1016/j.stem.2012.08.002 (2012).
43 Silva, J. & Smith, A. Capturing pluripotency. *Cell* **132**, 532-536, doi:10.1016/j.cell.2008.02.006 (2008).
44 Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274-1285, doi:10.1016/j.cell.2014.01.062 (2014).
45 Pan, X. *et al.* Site-specific Disruption of the Oct4-Sox2 Interaction Reveals Coordinated Mesendodermal Differentiation and the Epithelial-Mesenchymal Transition. *The Journal of biological chemistry*, doi:10.1074/jbc.M116.745414 (2016).

46 Jerabek, S., Merino, F., Scholer, H. R. & Cojocaru, V. OCT4: dynamic DNA binding pioneers stem cell pluripotency. *Biochimica et biophysica acta* **1839**, 138-154, doi:10.1016/j.bbagrm.2013.10.001 (2014).
47 Jauch, R. *et al.* Conversion of Sox17 into a pluripotency reprogramming factor by reengineering its association with Oct4 on DNA. *Stem cells (Dayton, Ohio)* **29**, 940-951, doi:10.1002/stem.639 (2011).
48 Tapia, N. *et al.* Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency. *Scientific reports* **5**, 13533, doi:10.1038/srep13533 (2015).
49 Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics* **38**, 431-440, doi:10.1038/ng1760 (2006).
50 Galonska, C., Ziller, M. J., Karnik, R. & Meissner, A. Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. *Cell stem cell* **17**, 462-470, doi:10.1016/j.stem.2015.07.005 (2015).
51 Kunarso, G. *et al.* Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature genetics* **42**, 631-634, doi:10.1038/ng.600 (2010).
52 Fort, A. *et al.* Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nature genetics* **46**, 558-566, doi:10.1038/ng.2965 (2014).
53 Schmidt, D. *et al.* Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148**, 335-348, doi:10.1016/j.cell.2011.11.058 (2012).
54 Kunath, T. *et al.* FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development (Cambridge, England)* **134**, 2895-2902, doi:10.1242/dev.02880 (2007).
55 Yuan, H., Corbi, N., Basilico, C. & Dailey, L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes & development* **9**, 2635-2645 (1995).
56 Tee, W. W., Shen, S. S., Oksuz, O., Narendra, V. & Reinberg, D. Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. *Cell* **156**, 678-690, doi:10.1016/j.cell.2014.01.009 (2014).
57 Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & development* **12**, 2048-2060 (1998).
58 Niwa, H., Ogawa, K., Shimosato, D. & Adachi, K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, 118-122, doi:10.1038/nature08113 (2009).
59 Martello, G., Bertone, P. & Smith, A. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *The EMBO journal* **32**, 2561-2574, doi:10.1038/emboj.2013.177 (2013).
60 Ye, S., Li, P., Tong, C. & Ying, Q. L. Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *The EMBO journal* **32**, 2548-2560, doi:10.1038/emboj.2013.175 (2013).

61 Chen, C. Y. *et al.* Bcl3 Bridges LIF-STAT3 to Oct4 Signaling in the Maintenance of Naive Pluripotency. *Stem cells (Dayton, Ohio)* **33**, 3468-3480, doi:10.1002/stem.2201 (2015).
62 Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-292 (2003).
63 Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523, doi:10.1038/nature06968 (2008).
64 Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590-604, doi:10.1016/j.cell.2012.03.026 (2012).
65 Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nature cell biology* **16**, 516-528, doi:10.1038/ncb2965 (2014).
66 Yeo, J. C. *et al.* Klf2 is an essential factor that sustains ground state pluripotency. *Cell stem cell* **14**, 864-872, doi:10.1016/j.stem.2014.04.015 (2014).
67 Qiu, D. *et al.* Klf2 and Tfcp2l1, Two Wnt/beta-Catenin Targets, Act Synergistically to Induce and Maintain Naive Pluripotency. *Stem cell reports* **5**, 314-322, doi:10.1016/j.stemcr.2015.07.014 (2015).
68 Kim, S. H. *et al.* ERK1 phosphorylates Nanog to regulate protein stability and stem cell self-renewal. *Stem cell research* **13**, 1-11, doi:10.1016/j.scr.2014.04.001 (2014).
69 Chen, H. *et al.* Erk signaling is indispensable for genomic stability and self-renewal of mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E5936-5943, doi:10.1073/pnas.1516319112 (2015).
70 Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H. & Young, R. A. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes & development* **22**, 746-755, doi:10.1101/gad.1642408 (2008).
71 Martello, G. *et al.* Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell stem cell* **11**, 491-504, doi:10.1016/j.stem.2012.06.008 (2012).
72 Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435 (2010).
73 Fazzio, T. G., Huff, J. T. & Panning, B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134**, 162-174, doi:10.1016/j.cell.2008.05.031 (2008).
74 Ding, L. *et al.* Systems Analyses Reveal Shared and Diverse Attributes of Oct4 Regulation in Pluripotent Cells. *Cell systems* **1**, 141-151, doi:10.1016/j.cels.2015.08.002 (2015).
75 Hu, G. *et al.* A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes & development* **23**, 837-848, doi:10.1101/gad.1769609 (2009).
76 Li, M., Liu, G. H. & Izpisua Belmonte, J. C. Navigating the epigenetic landscape of pluripotent stem cells. *Nature reviews. Molecular cell biology* **13**, 524-535 (2012).

