REVIEWS

Molecular features of cellular reprogramming and development

Zachary D. Smith*, Camille Sindhu* and Alexander Meissner

Abstract | Differentiating somatic cells are progressively restricted to specialized functions during ontogeny, but they can be experimentally directed to form other cell types, including those with complete embryonic potential. Early nuclear reprogramming methods, such as somatic cell nuclear transfer (SCNT) and cell fusion, posed significant technical hurdles to precise dissection of the regulatory programmes governing cell identity. However, the discovery of reprogramming by ectopic expression of a defined set of transcription factors, known as direct reprogramming, provided a tractable platform to uncover molecular characteristics of cellular specification and differentiation, cell type stability and pluripotency. We discuss the control and maintenance of cellular identity during developmental transitions as they have been studied using direct reprogramming, with an emphasis on transcriptional and epigenetic regulation.

Totipotent
Defines a cell that can autonomously contribute to all of the tissues of a developing organism, including extra-embryonic and placental tissues, as well as those of the embryo proper. This property is restricted during development to the zygote and the first two cleavage divisions.

Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA. *These authors contributed equally to this article. Correspondence to A.M. alexander_meissner@harvard.edu

During development, cells within a multicellular organism progressively differentiate into functionally and phenotypically distinct fates, a specialization that is enabled by cell type-specific gene expression. Once established, cellular states are remarkably stable and can be sustained over many cell divisions throughout an organism's lifespan. The process of differentiation proceeds from the totipotent zygote, which is itself formed at fertilization by the reactivation of early developmental programmes within the nuclei of two highly specialized gametes. The tremendous reprogramming potential of the ooplasm is highlighted by somatic cell nuclear transfer (SCNT) experiments, in which specialized cells from any somatic lineage are rapidly directed to totipotency by the cytoplasm of the enucleated oocyte (BOX 1). The feasibility of using SCNT to generate germline-competent organisms proved that developmental processes are imposed strictly by epigenetic mechanisms and are thus reversible. SCNT remained a major technology for studying the regulatory mechanisms behind this functional reset, until the demonstration that pluripotency could be accomplished *in vitro* by the ectopic expression of only four transcription factors — OCT4 (also known as POU5F1), SRY-box 2 (SOX2), Krüppel-like factor 4 (KLF4) and MYC, collectively referred to as OSKM — to produce induced pluripotent stem cells (iPSCs)[1] (FIG. 1). Whereas SCNT is technically challenging and utilizes a limited gametic resource, reprogramming by OSKM is experimentally tractable and can be performed *in vitro* on large populations of cells, enabling systematic biochemical and genomic characterization of the mechanisms that impose or surmount the epigenetic constraints governing cellular identity (see Supplementary information S1 (box)).

In this Review, we describe insights that have been uncovered using direct reprogramming as an experimental tool, beginning by defining cellular identity as a specific molecular state that is stabilized and maintained through cooperating transcriptional and epigenetic mechanisms, including many parameters that are unique to functional pluripotency (that is, the ability to contribute to the formation of all embryonic tissues). We also address how chromatin remodellers, transcription factors and various levels of epigenetic regulation are coordinated to re-establish developmental potential *in vitro*.

A molecular definition of cell state

Cellular identity and differentiation potential were originally assigned on the basis of phenotype and lineage history, as well as from the stability of these traits following transplantation of cells to ectopic sites within the developing embryo. Studies of molecular regulation of cellular identity initially focused on master transcription factors that are expressed at the inception of a cell lineage and that are often necessary or sufficient to direct its identity[2]. However, genomic expression profiling strategies revealed substantial redundancy in the function of specific master regulators within different cell types of the same lineage, or even across lineages, indicating that they perform context-specific activities within an identical genomic framework[3–5]. For instance, the transcription factor SOX2 participates in the regulation of pluripotency[6], specification to the neural lineage[7] and adult tissue homeostasis[8]. The non-overlapping spectra of genes targeted by this

Box 1 | **Global nuclear reprogramming during SCNT**

At fertilization, maternal factors must rapidly reconstitute (rechromatinize) the nucleosome-depleted, protamine-compacted paternal genome to form the totipotent zygote. This process is initiated within minutes of fertilization and proceeds through paternal genome decompaction by oocyte-specific linker histones and chaperones, including deposition of histone H3 variant 3.3 (H3.3)-containing nucleosomes by the chaperone HIRA, all of which occurs before DNA replication[197]. Blocking paternal rechromatinization by depleting HIRA arrests the assembly of a functional zygote at its earliest stages, before the assembly of nuclear structures, such as the envelope or the nucleolus, and transcription of rDNA[198,199]. Following this initial reprogramming, early fertilization and zygotic progression display high global rates of nucleosome turnover until after the two-cell stage, a regulatory transition that may correspond to the gradual loss of nuclear competence for high-efficiency cloning when using cells from subsequent divisions[200,201]. Many principles of global rechromatinization and zygotic genome activation have been applied to direct reprogramming, where supplementation with maternal histone variants or chaperones[202,203], or with DNA-binding factors that are specific to fertilization[183,204], improve the generation of induced pluripotent stem cells (iPSCs).

In somatic cell nuclear transfer (SCNT), nuclei of specialized cells from any somatic lineage are rapidly directed to totipotency by inserting them into the cytoplasm of enucleated oocytes. During SCNT, maternal factors are confronted with a fundamentally different, chromatinized somatic nucleus, but they seem to proceed with reprogramming through replacement strategies and kinetics that are highly similar to those observed during fertilization[205–207]. However, whereas fertilization is accompanied by massive chromatin remodelling of the paternal genome, the maternal genome remains chromatinized and is comparatively static[208]. A number of epigenetic mechanisms have emerged to protect maternal information during this phase of global paternal reprogramming, including recognition of maternal H3 Lys9 methylation (H3K9me) by the factor STELLA (also known as DPPA3) to maintain maternal imprinting and global DNA methylation[208,209]. In many cases, mechanisms that preserve maternal epigenetic states may counteract global epigenetic reprogramming during SCNT and propagate artefacts. For example, the efficiencies of SCNT-generated animals are notably lower for female versus male donor nuclei, and female cloned embryos exhibit preferential, nonrandom silencing of their epigenetically inactivated X chromosome[210,211]. This misdirected maintenance of an epigenetic state can be repaired by eliminating the X-inactive specific transcript (*Xist*) long non-coding RNA in donor cells before nuclear transfer[212]. Similarly, SCNT zygotes have global DNA demethylation levels comparable to those generated by fertilization, but with pronounced retention at repetitive elements that are repressed through SETDB1-mediated H3K9me3 (REF. 213). Careful analysis of transcriptional changes following SCNT indicates that H3K9 methylation in the donor nucleus represents a major epigenetic impediment to successful reprogramming[214]. Co-injection of oocytes with H3K9 demethylase *KDM4* mRNA rescues the resulting impaired gene expression and cloning deficiencies, as does injection of donor nuclei encoding a hypomorphic DNA methyltransferase 1 (*Dnmt1*) allele[214,215]. Altering global chromatin accessibility within the donor genome by increasing the fraction of H3.3-containing nucleosomes also improves SCNT efficiency[147]. Contemporary studies of SCNT demonstrate the extent and epigenetic limitations of global reprogramming to establish totipotency, which requires careful allocation of specific machineries to reset the paternal genome while maintaining maternal information.

Pluripotency
The ability of a cell to contribute to all embryonic tissues, including the germ line. Pluripotency is most stringently confirmed by the generation of germline-competent organisms after injection of cells into tetraploidized, embryo-deficient blastocysts.

trans-acting factor clearly influence the functional differences between pluripotent stem cells (PSCs) and neural progenitors, and they are regulated, in part, through lineage-specific cofactors[5,9].

Although the combinatorial expression of master transcriptional regulators constitutes a simple and plausible model for the control of diverse cell functions during development, the genomic distribution of transcription factor-binding sequences (motifs) or the co-expression of factors within a given cell type are surprisingly imprecise predictors of target loci or transcriptional output. These limitations hinder efforts to identify sets of factors that reprogramme one cell type to another, which instead generally requires systematically testing known regulators to identify a sufficient combination, as initially performed by Takahashi and Yamanaka[10–12]. During development, the genomic occupancy of master transcriptional regulators and their cofactors is often restricted to local 'nucleosome-free regions', in which *cis*-regulatory sequences are not occluded by chromatin, making it difficult to study the stepwise coordination of cell type-specific regulatory elements from static snapshots[13–15]. Recent efforts to comprehensively map regulatory networks across mammalian cell types demonstrate the complex interplay between transcriptional regulators and the local epigenetic environment as they cooperate to direct cellular identity[16,17]. This integration of transcriptional and chromatin state data encompasses empirically determined definitions of cell function, lineage and developmental potential to delineate a precise molecular genomic state.

Reprogramming the somatic cell state

Although epigenomic annotation efforts have compiled extensive information about the genomic abundance, location and function of regulatory sequences that underlie natural developmental transitions, they are inadequate to predict the potential of experimentally directed reprogramming. Ontogeny is sequential and spatiotemporally controlled, such that prior cellular states influence the activation of subsequent molecular programmes. By contrast, direct reprogramming of differentiated cells to pluripotency proves that, despite marked stability, cell fate is not irreversible and need not depend on lineage history. Classically, ectopic introduction of reprogramming factors in somatic cells only generates iPSCs after an extended latency of one or several weeks at a quantifiable, but low, frequency (FIG. 1). However, populations of intermediate cells can be isolated and assessed by molecular profiling or screened for emerging heterogeneity over the experimental time course. These experimental features have been repeatedly utilized to uncover insights into the nature of somatic cell identity by distinguishing ectopic factor activities that can function within pre-established, responsive regulatory programmes from those that must surmount the more complex, epigenetically constrained barriers that are imposed during differentiation.

Initial effects on the somatic epigenome. Early studies of the reprogramming process noted that ectopic introduction of OSKM into differentiated cells results in population-wide morphological changes and a loss of somatic identity markers, as well as a rapid reduction in cell size coupled with increased proliferation[18–23]. Notably, this initial 'de-differentiation' response appears to be largely unstable, such that induced cells will reactivate somatic state-specific markers if reprogramming factors are prematurely removed; similarly, isolated 'intermediate' cell states that divide continuously but are not pluripotent require persistent OSKM expression[19,24–26] (FIG. 2). The inability of OSKM to rapidly alter the differentiated state can be explained by epigenetic regulatory

mechanisms that preserve global features of cellular identity during mitosis in the absence of genetically targeted activating or repressive cues (see Supplementary information S2 (box)). Indeed, global profiling of epigenetic modifications such as DNA methylation and the trimethylation of histone H3 Lys4 and Lys27 (H3K4me3 and H3K27me3, respectively) indicate very limited

