Biological importance of OCT transcription factors in reprogramming and development

Kee-Pyo Kim[1,2], Dong Wook Han[3], Johnny Kim [4] and Hans R. Schöler[1,5]

Abstract

Ectopic expression of Oct4, Sox2, Klf4 and c-Myc can reprogram somatic cells into induced pluripotent stem cells (iPSCs). Attempts to identify genes or chemicals that can functionally replace each of these four reprogramming factors have revealed that exogenous Oct4 is not necessary for reprogramming under certain conditions or in the presence of alternative factors that can regulate endogenous Oct4 expression. For example, polycistronic expression of Sox2, Klf4 and c-Myc can elicit reprogramming by activating endogenous Oct4 expression indirectly. Experiments in which the reprogramming competence of all other Oct family members tested and also in different species have led to the decisive conclusion that Oct proteins display different reprogramming competences and species-dependent reprogramming activity despite their profound sequence conservation. We discuss the roles of the structural components of Oct proteins in reprogramming and how donor cell epigenomes endow Oct proteins with different reprogramming competences.

Introduction

The OCT protein family comprises eight transcription factors that bind to specific target sequences to regulate gene expression[1–3]. These factors play important roles in the maintenance of cellular identity in different tissues and in mediating cell fate decisions throughout embryonic development[4–6]. Because Oct proteins regulate the expression of hundreds of target genes that determine and maintain cellular identity, ectopic expression of Oct proteins is widely used as a means to redirect and reprogram cellular identity[7–13]. Well-studied examples include reprogramming of somatic cells into induced neural stem cells (iNSCs) with Oct9 or into induced pluripotent stem cells (iPSCs) with Oct4[7,10].

Arguably, the most studied Oct family member is Oct4, not only because of its essential physiological role in early development but also because of its unique transcriptional functions in reprogramming biology[5,14–17]. Oct4 is the only OCT protein that can induce pluripotency[18–21], which has been a somewhat surprising revelation since OCT family members display profound sequence conservation. Notwithstanding, these observations have fueled long-standing interest in understanding exactly how OCT4 evokes pluripotency and why other Oct family members do not have this effect.

Recent studies offer exciting new insights into the biology of reprogramming both with and, interestingly, without exogenously administered Oct4. It was shown that reprogramming somatic cells to pluripotency can in fact be achieved by completely abrogating ectopic Oct4 expression[22,23]. This result indicates that reprogrammed cells acquire and establish a self-sustaining pluripotent state that is evoked by endogenous Oct4 expression. Importantly, it was shown that reprogramming without the use of exogenous Oct4 significantly improves the overall quality of the iPSCs because exogenous Oct4 can disrupt imprinted gene expression and create off-target effects[23]. Indeed, iPSCs generated without the use of exogenous Oct4 display superior developmental potential

Correspondence: Johnny Kim (Johnny.Kim@mpi-bn.mpg.de) or
Hans R. Schöler (office@mpi-muenster.mpg.de)
[1]Department of Cell and Developmental Biology, Max Planck Institute for
Molecular Biomedicine, Röntgenstrasse 20, Münster 48149, Germany
[2]Department of Life Sciences, College of Medicine, The Catholic University of
Korea, 222 Banpo-daero Seocho-gu, Seoul 06591, Republic of Korea
Full list of author information is available at the end of the article
These authors contributed equally: Kee-Pyo Kim, Dong Wook Han

as determined by a tetraploid complementary assy. These observations, together with results from other recent studies, revealed that the induction of pluripotency can be achieved as long as endogenous Oct4 can be activated, either directly or indirectly[24–30]. Essentially, it has become clear that, while endogenous Oct4 expression is necessary, exogenous Oct4 is sufficient but not necessary for reprogramming somatic cells.

Along these lines, it was recently shown that virtually all OCT proteins harbor reprogramming competence, and they can activate the pluripotency network to different degrees under optimal conditions[31,32]; these conditions are differentially inferred by diverse donor cell epigenomes that differ not only between donor cell types but also between species. It appears that reprogramming competence not only can be significantly enhanced, for example, by chemical intervention, but can also be synthesized completely de novo via optimal combinations of chemically altered donor cell epigenomes and exogenously provided transcriptional regulators that activate the pluripotency network under defined culture conditions[31,32]. Here, we review our current understanding of how Oct proteins function in the reprogramming process and discuss the differential roles of exogenous and endogenous Oct4 in reprogramming biology. We provide an overview of factors that can replace Oct4 in murine and human reprogramming and discuss how these replacements enable the acquisition of pluripotency. We elaborate on recent findings showing how reprogramming can be achieved without administration of exogenous Oct4 and share our view on how different reprogramming competences of Oct proteins are differentially mediated in different species.

POU proteins: an overview

The POU (Pit-Oct-Unc) protein family comprises 15 transcription factors that bind to specific target sequences[33–36]. POU proteins have diverse roles in a wide range of cellular processes[4–6] and are categorized into six classes (POU I to POU VI) based on the sequence similarity of their DNA-binding domains (DBDs)[1,3]. Only proteins in the POU II, III and V classes, which predominantly bind to the octamer motif (consensus sequence ATGCAAAT) and its variants, are classified as octamer-binding (Oct) proteins. Proteins in other POU classes (POU I, IV and VI) display lower binding affinity for the octamer motif and thus are classified as nonoctamer-binding proteins.

The Oct protein group constitutes eight family members (Oct1, Oct2, Oct4, Oct6, Oct7, Oct8, Oct9 and Oct11), and the numbering of each protein is based on the position at which DNA probes used in electrophoretic mobility shift assays (EMSAs) are bound[2]. Oct proteins share a highly conserved bipartite DNA-binding domain (DBD), consisting of two structurally independent subdomains, a 75 amino acid N-terminal POU-specific domain (POU$_S$) and a 60 amino acid C-terminal POU homeodomain (POU$_H$)[1,3,4,37,38]. Each POU domain can bind a sequence of four base pairs in the major groove of the cognate DNA sequence, thereby placing each POU domain on either side of the DNA helix and effectively encircling target DNA sequences[37,38]. A linker, which varies in length and sequence between Oct proteins, is flanked by these two POU domains[4,6,37]. This linker also influences the binding specificity and conformation of Oct proteins to DNA sequences, although it does not appear to physically interact with the DNA sequence itself[39,40]. Posttranslational modifications of POU domains, including ubiquitylation, glycosylation, SUMOylation, phosphorylation and oxidation, also influence the binding ability of Oct proteins to cognate DNA sequences[41–45]. Oct proteins form homodimers or heterodimers with other partner proteins on specific target sequences[46–49], adding an additional layer of target gene control. Oct proteins harbor two transactivation domains (TADs) that are located on each side of the DBD[5]. In contrast to DBDs, TADs exhibit little sequence conservation and vary in length between Oct proteins. TADs are known to play important roles in the transcriptional stimulation of target genes by interacting with basal transcription machinery and other cofactors[50,51].