77 Gorkin, D. U., Leung, D. & Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell stem cell* **14**, 762-775, doi:10.1016/j.stem.2014.05.017 (2014).
78 Wei, Z. *et al.* Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell stem cell* **13**, 36-47, doi:10.1016/j.stem.2013.05.010 (2013).
79 Zhang, H. *et al.* Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell stem cell* **13**, 30-35, doi:10.1016/j.stem.2013.05.012 (2013).
80 Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295, doi:10.1016/j.cell.2013.04.053 (2013).
81 Rowe, H. M. *et al.* TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome research* **23**, 452-461, doi:10.1101/gr.147678.112 (2013).
82 Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & development* **16**, 919-932, doi:10.1101/gad.973302 (2002).
83 Cheng, B., Ren, X. & Kerppola, T. K. KAP1 represses differentiation-inducible genes in embryonic stem cells through cooperative binding with PRC1 and derepresses pluripotency-associated genes. *Molecular and cellular biology* **34**, 2075-2091, doi:10.1128/MCB.01729-13 (2014).
84 Sun, C. *et al.* Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. *Molecular and cellular biology* **29**, 4574-4583, doi:10.1128/MCB.01863-08 (2009).
85 Fujii, S. *et al.* Nr0b1 is a negative regulator of Zscan4c in mouse embryonic stem cells. *Scientific reports* **5**, 9146, doi:10.1038/srep09146 (2015).
86 Zhang, J. *et al.* Dax1 and Nanog act in parallel to stabilize mouse embryonic stem cells and induced pluripotency. *Nature communications* **5**, 5042, doi:10.1038/ncomms6042 (2014).
87 Ding, J. *et al.* Tex10 Coordinates Epigenetic Control of Super-Enhancer Activity in Pluripotency and Reprogramming. *Cell stem cell* **16**, 653-668, doi:10.1016/j.stem.2015.04.001 (2015).
88 Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nature structural & molecular biology* **20**, 311-316, doi:10.1038/nsmb.2510 (2013).
89 Yamaji, M. *et al.* PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell stem cell* **12**, 368-382, doi:10.1016/j.stem.2012.12.012 (2013).
90 Tu, S. *et al.* Co-repressor CBFA2T2 regulates pluripotency and germline development. *Nature* **534**, 387-390, doi:10.1038/nature18004 (2016).
91 Tiscornia, G. & Izpisua Belmonte, J. C. MicroRNAs in embryonic stem cell function and fate. *Genes & development* **24**, 2732-2741, doi:10.1101/gad.1982910 (2010).
92 Gangaraju, V. K. & Lin, H. MicroRNAs: key regulators of stem cells. *Nature reviews* **10**, 116-125 (2009).
93 Houbaviy, H. B., Murray, M. F. & Sharp, P. A. Embryonic stem cell-specific MicroRNAs. *Developmental cell* **5**, 351-358 (2003).