epigenetic remodelling following factor induction in mouse embryonic fibroblasts (MEFs) except in very specific contexts[18,26–29]. Thus, the general structure of the somatic genome is preserved in perpetually dividing, factor-dependent intermediate-state cells, further demonstrating the remarkable stability of epigenetic modifications over weeks-long experiments[18,26–28].
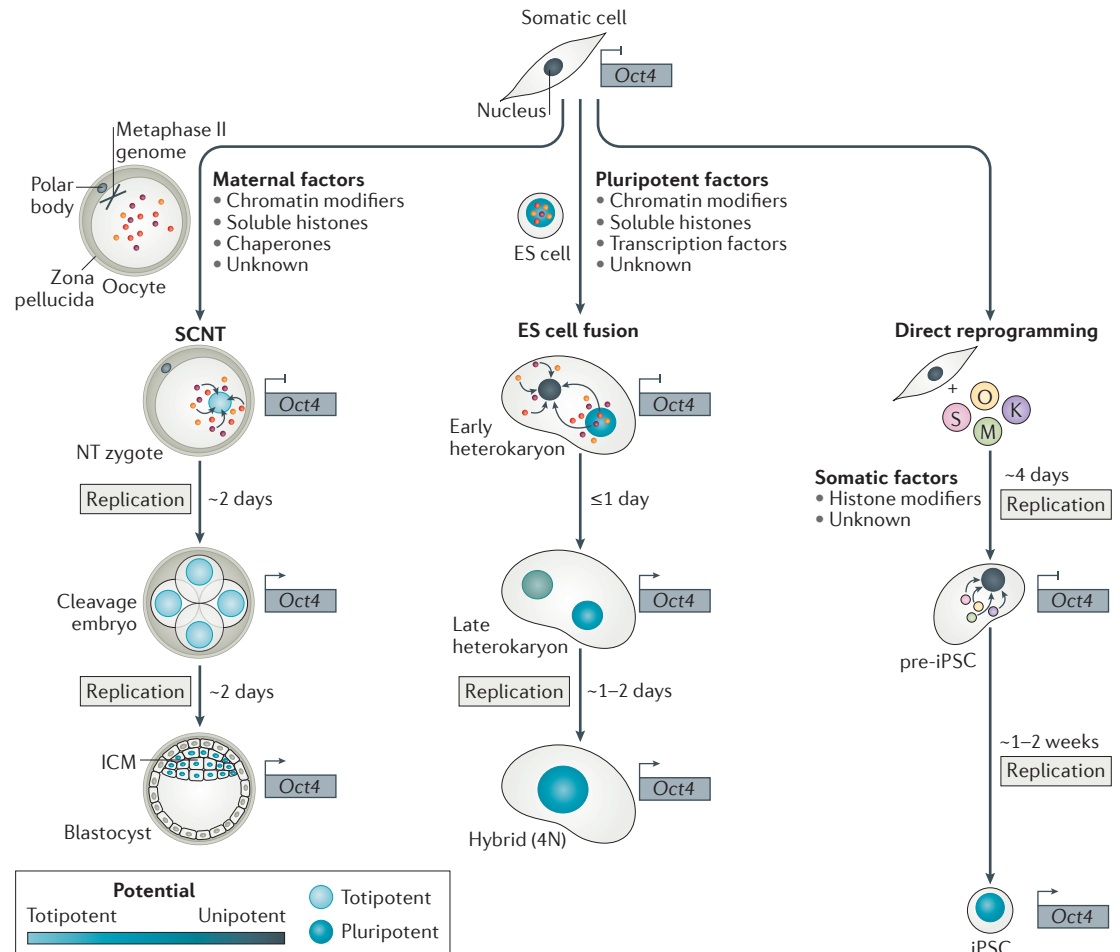

Figure 1 | **Approaches for reprogramming somatic nuclei.** A somatic nucleus can be directed to an early embryonic state through three general approaches: somatic cell nuclear transfer (SCNT), fusion with embryonic stem cells (ES cells) and ectopic factor expression (direct reprogramming). In somatic nuclei, pluripotency-specific genes such as *Oct4* are epigenetically silenced. Nuclear potential scales from totipotent (the ability to generate an entire organism, including extra-embryonic and embryonic tissues) to unipotent. In SCNT, somatic nuclei are transferred into enucleated oocytes. Upon oocyte activation, maternal factors rapidly and globally remodel the somatic genome (BOX 1). Successfully reprogrammed SCNT zygotes have restored totipotency and express select pluripotency factors, such as *Oct4*, that have developmental functions in both the early embryo and inner cell mass (ICM). After ~3–4 days, the ICM is formed in the blastocyst and developmental potency is restricted, producing pluripotent cells that contribute to the embryo proper, or from which ES cells can be derived. In experimentally induced ES cell fusion, a somatic cell is fused with a pluripotent cell. Reprogramming is initiated in the heterokaryon phase, during which both nuclei remain discrete, and includes global epigenetic remodelling that precedes the activation of pluripotent loci[216,217]. In the late heterokaryon phase, select loci are activated through a process that may require DNA replication[218–220]. The somatic and pluripotent nuclei fuse after cell division, and additional genes are then reprogrammed to consolidate the pluripotent network within the somatic genome[218,220]. During direct reprogramming, OCT4, SRY-box 2 (SOX2), Krüppel-like factor 4 (KLF4) and MYC (OSKM) are introduced into somatic cells, which respond by increasing proliferation and undergoing local changes to their epigenome. Shortly afterwards, KLF4 induces epithelial genes, and genes supporting the mesenchymal state are repressed. Early responses do not include activation of true pluripotency-associated regulators, including *Oct4*. The transition from a factor-dependent, non-pluripotent state to induced pluripotent stem cells (iPSCs) requires multiple divisions and an extended latency period under persistent ectopic OSKM expression. Complete reprogramming of the somatic nucleus occurs within a small percentage of responding cells and includes activation of the endogenous network, as well as consolidation of additional global epigenetic features of stable pluripotency. NT zygote, nuclear transfer zygote.

Direct reprogramming
Stable, experimentally induced changes in cellular state driven by a defined set of ectopic factors or conditions.

Chromatin remodellers
ATP-dependent proteins and complexes that change the relative positioning of nucleosomes to support either the activation or repression of a gene.
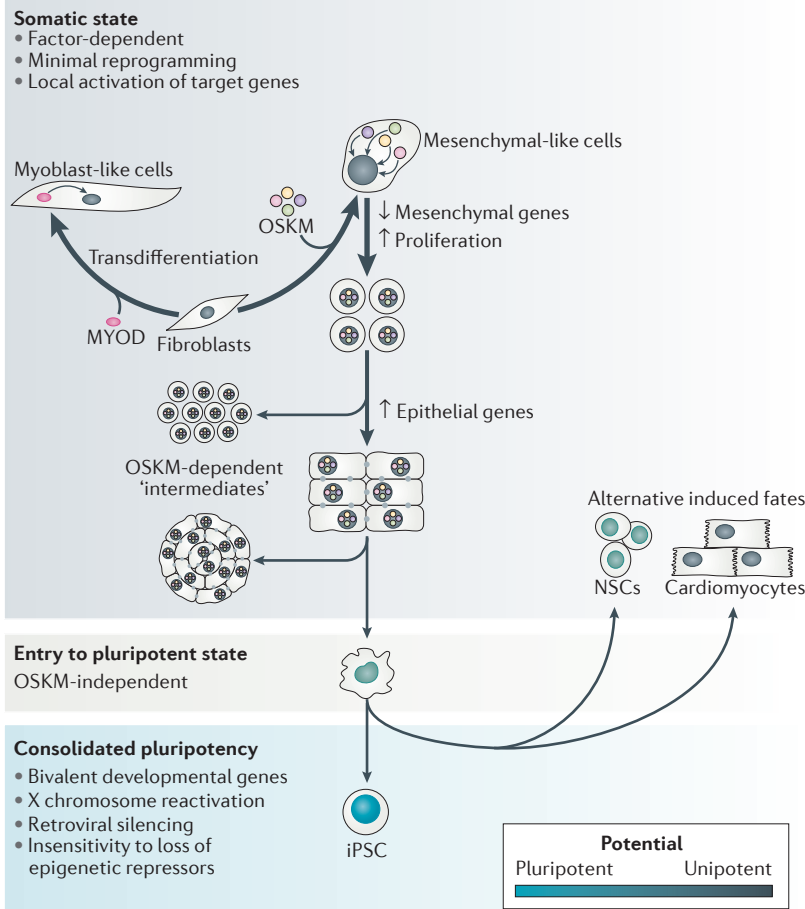
Figure 2 | Direct reprogramming traverses stable somatic and pluripotent state boundaries. The ectopic expression of lineage-altering transcription factors in somatic cells does not immediately result in the establishment of another cell state (known as transdifferentiation). For example, the ectopic expression of myoblast determination protein (MYOD) in fibroblasts can activate muscle-specific genes, but complete conversion to a muscle cell does not occur in all cell types[2]. Endogenous *MyoD* is not activated, suggesting that the myogenic programme is dependent on sustained expression of exogenous MYOD[221]. Moreover, genes expressed within the original cell type are not always fully suppressed, and their expression sometimes persists despite the acquisition of some myogenic markers. These 'myoblast-like cells' have distinct morphological features and gene expression programmes resulting from local reprogramming of select targets, and not from global epigenetic reprogramming, as is observed when directing cells to pluripotency. Initially, ectopic expression of OCT4, SRY-box 2 (SOX2), Krüppel-like factor 4 (KLF4) and MYC, collectively referred to as OSKM, induces unstable modulation of the original somatic state, with local rather than global epigenetic changes that depend on sustained OSKM expression and thus do not represent complete nuclear reprogramming. Mesenchymal genes are initially repressed by OCT4 and SOX2, and the cell cycle is accelerated by MYC, after which epithelial gene expression programmes are induced, in part by KLF4. At either of these stages, OSKM-dependent, non-pluripotent intermediates can be isolated as perpetually dividing cell lines. Neither transdifferentiation nor early direct reprogramming establish new state characteristics in addition to those prescribed as somatic. Reprogramming cells enter pluripotency after a sufficient number of key regulators have been activated, presumably downstream of a single switch-like event. During consolidation, cells undergo complete nuclear reprogramming and acquire many features that are specific to pluripotency, including complete embryonic potential, establishment of bivalent domains at promoters of developmental genes, silencing of retroviral vectors, insensitivity to loss of global repressors such as DNA methylation, and X chromosome reactivation in females. Alternative somatic fates can be induced by OSKM, including neural stem cells (NSCs) or cardiomyocytes, but require support from exogenous factors and transient passage through pluripotency, as measured by X chromosome reactivation and the silencing of viral vectors. iPSC, induced pluripotent stem cell.

Mechanisms for the mitotic inheritance of epigenetic modifications provide a robust, transcription factor-independent mode of maintaining nuclear homeostasis. As a consequence, foreign transcription factor activity is largely restricted to loci where stabilizing repressive mechanisms are either not in play or can be effectively reversed. Conversely, induction of OSKM expression in MEFs results in the rapid loss of thousands of distal, state-specific enhancers[27], presumably in response to the downregulation of corresponding transcription factors[18,30]. Immediate transcriptional responses to ectopic OSKM expression are also limited in scope, occurring largely at accessible, H3K4me3-modified promoters[18,27]. Nonspecific alteration of global somatic chromatin states can assist in removing epigenetic memory that buffers against the activities of OSKM. For example, the modifications H3K79me2 and H3K36me3 are canonically distributed over the bodies of transcribed genes and interfering with their regulation supports iPSC generation by eliminating a potent memory of prior transcriptional activity in the pre-induced somatic cell[31,32]. Similarly, the H3K27me3 demethylase UTX facilitates the induction of pluripotency-associated genes that are repressed by this modification during early development[33]. In each of these cases, a specific epigenetic pathway actively participates in the erasure of somatic memory at target genes. More globally, the broader epigenetic maintenance of somatic nuclear identity is preserved until iPSCs emerge, after which the majority of these modifications are reset by the unique regulatory characteristics of the pluripotent state (see below).

Transcriptional changes during early reprogramming. The epigenetic robustness of the somatic nuclear state initially limits reprogramming in response to ectopic OSKM, but transcriptional programmes that function as modules within multiple cell types during development seem to be particularly sensitive to induction. For instance, cellular proliferation and mitogen sensing are actively regulated in somatic cells, and promoters of genes involved in these processes are generally H3K4 trimethylated and primed for expression[34]. The earliest phenotypic and molecular changes following OSKM expression include an immediate increase in proliferation that is decoupled from normal homeostatic growth control[23]. MYC is broadly expressed across cell types, such that its role during reprogramming probably stems from its ectopic expression from a strong constitutive promoter, and participates in a largely separate regulatory network from OCT4 and SOX2 in pluripotent cells[28,35]. Upon OSKM induction, MYC preferentially targets open, accessible sites, primarily at core promoter sequences, where it promotes the transition from initiating to elongating forms of RNA polymerase II (Pol II) to enhance transcription at responsive cell cycle genes[27,28,36,37]. Similarly, the early response to OSKM includes a metabolic switch from oxidative phosphorylation to glycolysis that is directed, in part, through MYC and is a commonly observed feature of transformed cells, such as cancer cells, as well as of pluripotent and adult stem cells[38–42].