Biological functions of Oct proteins in development

Oct proteins of the POU II class include Oct1 (also known as Otf-1 and Pou2f1), Oct2 (also known as Pou2f2) and Oct11 (also known as Skn-1 and Pou2f3)[4–6]. Although they belong to the same class, their expression patterns and functions are grossly different during embryonic development. Oct1 is ubiquitously expressed and can be detected in almost all cell types[52,53], whereas the expression of Oct2 and Oct11 is largely restricted to specific cell types. Oct2 is highly expressed in B lymphocytes and plays an important role in normal germinal center reactions[54–56]. Oct11 is highly expressed in skin epithelial cells and taste cells and is critical for epidermal differentiation and the composition of taste receptor cells[57–59].

Oct4 belongs to the POU V class[4–6]. It is the best-characterized OCT member because of its profound biological importance in early embryonic development, germ cell maintenance, stem cell pluripotency and cellular reprogramming[5,14–17]. Oct4 is highly expressed in totipotent and pluripotent cells, including oocytes, early cleavage-stage embryos, inner cell mass (ICM) of blastocysts, epiblasts, embryonic stem cells (ESCs) and germ cells[60–62]. Oct4 expression is regulated by two cis-enhancer elements (distal and proximal enhancers) that are located 2 kb upstream of its transcriptional start site[63,64]. Interestingly, the activity of these enhancers differs between cell types. The proximal enhancer (PE) is

active in epiblasts of postimplantation embryos and epiblast stem cells (EpiSCs), whereas the distal enhancer (DE) is active in the ICM, ESCs and germ cells[64–66]. Reducing the expression levels of Oct4 in ESCs triggers their differentiation towards primitive endodermal, mesodermal and trophectodermal cell lineages[67,68]. Oct4-deficient embryos are viable to the blastocyst stage, but the ICM cannot be formed in these embryos[69]. These observations, together with the discovery of iPSC technology[10,70], have underscored the functional importance of Oct4 for both the maintenance and establishment of pluripotency.

POU III factors (Oct6, Oct7, Oct8 and Oct9) are prominently expressed in cells and tissues of the central nervous system (CNS)[1,71,72]. They play critical roles in neurogenesis and gliogenesis[73–77]. Interestingly, the deletion of any one of these genes in mice does not lead to severe phenotypes[74,75,78]. Distinct phenotypes are only observed when two or three of POU III factors are simultaneously deleted, indicating their functional redundancy. The DBDs in POU III factors are highly conserved, and consequently, these factors display similar DNA-binding characteristics and DNA-dependent dimerization features[18,20,79,80], providing a molecular explanation for their functional redundancy. Along this line, overexpression of any POU III factor can convert fibroblasts into neural lineage cells[7–9,11–13]. Thus, similar to Oct4, POU III factors are not only essential regulators of normal development but can also convert cellular phenotypes upon ectopic expression.

Oct6 (also known as SCIP, Tst-1 and Pou3f1) is expressed in Schwann cells, oligodendrocyte progenitor cells (OPCs) and keratinocytes[71,73–76,78,81]. Furthermore, Oct6 play an important role in the specification and differentiation of neuroectodermal lineages[82–85]. It was shown that Oct6 is expressed in ESCs[72,86,87]. However, we and others have recently failed to detect Oct6 expression in the ICM, iPSCs or ESCs[32,83,88–91]. Instead, its expression was detected in epiblasts of E5.5–E6.5 postimplantation embryos and EpiSCs[32,83,88–91]. Epiblasts and EpiSCs represent a more developmentally differentiated state compared to the ICM, iPSCs or ESCs and display gene expression characteristics of gastrula-stage ectoderm[66,92,93]. To date, knockout studies of Oct6 in epiblasts and EpiSCs have not been performed. Therefore, the specific role and functionality of Oct6 in epiblast formation and maintenance of the primed state of pluripotency remain incompletely understood. However, one can speculate that its role in embryonic development is not as critical as that of Oct4 because homozygous Oct6-knockout mice are born, although they die soon after birth because of respiratory defects[73,94].

Reprogramming of somatic cells to pluripotency

Yamanaka and Takahashi made a landmark discovery showing that ectopic expression of Oct4, Sox2, Klf4 and c-Myc can reprogram somatic cells into iPSCs[10,70]. These four reprogramming factors work interdependently to inactivate features of somatic cell identity and concurrently to activate properties of pluripotent stem cells[95–98]. During cell fate transition, cells undergo dramatic changes at the morphological, transcriptional and epigenetic levels. Notably, only a small proportion of these cells pass through all of these changes and eventually became bona fide iPSCs. Therefore, the reprogramming process is stochastic and inefficient[99].

Various gain- and loss-of-function screenings have led to the discovery of specific genes and molecular pathways that inhibit or enhance the reprogramming process[19,26,31,100–112]. For example, specific epigenetic modifications, including DNA methylation, H3K9 methylation and H3K79 methylation marks, and/or the related enzymes (e.g., DNMTs, HDACs, LSD1, and DOT1L) can act as barriers to the reprogramming process. Therefore, forced elimination of these epigenetic barriers either through genetic inactivation or chemical inhibition can enhance or improve the reprogramming process[31,100,102,103,106–109,112]. In contrast, forced expression of chromatin remodelers, such as Dppa2 and Dppa4, can reset the epigenome of somatic cells to a pluripotent configuration that enhances both reprogramming efficiency and kinetics[113]. In addition, other transcription factors, including Esrrb, Glis1, Nr5a2, Prdm14, Rarg, Sall4, Tbx3, Foxa2, Foxf1, Foxh1, Lhx1 and T, can each significantly enhance reprogramming when overexpressed together with Oct4, Sox2, Kfl4 and c-Myc[19,26,101,104,105,110,111,114]. Other studies have shown that altering culture conditions and/or using different donor cell types can result in completely different reprogramming outcomes[115–118]. For example, adding specific nutrients (e.g., vitamin C) to the culture medium or inducing hypoxic cell culture conditions can greatly enhance reprogramming efficiency[116,118]. Furthermore, significantly higher reprogramming efficiency can also be achieved with donor cells that contain inherent phenotypic plasticity and/or high replication potential[115,117].