94 Melton, C., Judson, R. L. & Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **463**, 621-626 (2010).
95 Li, M. & Izpisua Belmonte, J. C. Roles for noncoding RNAs in cell-fate determination and regeneration. *Nature structural & molecular biology* **22**, 2-4, doi:10.1038/nsmb.2946 (2015).
96 Gu, K. L. *et al.* Pluripotency-associated miR-290/302 family of microRNAs promote the dismantling of naive pluripotency. *Cell research* **26**, 350-366, doi:10.1038/cr.2016.2 (2016).
97 Sheik Mohamed, J., Gaughwin, P. M., Lim, B., Robson, P. & Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *Rna* **16**, 324-337, doi:10.1261/rna.1441510 (2010).
98 Loewer, S. *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature genetics* **42**, 1113-1117, doi:10.1038/ng.710 (2010).
99 Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J. A. & Kosik, K. S. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* **137**, 647-658, doi:10.1016/j.cell.2009.02.038 (2009).
100 Wang, Y. *et al.* Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Developmental cell* **25**, 69-80, doi:10.1016/j.devcel.2013.03.002 (2013).
101 Santoni, F. A., Guerra, J. & Luban, J. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. *Retrovirology* **9**, 111, doi:10.1186/1742-4690-9-111 (2012).
102 Lu, X. *et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature structural & molecular biology* **21**, 423-425, doi:10.1038/nsmb.2799 (2014).
103 Wang, J. *et al.* Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405-409, doi:10.1038/nature13804 (2014).
104 Salomonis, N. *et al.* Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 10514-10519, doi:10.1073/pnas.0912260107 (2010).
105 Gabut, M. *et al.* An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* **147**, 132-146, doi:10.1016/j.cell.2011.08.023 (2011).
106 Das, S., Jena, S. & Levasseur, D. N. Alternative splicing produces Nanog protein variants with different capacities for self-renewal and pluripotency in embryonic stem cells. *The Journal of biological chemistry* **286**, 42690-42703, doi:10.1074/jbc.M111.290189 (2011).
107 Lu, Y. *et al.* Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. *Cell stem cell* **15**, 92-101, doi:10.1016/j.stem.2014.04.002 (2014).
108 Han, H. *et al.* MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498**, 241-245, doi:10.1038/nature12270 (2013).

109 Jia, G., Fu, Y. & He, C. Reversible RNA adenosine methylation in biological regulation. *Trends in genetics : TIG* **29**, 108-115, doi:10.1016/j.tig.2012.11.003 (2013).
110 Batista, P. J. *et al.* m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell stem cell* **15**, 707-719, doi:10.1016/j.stem.2014.09.019 (2014).
111 Wang, Y. *et al.* N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nature cell biology* **16**, 191-198, doi:10.1038/ncb2902 (2014).
112 Morgani, S. M. *et al.* Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell reports* **3**, 1945-1957, doi:10.1016/j.celrep.2013.04.034 (2013).
113 Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57-63, doi:10.1038/nature11244 (2012).
114 Wu, J. & Izpisua Belmonte, J. C. Dynamic Pluripotent Stem Cell States and Their Applications. *Cell stem cell* **17**, 509-525, doi:10.1016/j.stem.2015.10.009 (2015).
115 Tonge, P. D. *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* **516**, 192-197, doi:10.1038/nature14047 (2014).
116 Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nature reviews. Molecular cell biology* **17**, 155-169, doi:10.1038/nrm.2015.28 (2016).
117 Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell stem cell* **13**, 351-359, doi:10.1016/j.stem.2013.06.004 (2013).
118 von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Molecular cell* **62**, 848-861, doi:10.1016/j.molcel.2016.04.025 (2016).
119 Murakami, K. *et al.* NANOG alone induces germ cells in primed epiblast in vitro by activation of enhancers. *Nature* **529**, 403-407, doi:10.1038/nature16480 (2016).
120 Factor, D. C. *et al.* Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. *Cell stem cell* **14**, 854-863, doi:10.1016/j.stem.2014.05.005 (2014).
121 Buecker, C. *et al.* Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell stem cell* **14**, 838-853, doi:10.1016/j.stem.2014.04.003 (2014).
122 Zhang, H. *et al.* MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. *Cell stem cell* **18**, 481-494, doi:10.1016/j.stem.2016.02.004 (2016).
123 Li, M., Suzuki, K., Kim, N. Y., Liu, G. H. & Izpisua Belmonte, J. C. A cut above the rest: targeted genome editing technologies in human pluripotent stem cells. *The Journal of biological chemistry* **289**, 4594-4599, doi:10.1074/jbc.R113.488247 (2014).
124 Suzuki, K. *et al.* Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell stem cell* **15**, 31-36, doi:10.1016/j.stem.2014.06.016 (2014).