OSKM induction also triggers immediate down-regulation of somatic identity genes, including those that are characteristic of mesenchymal cells (FIG. 2). The sensitivity of somatic genes to OSKM corresponds with the disassembly of distal enhancers and is a general characteristic of reprogramming cells, regardless of somatic origin, making it unclear whether this preliminary de-differentiation proceeds through the direct action of OSKM or by some other, indirect mechanism[18,27]. Following general, and reversible, gene suppression, mesenchymal cells activate a previously silent epithelial programme that shares some features with early embryonic cells and their *in vitro* counterparts[25,43]. A crucial step in gastrulation and mesodermal germ layer differentiation, as well as in later developmental stages, is the conversion of polarized, non-motile epithelial cells to a mesenchymal phenotype[44]. This epithelial-to-mesenchymal transition (EMT) is actively imposed by transcriptional regulators such as SNAIL1, SNAIL2, zinc-finger E box-binding homeobox 1 (ZEB1) and ZEB2, which repress promoters of epithelial genes such as *CDH1* (E-cadherin) to support a mesenchymal expression programme[44]. Additionally, signalling pathways such as the transforming growth factor-β (TGFβ) pathway promote EMT, in part by activating transcriptional regulators through SMAD phosphorylation[45]. During reprogramming, the loss of mesenchymal-supporting genes alleviates active repression of the epithelial programme, particularly by OCT4 and SOX2 inhibition of *Snail* transcription, as well as by inducing the expression of *Zeb2*-targeting microRNAs[25,46]. Simultaneously, KLF4 directly induces epithelial genes such as *Cdh1* (REF. 43). The somatic cell response to ectopic OSKM expression cooperates with exogenous growth factors: ectopic MYC downregulates TGFβ receptors, whereas bone morphogenetic proteins (BMPs) promote early induction of mesenchymal gene-suppressing microRNAs[25,43]. Inhibiting TGFβ signalling enhances reprogramming from mesenchymal cell types and can substitute for the role of MYC and SOX2 in suppressing EMT-supporting factors[47,48]. The early induction of epithelial genes therefore reflects diminishing repression by a continuously utilized genetic programme that is enforced through growth factors. Accordingly, epithelialization represents an inherently accessible potential that differs from the activation of the true pluripotency network. In keeping with this idea, many epithelial genes, including several low-stringency markers that are expressed in, but not specific to, pluripotent cells can be induced independently and with far higher efficiency during early reprogramming[20,25,43,49].

Induction and consolidation of pluripotency. Pluripotency itself is stabilized late in the reprogramming process and leads to independence from the expression of ectopic factors[50] (FIGS 1,2). Accumulating evidence suggests that this final transition occurs in a switch-like manner following activation of a few key genes, which subsequently re-establish a complete, self-sustaining regulatory network. Transcriptional analysis of single reprogramming cells divides the process into an initial, ectopic OSKM-dependent stochastic period, followed by a more deterministic phase established upon endogenous *Sox2* activation[51]. Subsequently, the stepwise re-establishment of the pluripotency network proceeds through the induction of a minimal set of genes and requires simultaneous silencing of ectopic OSKM expression[52]. The endogenous activation of this gene subset may therefore represent the theorized rate-limiting step that must be overcome in order for reprogramming cells to transition fully into a stably self-renewing pluripotent state[53].

Although the activation of key effectors may qualify as a determining event in the generation of iPSCs, it is insufficient to qualify cells as functionally pluripotent. Pluripotent cells not only self-renew but must also maintain the ability to respond to multiple developmental cues and generate all organismal cell types, as well as sustain an epigenetic memory of this potential as they divide. Establishing functional pluripotency appears to require additional downstream events beyond the primary induction of target genes, including molecular features that cannot be evaluated by transcriptional output alone[54–56]. These include the erasure of somatic DNA methylation signatures[18], activation of the silent X chromosome in female cells[57], and the re-establishment of bivalent histone modifications at developmental genes[18,58] (FIG. 2). Notably, the re-establishment of epigenetic modifications associated with pluripotency requires additional cell divisions in the absence of ectopic OSKM expression, such that alternative differentiation states may be acquired by culturing reprogramming cells in the presence of specific exogenous growth factors[59–63].

Molecular features of pluripotency

Successful reprogramming culminates in the establishment of stable, self-renewing and functionally pluripotent stem cell lines[58,64]. Although pluripotency exists only transiently during early embryonic development *in vivo*, the derivation of PSC lines that can be stably propagated *in vitro* has provided a powerful model for many developmental processes, including the mechanisms regulating downstream lineage restriction, commitment and eventual maintenance of a terminally differentiated state[65]. The crucial transcriptional regulators of pluripotency, OCT4 and NANOG, in cooperation with SOX2, were identified through genetic studies demonstrating their role in embryonic development and pluripotency maintenance both *in vivo* and *in vitro*[6,66–68]. Subsequent genome-wide localization and interaction analyses revealed that OCT4, SOX2 and NANOG bind to and regulate pluripotency-specific genes, often co-occupying the same target loci to form a regulatory circuit consisting of both feedforward and autoregulatory loops that maintain their own expression, as well as that of other key genes[50]. Unlike somatic cell states, this network must also maintain an extensive and unbiased differentiation potential but remain sensitive enough to integrate opposing differentiation cues efficiently and robustly. Active suppression of lineage specification is a key feature of pluripotency that is only partially controlled through the direct action of OCT4, SOX2 and NANOG, or their immediate effectors[50]. The consolidation of additional

sequence-specific transcription factors, signalling pathways and chromatin modifiers during the final stages of direct reprogramming can be carefully parsed into stepwise molecular pathways to understand the unique developmental potential of pluripotent cells and how it is restored.

Bivalence of developmental genes. Pluripotent cells are unique in that they must suppress multiple developmental pathways comprising thousands of genes while preserving their responsiveness to specific differentiation cues. Canonically, this dual regulation converges on the opposing functions of repressive Polycomb group (PcG) and transcription-associated Trithorax group (TrxG) proteins at CpG island-containing promoters, which establish chromatin with bivalent H3K27me3 and H3K4me3 modifications, respectively[69–72]. Within pluripotent cells, bivalent domains are prevalent at developmental gene promoters and provide a molecular analogy for cellular potential, as most subsequently resolve to either an expressed, TrxG-regulated or a repressed, PcG-regulated state, according to developmental trajectory[72–74]. Bivalent chromatin is functionally relevant to proper development: interfering with the machinery that maintains these dual modifications often results in aberrant or impaired differentiation[75]. In the G1 phase of the cell cycle, sensitivity to extracellular developmental cues favours imbalanced enrichment of H3K4me3 and, eventually, induction of gene expression[76]. In PSCs, many distal enhancers exist in a 'poised' state (marked by H3K4me1 or H3K4me2) and interact with cognate developmental gene promoters, subsequently acquiring repressive H3K27me3 owing to interactions with PcG at the CpG island[77,78]. When triggered to differentiate, p300 acetylates H3K27 (H3K27ac), which stabilizes Pol II and H3K4me3 at the promoter, destabilizes PcG, and activates the gene[77].

Bivalent signatures do not appear to be re-established at developmental genes until late in the reprogramming process[18], an observation consistent with the idea that they serve as molecular markers for functional pluripotency[79]. However, many lineage-specifying genes and those with dual roles in development and pluripotency gradually accumulate H3K4 methylation directly at their CpG island-containing promoters during reprogramming, resulting in local epigenetic remodelling at previously repressed, H3K27me3-only loci[27,56]. Evidence suggests that preliminary local remodelling may be carried out by binding of reprogramming factors to hypomethylated, distal *cis*-regulatory sequences, followed by H3K4me3 deposition at corresponding, constitutively unmethylated promoters, seemingly without Pol II recruitment or gene expression[27,80]. Local and reciprocal depletion of H3K27me3 is carried out by UTX, suggesting that even terminally bivalent genes may require transient activation to restore their inductive potential[27,33]. The assembly of bivalent domains during reprogramming remains incompletely understood, but it represents a valuable assay to molecularly characterize the establishment of unrestricted developmental potential that defines pluripotent cells.

Recent studies have added multiple tiers of regulatory parameters to bivalent domains beyond their original characterization as dually H3K4 and H3K27 methylated chromatin. PSCs cultured with two small-molecule kinase inhibitors (2i) and leukaemia inhibitory factor (LIF) are broadly depleted of H3K27me3 at bivalent promoters without affecting the repression or induction potential of the corresponding gene[81]. In these precise culture conditions, H3K27me3 appears to diffuse over the genome, possibly to compensate for global DNA hypomethylation (see below)[81,82]. Embryonic stem cells (ES cells) lacking the Polycomb repressive complex 2 (PRC2) factor EED also do not show impaired repression of developmental genes in the 2i/LIF condition, but they do exhibit spontaneous differentiation and mis-expression of early lineage-specifying factors in less-defined culture conditions or in ES cells derived from the post-implantation epiblast[83–85]. It now seems that the presence of dual modifications themselves may not sufficiently define functional bivalence (the ability to stably maintain multiple developmental pathways in a repressed but labile state) and may be more specific to later developmental stages of pluripotency that are primed for differentiation[85–87] (FIG. 3a,b). Alternatively, H3K4me3 and H3K27me3 may primarily serve as indicators of a larger regulatory module of poised repression with specific functions in the terminal stages of early lineage commitment[88–92] (see Supplementary information S3 (box)).

The prevalence of unmethylated CpG islands at poised promoters suggests that this feature may serve as the essential *cis*-regulatory template to instruct assembly of a bivalent epigenetic state[93,94]. Indeed, the PcG and TrxG complexes are recruited to CpG island-containing promoters of developmental genes by subunits that recognize unmethylated CpGs[94,95]. In the canonical model of PcG-based repression, PRC2 deposits H3K27me3, which functions as an epigenetic template for chromobox (CBX)-containing PRC1 complexes to trigger H2AK119 monoubiquitylation (H2AK119ub) and chromatin compaction[96]. However, H2AK119ub may also have a role in recruiting PRC2 and initiating *de novo* silencing, providing at least one alternative pathway of PcG-based repression that does not depend on upstream H3K27me3 deposition. In this model, non-canonical PRC1 complexes directly bind to unmethylated CpG sequences through the CXXC domain-containing protein Lys-specific demethylase 2B (KDM2B) to establish bivalency before PRC2 recruitment and activity[97–100] (FIG. 3b). How these mechanisms are reassembled during reprogramming, and in what order, is yet to be explored systematically.

Self-renewal in the absence of epigenetic repressors. The pluripotent state is also distinguished by the robust maintenance of its transcriptional network and self-renewal capacity in the absence of repressive chromatin modifications, whereas early post-implantation embryos or non-pluripotent cell types are inviable[79,101] (FIG. 3c). This unusual insensitivity is unique to one of two developmentally distinct classes of pluripotent states

---

p300
Co-activator protein with histone acetyltransferase activity, which associates with transcription factor-occupied enhancers that are actively engaged in promoting gene transcription.

Non-canonical PRC1 complexes
Whereas canonical Polycomb repressive complex 1 (PRC1) contains a chromobox subunit that recognizes PRC2-deposited epigenetic modifications, non-canonical PRC1 complexes are recruited to chromatin by cofactors such as Lys-specific demethylase 2B (KDM2B), which targets unmethylated CpG islands.

CXXC domain
Cysteine-rich zinc-finger domain found in numerous chromatin-modifying complexes that preferentially binds to unmethylated CpG-rich sequences such as CpG islands.

Naive or ground state cells
Pluripotent stem cells with properties of the inner cell mass or early epiblast, directed by culturing in media supplemented with leukaemia inhibitory factor and two kinase inhibitors (2i/LIF) that suppress fibroblast growth factor signalling and support WNT signalling.