Despite the significant advances that have been made to improve and understand the reprogramming process with Oct proteins, critical questions have remained unanswered, e.g., Is exogenous Oct4 expression truly essential for inducing pluripotency? More importantly, how are reprogramming competences of Oct proteins actually mediated and what are the structural bases that define them? Finally, how or why do reprogramming competences of Oct proteins differ, particularly among species?
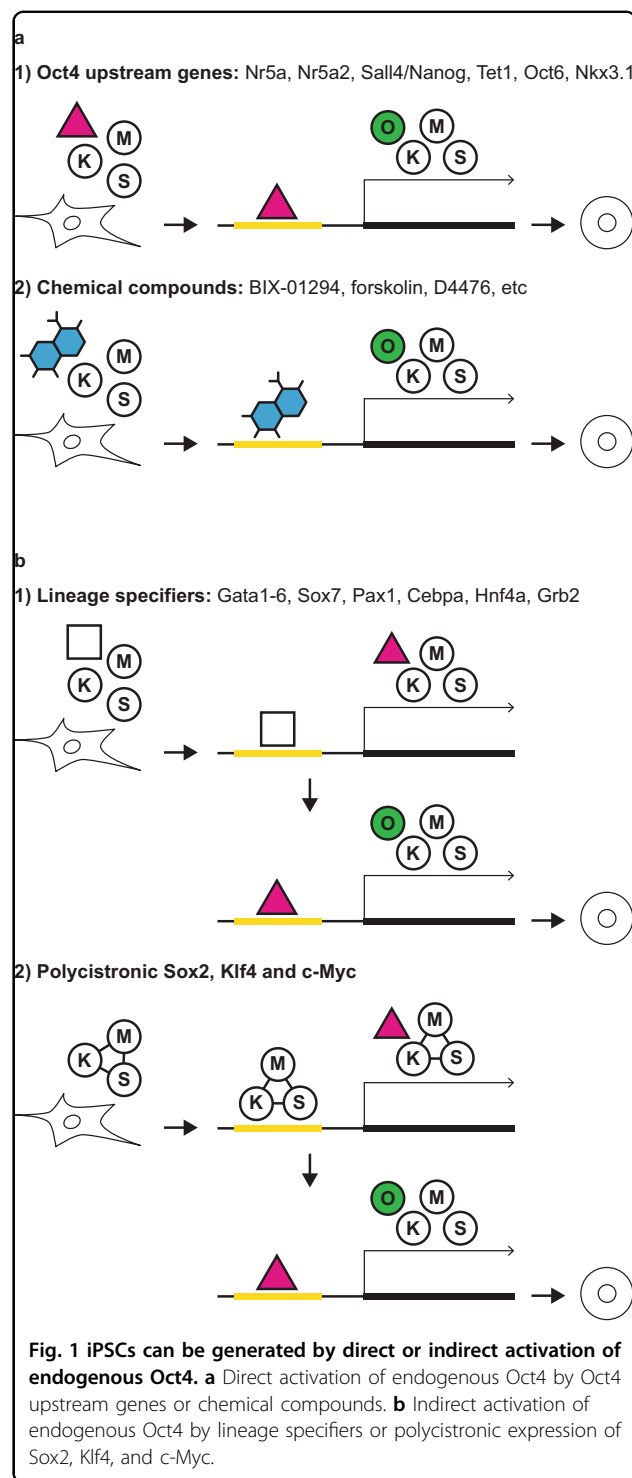
Exogenous Oct4, but not endogenous Oct4, is dispensable for reprogramming

For many years and for several reasons, it has been widely assumed that exogenously administered Oct4 is

the most important and, in fact, indispensable factor for the reprogramming process. First, while Sox2, Klf4 and c-Myc can each be replaced by their respective family members, Oct4 cannot be replaced by any of its paralogs[18–21]. Second, and again in contrast to Sox2, Klf4 or c-Myc, Oct4 alone can elicit reprogramming when specific pathways or molecules are inhibited by compounds (e.g., VPA, a HDAC inhibitor, or CHIR-99021, a GSK-3 inhibitor) or when a specific cell type is used as a donor cell (e.g., NSCs)[119–123]. Finally, perturbation of key Oct4 functions by deletion of structural components or by introduction of inactivating point mutations completely abolishes the reprogramming process[18,20,124,125].

However, the view that exogenous Oct4 is indispensable for reprogramming has recently changed. Undoubtedly, including Oct4 in reprogramming cocktails appears to be the most efficient means to generate iPSCs. However, recent studies have revealed that exogenous Oct4 can be entirely omitted from reprogramming cocktails or can be replaced by its family members, different transcription factors or small molecules[22,24–27,29–32,126–129]. For example, Nr5a1 (also known as Sf-1), Nr5a2 (also known as Lrh1) or Sall4/Nanog can each functionally replace Oct4 and elicit reprogramming together with Sox2, Klf4 and c-Myc[24,26]. Notably, these genes lie genetically upstream of Oct4, such that their ectopic expression can directly activate endogenous Oct4 expression, resulting in iPSC generation[24,130–132] (Fig. 1a). The DNA demethylase Tet1 can also functionally replace Oct4 in reprogramming[25]. Ectopic expression of Tet1 mediates demethylation at the regulatory regions of Oct4, consequently resulting in activation of endogenous Oct4. These studies clearly demonstrate that the induction of pluripotency in conjunction with Sox2, Klf4 and c-Myc can be mediated without exogenous Oct4 by the use of alternative factors that are capable of directly activating endogenous Oct4 expression.

Interestingly, genes that do not directly regulate the endogenous Oct4 locus can also elicit reprogramming in conjunction with Sox2, Klf4 and c-Myc (Fig. 1b). These genes include Gata1–6, Sox7, Pax1, Cebpa, Hnf4a and Grb2, most of which, in normal physiological processes, do not lie genetically upstream of Oct4 and are not expressed in ESCs or iPSCs. Most of these factors normally function as lineage-specific transcription factors and play important roles in cell fate determination of mesodermal and endodermal lineages during embryonic development[29,30]. A study aiming to decipher the molecular mechanisms of Gata factor-based reprogramming revealed that ectopic expression of Gata factors can activate Sall4, which in turn activates endogenous Oct4, consequently resulting in iPSC generation without exogenously administered Oct4[30]. Based on these observations, it has been speculated that activating endogenous

Fig. 1 iPSCs can be generated by direct or indirect activation of endogenous Oct4. a Direct activation of endogenous Oct4 by Oct4 upstream genes or chemical compounds. **b** Indirect activation of endogenous Oct4 by lineage specifiers or polycistronic expression of Sox2, Klf4, and c-Myc.

Oct4 expression indirectly may be a common way to elicit reprogramming by all other lineage-specific transcription factors. However, Nkx3.1 and Oct6, which are also lineage-specific transcription factors, do not seem to follow this mode of action. Indeed, we and others have recently revealed that Nkx3.1 and Oct6 can directly bind

to regulatory regions of Oct4 to regulate its expression[32,126]. Thus, these observations clearly suggest that the indirect mode of action might not apply to the reprogramming function of all other lineage-specific transcription factors.