125 Li, M. *et al.* Efficient correction of hemoglobinopathy-causing mutations by homologous recombination in integration-free patient iPSCs. *Cell research* **21**, 1740-1744 (2011).
126 Li, M. & Izpisua Belmonte, J. C. Looking to the future following 10 years of induced pluripotent stem cell technologies. *Nature protocols* **11**, 1579-1585, doi:10.1038/nprot.2016.108 (2016).
127 Goolam, M. *et al.* Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* **165**, 61-74, doi:10.1016/j.cell.2016.01.047 (2016).
128 Kolodziejczyk, A. A. *et al.* Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell stem cell* **17**, 471-485, doi:10.1016/j.stem.2015.09.011 (2015).
129 Angermueller, C. *et al.* Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nature methods* **13**, 229-232, doi:10.1038/nmeth.3728 (2016).
130 Guo, G. *et al.* Serum-Based Culture Conditions Provoke Gene Expression Variability in Mouse Embryonic Stem Cells as Revealed by Single-Cell Analysis. *Cell reports* **14**, 956-965, doi:10.1016/j.celrep.2015.12.089 (2016).
131 de Wit, E. *et al.* The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* **501**, 227-231, doi:10.1038/nature12420 (2013).
132 Apostolou, E. *et al.* Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell stem cell* **12**, 699-712, doi:10.1016/j.stem.2013.04.013 (2013).

## GLOSSARY **[Au: please add terms highlighted in purple to the glossary, providing a succinct 1-2 sentence definition for each]**

Inner cell mass (ICM): A small cluster of cells inside the early embryo termed blastocyst. These cells give rise to all the tissues of the future embryo, but not to extra-embryonic tissues (e.g. the placenta). The ICM may be isolated to generate embryonic stem cells.

Trophectoderm: First specialized lineage of cells forming the outer cell layer of a blastocyst. It gives rise to extra-embryonic tissues.

Mesendoderm: A layer of cell formed during early gastrulation that are destined to become mesoderm and endoderm.

Ectoderm: The outermost of the three germ layers formed during gastrulation of the early embryo. Ectoderm derived tissues include the nervous system, sensory organs and skin.

Stemness genes: genes that constitute the stem cell specific gene expression program.

Epithelial-mesenchymal transition (EMT): A process in which cells of an epithelial layer lose their polarity and cell-cell adhesion and become disorganized, migratory mesenchymal cells. EMT is an integral in normal developmental, wound healing and cancer development.

**[Au: please add competing interests statement.]**

COMPETING INTERESTS STATEMENT
None declared.

ONLINE ONLY

## Author Bios

**[Au: We need a brief (just 50 words or so) biography for each author, detailing your current job title, careers and interests. This will be an online-only feature of the journal.] ]**

Juan Carlos Izpisua Belmonte

Juan Carlos Izpisua Belmonte graduated from the University of Valencia, Spain, and received his Ph.D. from the University of Bologna, Italy, and the University of Valencia. He carried out postdoctoral studies at the EMBL and the University of California, Los Angeles, USA. He is currently a professor at the Salk Institute for Biological Studies, La Jolla, California, USA. His scientific interests include embryogenesis, stem cell biology, aging and regeneration.

Mo Li

Mo Li received his Ph.D. in cellular biology from the University of Georgia in Athens, Georgia, USA, in 2007. He worked as a postdoctoral researcher at the Salk Institute, where his research interests focus on somatic cell reprogramming and genetic correction of human monogenic diseases. He is now Assistant Professor of Bioscience at King Abdullah University of Science and Technology.

## Key points

**[Au: Please provide a list of up to 6 brief bullet points, each no more than 2 sentences long, highlighting the take-home messages of the Review.]**

- Pluripotency exists transiently in the early embryo and can be recapitulated *in vitro*.
- Pluripotency is stabilized by an interconnected network of pluripotency genes.
- Pluripotency gene network integrates external signals and exerts control over the decision between self-renewal and differentiation at the transcriptional, post-transcriptional and epigenetic levels.
- Diverse pathways including chromatin-mediated mechanisms, RNA-based regulation and 3D genome organization work in concert to maintain pluripotency.
- Recent evidence of alternative pluripotency states indicates regulatory flexibility of this network.

TOC blurb