Primed state cells
Pluripotent stem cells that require fibroblast growth factor and activin signalling for continuous self-renewal. Associated with non-murine (including human) embryonic stem cells, with phenotypic and molecular features of the post-implantation epiblast.
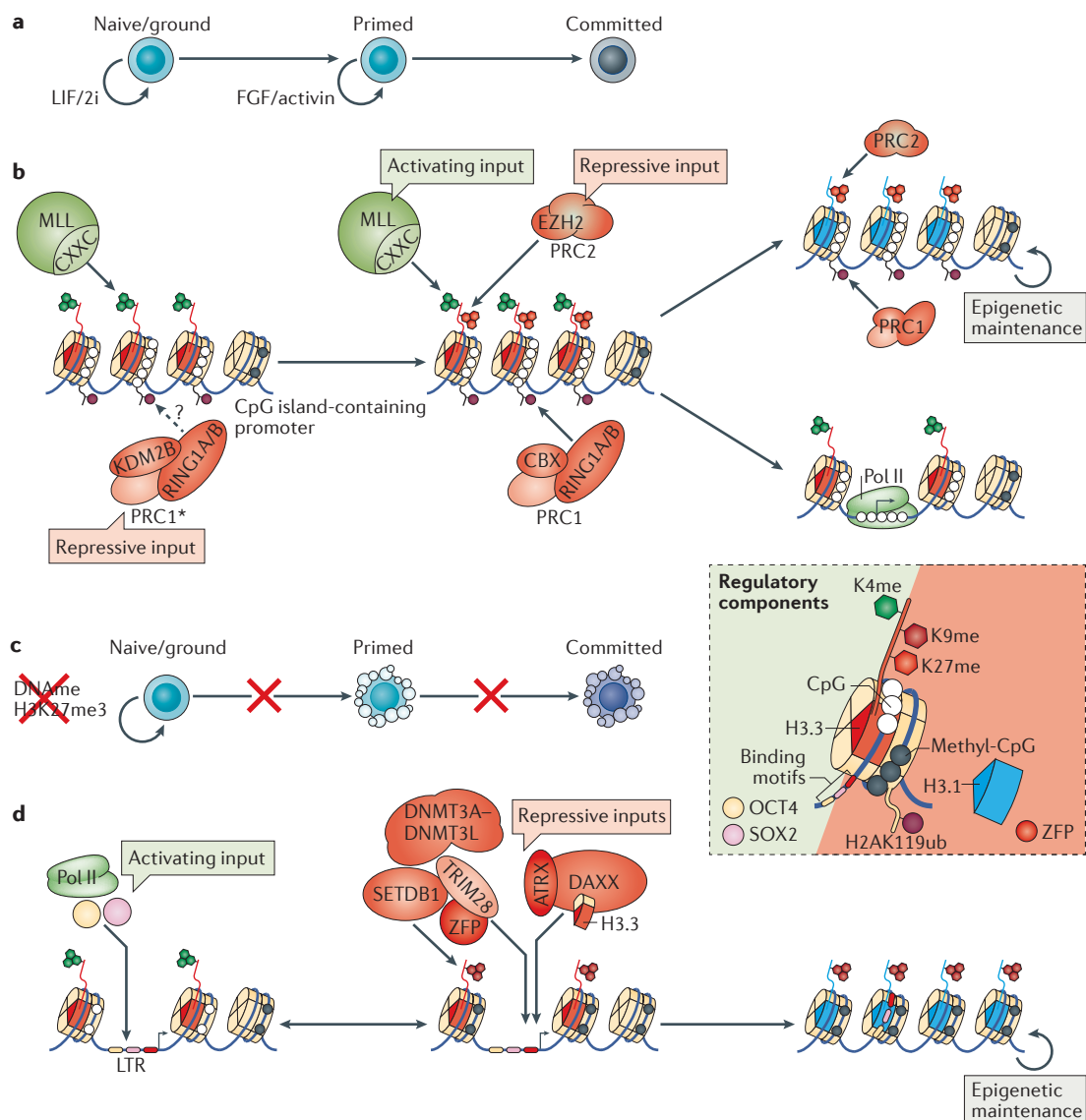
Figure 3 | **Molecular features of pluripotent cells. a** | Pluripotent cells include several distinguishable molecular states: naive or ground state cells propagated in 2i/LIF (two small-molecule kinase inhibitors and leukaemia inhibitory factor) share features with the inner cell mass or early epiblast, whereas primed state cells cultured in fibroblast growth factor (FGF) and activin resemble the late epiblast. Both states are functionally pluripotent, but they can be distinguished at the molecular level by multiple criteria and by their supporting growth conditions. **b** |Pluripotent cells exhibit unbiased developmental potential, which is epigenetically encoded by balancing activating and repressive inputs at developmental gene promoters. Mixed lineage leukaemia (MLL) complexes (which are the mammalian orthologues of Trithorax) and Polycomb repressive complex 2 (PRC2) deposit histone 3 Lys4 trimethylation (H3K4me3) and H3K27me3, respectively, to hold promoters in a facultative state (middle). H3K27me3 recruits canonical PRC1 complexes to monoubiquitylate H2AK119 (H2AK119ub). As pluripotent cells differentiate, such bivalent domains are modified to favour either the H3K4me3-only or H3K27me3-only states (right). H3K27me3 is more diffusely distributed over the genome in naive state cells (not shown). However, bivalent domains can be established *de novo* by the direct recruitment of non-canonical PRC1 complexes (PRC1*) to unmethylated CpG islands by the CXXC domain-containing subunit Lys-specific demethylase 2B (KDM2B), which may maintain local repression under these conditions (left). **c** | Naive cells can be distinguished from primed cells by their capacity for stable self-renewal in the absence of genome-wide epigenetic repressors, such as DNA or H3K27 methylation, although differentiation is then impeded. When these repressors are removed, non-naive cells acutely die. **d** | Many species-specific endogenous retroviruses (ERVs) are dynamically regulated in pluripotent cells. Left: some retroelements are expressed through the binding of OCT4 and SRY-box 2 (SOX2) to *cis*-regulatory motifs within their long terminal repeat (LTR) promoter. Middle: sequence-specific Krüppel-associated box (KRAB) zinc-finger proteins (ZFPs) interact with tripartite motif-containing protein 28 (TRIM28) and SETDB1 to establish repressive H3K9me3, which is subsequently recognized by the H3.3 chaperone complex ATRX–DAXX to support silencing. Many ERVs exhibit dynamic DNA methylation carried out by the *de novo* DNA methyltransferase DNMT3A and by the non-catalytic cofactor DNMT3L. Sequence-specific *de novo* ERV repression is largely specific to the pluripotent state. Right: following lineage commitment, maintenance of gene silencing is carried out epigenetically. Pol II, RNA polymerase II.

that, until recently, had only been isolated in mouse[102]. Naive or ground state cells cultured in 2i/LIF conditions broadly resemble cells of the inner cell mass (ICM) or early epiblast and are canonically associated with mouse ES cells. Alternatively, primed state cells more generally exhibit features associated with the post-implantation epiblast and share molecular similarities with human ES cells[103,104]. The unique tolerance of pluripotent naive cells for the loss of epigenetic repressors can be used to select for iPSCs by eliminating cells that have not been successfully reprogrammed, or to facilitate late-stage reprogramming events by establishing pluripotency-like global chromatin features[18,105]. Notably, repression-deficient PSCs do not show substantial changes to their self-sustaining transcriptional network, but their developmental potential is severely affected and hence their functional pluripotency is lost[83,106–111]. When DNA methylation is globally depleted in mouse ES cells, they cannot differentiate into embryonic germ layers but gain the capacity to differentiate into extra-embryonic tissues[112,113]. Similarly, mouse ES cells lacking the histone H3K9 methyltransferase SETDB1 are unstable by virtue of spontaneous extra-embryonic differentiation, which may reflect derepression of endogenous retroviral elements that function as enhancers or promoters for placental genes[114–117]. Despite the widespread and profound epigenetic changes that are induced by loss of epigenetic repressors, the decoupling of functional developmental potential from self-renewal can be reversible; restoring repressors often rescues the ability to differentiate, further highlighting the necessity of epigenetic silencing mechanisms in this process[108,109].

PSCs cultured in 2i/LIF also seem to lose or redistribute repressive epigenetic modifications, specifically global DNA methylation and H3K27me3 (REFS 81,84,118–121), although how these changes affect differentiation is not yet completely understood. Some of the earliest transcriptional responses to 2i/LIF withdrawal include induction of genes encoding epigenetic repressors such as DNA methyltransferase 3b (*Dnmt3b*), *Dnmt3l* and jumonji and AT-rich interaction domain-containing 2 (*Jarid2*), indicating the rapid establishment of a differentiation-competent epigenome[81,84,118–121]. Consistent with the requirement for epigenetic repression in somatic cells, reprogramming in the presence of 2i from the outset is severely limited, whereas switching to 2i conditions at later time points can promote iPSC generation, similar to what is observed following the targeted depletion of global repressors[122]. Culturing reprogramming cells with the glycogen synthase kinase-β (GSK3β) inhibitor component of 2i, and with ascorbic acid, which promotes global DNA demethylation, results in rapid, homogeneous iPSC generation from somatic or factor-dependent intermediate states[123,124]. Intriguingly, human ES cells or mouse epiblast stem cells do not tolerate the global loss of DNA methylation, indicating that there are fundamental regulatory differences between PSCs that correspond to unique developmental periods, despite their similar self-renewal properties and shared expression of many canonical pluripotency regulators[85,87].

Dynamically regulated retrotransposons. In their original reprogramming screen, Takahashi and Yamanaka introduced transcription factors into fibroblasts using retroviral vectors derived from the Moloney murine leukaemia virus (M-MuLV)[1]. This approach specifically exploited a key distinguishing feature between somatic and pluripotent cell states: the targeted silencing of retroviral long terminal repeats (LTRs) in early embryos and ES cells[125] (FIG. 3d; see Supplementary information S4 (box)). With this strategy, emerging iPSC colonies intrinsically switch off transgene expression and propagate indefinitely without further support from the OSKM factors.

The targeted silencing of LTRs in pluripotent cells has only recently been appreciated as reflecting a fundamental principle of genome regulation during mammalian development[126]. Cumulatively, approximately 40% of the mouse and human genomes are of retrotransposon origin, primarily derived from long and short interspersed nuclear elements (LINEs and SINEs, respectively) or LTR-containing endogenous retroviruses (ERVs)[127,128]. Although each class exhibits unique modes of retrotransposition and has evolved exceedingly divergent, species-specific elements, their regulation seems to be largely conserved[129]. Generally, primary silencing is initiated by sequence-specific, Krüppel-associated box domain-containing zinc-finger proteins (KRAB-ZFPs), which interface with SETDB1 to direct repression[130]. Interestingly, KRAB-ZFPs evolve continuously to counteract the emergence of new retro-elements and represent a general mechanism for genomic surveillance[117,130–133]. Downstream of primary targeting, numerous germline-associated genes expressed in mouse ES cells establish epigenetically heritable repressive states that can be maintained in differentiated cells, where sequence-specific regulators may be absent[113].

Although the pluripotent state is usually associated with the active repression of retrotransposons, their expression is surprisingly dynamic and involves interactions between activating and repressive inputs that are not restored during reprogramming until after iPSCs emerge[134,135]. In both mouse and human, *OCT4*, *SOX2* or *NANOG cis*-regulatory sequences encoded within the promoters of species-specific ERVs restrict their activity to early embryonic states, which supports their progressive radiation to new genomic positions through the host germ line[136,137]. As new integrations arise, these elements can be subsequently co-opted to function as enhancers or alternative promoters for numerous genes, including a majority of state-specific non-coding RNAs that can be essential to pluripotency and experimentally included to improve reprogramming efficiency[135,138–140]. DNA is only heterogeneously methylated at loci of many repetitive element classes, and this methylation depends on the continuous activity of *de novo* DNMTs[113]. In mouse ES cells, specific classes of ERVs are expressed within a small population of cells that also exhibit low global levels of repressive chromatin modifications and demonstrate an expanded extra-embryonic potential when injected into pre-implantation embryos[141]. Notably, this extra-embryonic-contributing, retrotransposon-expressed

Endogenous retroviruses
(ERVs). Genomic retrotransposons originating from exogenous retroviruses, but which propagate intracellularly within germline-competent cell states to enable inheritance to the subsequent generation.

Krüppel-associated box domain-containing zinc-finger proteins
(KRAB-ZFPs). Zinc-finger proteins that contain an amino-terminal KRAB domain, which interfaces with the tripartite motif-containing protein 28 (TRIM28)–SETDB1 complex to direct repressive histone H3 Lys9 trimethylation, and a variable number of rapidly evolving zinc-finger domains that can confer sequence specificity to emerging repetitive elements.

state and the embryonic-restricted, retrotransposon-repressed state are reversible, similar to the dynamic cellular heterogeneity of many pluripotency-associated transcription factors[142–146]. This reversibility appears to reflect the fluidity of locus-specific chromatin architecture in pluripotent cells, as ERV induction corresponds to higher global levels of DNA replication-independent histone H3 variant 3.3 (H3.3)-containing chromatin[147]. Reciprocally, ERV repression is, in part, directed though the H3.3-specific ATRX–DAXX chaperone complexes (FIG. 3d), indicating that DNA replication-independent histone exchange and retrotransposon expression may be required to initiate silencing[148–150]. Thus, the dynamic regulation of genomic repetitive elements represents another specific feature of the pluripotent state that has a direct impact on the developmental potential of each cell within the population.