In addition to ectopic expression of genetic factors, specific chemicals can also functionally replace exogenous Oct4 and induce pluripotency in conjunction with Sox2, Klf4 and c-Myc[27,28]. For instance, NSCs transduced with Sox2, Klf4 and c-Myc and subsequently cultured in the presence of BIX-01294, an inhibitor of the histone methyltransferase G9a, give rise to iPSC colonies[28] (Fig. 1a). G9a is critical for de novo DNA methylation at regulatory regions of Oct4[133,134]. Thus, BIX-01294 treatment results in demethylation at regulatory regions of Oct4, leading to activation of endogenous Oct4 expression[28]. BIX-01294-mediated reprogramming does not seem to be effective with mouse embryonic fibroblasts (MEFs), indicating that inhibition of different molecular pathways is required for eliciting reprogramming of different types of donor cells. In MEFs, exogenous Oct4 can be functionally replaced by various other small molecules, including forskolin (a cAMP agonist), 2-methyl-5-hydroxytryptamine (a 5-HT3 agonist) or D4476 (a casein kinase 1 inhibitor)[27] (Fig. 1a). The precise molecular mechanism by which each of these chemicals induces pluripotency in conjunction with Sox2, Klf4 and c-Myc remains elusive. However, it is likely that some, or perhaps all, of these chemicals induce endogenous Oct4 expression either directly or indirectly since iPSC colonies would otherwise not be formed. With these studies, it has been established that exogenous Oct4 is dispensable for reprogramming, but endogenous Oct4 is not, which is in agreement with the fact that endogenous Oct4 is absolutely required for the establishment of pluripotency in preimplantation-stage embryos[69].

Reprogramming with Sox2, Klf4 and c-Myc alone produces high-quality iPSCs

We and others recently showed that iPSC generation is possible without exogenous administration of Oct4 or any of its known replacers (Oct4 upstream genes, lineage-specification transcription factors or chemicals) when Sox2, Klf4 and c-Myc are connected together by the 2 A peptide and ectopically expressed from a single polycistronic vector in MEFs[22,128]. Integrative ChIP-seq and RNA-seq analyses of cells undergoing reprogramming with Sox2, Klf4 and c-Myc alone revealed that Sox2 and Klf4 can directly bind and mediate the activation of pluripotency genes, including but not limited to Sall4 and Nanog, which in turn mediate the activation of endogenous Oct4 expression[22,128]. Interestingly, reprogramming by Sox2, Klf4 and c-Myc alone is context-dependent for various reasons. First, only polycistronic expression of Sox2, Klf4 and c-Myc can elicit reprogramming[22,128].

Monocistronic expression of Sox2, Klf4 and c-Myc does not yield iPSC colonies, indicating that stoichiometric cooperativity with these three reprogramming factors is necessary for eliciting reprogramming. Second, iPSC generation is possible by lentiviral and episomal delivery of Sox2, Klf4 and c-Myc but not by retroviral delivery due to rapid retroviral silencing that occurs in the early phase of reprogramming. Interestingly, the simultaneous expression of Sox2 and Klf4 activates retrovirus silencing machinery (e.g., Trim28, Setdb1 and Chaf1a/b), which in turn mediates retroviral silencing[128]. Finally, reprogramming with Sox2, Klf4 and c-Myc alone is species-dependent, such that reprogramming in this manner can be achieved in mouse cells but not in human cells[128]. In future studies, elucidating exactly why or how the function of these three factors differs between species will provide important mechanistic insights that will further help to define the molecular impediments to the reprogramming process.

The poor quality of iPSCs yields from reprogramming is a key limitation that precludes the development of iPSC-based applications in regenerative biomedicine. It has been shown that the quality of iPSCs can be directly influenced by the stoichiometric expression of reprogramming factors (e.g., $Oct4^{high}/Klf4^{high}/Sox2^{low}/c-Myc^{low}$), supplements of cell culture media (e.g., vitamin C) and/or the choice of reprogramming factors (e.g., Sall4/Nanog/Esrrb/Lin28a or Oct4/Sox2/Klf4)[135–137]. Although reprogramming with c-Myc together with Oct, Sox2 and Klf4 provides larger yields of iPSCs, c-Myc exerts a detrimental effect on imprinted loci, resulting in significantly reduced developmental potential[135]. Intriguingly, we recently discovered that the quality of iPSCs is dramatically reduced by exogenously expressed Oct4[128]. In fact, iPSC lines generated by polycistronic expression of Sox2, Klf4 and c-Myc without Oct4 produced all-iPSC mice much more efficiently than iPSC lines generated by conventional polycistronic expression of the four factors. By comparing transcriptomes and epigenomes between a series of iPSC lines generated with the two different reprogramming cocktails, i.e., with and without exogenous Oct4, we came to the decisive conclusion that ectopic Oct4 expression reduces iPSC quality by anomalously activating off-target genes unrelated to pluripotency and disrupting imprinted gene expression[128]. Considering the immense usage of iPSCs in various biomedical applications, investigations into whether iPSC lines generated without exogenous Oct4 by alternative factors are of good quality are warranted.

It has been shown that obtaining high-quality iPSCs is strongly correlated with lower reprogramming efficiency[135,136]. Consistently, reprogramming without Oct4 but with Sox2, Klf4 and c-Myc produces fewer iPSC colonies with slower reprogramming kinetics, but the

majority of the iPSC lines derived from these colonies are of superior quality, as assessed by molecular profiling and their contribution to tetraploid complementation[128]. Thus, the theme that "less is more" appears to hold true for reprogramming biology. Reprogramming systems that yield iPSCs with high efficiency and fast kinetics come with the price of generating a high proportion (too high) of low-quality iPSCs.

Transactivation domains and donor-cell epigenomes are tightly linked to reprogramming competences

Despite the iPSC quality drawbacks that can occur by reprogramming with exogenous Oct4, it is by far the strongest pluripotency inducer. The observation that Oct4 cannot be functionally replaced by other Oct family members in murine reprogramming[18,20,21,31], although Oct proteins share profound sequence conservation, has fueled interest in understanding what makes Oct4 uniquely competent in inducing pluripotency. We and others have compared the structures and biochemical properties of Oct proteins to better understand the molecular basis of the reprogramming processes. Interestingly, despite their high similarities in protein sequences, the DNA-binding profiles of Oct proteins and their DNA-dependent dimerization with Sox2 are remarkably different[18,20,80]. Specifically, Oct4 preferentially forms heterodimers with Sox2 through the canonical SoxOct motif (CATTGTTATGCAAAT), which is highly enriched in enhancers of pluripotency genes, including Nanog, Fgf4, Utf1 and Pou5f1[18,46,47,80]. In contrast, other Oct proteins, such as Oct6 and Oct7, preferentially form homodimers through the more palindromic octamer recognition element (MORE; ATGCATATGCAT) motif, which is predominantly found in the regulatory regions of neural genes, including Neurog1, Olig1 and Ascl1[18,47–49]. The formation of Oct4-Sox2 heterodimers through SoxOct motifs appears to be necessary for inducing pluripotency because mutations that disrupt Oct4-Sox2 heterodimer formation abolish reprogramming competence[18,138,139]. Oct6 and Oct7 can also bind to the SoxOct motif in EMSAs, albeit with low efficiency[18,80]. However, ChIP assays with ESCs, neural precursor cells (NPCs) and cells undergoing reprogramming failed to detect Oct6 and Oct7 binding to the SoxOct motif[13,20,32,77,80,140]. The discrepancy of these results might be due to the different experimental settings and sensitivities of these two different assays. Notwithstanding, these studies clearly indicate that Oct proteins display differential DNA-dependent dimerization properties and DNA-binding profiles, providing an explanation for why they display different reprogramming competences and, more specifically, why Oct4 cannot be functionally replaced by its family members in murine reprogramming.

