Peering into the Black Box of Reprogramming to the Pluripotent State

Steven A. Jackson · Rupa Sridharan

Published online: 30 March 2013
© Springer Science+Business Media New York 2013

Abstract Somatic cells can be de-differentiated and acquire the properties of pluripotent cells by the overexpression of a small set of proteins. This process has revolutionized not only the prospect of using pluripotent cells for therapeutic purposes but also our understanding of the maintenance of cell fate. Delineating the mechanism of reprogramming has been challenging because of the low efficiency of the process and variable latency in the conversion of somatic cells to the induced pluripotent stem cell state. Prospective isolation of cells undergoing faithful reprogramming has not been possible. In this review we focus on advances made in the past year, which include proteomics and single cell transcriptomics of cells undergoing reprogramming. These studies have revealed the timing of major transcriptional and phenotypic changes, barriers to the process at early and intermediate stages and the hierarchical activation of the pluripotency network.

Keywords Induced pluripotent stem cells (iPSCs) · Metabolism · Single cell analysis · MET · Reprogramming · Transition to pluripotent state

Introduction

Embryonic stem cells (ESCs) possess the unique properties of pluripotency—they can both self renew and differentiate into a myriad of other cell types, including an entire

animal, upon receiving the appropriate stimulus. Remarkably, these same properties can be conferred upon somatic cells by the overexpression of a few proteins to generate induced pluripotent stem cells (iPSCs) [1]. The observation that cell fate can be reversed to a “blank” state has multiple ramifications for regenerative therapy. If ESCs and iPSCs are truly equivalent, they can be used to overcome the practical and ethical concerns of using ESCs for such therapy. Much progress has been made in delivering the reprogramming factors by integration free methods to make the process more translatable for human therapeutic use [2–4]. However, the mechanism of reprogramming remains elusive, mainly because even in a culture of cells where each individual cell has incorporated the reprogramming factors, only about 2–4 % complete the process and at variable timescales [5–8]. This inherent heterogeneity of the process makes it challenging to study.

The reprogramming of somatic cells to iPSCs is a profound change in cell identity, encompassing erasure of the somatic cell program including the mesenchymal-to-epithelial transition (MET), [9, 10], cell cycle [11], metabolism [12, 13], chromatin structure [14] and activation of the pluripotency network [9, 10, 15–17]. Herein, we discuss the significant strides that have been made in elucidating when and whether these changes occur throughout reprogramming to specify the acquisition of the pluripotency gene activation.

The Order of Events: Tools and Timescales

Almost all mechanistic studies thus far have been carried out with mouse embryonic fibroblasts (MEFs) as the starting somatic cells for reprogramming. While these studies have improved the overall understanding of somatic cell reprogramming, the lack of studies using diverse cell

S. A. Jackson · R. Sridharan (✉)
Wisconsin Institute for Discovery, 330 N Orchard St, Madison,
WI 53706, USA
e-mail: rsridharan@discovery.wisc.edu

S. A. Jackson
e-mail: sajackson2@wisc.edu

types means some of this knowledge may be specific to MEFs (e.g. role of MET in the reprogramming of mesenchymal cells). These studies usually take two broad approaches: 1. The molecular profiling of either heterogeneous reprogramming cultures or clonal intermediates, and 2. loss or gain of function of specific candidates or libraries of genes to test if they affect reprogramming efficiency [9, 18, 19, 20, 21–26]. In the past year these approaches have been extended to purified intermediate reprogramming populations [27, 28, 29, 30], early stages of human reprogramming, and single cell expression profiling of a cohort of important genes.

Previous studies have established a timescale for the sequence of events in reprogramming. Early in the process almost 80 % of the initial MEFs lose somatic cell surface markers (Thy1) [15, 16]. During the middle phase of the process, cells undergoing reprogramming gain intermediate markers of pluripotency such as SSEA1 [15, 16] followed by activation of a pluripotency gene such as Oct4 or Nanog. While these initial studies were performed with the reprogramming factors transduced as lentiviruses or retroviruses, more recent studies have taken advantage of secondary systems [5–8, 31–33] in which every cell in the reprogramming culture can inducibly express the reprogramming factors. Using the methods and signposts of reprogramming described above, in the last year a temporal picture has emerged for when changes in MET, cell cycle, metabolism, the epigenome and activation of pluripotency loci occur and are discussed in greater detail below (Fig. 1).