Developmental barriers to reprogramming

Differentiation involves inactivation of the pluripotency transcriptional network and activation of lineage-specifying transcriptional programmes. As cells proceed through this process, they become responsive to extracellular cues to specify early embryonic lineages. As part of this transition, distally located, poised enhancers are activated by cofactor-directed recruitment of the histone acetyltransferase p300 to instruct the resolution of bivalent promoters[77,78]. Simultaneously, promoters and enhancers of the pluripotency state are shut down and heterochromatinized[151,152]. Unlike bivalent, differentiation-associated genes, which are held in an activation-poised state, many pluripotency-associated genes have different promoter features and utilize distinct silencing mechanisms that affect their reactivation potential during reprogramming[79].

The removal of exogenous pluripotency-supporting culture conditions triggers the downregulation of key factors and destabilizes the self-sustaining pluripotency network, thereby initiating the major transition through which cells become committed to differentiate[153]. Most bivalent promoters resolve to either active H3K4me3-only or repressive H3K27me3-only states, according to their specific regulatory programme[72]. Alternatively, stable shutdown of the pluripotency network is ensured by silencing core regulators such as *Oct4* and *Nanog* through targeted H3K9 methylation followed by DNA methylation[113,151,154]. The more permanent modes of epigenetic silencing at specific pluripotency gene promoters, including many germline genes, may guard against their accidental re-activation in somatic tissues, an event with potentially oncogenic consequences[155]. Alternatively, genes such as *Sox2* and *Klf4*, which may remain expressed in a lineage-dependent manner or be re-induced in later developmental programmes, are instead silenced by deposition of H3K27me3 at their CpG island-containing promoters[8,156,157]. DNA methylation is highly dynamic during differentiation in mouse and human ES cells, in particular during the first transition from pluripotency to a lineage-committed state[73,74,158]. Early commitment is also accompanied by a global shift in nuclear organization, including focal accumulation of heterochromatic

protein 1 (HP1) and H3K9me3-modified heterochromatin, leading to extensive chromatin compaction and a reduction in nucleosome turnover[159]. Substantial changes in metabolic programmes and cell cycle regulation are also coordinated during this transition, resulting in a prolonged G1 phase[42].

Dynamic regulation and disassembly of active enhancers. In PSCs, the activity of OCT4, SOX2 and NANOG is concentrated at super-enhancers that interface with target genes through the Mediator and cohesin complexes to control their expression[70,160,161] (FIG. 4a). The super-enhancer architecture is acutely sensitive to cellular state, and disruption of Mediator activity or downregulation of master pluripotency factors leads to rapid downregulation of associated genes[161]. Transcriptionally permissive chromatin at pluripotency-associated enhancers is progressively disassembled in a stepwise manner, beginning with the removal of histone modifications, followed by encroachment of nucleosomes, and culminating in the methylation of repression-associated histone residues and of DNA[152,162] (FIG. 4b,c).

The sensitivity of pluripotency enhancers to disassembly seems to reflect opposing inputs between activating, transcription factor-guided recruitment of histone acetyltransferases and the repressive activity of the nucleosome remodelling and deacetylase (NuRD) complex[163–165]. Diminished OCT4 binding in the earliest stages of differentiation favours decommissioning of pluripotency-associated enhancers through the activity of the NuRD subunit LSD1, an H3K4 demethylase[162]. The NuRD complex has a central role in the removal of permissive chromatin marks and in local chromatin remodelling, which temporally precedes the establishment of heterochromatic modifications[166,167]. Methyl CpG-binding domain protein 3 (MBD3) seems to be an essential component for NuRD assembly and recruitment[162,168]. Following the loss of activating epigenetic modifications, ATP-dependent chromatin remodellers — either NuRD or members of the BRG1-associated factor (BAF) chromatin remodelling complex (which is the mammalian equivalent of SWI/SNF) — change the architecture of the enhancer to occlude previously nucleosome-free DNA into chromatin as a template for repressive modifications, such as methylation of H3K9 and DNA[152,169].

When initially introduced into somatic cells, OSKM seem to be insufficient to re-establish the balance between the activity of repressive and permissive chromatin modifiers that regulate target loci in pluripotent cells. The ability of regulatory elements to support the transcriptional activity of pluripotency genes is initiated by OSK, which can find and engage select sequences in closed chromatin (see below). However, epigenetic repressors that are expressed in both somatic and pluripotent cells, specifically NuRD, may be inadvertently recruited and impede transcriptional activation of target loci during early reprogramming[170]. Indeed, perturbing the NuRD complex can substantially accelerate the kinetics and improve the efficiency of iPSC generation, presumably by eliminating counterproductive and premature recruitment of this repressive input at a stage when opposing activators


Histone H3 variant
Histone H3 variants include H3.1 and H3.2, which are typically incorporated into chromatin during DNA replication, as well as H3.3, which is directed to loci in a replication-independent manner by specific histone chaperones.

Super-enhancers
The topological coordination of multiple, spatially discrete enhancers to direct the expression of a gene through the Mediator complex. Often, super-enhancers are necessary for the expression of essential cell type-specific genes.

Mediator
A large, multi-subunit complex that interacts with and spatially juxtaposes transcription factors at promoters and enhancers to coordinate transcription.

---
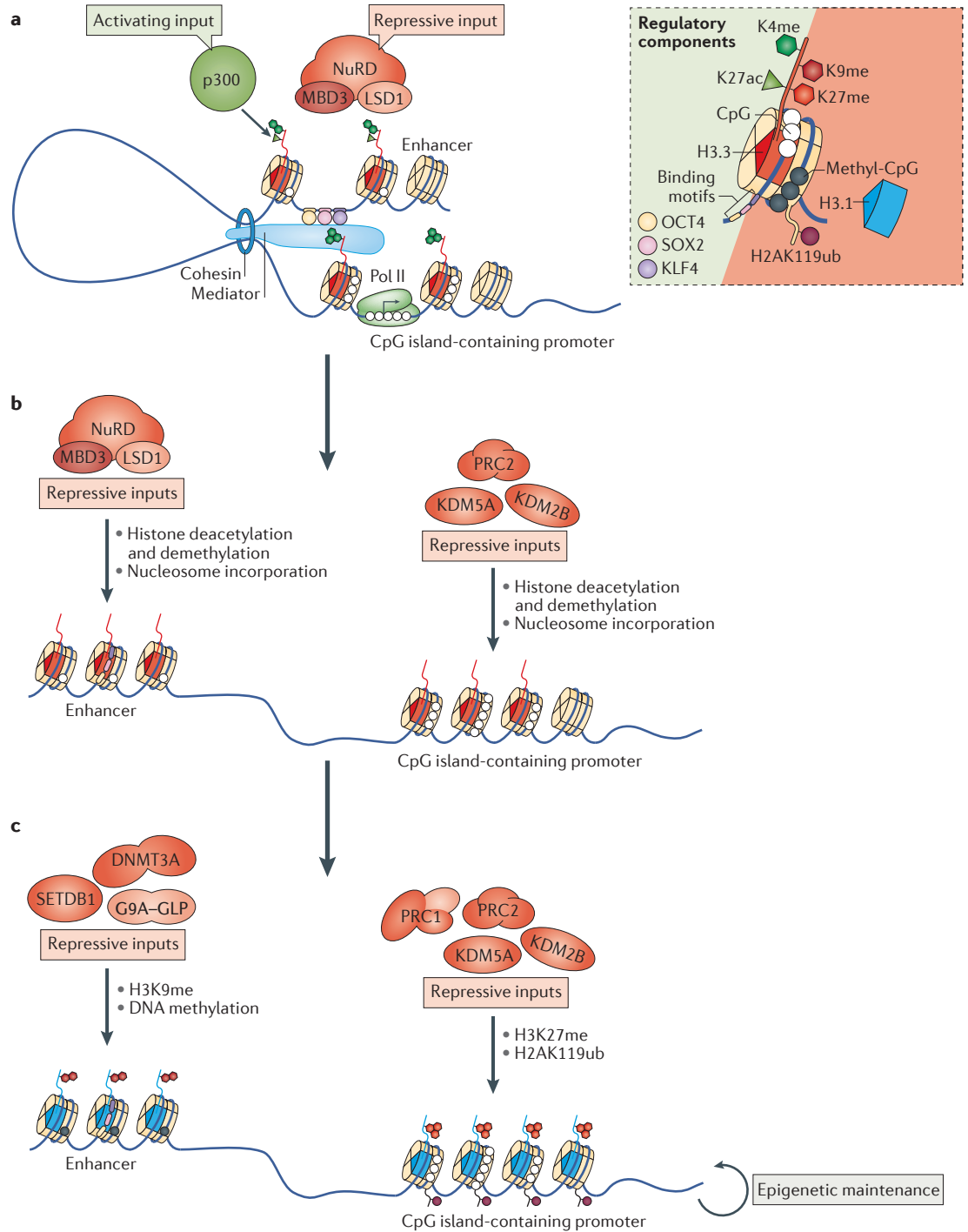
remain absent[170–173]. Once pluripotency is acquired, additional essential cofactors may override this repression to support the assembly of active enhancers and stably maintain target gene expression.

Transcription factor binding at inert chromatin. In stably propagating cell types, most transcription factors bind within open chromatin, indicated by the presence of a nucleosome-free region surrounded by highly dynamic, phased nucleosomes[174–176]. However, pioneer factors can directly bind to their cognate DNA motifs, even in

compact chromatin, to evict nucleosomes and initiate enhancer activation, whereas additional cofactors may be required to induce transcription following primary locus engagement[177]. During normal developmental transitions, multiple DNA-binding factors, including those with pioneer activities, are expressed in a coordinated fashion, often leading to near-simultaneous binding, chromatin modification and transcriptional changes that complicate the molecular assignment of each contributing factor into a clear, linear pathway[178]. Alternatively, direct reprogramming introduces a minimal set of

Pioneer factors
Sequence-specific DNA-binding factors that can engage compact chromatin to initiate the formation of nucleosome-free regions.

◀ Figure 4 | **Differentiation establishes epigenetic barriers to reprogramming.**
a | Within pluripotent cells, multiple enhancers bound by OCT4, SRY-box 2 (SOX2) and NANOG form a super-enhancer that engages a target promoter and is held in this topological configuration by the Mediator and cohesin complexes. Transcription factor binding preserves a nucleosome-free region, with adjacent phased nucleosomes acetylated by p300. In the pluripotent state, these enhancers also recruit the repressive nucleosome remodelling and deacetylase (NuRD) complex, which contains the histone H3 Lys4 (H3K4) demethylase LSD1, the core component methyl CpG-binding domain protein 3 (MBD3) and histone deacetylases. This balance of activating and repressive inputs favours gene activation and preserves the expression of the target gene. **b** | Removal of exogenous pluripotency-supporting conditions leads to the downregulation of OCT4, SOX2 and NANOG and triggers super-enhancer disassembly. The loss of transcription factor binding strongly favours the activity of repressive inputs. At the enhancer, NuRD (or possibly other ATP-dependent chromatin remodellers) establishes nucleosomes at the previous nucleosome-free region and removes histone acetylation and methylation modifications associated with transcriptional activity. Loss of enhancer–promoter association results in a similar disassembly of promoter-associated activating modifications, such as H3K4 trimethylation (H3K4me3) and H3K36me3, by Polycomb repressive complex 2 (PRC2)-associated histone demethylases KDM5A and KDM2B (see also Supplementary information S3 (box)). **c** | Repressive epigenetic modifications are established following the erasure of the previous chromatin state. At enhancers, these repressive modifications include H3K9me2 and possibly H3K9me3 deposition by G9A in complex with G9A-like protein (GLP) and SETDB1, respectively, as well as DNA methylation by DNA methyltransferase 3A (DNMT3A). At promoters, PRC2 deposits H3K27me3, which recruits the canonical PRC1 complex to monoubiquitylate H2AK119 (H2AK119ub) and initiate chromatin compaction. These new modifications then serve as epigenetic templates to preserve repression during subsequent cell divisions. The schematic depicts silencing pathways both at a distal regulatory element and in the associated CpG island-containing promoter of a pluripotency gene, such as *Sox2* or Krüppel-like factor 2 (*Klf2*). Many essential pluripotency genes, such as *Oct4* and *Nanog*, do not have CpG island-containing promoters and, in these cases, the mechanisms of promoter silencing largely resemble those observed for low-CpG density, distal regulatory elements. K27ac, Lys27 acetylation; Pol II, RNA polymerase II.

transcription factors into a nuclear environment in which the majority of pluripotent state-specific enhancers are chromatinized and their target loci repressed, which allows intermediate molecular states of locus activation to be isolated and characterized[26–28,56,173].