To support this idea, through structural modeling, we found that methionine at position 151 is critical for the homodimerization of Oct proteins through the MORE motif[18]. This amino acid is contained in POU III factors (e.g., Oct6) but not in Oct4. Therefore, amino acid substitution of methionine 151 with serine (Oct4^{S151M}) increases its ability to form homodimers through the MORE motif[18]. In contrast, amino acid substitution of methionine to serine in residue 151 of Oct6 (Oct6^{M151S}) reduced its ability to form homodimers through the MORE motif[18]. Although the binding propensities of these two mutants to the MORE motif were significantly altered, their respective binding propensity for the SoxOct motif was not substantially changed compared to that of their wild-type counterparts[18]. As such, Oct4^{S151M} still produced iPSC colonies similar to wild-type Oct4 and Oct6^{M151S} was unable elicit reprogramming similar to wild-type Oct6[18]. These findings indicate the binding preference of Oct proteins for MORE is not the key to their reprogramming competences. Previously, it was shown that two amino acids at residues 7 and 22 in the POU-specific domain of Oct4 are essential to form an Oct4-Sox2-DNA ternary structure[141]. Interestingly, changing these amino acids by substituting aspartic acid with lysine at residue 7 and lysine with threonine at residue 22 in Oct6^{M151S} (OCT6 $^{M151S/D7K,K22T}$) enables this engineered Oct6 mutant to bind to the SoxOct motif. Consequently, it is capable of producing iPSC colonies[18]. However, the reprogramming efficiency of this mutant was extremely low. Thus, other structural features in addition DNA-dependent dimerization are required to mediate Oct6-mediated reprogramming efficiently in murine cells.

In mouse ESCs, Oct4 mutants in which both N-TAD and C-TAD are deleted show severe phenotypes and cannot maintain a pluripotent state[142]. Interestingly, only one TAD in Oct4 is needed to maintain the functionality of Oct4, since mouse ESCs in which either the N-TAD or C-TAD of Oct4 is deleted maintain pluripotency[142]. However, when the reprogramming competences of Oct6-Oct4 domain-swapped chimeras were tested, the chimeric proteins (O466 and O664) in which either N- or C-TADs of Oct4 was introduced into the corresponding site of Oct6 did not acquire reprogramming activity[32]. However, when both N- and C-TADs of Oct4 were introduced into corresponding sites of Oct6, the chimeric protein (O464) induced the formation of iPSC colonies[32]. Overall, these observations reveal interconnected functional features of TAD and DBD, which together are crucial for mouse cell reprogramming and pluripotency maintenance: differential DNA-dependent binding propensities through the DBD and functional features of the TAD that are yet to be discovered.

OCT6 is a pluripotency inducer and acts as a pioneer transcription factor in human reprogramming

Oct4 has been used for iPSC generation from cells obtained from a number of different species, including mice, humans, rats, monkeys and gorillas[10,70,143–145], underscoring the universality and conserved fundamental function of Oct4 for inducing pluripotency. Notably, most studies addressing the possibility of replacing Oct4 for reprogramming have been performed with mouse cells[24–26,29,30]. However, because of its ever more increasing significance for biomedicine, there is great interest in determining whether Oct4 replacers can function similarly in cells from other species, especially humans. Until recently, Nxk3.1 was the only known factor that can induce pluripotency in addition to Oct4 in both human and mouse cells[126]. Nr5a1, Nr5a2, Tet1, Sall4, Nanog and Gata3 can all replace Oct4 in reprogramming murine cells to pluripotency. However, these paralogs fail to produce iPSC colonies in conjunction with Sox2, Klf4 and c-Myc in human cells[32], clearly indicating their mouse-specific reprogramming competence. A potential explanation for these divergence outcomes might involve the differences in the chromatin structures and epigenomes in human and mouse cells. This is because the accessibility of reprogramming factors to target sites differ between human and mouse cells[146–149]. Thus, the reprogramming competences of these Oct4 replacers are likely differentially regulated depending on the epigenetic state of the donor cells. In agreement with this concept, we recently showed that creating an epigenetically permissive environment for human donor cells through a combinational treatment of compounds that affect chromatin modifications, including DNA methylation and histone methylation, enabled the generation of iPSC colonies upon administration of Nr5a1, Nr5a2, Tet1, Sall4, Nanog and Gata3[31]. Together, these findings triggered our hypothesis that other factors which can replace Oct4 specifically in human reprogramming might exist.

By executing a functional screening of 100 tested candidate genes, we surprisingly found that Oct6 can induce pluripotency in conjunction with Sox2, Klf4 and c-Myc in human cells[32]. This observation was especially intriguing because Oct6 cannot generate iPSC colonies from mouse cells[18,20,21]. Notably, human and mouse Oct6 proteins share 98.89% identity, and both proteins can elicit reprogramming of human cells but not of mouse cells[32]. Therefore, its species-dependent reprogramming competence is not mediated by sequence variations of orthologous Oct6 proteins but by the differences in mouse and human cell epigenomes. Indeed, if the levels of specific epigenetic modifications, including H3K27 methylation, H3K79 methylation, H3K4 methylation and H3K9 methylation, are modulated, ectopic Oct6 expression enables iPSC generation from mouse cells[31].

Although Oct4 and Oct6 are functionally interchangeable for inducing human pluripotency, OCT6-mediated reprogramming is significantly slower and less efficient than that mediated by OCT4[32]. Comparative RNA-seq analysis of TRA-1–60+/CD31− cells collected over the time course of reprogramming with Oct4 and Oct6 revealed that the mesenchymal-to-epithelial transition (MET) was not different. This finding supports the argument that Oct proteins do not affect or delay the MET process during reprogramming. However, the activation of pluripotency genes is remarkably delayed in Oct6-mediated reprogramming[32]. Integrative ChIP-seq and RNA-seq analyses revealed that the delayed activation of pluripotency genes in Oct6-based reprogramming is largely due to the delayed binding of Oct6 to regulatory regions of corresponding genes[32]. Consequently, Oct6-mediated reprogramming is relatively inefficient and slow. Importantly, Oct6 can directly bind to the Oct4 locus, underscoring the concept that endogenous Oct4 expression is essential to the reprogramming process. Therefore, Oct6 lies genetically upstream of Oct4, at least in humans. In mice, Oct4 and Oct6 are coexpressed in EpiSCs and epiblasts[32,83,88–91]. Therefore, in future studies, it will be important to investigate whether and how Oct4 is regulated by Oct6, or vice versa, in these specific cell types to better understand the species-specific impediments of reprogramming with these factors.