The Mesenchymal-to-Epithelial Transition

During development, the cells of the embryo undergo an epithelial-to-mesenchyme (EMT) transition to delaminate from the epiblast so as to traverse the primitive streak and differentiate into mesoderm and endoderm [34, 35]. Since embryonic stem cells (ESCs) are isolated from the embryo before this event occurs, they maintain an epithelial nature in culture including expression of Cdh1 (E-cadherin) on their cell surface. However, MEFs are mesenchymal in nature, and a reversal to the iPSCs state involves a mesenchymal-to-epithelial transition (MET). MET is described as one of the first events to occur in the process of forming iPSCs [15, 16]. The Hochedlinger and Krijgsvelde groups [27, 28] isolated populations of cells that expressed different cell surface markers (Thy1, SSEA1) or were positive for Oct4-GFP expression at various time-points [day 3 (d3), d6, d9 and d12] and performed transcriptomic, ChIP-Seq and proteomic analysis to thoroughly characterize these intermediate stages. At both the transcriptional and proteomic levels, they observed that genes associated with maintaining the mesenchymal state such as

Snai1 are down-regulated by day 3 of reprogramming while epithelial associated genes such as Cdh1 were up-regulated [27, 28]. The Jaenisch group [29] took a complementary approach. Instead of characterizing the purified intermediate populations, they chose select genes involved in MET, cell cycle and pluripotency and performed a single cell transcriptome analysis to record reprogramming in individual cells. They demonstrated that at the earliest time point they analyze (d6), cells contain both Snai1 and Cdh1 transcripts, indicating the process is a transition and a cell does not have to completely lose mesenchymal gene expression before gaining epithelial gene expression [29]. While there is no apparent difference in the number of Cdh1 transcripts per cell when comparing between the d6 and d12 single cell transcriptomes, by d12, no cell expressing Cdh1 has transcripts for Snai1, indicating a complete silencing of the mesenchymal somatic cell program at this time [29].

Interestingly, a ChIP-Seq analysis of the reprogramming factors in human cells within 48 h of their induction showed that binding sites were enriched for the functional category of MET [30], implying that MET genes are direct targets of the reprogramming factors.

Cell Cycle

The cell cycle of pluripotent stem cells is uniquely different from that of somatic cells (reviewed by White and Dalton [36]). Exhibiting a short G1 and G2/M phase without reducing the time taken to transit through S phase, pluripotent cells rapidly proliferate and divide symmetrically to produce equivalent daughter cells [36], while somatic cells exhibit extended division times. In the process of reprogramming somatic cells to iPSCs, a reversion to the pluripotent cell cycle occurs, with iPSCs exhibiting highly similar cell cycle profiles to those of embryonic stem cells [11, 37]. It is known that rapid cell cycling is important for the efficient formation of iPSCs from human somatic cells, and that inducing cell cycle arrest prevents the formation of iPSCs [38]. In a testament to the importance of cell cycle to reprogramming the removal or reduction in p53 levels or the loss of Ink4/Arf improve the number of reprogrammed cells that are obtained [21–26]. However, this may be because a greater number of cell divisions give the necessary chance for reprogramming relevant stochastic changes to take place [39]. Like MET, cell cycle changes seem to occur early in the reprogramming process. Ccnb1 (Cyclin B1), which is an essential gene for the control of the G2/M phase of cell division, Cdk1 (serine/threonine kinase essential for G1/S and G2/M phase transitions) and Plk1 (serine/threonine protein kinase performing numerous functions throughout the M phase of

Fig. 1 Major events and barriers in reprogramming events that take place in the early (day 0–3), middle (day 3–9) and late (day 9 onwards) stages of reprogramming are indicated. Days for each phase are according to Polo et al. [27••] Typical cell populations identified in each phase include: Early 30 % Thy1– [15], Intermediate up to 55 % Thy1– with 27 % SSEA1+, and Late 90–99 % Thy1– and SSEA expression variable depending on establishment of iPSC [15, 27••]. Events or gene (*italicized*) expression recorded in red are down-regulated, while events recorded in green are up-regulated. Events spanning multiple stages of reprogramming are contained

the cell cycle) are all rapidly upregulated in all cells, including the ones that will remain non-reprogrammed. However, at later time points in the reprogramming process (d6, d9, d12 and iPSCs) these genes are clearly expressed at higher levels than “non-reprogramming” counterparts [27••]. Whether the down-regulation of cell cycle genes is a root-cause of reprogramming failure or a consequence of the failure remains to be examined.

These results are also corroborated in the single cell studies where a trifecta of genes that interact for mitotic maintenance Bub1, Cdc20 and Mad211 are upregulated by day 2–4 compared to MEFs [29••]. Gene ontology analysis of both transcriptional changes and proteins upregulated at d3 include a plethora of terms associated with cell cycle, cell division, mitosis, DNA replication and repair [27••, 28••]. c-Myc binds to Cdkn2d (p19), which is a potent inhibitor of Cdk4 and Cdk6 and results in negative regulation of the cell cycle and reduction of proliferative capacity [30••]. At the transcriptional level, there is an immediate down-regulation of Cdkn2b (p15), belonging to the same family as Cdkn2d and affecting cell cycle in a similar manner [27••].

within triangles indicating when they occur. Underlined chromatin events occur at select loci, while non-underlined events are global. DNAme is both up- and down-regulated. MET, cell cycle and metabolism data summarized from references [27••, 28••, 29••], Chromatin, data summarized from references [20•, 27••, 30••, 52, 59], activation of the pluripotency network from references [29••, 63••]. MET; mesenchymal-to-epithelial transition, MEF; mouse embryonic fibroblast, iPSC; induced pluripotent stem cells, Pre-iPSC; pre-induced pluripotent stem cells, DNAme; DNA methylation, GDNF/RET; glial cell line derived neurotrophic factor/RET pathway

Interestingly, as mentioned above, knockdown of another Cdkn2 family member Cdkn2a, also known as Ink4a, improves reprogramming efficiency [24]. It is interesting to speculate that c-Myc binding to Cdkn2 family members reduces their expression which subsequently removes the block in cell division that Cdkn2b and Cdkn2d impose resulting in increased proliferation. In addition, transcription factor binding data indicate that c-Myc could be partially responsible for removing the blocks imposed in somatic cells on cell cycle [17]. However, reprogramming is also successful when c-Myc is omitted from the reprogramming cocktail. While these studies utilized all four of the Yamanaka reprogramming factors it will be interesting to perform these experiments with just Oct4, Sox2 and Klf4 to test if similar timelines are maintained.

Metabolism

The role of metabolic changes in reprogramming is only now being explored. Like the MET transitions, changes in

metabolism also seem to be initiated early and accumulate in the middle and late stages during reprogramming. Both the transcriptomic and proteomic studies identified a number of changes that suggest a decrease in oxidative phosphorylation in the earliest stages of reprogramming [27•, 28•], followed by the up-regulation of glycolysis during the middle phase of reprogramming [28•]. There is no indication why these proteins accumulate rather than appear abruptly after transcription initiation. However, the changes that must occur in rearranging mitochondria into a state for glycolysis are extensive [13]. To date there is still no evidence on what drives the changes in metabolism, or which factor or combination of factors results in the dramatic reversal of metabolism seen in the formation of iPSCs. This undoubtedly will be the focus of future investigations. Another facet of metabolism is the emerging evidence of dependence of ESCs on the amino acid threonine [40]. Threonine catabolism is important for the regeneration of *S*-adenosylmethionine (SAM), a necessary cofactor for protein methylation. Similarly, Acetyl CoA, which is a metabolite, involved in producing reducing equivalents such as NAD, is also produced from threonine. Unlike the glycolytic enzymes, the enzymes involved in threonine metabolism are upregulated late in the reprogramming process (d9–12 at the protein level) [28•, 41]. Metabolic changes may not only be related to a rapid production of energy but also provide intermediates for pluripotent related chromatin changes [42].

Epigenetic Changes

Histone Modifications

The epigenome of a cell refers to inheritable changes that do not involve changes in genomic sequence but affect the phenotype of a cell. These epigenomic marks include post-translational modifications (PTMs) on histones and covalent methylation of DNA. It is well established that iPSCs regain an ESC like epigenome upon reprogramming [14]. ESCs are also known to have fewer heterochromatic spots in their nuclei—i.e. a different and more permissive chromatin architecture [43] and a more transcriptionally active genome [44], which is thought to increase their plasticity to be able to respond to differentiation cues. That chromatin plays an important role in reprogramming is confirmed by experiments where modulating the levels and functions of enzymes and proteins that participate in chromatin structure affects reprogramming efficiency [45–51]. Changes in levels of histone modifications occur at both the global and locus-specific levels as measured by chromatin immunoprecipitation. Various studies over the last decade have established that specific histone modifications at regulatory

regions of genes correlate with their expression or lack thereof. It has previously been established that one early epigenomic event is the marking of pluripotency specific enhancers with histone H3 lysine [K] 4 dimethylation [me2] much before the corresponding genes are expressed [52]. Two well-studied modifications that change during reprogramming are the “activating” histone H3K4me3 and “repressive” H3K27me3. Loci that contain both of these modifications are called bivalent and are thought to be marks of “poised” chromatin, since upon receiving differentiation cues these bivalent genes could rapidly retain one or the other mark consolidating the gene expression pattern [53]. It is important to note that bivalent genes are more prevalent in, but not unique to, pluripotent cells [53, 54]. Several genes that are marked with H3K4me3 in MEFs gain H3K27me3 and acquire bivalency after transitioning to the iPSCs state [14]. Polo et al. [27•] found that there is a sudden accumulation of bivalent marks in the initial stage of reprogramming followed by gradual accumulation from day 3 through the middle and late stages. This may enhance the plasticity of these intermediate stages. Taking a more global approach, Mattout et al. performed immunofluorescence for various histone modifications and found that those usually associated with active transcription, such as H3K9ac increase as reprogramming progresses [55]. Supporting this observation, general histone deacetylase inhibitors have been shown to improve reprogramming efficiency [45, 46, 50, 51]. While at the global level, Mattout et al. did not find changes in H3K27me3 levels [55], it is interesting that the loss of a H3K27me3 demethylase, *Utx*, prevents reprogramming [56]. This was partially due to failure of some pluripotency genes such as *Sall4* and *Utf1* to lose their repressive H3K27me3. This suggests that for some chromatin changes, locus specific effects rather than global effects are required for reprogramming to occur.

There are now several reports suggesting that modulating the levels of enzymes associated with histone modifications improve reprogramming [47–49]. The loss of *Dot1L*, which performs H3K79 methylation, which is associated with transcription elongation, increases reprogramming efficiency [18]. This is a surprising finding given that pluripotent cells are usually thought to be transcriptionally hyperactive [44]. High H3K9me3 levels form a barrier both for initial docking of reprogramming factors at pluripotency loci and also at intermediate stages [20•, 30•]. Taken together, these results suggest that epigenomic changes both at global and local levels affect the reprogramming process.

DNA Methylation

An important feature of the epigenome is direct covalent methylation of DNA. While reprogramming from

intermediate stages is improved by the addition of 5-azacytidine, which decreases DNA methylation [57], the two *de novo* DNA methyltransferases are dispensable for acquisition of an iPSCs state [58]. Nonetheless, the loss or gain of DNA methylation at specific promoters is a late event in reprogramming [15]. Recently, a new modification in DNA, 5-hydroxymethylcytosine (5hmC), has been described. This offers a new route to demethylate DNA by the conversion of 5-methylcytosine (5mC) to 5hmC. The enzymes that are responsible for this modification are called Tet1, Tet2 and Tet3. At early stages of reprogramming, 5hmC levels increase at the regulatory regions of Nanog, an important pluripotency gene and the loss of Tet2 abolished reprogramming [59] implying an important role for this modification. It has also been suggested that the long time frame of reprogramming is required for passive DNA demethylation. The discovery of 5hmC raises the intriguing possibility that there may be an active conversion of 5mC to 5hmC during the reprogramming process.

Activation of the Pluripotency Network

Activation of the naïve pluripotency network, with a cessation in reliance upon the exogenous reprogramming factors, is one of the final steps to occur in the formation of iPSCs. The early and middle events set the stage for the activation of the regulatory network to occur. Polo et al. [27•] observed large transcriptional changes at day 3 (compared to MEFs) and day 12 (compared to d9) reprogramming cells. During the intervening 9 days there were no dramatic changes in transcriptional activity until activation of the pluripotency network. However, while there were no dramatic transcriptional changes during the middle phase of reprogramming, Polo and colleagues observed changes in transcript levels of various differentiation related genes [27•]. Interestingly, a recent paper reports the reprogramming factors can convert human fibroblasts into a plastic state where they can subsequently be directed to differentiate toward angioblast-like cells prior establishing pluripotency [60•]. It is interesting to speculate that the authors were able to capitalize on the erasure of the somatic epigenome, accumulation of bivalent genes and changeability in the expression of differentiation related genes to generate the plasticity observed in their study. Other studies have established intermediates that are stalled at this intervening stage where the somatic genome is silenced but activation of pluripotency network has not occurred are commonly recovered from reprogramming cultures starting with MEFs, NPCs and B cells [17, 57, 61, 62]. These studies reveal the activation of the pluripotency network can be a barrier to reprogramming and that there are ways for this barrier to be overcome.

Previous reports suggested a two-step process in the activation of the pluripotency network, termed “maturation” and “stabilization”, primarily involving the activation of early and late markers of pluripotency [9]. Maturation involved the upregulation of expression of Nanog, Sall4, Esrrb and Rex1 while stabilization included the upregulation of Lin28, Utf1 and Pecam [9]. A recent report by Golipour and colleagues delves further into the intricacies of the transition of cells from maturation to the stabilization phase [63•]. During this transition, reprogrammed cells become independent of the transgenes, while failure to down-regulate the transgenes results in a failure to proceed to the stabilization phase. Interestingly, siRNA screening revealed that there were two groups of genes necessary for the transition from the maturation to stabilization phase. The first group consists of genes typically associated with pluripotency including Oct4, Nanog and Sox2. Loss of these genes disrupts maintenance of pluripotency but also prevents the transition to the stabilization phase in reprogramming cells. A second cohort of genes was identified whose expression is required for the transition to occur, but which are not typically associated with maintenance of pluripotency. These genes appear to form a separate but overlapping network which is enriched for processes such as signaling pathways (GDNF/RET) and cell cycle. These studies imply that maturation and stabilization represent real checkpoints in the reprogramming process and certain processes need to be upregulated to allow the final transition between the two. It is interesting to note that cell cycle changes continue to be of vital importance well into the late stages of reprogramming.

Past investigations have relied on the activation of reporters linked to Nanog or Oct4 regulatory regions, but recent studies have revealed that there may be a hierarchy of events in the activation of pluripotency genes and Oct4 or Nanog activation may not represent the completion of reprogramming. For example, the Oct4 locus is activated in single cells undergoing reprogramming, even in clones that remain partially reprogrammed [29•]. Other pluripotency genes such as Esrrb, Utf1 and Sall4 are upregulated in a subset of cells that have undergone MET as early as d6 of reprogramming [29•], and could represent early markers of cells that will actually complete the process. However, Esrrb and Utf1 represent better markers of reprogramming cells than Sall4, as Sall4 expression is found in a number of cells that have not upregulated Cdh1 expression, a necessary step for reprogramming. This is valuable insight because detection of both Esrrb and Utf1 at such an early time during reprogramming may represent an important landmark of the intermediate stage of the process. This is especially intriguing given the fact that Esrrb can functionally replace Nanog [64, 65].

Work conducted by Buganim et al., seems to indicate that establishing endogenous Sox2 expression may initiate the transition to the stabilization phase and ensuing activation of the pluripotency network. Establishing Sox2 expression leads to subsequent Lin28, Sall4, Fgf4, Fbx15 and Dnmt3b upregulation, placing Sox2 at the top of a number of genes associated with the endogenous pluripotent network [29•]. While events prior to Sox2 activation appear to occur stochastically, once activated, it seems to drive a feedforward pathway. Thus, activation of Sox2 can lead to the establishment of the endogenous pluripotent network and may represent cells that have entered the stabilization phase. This study also highlights the power of single cell studies. While Golipour et al. [63•] were able to identify Sox2 as necessary for the transition to stabilization to occur in their assessment of bulk reprogramming cultures, the studies by Buganim and colleagues was able to place Sox2 at the top of a hierarchy of genes leading to establishment of the naïve pluripotency network.

Conclusion

The recent transcriptome, proteome and epigenome analyses presented here have increased our understanding of the temporal specifics of the reprogramming process. However, a number of questions still remain: For example, while MET is an important process for reprogramming from MEFs it will be interesting to explore how cells that are already epithelial in nature traverse the landscape to gain pluripotency. Similarly, it is surprising that the single cell studies mark activation of Sox2 as an important trigger, given that reprogramming from NPCs which express endogenous Sox2 results in rapid accumulation of pre-iPSCs [61]. It will be interesting to investigate whether the activation of Sox2 from MEFs correlates with a particular epigenetic change that cannot be detected by transcription profiling, and is the dominant event causing entry into the stabilization phase.

Do the metabolic changes have to precede the activation of pluripotency loci or chromatin changes? Given that many important metabolic intermediates are utilized by histone modifying enzymes as substrates, these connections are worth exploring. The data sets from the reprogramming intermediates are likely to prove treasure troves that could be mined further with the potential to find better markers of faithful reprogramming than currently reported. These exciting recent studies pave the way for many more investigations into the mechanism of reprogramming.

Disclosure S.A. Jackson and R. Sridharan declare that they have no conflict of interest.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4):663–676
2. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F et al (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 7(5):618–630
3. Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II et al (2009) Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 324(5928):797–801
4. Okita K, Hong H, Takahashi K, Yamanaka S (2010) Generation of mouse-induced pluripotent stem cells with plasmid vectors. *Nat Protoc* 3:418–428
5. Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458(7239):771–775
6. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hämläinen R et al (2009) PiggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458(7239):766–770
7. Carey BW, Markoulaki S, Beard C, Hanna J, Jaenisch R (2010) Single-gene transgenic mouse strains for reprogramming adult somatic cells. *Nat Methods* 7(1):56–98
8. Stadtfeld MM, Maherali NN, Borkent MM, Hochedlinger KK (2010) A reprogrammable mouse strain from gene-targeted embryonic stem cells. *Nat Methods* 7(1):53–55
9. Samavarchi-Tehrani P, Golipour A, David L, Sung H-K, Beyer TA, Datti A et al (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7(1):64–77
10. Li R, Liang J, Ni S, Zhou T, Qing X, Li H et al (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7(1):51–63
11. Ghule PN, Medina R, Lengner CJ, Mandeville M, Qiao M, Dominski Z et al (2011) Reprogramming the pluripotent cell cycle: restoration of an abbreviated G1 phase in human induced pluripotent stem (iPS) cells. *J Cell Physiol* 226(5):1149–1156
12. Kondoh H, Leonart ME, Nakashima Y, Yokode M, Tanaka M, Bernard D et al (2007) A high glycolytic flux supports the proliferative potential of murine embryonic stem cells. *Antioxid Redox Signal* 9(3):293–299
13. Prigione A, Fauler B, Lurz R, Lehrach H, Adjaye J (2010) The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. *Stem Cells* 28(4):721–733
14. Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K et al (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1(1):55–70
15. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K (2008) Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2(3):230–240
16. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H et al (2008) Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2(2):151–159

17. Sridharan R, Tchieu J, Mason MJ, Yachechko R, Kuoy E, Horvath S et al (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136(2):364–377
18. Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM et al (2012) Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483(7391):598–602
19. Bhutani N, Decker MN, Brady JJ, Bussat RT, Burns DM, Corbel SY et al (2013) A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J* 27(3):1107–1113
20. • Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet.* 2012 Dec 2;45(1):34–42. *Identifies H3K9me3 of pluripotency loci, mediated by BMP pathway acting via H3K9 methyltransferases, as a major barrier to reprogramming pre-IPSC to iPSCs*
21. Hong H, Takahashi K, Ichisaka T, Aoi T, Kanagawa O, Nakagawa M et al (2009) Suppression of induced pluripotent stem cell generation by the p53–p21 pathway. *Nature* 460(7259):1132–1135
22. Kawamura T, Suzuki J, Wang YV, Menendez S, Morera LB, Raya A et al (2009) Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460(7259):1140–1144
23. Marión RM, Strati K, Li H, Murga M, Blanco R, Ortega S et al (2009) A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 460(7259):1149–1153
24. Li H, Collado M, Villasante A, Strati K, Ortega S, Cañamero M et al (2009) The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* 460(7259):1136–1139
25. Utikal J, Polo JM, Stadtfeld M, Maherali N, Kulalert W, Walsh RM et al (2009) Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 460(7259):1145–1148
26. Banito A, Rashid ST, Acosta JC, Li S, Pereira CF, Geti I et al (2009) Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev* 23(18):2134–2139
27. •• Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, et al. A molecular roadmap of reprogramming somatic cells into iPS Cells. *Cell.* 2012 Dec 21;151(7):1617–32. *RNA-Seq analysis timecourse of purified reprogramming populations. Equivalent populations to Hansson et al. Identifies two waves of gene expression changes associated with the reprogramming process including early MET, cell cycle changes and switch from oxidative to glycolytic metabolism*
28. •• Hansson J, Rafiee MR, Reiland S, Polo JM, Gehring J, Okawa S, et al. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep.* 2012 Dec 27;2(6):1579–92. *Proteomics analysis of timecourse of purified reprogramming populations. Equivalent populations to Polo et al. Identifies protein up- and down-regulation in number of processes including MET, cell cycle, oxidative and glycolytic metabolism in a stage specific manner*
29. •• Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchical phase. *Cell.* 2012 Sep 14;150(6):1209–22. *This article is the first to provide an in-depth look at the transcriptional changes occurring throughout reprogramming at the single cell level. It identifies early and late markers of reprogramming cells*
30. •• Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell.* 2012 Nov 21;151(5):994–1004. *Comprehensive study of the binding of the four reprogramming factors during the first 48 hours of expression in human fibroblasts. Established binding is enriched at MET sites, KM bind promoters, OSKM bind distal elements and knockdown of DNA methylases results in increased OSKM binding in DNA otherwise refractory to binding*
31. Wernig M, Lengner CJ, Hanna J, Lodato MA, Steine E, Foreman R et al (2008) A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nat Biotechnol* 26(8):916–924
32. Maherali N, Ahfeldt T, Rigamonti A, Utikal J, Cowan C, Hochdinger K (2008) A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 3(3):340–345
33. Markoulaki S, Hanna J, Beard C, Carey BW, Cheng AW, Lengner CJ et al (2009) Transgenic mice with defined combinations of drug-inducible reprogramming factors. *Nat Biotechnol* 27(2):169–171
34. Burdsal CAC, Damsky CHC, Pedersen RAR (1993) The role of E-cadherin and integrins in mesoderm differentiation and migration at the mammalian primitive streak. *Development* 118(3):829–844
35. Ciruna BB, Rossant JJ (2001) FGF Signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev Cell* 1(1):3–13
36. White J, Dalton S (2005) Cell cycle control of embryonic stem cells. *Stem Cell Rev* 1(2):131–138
37. Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B et al (2010) Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genes Dev* 20(2):155–169
38. Ruiz S, Panopoulos AD, Herreras A, Bissig K-D, Lutz M, Berggren WT et al (2011) A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. *Curr Biol* 21(1):45–52
39. Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creighton MP et al (2009) Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 462(7273):595–601
40. Wang J, Alexander P, Wu L, Hammer R, Cleaver O, McKnight SL (2009) Dependence of mouse embryonic stem cells on threonine catabolism. *Science* 325(5939):435–439
41. Shyh-Chang N, Locasale JW, Lyssiotis CA, Zheng Y, Teo RY, Ratanasirintrao S et al (2013) Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science* 339(6116):222–226
42. Sassone-Corsi P (2013) When metabolism and epigenetics converge. *Science* 339(6116):148–150
43. Meshorer EE, Yellajoshula DD, George EE, Scambler PJP, Brown DTD, Misteli TT (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 10(1):2–12
44. Efroni SS, Dutttagupta RR, Cheng JJ, Dehghani HH, Hoepfner DJD, Dash CC et al (2008) Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* 2(5):437–447
45. Liang G, Taranova O, Xia K, Zhang Y (2010) Butyrate promotes induced pluripotent stem cell generation. *J Biol Chem* 285(33):25516–25521
46. Mali P, Chou B-K, Yen J, Ye Z, Zou J, Doney S et al (2010) Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* 28(4):713–720
47. Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A et al (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* 460(7257):863–868
48. Ang Y-S, Tsai S-Y, Lee D-F, Monk J, Su J, Ratnakumar K et al (2011) Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145(2):5–15

49. Liang G, He J, Zhang Y (2012) Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nat Cell Biol* 14(5):457–466
50. Huangfu DD, Maehr RR, Guo WW, Eijkelenboom AA, Snitow MM, Chen AEA et al (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* 26(7):795–797
51. Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S et al (2008) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* 26(11):1269–1275
52. Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A et al (2011) Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* 8(1):96–105
53. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2):315–326
54. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153):553–560
55. Mattout A, Biran A, Meshorer E (2011) Global epigenetic changes during somatic cell reprogramming to iPS cells. *J Mol Cell Biol* 3(6):341–350
56. Mansour AA, Gafni O, Weinberger L, Zviran A, Ayyash M, Rais Y et al (2012) The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* 488(7411):409–413
57. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P et al (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454(7200):49–55
58. Pawlak M, Jaenisch R (2011) De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev* 25(10):1035–1040
59. Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R et al (2012) Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488(7413):652–655
60. • Kurian L, Sancho-Martinez I, Nivet E, Aguirre A, Moon K, Pendaries C, et al. Conversion of human fibroblasts to angioblast-like progenitor cells. *Nature methods*. 2013 Jan;10(1):77–83. *Demonstrates cells exposed to reprogramming factors obtain plasticity and ability to transdifferentiate before the establishment of pluripotency. Findings have implications for epigenetic status of reprogramming cells*
61. Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A (2008) Promotion of reprogramming to ground state pluripotency by signal inhibition. Goodell MA, editor. *PLoS Biol* 6(10):e253
62. Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D et al (2009) A small-molecule inhibitor of Tgf- β signaling replaces Sox2 in reprogramming by inducing nanog. *Cell Stem Cell* 5(5):491–503
63. • Golipour A, David L, Liu Y, Jayakumaran G, Hirsch CL, Trcka D, et al. A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell*. 2012 Dec 7;11(6):769–82. *This article reports on the regulatory network that controls completion of pluripotency*
64. Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S et al (2012) Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* 11(4):491–504
65. Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D et al (2012) Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* 11(4):477–490