Ectopically expressed OSKM must engage a somatic genome in which the majority of their target enhancers are epigenetically silenced (FIG. 5a). Consequently, OSKM initially only bind to a minimal number of pluripotency-associated targets, and they seem to do so cooperatively at nucleosomal DNA that lacks obvious histone modifications and contains recognition motifs for OCT4, SOX2 and KLF4 (REFS 28,179). Similar loci containing repressive modifications seem to be intransigent to factor binding[179]. Canonical pioneer factors like Forkhead box protein A2 (FOXA2; also known as HNF3β) identify their binding motifs within chromatin and possess a DNA-binding domain with structural similarity to linker histones that outcompetes nucleosomes to initiate a region of open chromatin[177]. Although OSKM lack similar structural features, recent evidence indicates that OCT4, SOX2 and KLF4 co-bind to shared somatic targets through a cooperative, pioneer-like activity involving combinatorial binding to outwardly facing partial motif sequences within the nucleosome[180] (FIG. 5b). Binding initiates preliminary chromatin remodelling through the deposition of H3K4me1 and H3K4me2 and, through additional steps that are not yet clear, eventual induction of target genes[179,180].

The majority of reprogramming-related regulatory regions that are targeted by OCT4, SOX2 and KLF4 in PSCs remain unbound during most of the reprogramming process[28,179]. It remains to be seen whether the few loci at which cooperative binding is observed represent a sufficient subset to permit entry into the pluripotent state upon gene induction, or how enhancers that do not contain an appropriate configuration of motifs are nevertheless activated. A cooperative pioneering activity could explain notable improvements to reprogramming efficiencies when additional pluripotency-associated transcription factors (such as NANOG, SAL-like protein 4 (SALL4), oestrogen-related receptor-β (ESRRB), nuclear receptor subfamily 5 group A member 2 (NR5A2) or zinc-finger protein GLIS1) are ectopically expressed, which may broaden the number of *cis*-regulatory elements that can be accessed in somatic cells[53,181–185]. Additionally, this model may explain why combinations of different members of the POU, SOX or KLF transcription factor families can adequately replace their canonical member, as some share highly similar recognition motifs[180,186,187].

Initial binding by OSK within somatic cells appears to be population-wide and represents only the earliest step in re-activating the endogenous pluripotency network. Moreover, the extended latency between binding and the induction of cognate genes indicates that these preliminary interactions are themselves insufficient. Transcription factor-bound loci can recruit chromatin remodellers such as cell-type specific BAF complexes that phase nucleosomes around the site of transcription factor binding to stabilize a nucleosome-depleted region[169,177,188]. The constitution of these complexes varies considerably among cell types, and overexpressing components of the ES cell-specific BAF (esBAF) complex during reprogramming enhances iPSC generation[189–192]. There is also some evidence that this primary genomic engagement by OSK may still depend on underlying chromatin status, as OCT4 binding in somatic cells occurs preferentially at distal *cis*-regulatory sequences that lack DNA methylation but are nevertheless nucleosomal[152]. The constrained contexts that dictate nucleosomal OSK binding in differentiated cells resemble priming by 'fragile nucleosomes' at *cis*-regulatory elements in yeast, which can be rapidly displaced by transcription factors to direct immediate transcriptional responses[193]. Similarly, suppressing chromatin assembly factor 1 (CAF1), which is an H3.1-specific, DNA replication-dependent chaperone complex, improves the efficiency and kinetics of iPSC generation[194]. Presumably, the diminishing presence of H3.1 within heterochromatin expands the number of *cis*-regulatory sequences that are occluded within more labile H3.3-containing nucleosomes, which must be nonspecifically incorporated as compensation. Following binding, OCT4 initiates chromatin modification at the enhancer and seems to interact with corresponding CpG island-containing promoters to direct targeted H3K4 methylation and H3K27 demethylation[27,33,80,195] (FIG. 5b). The exact nature of these interactions is still unclear, but they may correspond to a key intermediate state of enhancer assembly within reprogramming populations that precedes transcriptional activation in the few cells that do generate iPSCs[27,179,196].

These preliminary steps in enhancer activation must be followed by the recruitment of co-regulators, including histone methyltransferases and acetyltransferases, the assembly of super-enhancers and the Mediator-dependent topological juxtaposition of enhancers and promoters, although not necessarily in that order (FIG. 5c). Many of these downstream events may require support from DNA-binding factors that are not present until late in the reprogramming process[49]. Eventually, a full complement of factors must recruit and assemble Pol II pre-initiation complexes and proceed to the stabilization of transcription elongation. Much of the transcriptional reprogramming that is required to consolidate pluripotency post-induction remains to
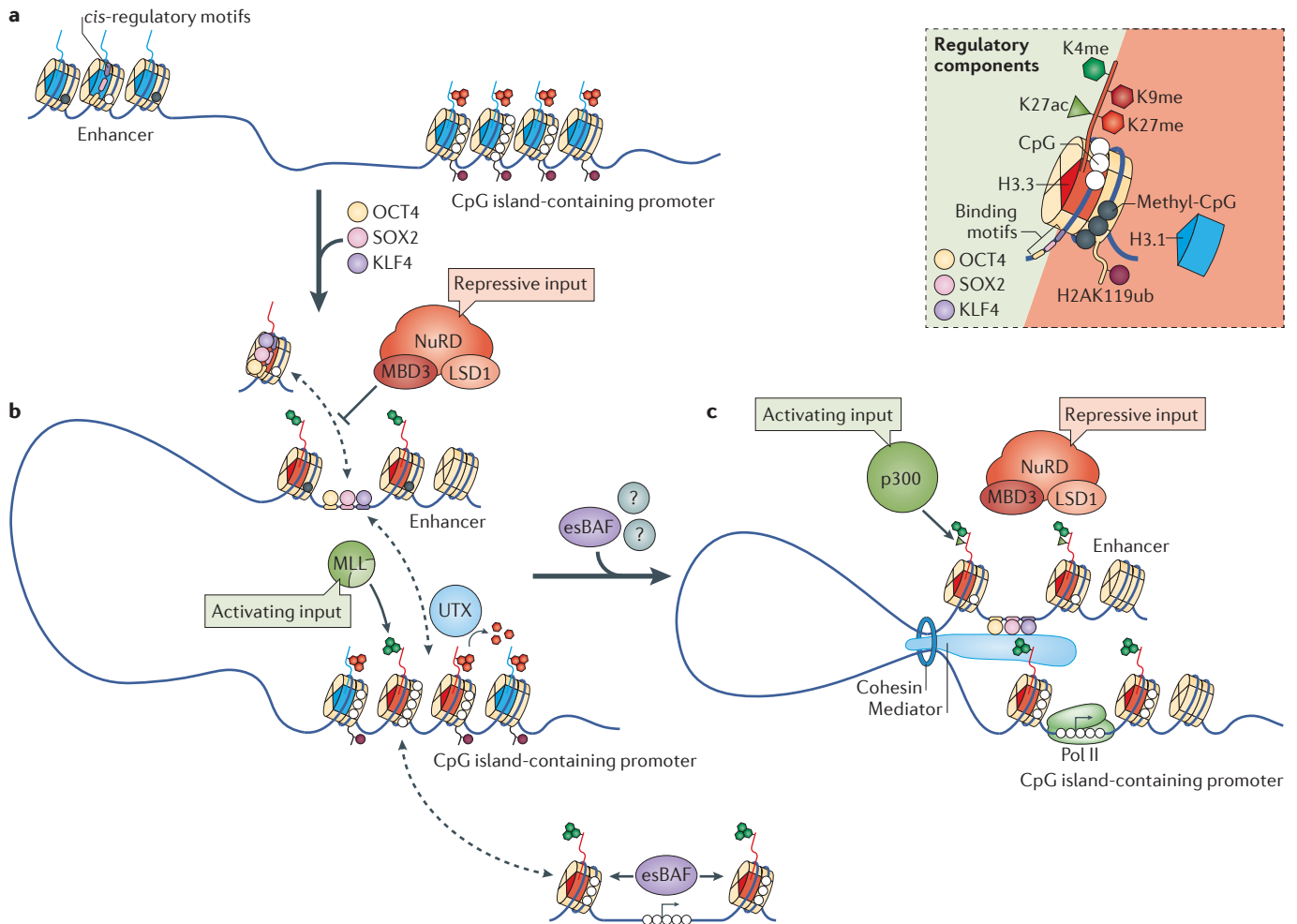
Figure 5 | **Transcriptional activation of silent pluripotency genes.** **a** | A schematic example of a CpG island-containing, pluripotency-associated gene as it is activated by a cognate distal regulatory element (enhancer) that includes recognition motifs for OCT4, SRY-box 2 (SOX2) and Krüppel-like factor 4 (KLF4), collectively known as OSK. In somatic cells, CpG island-containing promoters of some pluripotency genes are epigenetically modified with histone H3 Lys27 trimethylation (H3K27me3) and H2AK119 monoubiquitylation (H2AK119ub). Many distal regulatory elements are highly DNA methylated, whereas others have lower methylation levels but are nonetheless occluded with nucleosomes that lack obvious histone modifications associated with repression. **b** | At select enhancers, OCT4, SOX2 and KLF4 cooperate to bind to partial motif sequences to direct pioneer factor-like activity. The exact kinetics of nucleosome exchange or identity of additional cofactors is not fully understood, but these sites do undergo preliminary epigenetic remodelling, including H3K4me1 and H3K4me2 deposition. Initial binding also increases H3K4 methylation at the promoter through interactions with the mixed lineage leukaemia (MLL) component WD repeat-containing protein 5 (WDR5) and co-occurs with local depletion of H3K27me3 by the histone demethylase UTX. During this stage, transcriptional induction of the target gene is not observed, in part because OSK factors

interact with the nucleosome remodelling and deacetylase (NuRD) repressor complex, which suppresses efficient reprogramming in the absence of sufficient activating inputs by inhibiting nucleosome eviction or preventing stable acetylation. The nature of the topological interactions between enhancer and promoter during cell reprogramming, or of the stability of preliminary, OSK-bound nucleosome-free regions, are unknown and are highlighted by dashed bidirectional arrows. **c** | In the final steps of gene activation, additional, unknown factors cooperate with OSK to stably evict nucleosomes and establish a canonical enhancer signature, which includes H3K4me2 and H3K27 acetylation (H3K27ac) and tight topological linkage to the cognate promoter through the Mediator and cohesin complexes. Nucleosome eviction may be stabilized by embryonic stem cell-specific BAF (esBAF), which is a pluripotent state-specific version of the ATP-dependent chromatin remodelling BRG1-associated factor (BAF) complex. At this point, activating inputs override repressive inputs, although the NuRD complex remains associated with enhancers in pluripotent cells. Once select enhancers are sufficiently activated and direct the expression of their target genes, the reprogramming process follows a deterministic consolidation. LSD1, Lys-specific demethylase 1; MBD3, methyl CpG-binding domain protein 3; Pol II, RNA polymerase II.

be characterized, although current efforts have illuminated a number of notable molecular features that were previously opaque.

Conclusion

The application of genomic technologies to the study of direct reprogramming has raised substantial new considerations that must be taken into account when defining or manipulating cell states. Whereas cellular identity was previously empirically determined according to functional criteria, the ability to modulate it in a controlled and measurable fashion has reframed these phenotypic observations into a precise molecular definition that includes the composition of regulators as they control or constrain specific genetic programmes. Although reprogramming experiments tend to focus on the reversibility of somatic identity, their results reveal the remarkable resistance of cell state to perturbation: substantial barriers are imposed by cooperative interactions between self-sustaining transcriptional networks as they operate within stable epigenetic landscapes. Similarly,

measurable improvements to reprogramming kinetics or efficiency following the modulation of epigenetic features highlight the extent to which chromatin state influences chromatin–transcription factor interactions and target gene expression. Early SCNT experiments revealed the role of epigenetic regulation in determining reprogramming outcomes, but the use of ectopic transcription factors *in vitro* has provided a more dynamic description of the regulators that coordinate the induction of silent genes. Additionally, the imposed changes to developmental potential that occur during the generation of iPSCs can be used to molecularly distinguish the pluripotent state according to multiple parameters and measure them following the endogenous activation of a sufficient set of master regulators. Although complete descriptions of how these additional features are first initiated and subsequently consolidated have yet to be achieved, direct reprogramming has proven to be an extraordinarily tractable and insightful tool for the systematic dissection of developmental potential, differentiation and cellular states.