Oct4 fulfills the criteria of being a pioneering transcription factor, such that it can engage nucleosomal DNA and thereby initiate regulation of its target genes that promote a change in cell fate[148–150]. To date, exactly how this pioneering activity is achieved remains poorly understood. We recently showed that the pioneering activity of Oct4 depends on functional synergism between Oct4 C-TAD and Oct4 DBD[32]. The absence of Oct4 C-TAD significantly abolishes the ability of Oct4 to bind pluripotency gene enhancers, and more importantly, its function is only possible when it is connected to the Oct4 DBD. In fact, introducing Oct4 C-TAD to corresponding sites of Oct6 does not change the binding of Oct6 to pluripotency gene enhancers and its reprogramming competence[32]. Intriguingly, in contrast to Oct4, Oct6 displays its pioneering synergistic activity with its DBD through the N-TAD. Interestingly, the Oct4 C-TAD can be functionally replaced by the Oct6 N-TAD but not the Oct6 C-TAD[32]. Similarly, when the N-TAD of Oct6 is connected with Oct4 DBD, the binding ability of these chimeras to enhancers of pluripotency genes dramatically increases. It has been shown that the N-TAD of Gata3 is critical for Gata3-mediated chromatin remodeling and its pioneering activity[151]. The C-TAD of Ebf is also required for its pioneering activity and B cell reprogramming[152]. Furthermore, the C-TAD of Hnf3 (also known as FoxA) is critical for its pioneering function[153]. Overall, these

observations support our conclusion that the C-TAD of Oct4 and the N-TAD of Oct6 are critical for their pioneering activities and reprogramming competences.

Concluding remarks

Reprogramming somatic cells to pluripotency is, in fact, a complicated process, and this process does not rely on a single molecular pathway. Our recent findings together with those of other studies clearly suggest that molecular pathways and roadmaps to reach pluripotency can be diverse, depending on the epigenetic state of the donor cells and the exogenously provided transcription factors. Therefore, modulating epigenomes by chemical intervention and introducing different combinations of transcription factors in different donor cell types and also in different species might further enhance our understanding of reprogramming mechanisms.

Author details
¹Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, Münster 48149, Germany. ²Department of Life Sciences, College of Medicine, The Catholic University of Korea, 222 Banpo-daero Seocho-gu, Seoul 06591, Republic of Korea. ³Guangdong Provincial Key Laboratory of Large Animal models for Biomedicine, Wuyi University, Jiangmen 529020, China. ⁴Department of Cardiac Development and Remodeling, Max-Planck-Institute for Heart and Lung Research, Bad Nauheim 61231, Germany. ⁵Medical Faculty, University of Münster, Münster, Germany

Conflict of interest
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Rosenfeld, M. G. POU-domain transcription factors: pou-er-ful developmental regulators. *Genes Dev.* **5**, 897–907 (1991).
2. Schöler, H. R. Octamania: the POU factors in murine development. *Trends Genet.* **7**, 323–329 (1991).
3. Wegner, M., Drolet, D. W. & Rosenfeld, M. G. POU-domain proteins: structure and function of developmental regulators. *Curr. Opin. Cell Biol.* **5**, 488–498 (1993).
4. Herr, W. & Cleary, M. A. The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev.* **9**, 1679–1693 (1995).
5. Kang, J., Shakya, A. & Tantin, D. Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem Sci.* **34**, 491–499 (2009).
6. Malik, V., Zimmer, D. & Jauch, R. Diversity among POU transcription factors in chromatin recognition and cell fate reprogramming. *Cell Mol. Life Sci.* **75**, 1587–1612 (2018).
7. Han, D. W. et al. Direct reprogramming of fibroblasts into neural stem cells by defined factors. *Cell Stem Cell* **10**, 465–472 (2012).
8. Lujan, E., Chanda, S., Ahlenius, H., Sudhof, T. C. & Wernig, M. Direct conversion of mouse fibroblasts to self-renewing, tripotent neural precursor cells. *Proc. Natl Acad. Sci. USA* **109**, 2527–2532 (2012).
9. Sheng, C. et al. Direct reprogramming of Sertoli cells into multipotent neural stem cells by defined factors. *Cell Res.* **22**, 208–218 (2012).
10. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
11. Tsunemoto, R. et al. Diverse reprogramming codes for neuronal identity. *Nature* **557**, 375–380 (2018).
12. Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
13. Wapinski, O. L. et al. Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* **155**, 621–635 (2013).
14. Pan, G. J., Chang, Z. Y., Scholer, H. R. & Pei, D. Stem cell pluripotency and transcription factor Oct4. *Cell Res.* **12**, 321–329 (2002).
15. Shi, G. & Jin, Y. Role of Oct4 in maintaining and regaining stem cell pluripotency. *Stem Cell Res Ther.* **1**, 39 (2010).
16. Tantin, D. Oct transcription factors in development and stem cells: insights and mechanisms. *Development* **140**, 2857–2866 (2013).
17. Wu, G. & Scholer, H. R. Role of Oct4 in the early embryo development. *Cell Regen.* **3**, 7 (2014).
18. Jerabek, S. et al. Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. *EMBO Rep.* https://doi.org/10.15252/embr.201642958 (2016).
19. Maekawa, M. et al. Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–229 (2011).
20. Malik, V. et al. Pluripotency reprogramming by competent and incompetent POU factors uncovers temporal dependency for Oct4 and Sox2. *Nat. Commun.* **10**, 3477 (2019).
21. Nakagawa, M. et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).
22. An, Z. et al. Sox2 and Klf4 as the functional core in pluripotency induction without exogenous Oct4. *Cell Rep.* **29**, 1986–2000 e1988 (2019).
23. Velychko, S. et al. Fusion of reprogramming factors alters the trajectory of somatic lineage conversion. *Cell Rep.* **27**, 30–39 e34 (2019).
24. 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).
25. Gao, Y. et al. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
26. Heng, J. C. et al. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* **6**, 167–174 (2010).
27. Hou, P. et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651–654 (2013).
28. Shi, Y. et al. A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* **2**, 525–528 (2008).
29. Shu, J. et al. Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**, 963–975 (2013).
30. Shu, J. et al. GATA family members as inducers for cellular reprogramming to pluripotency. *Cell Res.* **25**, 169–180 (2015).
31. Kim, K. P. et al. Permissive epigenomes endow reprogramming competence to transcriptional regulators. *Nat. Chem. Biol.* https://doi.org/10.1038/s41589-020-0618-6 (2020).
32. Kim, K. P. et al. Reprogramming competence of OCT factors is determined by transactivation domains. *Sci. Adv.* **6**, https://doi.org/10.1126/sciadv.aaz7364 (2020).
33. Falkner, F. G. & Zachau, H. G. Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. *Nature* **310**, 71–74 (1984).
34. Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. Structure of the 5′ ends of immunoglobulin genes: a novel conserved sequence. *Proc. Natl Acad. Sci. USA* **81**, 2650–2654 (1984).
35. Herr, W. et al. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. *Genes Dev.* **2**, 1513–1516 (1988).

36. Lenardo, M. J. et al. Repression of the IgH enhancer in teratocarcinoma cells associated with a novel octamer factor. *Science* **243**, 544–546 (1989).

37. Phillips, K. & Luisi, B. The virtuoso of versatility: POU proteins that flex to fit. *J. Mol. Biol.* **302**, 1023–1039 (2000).

38. Verrijzer, C. P. et al. The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO J.* **11**, 4993–5003 (1992).

39. Kong, X. et al. Functional interplay between the RK motif and linker segment dictates Oct4-DNA recognition. *Nucleic Acids Res.* **43**, 4381–4392 (2015).

40. van Leeuwen, H. C., Strating, M. J., Rensen, M., de Laat, W. & van der Vliet, P. C. Linker length and composition influence the flexibility of Oct-1 DNA binding. *EMBO J.* **16**, 2043–2053 (1997).

41. Brumbaugh, J. et al. Phosphorylation regulates human OCT4. *Proc. Natl Acad. Sci. USA* **109**, 7162–7168 (2012).

42. Constable, S., Lim, J. M., Vaidyanathan, K. & Wells, L. O-GlcNAc transferase regulates transcriptional activity of human Oct4. *Glycobiology* **27**, 927–937 (2017).

43. Li, S. et al. Disruption of OCT4 ubiquitination increases OCT4 protein stability and ASH2L-B-mediated H3K4 methylation promoting pluripotency acquisition. *Stem Cell Rep.* **11**, 973–987 (2018).

44. Marsboom, G. et al. Glutamine metabolism regulates the pluripotency transcription factor OCT4. *Cell Rep.* **16**, 323–332 (2016).

45. Wei, F., Scholer, H. R & Atchison, M. L. Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *J. Biol. Chem.* **282**, 21551–21560 (2007).

46. Ambrosetti, D. C., Basilico, C. & Dailey, L. Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.* **17**, 6321–6329 (1997).

47. Remenyi, A. et al. Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17**, 2048–2059 (2003).

48. Rhee, J. M., Gruber, C. A., Brodie, T. B., Trieu, M. & Turner, E. E. Highly cooperative homodimerization is a conserved property of neural POU proteins. *J. Biol. Chem.* **273**, 34196–34205 (1998).

49. Tomilin, A. et al. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* **103**, 853–864 (2000).

50. Mitchell, P. J. & Tjian, R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371–378 (1989).

51. Seipel, K., Georgiev, O. & Schaffner, W. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11**, 4961–4968 (1992).

52. Fletcher, C., Heintz, N. & Roeder, R. G. Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. *Cell* **51**, 773–781 (1987).

53. Sturm, R. A., Das, G. & Herr, W. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev.* **2**, 1582–1599 (1988).

54. Hodson, D. J. et al. Regulation of normal B-cell differentiation and malignant B-cell survival by OCT2. *Proc. Natl Acad. Sci. USA* **113**, E2039–E2046 (2016).

55. Muller, M. M., Ruppert, S., Schaffner, W. & Matthias, P. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature* **336**, 544–551 (1988).

56. Scheidereit, C. et al. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homoeobox protein. *Nature* **336**, 551–557 (1988).

57. Andersen, B. et al. Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors expressed in epidermis. *Science* **260**, 78–82 (1993).

58. Andersen, B. et al. Functions of the POU domain genes Skn-1a/i and Tst-1/Oct-6/SCIP in epidermal differentiation. *Genes Dev.* **11**, 1873–1884 (1997).

59. Matsumoto, I., Ohmoto, M., Narukawa, M., Yoshihara, Y. & Abe, K. Skn-1a (Pou2f3) specifies taste receptor cell lineage. *Nat. Neurosci.* **14**, 685–687 (2011).

60. Okamoto, K. et al. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**, 461–472 (1990).

61. Rosner, M. H. et al. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345**, 686–692 (1990).

62. Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K. & Gruss, P. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**, 435–439 (1990).

63. Nordhoff, V. et al. Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm. Genome* **12**, 309–317 (2001).

64. Yeom, Y. I. et al. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894 (1996).

65. Choi, H. W. et al. Distinct enhancer activity of Oct4 in naive and primed mouse pluripotency. *Stem Cell Rep.* **7**, 911–926 (2016).

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

67. Loh, Y. H. et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431–440 (2006).

68. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).

69. 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).

70. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).

71. He, X. et al. Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* **340**, 35–41 (1989).

72. Schöler, H. R., Hatzopoulos, A. K, Balling, R., Suzuki, N. & Gruss, P. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. *EMBO J.* **8**, 2543–2550 (1989).

73. Bermingham, J. R. Jr. et al. Tst-1/Oct-6/SCIP regulates a unique step in peripheral myelination and is required for normal respiration. *Genes Dev.* **10**, 1751–1762 (1996).

74. Friedrich, R. P., Schlierf, B., Tamm, E. R., Bosl, M. R. & Wegner, M. The class III POU domain protein Brn-1 can fully replace the related Oct-6 during schwann cell development and myelination. *Mol. Cell. Biol.* **25**, 1821–1829 (2005).

75. Jaegle, M. et al. The POU proteins Brn-2 and Oct-6 share important functions in Schwann cell development. *Genes Dev.* **17**, 1380–1391 (2003).

76. Jaegle, M. et al. The POU factor Oct-6 and Schwann cell differentiation. *Science* **273**, 507–510 (1996).

77. 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).

78. Schreiber, J. et al. Redundancy of class III POU proteins in the oligodendrocyte lineage. *J. Biol. Chem.* **272**, 32286–32293 (1997).

79. Chang, Y. K. et al. Quantitative profiling of selective Sox/POU pairing on hundreds of sequences in parallel by Coop-seq. *Nucleic Acids Res.* **45**, 832–845 (2017).

80. Mistri, T. K. et al. Selective influence of Sox2 on POU transcription factor binding in embryonic and neural stem cells. *EMBO Rep.* **16**, 1177–1191 (2015).

81. Faus, I., Hsu, H. J. & Fuchs, E. Oct-6: a regulator of keratinocyte gene expression in stratified squamous epithelia. *Mol. Cell. Biol.* **14**, 3263–3275 (1994).

82. Iwafuchi-Doi, M. et al. Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development* **139**, 3926–3937 (2012).

83. Kamiya, D. et al. Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* **470**, 503–509 (2011).

84. Yasuhara, N. et al. Triggering neural differentiation of ES cells by subtype switching of importin-alpha. *Nat. Cell Biol.* **9**, 72–79 (2007).

85. Zhu, Q. et al. The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. *Elife* **3**, https://doi.org/10.7554/eLife.02224 (2014).

86. Meijer, D. et al. The octamer binding factor Oct6: cDNA cloning and expression in early embryonic cells. *Nucleic Acids Res.* **18**, 7357–7365 (1990).

87. Suzuki, N., Rohdewohld, H., Neuman, T., Gruss, P. & Schöler, H. R. Oct-6: a POU transcription factor expressed in embryonal stem cells and in the developing brain. *EMBO J.* **9**, 3723–3732 (1990).

88. Acampora, D. et al. Loss of the Otx2-binding site in the Nanog promoter affects the integrity of embryonic stem cell subtypes and specification of inner cell mass-derived epiblast. *Cell Rep.* **15**, 2651–2664 (2016).

89. Kurek, D. et al. Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Rep.* **4**, 114–128 (2015).

90. Li, G. et al. Dysregulation of the SIRT1/OCT6 axis contributes to environmental stress-induced neural induction defects. *Stem Cell Rep.* **8**, 1270–1286 (2017).

91. Yasuhara, N. et al. Importin alpha subtypes determine differential transcription factor localization in embryonic stem cells maintenance. *Developmental Cell* **26**, 123–135 (2013).

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

93. Kojima, Y. et al. The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* **14**, 107–120 (2014).

94. DeVeale, B. et al. Oct4 is required ~E7.5 for proliferation in the primitive streak. *PLoS Genet* **9**, e1003957 (2013).

95. Adachi, K. & Scholer, H. R. Directing reprogramming to pluripotency by transcription factors. *Curr. Opin. Genet. Dev.* **22**, 416–422 (2012).

96. Buganim, Y., Faddah, D. A. & Jaenisch, R. Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet* **14**, 427–439 (2013).

97. Hochedlinger, K. & Plath, K. Epigenetic reprogramming and induced pluripotency. *Development* **136**, 509–523 (2009).

98. Papp, B. & Plath, K. Epigenetics of reprogramming to induced pluripotency. *Cell* **152**, 1324–1343 (2013).

99. Hanna, J. et al. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).

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

101. Chia, N. Y. et al. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468**, 316–320 (2010).

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

103. Ebrahimi, A. et al. Bromodomain inhibition of the coactivators CBP/EP300 facilitate cellular reprogramming. *Nat. Chem. Biol.* **15**, 519–528 (2019).

104. Feng, B. et al. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* **11**, 197–203 (2009).

105. Han, J. et al. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* **463**, 1096–1100 (2010).

106. Hu, X. et al. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* **14**, 512–522 (2014).

107. Mali, P. et al. Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* **28**, 713–720 (2010).

108. Onder, T. T. et al. Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).

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

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

111. 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).

112. Zhang, Z., Xiang, D. & Wu, W. S. Sodium butyrate facilitates reprogramming by derepressing OCT4 transactivity at the promoter of embryonic stem cell-specific miR-302/367 cluster. *Cell Reprogram* **16**, 130–139 (2014).

113. Hernandez, C. et al. Dppa2/4 facilitate epigenetic remodeling during reprogramming to pluripotency. *Cell Stem Cell* **23**, 396–411 e398 (2018).

114. Takahashi, K. et al. Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nat. Commun.* **5**, 3678 (2014).

115. Aasen, T. et al. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* **26**, 1276–1284 (2008).

116. Esteban, M. A. et al. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**, 71–79 (2010).

117. Panopoulos, A. D. et al. Rapid and highly efficient generation of induced pluripotent stem cells from human umbilical vein endothelial cells. *PloS ONE* **6**, e19743 (2011).

118. Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T. & Yamanaka, S. Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell* **5**, 237–241 (2009).

119. Kim, J. B. et al. Direct reprogramming of human neural stem cells by OCT4. *Nature* **461**, 649–643 (2009).

120. Kim, J. B. et al. Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**, 411–419 (2009).

121. Li, Y. et al. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res.* **21**, 196–204 (2011).

122. Yuan, X. et al. Brief report: combined chemical treatment enables Oct4-induced reprogramming from mouse embryonic fibroblasts. *Stem Cells* **29**, 549–553 (2011).

123. Zhu, S. et al. Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell* **7**, 651–655 (2010).

124. Esch, D. et al. A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat. Cell Biol.* **15**, 295–301 (2013).

125. Jin, W. et al. Critical POU domain residues confer Oct4 uniqueness in somatic cell reprogramming. *Sci. Rep.* **6**, 20818 (2016).

126. Mai, T. et al. NKX3-1 is required for induced pluripotent stem cell reprogramming and can replace OCT4 in mouse and human iPSC induction. *Nat. Cell Biol.* **20**, 900–908 (2018).

127. Montserrat, N. et al. Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* **13**, 341–350 (2013).

128. Velychko, S. et al. Excluding Oct4 from Yamanaka cocktail unleashes the developmental potential of iPSCs. *Cell Stem Cell* **25**, 737–753 e734 (2019).

129. Zhao, Y. et al. A XEN-like state bridges somatic cells to pluripotency during chemical reprogramming. *Cell* **163**, 1678–1691 (2015).

130. Barnea, E. & Bergman, Y. Synergy of SF1 and RAR in activation of Oct-3/4 promoter. *J. Biol. Chem.* **275**, 6608–6619 (2000).

131. Gu, P. et al. Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol. Cell. Biol.* **25**, 3492–3505 (2005).

132. Zhang, J. et al. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat. cell Biol.* **8**, 1114–1123 (2006).

133. 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).

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

135. Buganim, Y. et al. The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. *Cell Stem Cell* **15**, 295–309 (2014).

136. Carey, B. W. et al. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* **9**, 588–598 (2011).

137. Chen, J. et al. Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. *Cell Res.* **21**, 884–894 (2011).

138. Jauch, R. et al. Conversion of Sox17 into a pluripotency reprogramming factor by reengineering its association with Oct4 on DNA. *Stem Cells* **29**, 940–951 (2011).

139. Tapia, N. et al. Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency. *Sci. Rep.* **5**, 13533 (2015).

140. Urban, S. et al. A Brn2-Zic1 axis specifies the neuronal fate of retinoic-acid-treated embryonic stem cells. *J. Cell Sci.* **128**, 2303–2318 (2015).

141. Nishimoto, M. et al. Oct-3/4 maintains the proliferative embryonic stem cell state via specific binding to a variant octamer sequence in the regulatory region of the UTF1 locus. *Mol. Cell. Biol.* **25**, 5084–5094 (2005).

142. Niwa, H., Masui, S., Chambers, I., Smith, A. G. & Miyazaki, J. Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. *Mol. Cell. Biol.* **22**, 1526–1536 (2002).

143. Fang, R. et al. Generation of naive induced pluripotent stem cells from rhesus monkey fibroblasts. *Cell Stem Cell* **15**, 488–497 (2014).

144. Li, W. et al. Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. *Cell Stem Cell* **4**, 16–19 (2009).

145. Wunderlich, S. et al. Primate iPS cells as tools for evolutionary analyses. *Stem Cell Res* **12**, 622–629 (2014).

146. Chronis, C. et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459 e420 (2017).

147. Fu, K. et al. Comparison of reprogramming factor targets reveals both species-specific and conserved mechanisms in early iPSC reprogramming. *BMC Genomics* **19**, 956 (2018).

148. 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).

149. Soufi, A. et al. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).

150. Zaret, K. S. & Mango, S. E. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr. Opin. Genet. Dev.* **37**, 76–81 (2016).

151. Takaku, M. et al. GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. *Genome Biol.* **17**, 36 (2016).

152. Boller, S. et al. Pioneering activity of the C-terminal domain of EBF1 shapes the chromatin landscape for B cell programming. *Immunity* **44**, 527–541 (2016).

153. Cirillo, L. A. et al. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002).