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
2. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
3. Su, A. I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl Acad. Sci. USA* **99**, 4465–4470 (2002).
4. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
5. Kondoh, H. & Kamachi, Y. SOX-partner code for cell specification: regulatory target selection and underlying molecular mechanisms. *Int. J. Biochem. Cell Biol.* **42**, 391–399 (2010).
6. Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.* **9**, 625–635 (2007).
7. Miyagi, S. *et al.* Consequence of the loss of Sox2 in the developing brain of the mouse. *FEBS Lett.* **582**, 2811–2815 (2008).
8. Sarkar, A. & Hochedlinger, K. The Sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* **12**, 15–30 (2013).
9. Lodato, M. A. *et al.* SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. *PLoS Genet.* **9**, e1003288 (2013).
10. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
11. Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
12. D'Alessio, A. C. *et al.* A systematic approach to identify candidate transcription factors that control cell identity. *Stem Cell Rep.* **5**, 763–775 (2015).
13. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
14. Arvey, A., Agius, P., Noble, W. S. & Leslie, C. Sequence and chromatin determinants of cell-type-specific transcription factor binding. *Genome Res.* **22**, 1723–1734 (2012).
15. Biggin, M. D. Animal transcription networks as highly connected, quantitative continua. *Dev. Cell* **21**, 611–626 (2011).
16. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
17. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

18. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
19. Stadtfeld, M., Maherali, N., Breault, D. & Hochedlinger, K. Defining molecular cornerstones during early iPS cell reprogramming in mouse. *Cell Stem Cell* **2**, 230–240 (2008).
20. Zunder, E. R., Lujan, E., Goltsev, Y., Wernig, M. & Nolan, G. P. A continuous molecular roadmap to iPSC reprogramming through progression analysis of single-cell mass cytometry. *Cell Stem Cell* **16**, 323–337 (2015).
21. Hansson, J. *et al.* Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep.* **2**, 1579–1592 (2012).
22. Guo, S. *et al.* Nonstochastic reprogramming from a privileged somatic state. *Cell* **156**, 649–662 (2014).
23. Smith, Z. D., Nachman, I., Regev, A. & Meissner, A. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat. Biotechnol.* **28**, 521–526 (2010).
24. Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151–159 (2008).
25. Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
26. Hussein, S. M. *et al.* Genome-wide characterization of the routes to pluripotency. *Nature* **516**, 198–206 (2014).
27. Koche, R. P. *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* **8**, 96–105 (2011).
28. Sridharan, R. *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136**, 364–377 (2009).
29. Lee, D. S. *et al.* An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. *Nat. Commun.* **5**, 5619 (2014).
30. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
31. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
32. Liang, G., He, J. & Zhang, Y. Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nat. Cell Biol.* **14**, 457–466 (2012).
References 31 and 32 identify the H3K79 methyltransferase DOT1L and the H3K36 demethylase KDM2B as participating in the early

disassembly of epigenetic modifications associated with transcription elongation at somatically active genes.
33. Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413 (2012). **This work reveals that the H3K27 demethylase UTX participates in the activation of pluripotency-associated genes by removing differentiation-associated repressive modifications at CpG island-containing promoters.**
34. Ramirez-Carrozzi, V. R. *et al.* A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* **138**, 114–128 (2009).
35. Kim, J. *et al.* A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* **143**, 313–324 (2010). **Kim *et al.* demonstrate that transcriptional signatures shared between ES cells and hyperproliferative cancer cell models reflect a discrete MYC-driven genetic module that is separate from that of OCT4 and SOX2 in pluripotent cells and uniquely present in proliferating, OSKM-dependent intermediates.**
36. Nie, Z. *et al.* c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**, 68–79 (2012).
37. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010). **Rahl *et al.* demonstrate that MYC stabilizes transcription elongation, which explains its contribution to the early phenotypic response to direct reprogramming and its more universal binding to the promoters of cell cycle genes at this stage.**
38. Mathieu, J. *et al.* Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency. *Cell Stem Cell* **14**, 592–605 (2014).
39. Panopoulos, A. D. *et al.* The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. *Cell Res.* **22**, 168–177 (2012).
40. Folmes, C. D. *et al.* Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab.* **14**, 264–271 (2011).
41. Cao, Y. *et al.* miR-290/371-Mbd2-Myc circuit regulates glycolytic metabolism to promote pluripotency. *EMBO J.* **34**, 609–623 (2015).
42. Zhang, J., Nuebel, E., Daley, G. Q., Koehler, C. M. & Teitell, M. A. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. *Cell Stem Cell* **11**, 589–595 (2012).

43. Li, R. *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* **7**, 51–63 (2010).

44. Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).

45. Xu, J., Lamouille, S. & Derynck, R. TGF-β-induced epithelial to mesenchymal transition. *Cell Res.* **19**, 156–172 (2009).

46. Wang, G. *et al.* Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. *Proc. Natl Acad. Sci. USA* **110**, 2858–2863 (2013).

47. Ichida, J. K. *et al.* A small-molecule inhibitor of Tgf-β signaling replaces *Sox2* in reprogramming by inducing *Nanog*. *Cell Stem Cell* **5**, 491–503 (2009).

48. Maherali, N. & Hochedlinger, K. Tgfβ signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Curr. Biol.* **19**, 1718–1723 (2009).

49. Lujan, E. *et al.* Early reprogramming regulators identified by prospective isolation and mass cytometry. *Nature* **521**, 352–356 (2015).
Lujan et al. identified transcriptional activators that are associated with the induction of pluripotency but are not necessary for the maintenance of self-renewal.

50. Jaenisch, R. & Young, R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567–582 (2008).

51. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–1222 (2012).
This study demonstrates that the consolidation of pluripotency proceeds deterministically following the induction of endogenous SOX2 expression.

52. Golipour, A. *et al.* A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* **11**, 769–782 (2012).
This paper demonstrates that the terminal steps of reprogramming require withdrawal of exogenous OSKM factors to consolidate autonomous self-renewal and reactivate the endogenous pluripotency network.

53. Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).
This work demonstrates that all perpetually dividing reprogramming cell lineages will eventually generate iPSCs, supporting a single switch-like transition into deterministic reprogramming.

54. Araki, R. *et al.* Conversion of ancestral fibroblasts to induced pluripotent stem cells. *Stem Cells* **28**, 213–220 (2010).
Araki et al. characterize the morphological, proliferative and molecular properties of iPSC-forming cell lineages from the time of OSKM induction using continuous live-cell imaging.

55. O'Malley, J. *et al.* High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* **499**, 88–91 (2013).

56. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617–1632 (2012).

57. Maherali, N. *et al.* Global epigenetic remodeling in directly reprogrammed fibroblasts. *Cell Stem Cell* **1**, 55–70 (2007).

58. Wernig, M. *et al.* *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).

59. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* **28**, 848–855 (2010).

60. Tonge, P. D. *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* **516**, 192–197 (2014).

61. Pasque, V. *et al.* X chromosome reactivation dynamics reveal stages of reprogramming to pluripotency. *Cell* **159**, 1681–1697 (2014).

62. Maza, I. *et al.* Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. *Nat. Biotechnol.* **33**, 769–774 (2015).

63. Bar-Nur, O. *et al.* Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage. *Nat. Biotechnol.* **33**, 761–768 (2015).

64. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).

65. Posfai, E., Tam, O. H. & Rossant, J. Mechanisms of pluripotency *in vivo* and *in vitro*. *Curr. Top. Dev. Biol.* **107**, 1–37 (2014).

66. 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).

67. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).

68. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).

69. Bernstein, B. E., Meissner, A. & Lander, E. S. The mammalian epigenome. *Cell* **128**, 669–681 (2007).

70. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).

71. Azuara, V. *et al.* Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538 (2006).

72. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

73. Gifford, C. A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).

74. Ziller, M. J. *et al.* Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* **518**, 355–359 (2015).

75. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).

76. Singh, A. M. *et al.* Cell-cycle control of bivalent epigenetic domains regulates the exit from pluripotency. *Stem Cell Rep.* **5**, 323–336 (2015).

77. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).

78. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107**, 21931–21936 (2010).

79. Meissner, A. Epigenetic modifications in pluripotent and differentiated cells. *Nat. Biotechnol.* **28**, 1079–1088 (2010).

80. Taberlay, P. C. *et al.* Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* **147**, 1283–1294 (2011).
Taberlay et al. demonstrate that OCT4 binding to the *MyoD* enhancer results in H3K4 methylation and H3K27 demethylation at the promoter, which is indicative of a transient topological interaction that precedes transcription.

81. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).

82. Brinkman, A. B. *et al.* Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138 (2012).

83. Chamberlain, S. J., Yee, D. & Magnuson, T. Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* **26**, 1496–1505 (2008).

84. 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 (2015).

85. Geula, S. *et al.* m⁶A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **347**, 1002–1006 (2015).
Geula et al. identified a requirement for DNA, H3K27 and mRNA methylation for self-renewal in mouse epiblast stem cells but not for stem cells cultured in 2i/LIF.

86. Chan, Y. S. *et al.* A PRC2-dependent repressive role of PRDM14 in human embryonic stem cells and induced pluripotent stem cell reprogramming. *Stem Cells* **31**, 682–692 (2013).

87. Liao, J. *et al.* Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* **47**, 469–478 (2015).

88. Voigt, P., Tee, W. W. & Reinberg, D. A double take on bivalent promoters. *Genes Dev.* **27**, 1318–1338 (2013).

89. Landeira, D. *et al.* Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA polymerase II to developmental regulators. *Nat. Cell Biol.* **12**, 618–624 (2010).

90. Jia, J. *et al.* Regulation of pluripotency and self-renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell* **151**, 576–589 (2012).

91. Pasini, D. *et al.* Regulation of stem cell differentiation by histone methyltransferases and demethylases. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 253–263 (2008).

92. Joshi, O. *et al.* Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. *Cell Stem Cell* **17**, 748–757 (2015).

93. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).

94. Mendenhall, E. M. *et al.* GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* **6**, e1001244 (2010).

95. Thomson, J. P. *et al.* CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082–1086 (2010).

96. Simon, J. A. & Kingston, R. E. Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10**, 697–708 (2009).

97. He, J. *et al.* Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat. Cell Biol.* **15**, 373–384 (2013).

98. Farcas, A. M. *et al.* KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife* **1**, e00205 (2012).

99. Wu, X., Johansen, J. V. & Helin, K. Fbxl10/Kdm2b recruits Polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. *Mol. Cell* **49**, 1134–1146 (2013).

100. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459 (2014).
This study identified a specific complex, normally recruited to CpG islands by the H3K36 demethylase KDM2B, that is sufficient to recruit PRC2 and establish H3K27me3 downstream of PRC1.

101. Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.* **3**, 662–673 (2002).

102. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* http://dx.doi.org/10.1038/nrm.2015.28 (2016).

103. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).

104. Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).

105. Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–797 (2008).

106. Fouse, S. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**, 1–10 (2008).

107. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).

108. Jackson, M. *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell. Biol.* **24**, 8862–8871 (2004).

109. Leeb, M. *et al.* Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* **24**, 265–276 (2010).

110. Pasini, D. *et al.* Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev.* **22**, 1345–1355 (2008).

111. Dovey, O. M., Foster, C. T. & Cowley, S. M. Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proc. Natl Acad. Sci. USA* **107**, 8242–8247 (2010).

112. Ng, R. K. *et al.* Epigenetic restriction of embryonic cell lineage fate by methylation of *Elf5*. *Nat. Cell Biol.* **10**, 1280–1290 (2008).

113. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).

114. Yuan, P. *et al.* Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev.* **23**, 2507–2520 (2009).

115. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* **23**, 2484–2489 (2009).

116. Fei, Q. *et al.* SETDB1 modulates PRC2 activity at developmental genes independently of H3K9 trimethylation in mouse ES cells. *Genome Res.* **25**, 1325–1335 (2015).

117. Karimi, M. M. *et al.* DNA methylation and SETDB1/ H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**, 676–687 (2011).

118. 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 (2013).

119. Habibi, E. *et al.* Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).

120. Hackett, J. A. *et al.* Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Rep.* **1**, 518–531 (2013).

121. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).

122. Theunissen, T. W. *et al.* Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. *Curr. Biol.* **21**, 65–71 (2011).

123. Tran, K. A. *et al.* Collaborative rewiring of the pluripotency network by chromatin and signalling modulating pathways. *Nat. Commun.* **6**, 6188 (2015).

124. Bar-Nur, O. *et al.* Small molecules facilitate rapid and synchronous iPSC generation. *Nat. Methods* **11**, 1170–1176 (2014).

125. Jahner, D. *et al.* *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298**, 623–628 (1982).

126. Gifford, W. D., Pfaff, S. L. & Macfarlan, T. S. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* **23**, 218–226 (2013).

127. Mouse Genome Sequencing Consortium *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).

128. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

129. Smith, Z. D. *et al.* DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).

130. Rowe, H. M. *et al.* KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010).

131. Thomas, J. H. & Schneider, S. Coevolution of retroelements and tandem zinc finger genes. *Genome Res.* **21**, 1800–1812 (2011).

132. Jacobs, F. M. *et al.* An evolutionary arms race between KRAB zinc-finger genes *ZNF91/93* and SVA/L1 retrotransposons. *Nature* **516**, 242–245 (2014).

133. Castro-Diaz, N. *et al.* Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev.* **28**, 1397–1409 (2014).

134. Friedli, M. *et al.* Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. *Genome Res.* **24**, 1251–1259 (2014).

135. Ohnuki, M. *et al.* Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc. Natl Acad. Sci. USA* **111**, 12426–12431 (2014).

136. Bourque, G. *et al.* Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* **18**, 1752–1762 (2008).

137. Kunarso, G. *et al.* Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634 (2010). **Genomic analysis of OCT4 and NANOG binding in mouse and human ES cells revealed substantial divergence that can be partially explained by the species-specific radiation of repetitive elements with OCT4- and NANOG-binding motifs that function as enhancers.**

138. Loewer, S. *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* **42**, 1113–1117 (2010).

139. Lu, X. *et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* **21**, 423–425 (2014). **References 135 and 139 demonstrate that endogenous retroviruses of the LTR7 subfamily are specifically bound by OCT4, SOX2 and KLF4 in humans, and that their transcription supports generation of iPSCs.**

140. Durruthy-Durruthy, J. *et al.* The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nat. Genet.* **48**, 44–52 (2015).

141. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012). **Macfarlan *et al.* find that specific retrotransposons are dynamically expressed in a small population of mouse ES cells that globally resemble early cleavage stage embryos and have an unusual extra-embryonic potential.**

142. Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K. & Niwa, H. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**, 909–918 (2008).

143. Singh, A. M., Hamazaki, T., Hankowski, K. E. & Terada, N. A heterogeneous expression pattern for Nanog in embryonic stem cells. *Stem Cells* **25**, 2534–2542 (2007).

144. Singer, Z. S. *et al.* Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Mol. Cell* **55**, 319–331 (2014).

145. Hayashi, K., Lopes, S. M., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).

146. Papatsenko, D. *et al.* Single-cell analyses of ESCs reveal alternative pluripotent cell states and molecular mechanisms that control self-renewal. *Stem Cell Rep.* **5**, 207–220 (2015).

147. Ishiuchi, T. *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**, 662–671 (2015).

148. Elsasser, S. J., Noh, K. M., Diaz, N., Allis, C. D. & Banaszynski, L. A. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**, 240–244 (2015).

149. Sadic, D. *et al.* Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep.* **16**, 836–850 (2015).

150. Yang, B. X. *et al.* Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell* **163**, 230–245 (2015).

151. Feldman, N. *et al.* G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* **8**, 188–194 (2006).

152. You, J. S. *et al.* OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. *Proc. Natl Acad. Sci. USA* **108**, 14497–14502 (2011). **You *et al.* demonstrate that OCT4 can bind to nucleosomal DNA at DNA-hypomethylated enhancers in somatic cells to initiate a preliminary nucleosome-free region.**

153. MacArthur, B. D. *et al.* Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. *Nat. Cell Biol.* **14**, 1139–1147 (2012).

154. Epsztejn-Litman, S. *et al.* *De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).

155. Hochedlinger, K., Yamada, Y., Beard, C. & Jaenisch, R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121**, 465–477 (2005).

156. Shields, J. M., Christy, R. J. & Yang, V. W. Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J. Biol. Chem.* **271**, 20009–20017 (1996).

157. Garrett-Sinha, L. A., Eberspaecher, H., Seldin, M. F. & de Crombrugghe, B. A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J. Biol. Chem.* **271**, 31384–31390 (1996).

158. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).

159. Meshorer, E. & Misteli, T. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat. Rev. Mol. Cell Biol.* **7**, 540–546 (2006).

160. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).

161. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).

162. Whyte, W. A. *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482**, 221–225 (2012). **References 160–162 demonstrate that super-enhancers regulate cell-type specific genes. The H3K4 demethylase LSD1 interacts with these enhancers in mouse ES cells and controls differentiation by removing epigenetic modifications associated with their activity.**

163. Wang, Y. *et al.* LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell* **138**, 660–672 (2009).

164. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol. Cell. Biol.* **30**, 4851–4863 (2010).

165. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).

166. Macfarlan, T. S. *et al.* Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* **25**, 594–607 (2011).

167. Rudolph, T. *et al.* Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol. Cell* **26**, 103–115 (2007).

168. Reynolds, N. *et al.* NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell Stem Cell* **10**, 583–594 (2012).

169. Lessard, J. A. & Crabtree, G. R. Chromatin regulatory mechanisms in pluripotency. *Annu. Rev. Cell Dev. Biol.* **26**, 503–532 (2010).

170. Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).

171. dos Santos, R. L. *et al.* MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell* **15**, 102–110 (2014).

172. Luo, M. *et al.* NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells* **31**, 1278–1286 (2013).

173. Cacchiarelli, D. *et al.* Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* **162**, 412–424 (2015).

174. Jin, C. *et al.* H3.3/H2A. Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941–945 (2009).

175. Mito, Y., Henikoff, J. G. & Henikoff, S. Histone replacement marks the boundaries of *cis*-regulatory domains. *Science* **315**, 1408–1411 (2007).

176. He, H. H. *et al.* Nucleosome dynamics define transcriptional enhancers. *Nat. Genet.* **42**, 343–347 (2010).

177. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).

178. Sherwood, R. I. *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014).

179. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).

180. Soufi, A. *et al.* Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015). **This article demonstrates that the pioneer factor-like activity of OCT4, SOX2 and KLF4 is achieved by cooperatively binding to partially exposed recognition motifs within nucleosomal DNA.**

181. Heng, J. C. *et al.* The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotency. *Cell Stem Cell* **6**, 167–174 (2010).

182. Hutchins, A. P. *et al.* Co-motif discovery identifies an Esrrb-Sox2-DNA ternary complex as a mediator of transcriptional differences between mouse embryonic and epiblast stem cells. *Stem Cells* **31**, 269–281 (2013).

183. Maekawa, M. *et al.* Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–229 (2011).

184. Tsubooka, N. *et al.* Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. *Genes Cells* **14**, 683–694 (2009).

185. Wang, W. *et al.* Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. *Proc. Natl Acad. Sci. USA* **108**, 18283–18288 (2011).

186. Jolma, A. *et al.* DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339 (2013).

187. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).

188. Kingston, R. E. & Narlikar, G. J. ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* **13**, 2339–2352 (1999).

189. Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl Acad. Sci. USA* **106**, 5181–5186 (2009).

190. Singhal, N. *et al.* Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* **141**, 943–955 (2010).
Singhal *et al.* demonstrate a requirement for the ES cell-specific BAF (esBAF) complex to reactivate the endogenous *Oct4* locus in somatic nuclei during reprogramming.

191. van den Berg, D. L. *et al.* An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369–381 (2010).

192. Pardo, M. *et al.* An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**, 382–395 (2010).

193. Floer, M. *et al.* A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell* **141**, 407–418 (2010).

194. Cheloufi, S. The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* **528**, 218–224 (2015).

195. Ang, Y. S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).

196. Apostolou, E. *et al.* Genome-wide chromatin interactions of the *Nanog* locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* **12**, 699–712 (2013).

197. Yang, P., Wu, W. & Macfarlan, T. S. Maternal histone variants and their chaperones promote paternal genome activation and boost somatic cell reprogramming. *Bioessays* **37**, 52–59 (2015).

198. Lin, C. J., Koh, F. M., Wong, P., Conti, M. & Ramalho-Santos, M. Hira-mediated H3.3 incorporation is required for DNA replication and ribosomal RNA transcription in the mouse zygote. *Dev. Cell* **30**, 268–279 (2014).

199. Inoue, A. & Zhang, Y. Nucleosome assembly is required for nuclear pore complex assembly in mouse zygotes. *Nature Struct. Mol. Biol.* **21**, 609–616 (2014).

200. Boskovic, A. *et al.* Higher chromatin mobility supports totipotency and precedes pluripotency *in vivo*. *Genes Dev.* **28**, 1042–1047 (2014).

201. Oback, B. & Wells, D. N. Donor cell differentiation, reprogramming, and cloning efficiency: elusive or illusive correlation? *Mol. Reprod. Dev.* **74**, 646–654 (2007).

202. Gonzalez-Munoz, E., Arboleda-Estudillo, Y., Otu, H. H. & Cibelli, J. B. Histone chaperone ASF1A is required for maintenance of pluripotency and cellular reprogramming. *Science* **345**, 822–825 (2014).

203. Shinagawa, T. *et al.* Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell* **14**, 217–227 (2014).

204. Miyamoto, K. *et al.* Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development. *Science* **341**, 1002–1005 (2013).

205. Nashun, B., Akiyama, T., Suzuki, M. G. & Aoki, F. Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos. *Epigenetics* **6**, 1489–1497 (2011).

206. Wen, D. *et al.* Histone variant H3.3 is an essential maternal factor for oocyte reprogramming. *Proc. Natl Acad. Sci. USA* **111**, 7325–7330 (2014).

207. Gao, S. *et al.* Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice. *Dev. Biol.* **266**, 62–75 (2004).

208. Burton, A. & Torres-Padilla, M. E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–734 (2014).

209. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).

210. Eggan, K. *et al.* X-chromosome inactivation in cloned mouse embryos. *Science* **290**, 1578–1581 (2000).

211. Pasque, V., Gillich, A., Garrett, N. & Gurdon, J. B. Histone variant macroH2A confers resistance to nuclear reprogramming. *EMBO J.* **30**, 2373–2387 (2011).

212. Inoue, K. *et al.* Impeding *Xist* expression from the active X chromosome improves mouse somatic cell nuclear transfer. *Science* **330**, 496–499 (2010).

213. Chan, M. M., Smith, Z. D., Egli, D., Regev, A. & Meissner, A. Mouse ooplasm confers context-specific reprogramming capacity. *Nat. Genet.* **44**, 978–980 (2012).

214. Matoba, S. *et al.* Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. *Cell* **159**, 884–895 (2014).
This work identified H3K9 methylation as a significant epigenetic impediment to successful zygotic genome activation and progression of embryonic development during SCNT.

215. Blelloch, R. *et al.* Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells* **24**, 2007–2013 (2006).

216. Pereira, C. F. *et al.* ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* **6**, 547–556 (2010).

217. Piccolo, F. M. *et al.* Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell* **49**, 1023–1033 (2013).

218. Foshay, K. M. *et al.* Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Mol. Cell* **46**, 159–170 (2012).

219. Han, D. W. *et al.* Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. *Stem Cells* **26**, 445–454 (2008).

220. Looney, T. J. *et al.* Systematic mapping of occluded genes by cell fusion reveals prevalence and stability of *cis*-mediated silencing in somatic cells. *Genome Res.* **24**, 267–280 (2014).

221. Weintraub, H. *et al.* Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl Acad. Sci. USA* **86**, 5434–5438 (1989).

SUPPLEMENTARY INFORMATION
See online article: S1 (box) | S2 (box) | S3 (box) | S4 (box)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF