Published in final edited form as:

Stem Cells. 2013 June ; 31(6): 1033–1039. doi:10.1002/stem.1352.

The Sox2-Oct4 Connection: Critical players in a much larger interdependent network integrated at multiple levels

Angie Rizzino^{1,2}

¹Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, 985950 Nebraska Medical Center, Omaha, Nebraska 68198-5950

Abstract

The transcription factors Sox2 and Oct4 have been a major focus of stem cell biology since the discovery, more than 10 years ago, that they play critical roles during embryogenesis. Early work established that these two transcription factors work together to regulate genes required for the self-renewal and pluripotency of embryonic stem cells (ESC). Surprisingly, small changes (~2-fold) in the levels of either Oct4 or Sox2 induces the differentiation of ESC. Consequently, ESC must maintain the levels of these two transcription factors within narrow limits. Genome-wide binding studies and unbiased proteomic screens have been conducted to decipher the complex roles played by Oct4 and Sox2 in the transcriptional circuitry of ESC. Together, these and other studies provide a comprehensive understanding of the molecular machinery that sustains the self-renewal of ESC and restrains their differentiation. Importantly, these studies paint a landscape in which Oct4 and Sox2 are part of a much larger interdependent network composed of many transcription factors that are interconnected at multiple levels of function.

Keywords

embryonic stem cells; iPS cells; Sox2; Oct4; transcription factors; proteomics; gene regulatory networks; transcriptional circuitry; DNA repair; reprogramming; self-renewal

Introduction

Embryonic stem cells (ESC) have had a major impact on the field of stem cell biology since they were first established in 1981. ESC have not only helped revolutionize our understanding of mammalian embryogenesis (1), they have also provided an efficient means of genetically engineering mice (2). Furthermore, their study laid the critical foundation for the discovery of induced pluripotent stem cells (iPSC) (3). ESC are arguably the most extensively characterized mammalian cell, and their study has provided the most coherent understanding of any mammalian cell. Thus far, the study of Sox2 (Sry-box containing gene 2), Oct4 (Pou5f1 - POU domain, class 5, transcription factor 1) and a small group of other essential transcription factors has provided major insights into the transcriptional circuitry responsible for sustaining the self-renewal and pluripotency of ESC. The picture that has emerged from these studies is that Sox2 and Oct4, along with a cadre of many other essential transcription factors, work together as part of a highly integrated network. This interdependent network is responsible for not only sustaining the self-renewal of ESC, but

²Corresponding Author Telephone: 402-559-6338, Fax: 402-559-2813, arizzino@unmc.edu.

Disclosure of Potential Conflicts of Interest

The author indicates no potential conflicts of interest.

Author contribution: AR researched and wrote this review.

also restraining their propensity to differentiate into the cornucopia of different cell types that make up the three embryonic germ layers.

This review provides a concise overview of some of the key advances made in the course of studying Sox2 and Oct4 in pluripotent stem cells, and discusses their participation in a network composed of many essential transcription factors. Much of the discussion revolves around genome-wide binding studies and unbiased proteomic screens conducted with transcription factors that control the fate of ESC. Together, these studies provide a new perspective on the core transcriptional circuitry in ESC and reveal multiple levels of integration (Figure 1). First, many of the transcription factors in this network not only associate with one another, they also associate with many of the same proteins in the network. Second, the transcription factors in this network help regulate the transcription of a high percentage of the genes that code for the proteins with which they associate. Third, new and exciting studies argue that the transcription factors in this network, in particular Sox2 and Oct4, not only regulate gene expression, they are also significant players in DNA repair and/or replication. A fourth level of integration, not reviewed here, is the important connection between signal transduction and the regulation of essential transcription factors. Readers interested in this topic are referred to several excellent reviews (4-6). Given this high degree of integration, it is understandable that the fate of ESC is dramatically altered by small changes in the levels of master regulators, such as Sox2 and Oct4.

Sox2 and Oct4 Levels Strongly Influence the Fate of ESC and the formation of iPSC

Early interest in the cooperative interactions between Sox2 and Oct4 in ESC grew out of a series of studies showing that Sox2 and Oct4 bind cooperatively to two adjacent *cis*-regulatory elements located within a powerful enhancer of the FGF4 gene (7). Subsequent knockout studies demonstrated that FGF4, Oct4 and Sox2 are each essential for mammalian embryogenesis (8-10). However, unlike Sox2 and Oct4, FGF4 is not required for the self-renewal of mouse ESC (11). Instead, FGF4 influences the differentiation of mouse ESC (11). The excitement surrounding Oct4 grew substantially with the surprising discovery that small increases in Oct4 levels (~2-fold) induce ESC to differentiate into cells that express markers of endoderm and mesoderm; whereas, decreasing Oct4 promotes the formation of trophectoderm-like cells (12). Similarly, knocking down Sox2 in ESC promotes their differentiation into trophectoderm-like cells (13), and increasing Sox2 in ESC (~2-fold) induces their differentiation into cells that express markers of ectoderm, mesoderm and trophectoderm, but not endoderm (14).

Collectively, these and other studies made it abundantly clear that both Oct4 and Sox2 behave as master regulators during early mammalian development and that their expression must be tightly controlled in ESC. A major unanswered question is how ESC maintain the expression of these key transcription factors within narrow limits. Answering this question will require significantly more study.

Currently, it is also unclear how small increases in Sox2 and Oct4 trigger the differentiation of ESC cells. One likely explanation is the induction of genes that drive differentiation. For example, the Sox21 gene turns on within 3 hours after Sox2 levels are elevated in ESC (14). Although Sox21 expression does not appear to be the sole reason why ESC differentiate under these conditions, even low levels of Sox21 expression induce the differentiation of ESC (15). Other possible explanations have been examined, including the possibility that elevating either Sox2 or Oct4 in ESC changes the critical ratio between these two master regulators. This was examined by engineering ESC that simultaneously elevate the expression of both Sox2 and Oct4 from an inducible promoter (i-OS-ESC). In contrast to the induction of differentiation observed within 24-36 hrs after elevating the level of Sox2 in ESC, increasing both Sox2 and Oct4 ~2-fold in i-OS-ESC does not rapidly induce

differentiation. However, after 2-3 passages, elevated levels of Sox2 and Oct4 in i-OS-ESC cause a substantial reduction in their cloning efficiency and the appearance of differentiated cells that upregulate a wide-range of markers expressed by cells derived from each of the three embryonic germ layers (16). Thus, changing the relative ratio between Sox2 and Oct4 does not appear to be the primary reason why ESC differentiate when either master regulator is overexpressed; however, it remains to be determined whether it is a contributing factor. Interestingly, ESC engineered to express Oct4, Sox2, Klf4 and c-Myc from an inducible transgene (i-OSKM-ESC) self-renew and do not differentiate when these four factors are each elevated ~2-fold for at least 5 passages (16). Moreover, there were remarkably few changes in the transcriptome of i-OSKM-ESC when the levels of the four transcription factors were increased. Although the roles of Klf4 and c-Myc in preventing the differentiation i-OSKM-ESC have not yet been examined, simply elevating Sox2 and Oct4 expression in ESC does not appear to be the reason why ESC differentiate when the levels of either Sox2 or Oct4 are increased ~2-fold.

The levels at which Sox2 and Oct4 are expressed during reprogramming of somatic cells to iPSC is also critical. Several studies have examined how changes in the stoichiometry of reprogramming factors, in particular the ratio of Sox2 to Oct4, influence both the efficiency of producing iPSC and their quality. Papapetrou et al. (17) observed that increasing the levels of Oct4 (~3-fold), while holding the levels of Sox2, Klf4 and c-Myc constant, slightly increased the frequency of reprogramming; whereas decreasing Oct4 levels (~3-fold) led to a dramatic decrease in the frequency of reprogramming. In contrast, increasing the levels of Sox2 decreased the efficiency of reprogramming. Other investigators reported that decreasing Sox2 levels increased the frequency of producing partially reprogrammed iPSC cells (18). Furthermore, Carey and coworkers (19) demonstrated that elevating Oct4 and decreasing Sox2 substantially improves the ability to generate iPSC able to form “all iPSC mice” after injection into tetraploid blastocysts. This highly stringent test requires that all embryonic tissues are generated from the injected iPSC. Collectively, these studies make it clear that the levels of Sox2 and Oct4 need to be controlled carefully for optimal production of iPSC. However, it remains to be determined how the levels of these transcription factors effect the molecular events that control the efficiency of reprogramming. Given the strict requirement for Sox2 and Oct4 during development, their key roles in ESC and the pronounced differences in reprogramming when their levels are not optimized, additional efforts should be made to determine why small changes in the levels of these two master regulators alter the behavior of pluripotent stem cells.

Sox2 and Oct4 are Part of a Highly Integrated Protein-Protein Interaction Network

Transcription factors do not work in isolation, but function together in complexes composed of a wide array of other proteins. Several studies have shown that Sox2, Nanog, Oct4, Dax1, Zfp281, and Nac1 are each part of large protein complexes (some >1 MDa) (20,21) that contain many other transcription factors, chromatin-modifying machinery and co-activators. Unbiased proteomic screens conducted in mouse ESC have begun to identify the proteins present in these complexes. Wang and coworkers (21) initially performed an unbiased proteomic screen in mouse ESC using epitope-tagged Nanog as bait and identified 17 Nanog-associated proteins. As part of this study, they also identified proteins present in complexes of five Nanog-associated proteins (Oct4, Dax1, Nac1, Zfp281 and Rex1). The virtual protein-protein interaction landscape (“mini-interactome”) generated in this study led to an exciting conclusion. Nanog, Oct4, Dax1, Nac1, Zfp281 and Rex1 not only associate with one another in ESC, they also associate with other transcription factors in common. Thus, there is a high degree of integration at the protein-protein interaction level among transcription factors required for the self-renewal of ESC. More recently, significant support

for this conclusion has been provided by proteomic screens conducted with Sox2, as well as additional Oct4 proteomic screens (16,22-24).

An unbiased proteomic screen identified a set of ~70 proteins that associate with Sox2 in mouse ESC (16). As in the case of the interactomes of Nanog and other transcription factors in ESC (21), Sox2-associated proteins comprise a wide array of proteins, including other transcription factors (e.g. Oct4, Nanog, Klf4, Sall4 and Esrrb), multiple subunits of chromatin-remodeling complexes [e.g. Smarca4, Smarca5 and Smarcd1 of the esBAF complex (25); Gatad2b, Mat1, and Mat2 of the Nurd complex], and novel co-activators (discussed below) (16). In addition to the Oct4 proteomic screen conducted by Wang and coworkers (21), three other proteomic studies have focused on the Oct4-interactome in mouse ESC (22-24). Interestingly, these three studies identified a total of 281 Oct4-associated proteins. Surprisingly, only 18 Oct4-associated proteins were identified in all three studies and only two (Sall1 and Sall4) were also identified in the study by Wang and coworkers. The reasons for such varied results has not been examined. However, differences in experimental designs, in particular the methods used to isolate Oct4 complexes and differences in the mass spectrometry platforms used to identify proteins present in the Oct4 complexes, were likely to be significant contributing factors.

To help understand the interrelationships that exist between the proteins identified in the proteomic studies discussed above, the interactomes of Oct4, Sox2, Nanog, Sall4, Tcfcp2l1, Esrrb, Nac1, Dax1, Zfp281 and Rex1 were integrated to produce a more complete picture of the protein-protein interaction landscape in ESC. A simplified version of the virtual protein-protein interaction landscape is provided here (Figure 2) and a more complete landscape is provided by Gao and coworkers (16). Inspection of this network, and the one generated by Wang and coworkers (21), argue that these 10 nuclear proteins not only associate with one another, they also associate with many of the same proteins in ESC. For example, 28 Sox2-associated proteins were identified in the interactome of at least one other factor, and several Sox2-associated proteins, including Smarcd1, Sall4, Sall1, Esrrb, Gatad2b, Mat1, and Mat2, were identified in at least three other interactomes. This high degree of integration between these 10 transcription factors leads to the conclusion that these 10 nuclear proteins, and undoubtedly many others that have not yet been studied, are likely to heavily influence the function of each other in ESC. Thus, it is not surprising that a small change in the levels of a master regulator, such as Oct4 or Sox2, blocks the self-renewal of ESC and induces their differentiation (12-14).

The protein-protein interaction landscape in ESC has not been adequately mined. Many of the understudied proteins detected in these proteomic screens warrant further investigation, in particular those that associate with several essential transcription factors in ESC. In this regard, we predicted that proteins interacting with more than one essential transcription factor in ESC, such as Smarcd1, are also likely to be required for the self-renewal of ESC, because they are also a key component of the interconnected network (26). Indeed, knocking down Smarcd1 disrupts the self-renewal of ESC and induces differentiation (16). As discussed below, continued mining of these protein-protein interaction landscapes will surely help identify other proteins that are required for the self-renewal and pluripotency of ESC.

Two Additional Levels of Network Integration

Transcription factors required for the self-renewal of ESC are not only part of a highly integrated protein-protein interaction network, these transcription factors are also interconnected with the proteins in this network at two other important levels. In 2005, Boyer and coworkers examined the genome-wide binding of OCT4, SOX2 and NANOG in human ESC (27). Subsequently, the genome-wide binding of a larger set of transcription

factors was examined in mouse ESC. One of these studies focused on the genome-wide binding of Oct4, Sox2, Nanog, Klf4, c-Myc, Dax1, Rex1, Nac1 and Zpf281, as well as the genome-wide distribution of two histone modifications, histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 27 trimethylation (H3K27me3), which are associated with actively transcribed genes and repressed genes, respectively (28). Another study conducted with mouse ESC examined the genome-wide binding of Oct4, Sox2, Nanog, Stat3, Smad1, c-Myc, n-Myc, Esrrb, Tcfcp2l1, Zfx, and E2f1, as well as two transcriptional regulators, the co-activator p300 and the insulator protein CTCF (29).

Collectively, these genome-wide binding studies demonstrated that the core transcriptional circuitry which controls the fate of ESC involves a large set of transcription factors that work together closely. The study by Boyer et al., which used ChIP-Chip, initially identified ~350 genes (out of the ~18,000 genes interrogated) in human ESC that are bound by all three transcription factors - OCT4, SOX2 and NANOG (27). This number increased to more than 600 genes bound by Oct4, Sox2 and Nanog when ChIP-seq was used by Chen and coworkers (29) to determine their genome-wide binding in mouse ESC. One of the reasons for the binding of multiple transcription factors to a large set of genes is the stereo-alignment of DNA binding sites within these genes. Chen and coworkers used a motif-discovery algorithm to identify “enriched DNA sequence motifs” (similar to consensus DNA sequences) for the 13 factors tested in their ChIP-seq studies (29). The enriched DNA sequence motif for Sox2 was found to be very similar to the Oct4 enriched DNA sequence motif. Moreover, the enriched Sox2 motif and the enriched Oct4 motif each contained adjacent Sox2 and Oct4 consensus DNA binding.

Chen and coworkers also reported that ~3,500 genes possess DNA loci (referred to as multiple transcription factor-binding loci - MTL) that bind in close proximity at least 4 of the 13 transcription factors examined in their study (29). Moreover, ~40% of the MTLs that bind at least four transcription factors are located far from the promoters that they are believed to regulate. They also identified ~100 genes with MTLs that bind at least 8 of the 13 transcription factors examined. In the case of the Oct4 gene, 11 of the 13 transcription factors studied (29), including Sox2 and Oct4, bind to one of its regulatory regions (13). More recently, similar conclusions were reached by Martello and coworkers who interrogated publically available data sets (30). They noted that more than 3,100 genomic regions in mouse ESC are occupied by five transcription factors (Oct4, Sox2, Nanog, Tcf3 and Esrrb) known to control the fate of ESC. Further analysis of the data sets generated by these genome-wide binding studies and recent results from the ENCODE Consortium (31) offer excellent opportunities for identifying uncharacterized regulatory regions that control the expression of essential genes in ESC.

Analysis of genome-wide binding studies of Sox2, Oct4 and Nanog leads to another remarkable conclusion and reveals a second level of integration. Sox2, Oct4 and Nanog help regulate a high percentage of the genes that code for the proteins in their interactomes. More specifically, Sox2, Oct4 and Nanog each bind to the regulatory regions of many of the genes that code for the proteins in their respective interactomes (16,21-24). This occurs at levels that are far in excess of what would occur by chance. For example, Sox2 binds to the regulatory regions of >50% of the genes that code for Sox2-associated proteins (16). Furthermore, the regulatory regions of nearly 40% of the genes that code for Sox2-associated proteins are bound by both Sox2 and Oct4, and ~25% of these genes are bound by Sox2, Oct4 and Nanog (16). Similar observations have been made for genes that code for Oct4-associated proteins and Nanog-associated proteins (21-24).

Efforts to mine the protein-protein interaction landscape in ESC have unearthed a third level of integration, in this case between pluripotency-associated transcription factors and other

critical processes in ESC, including DNA repair. The Sox2-interactome in ESC included all three subunits of the Replication Protein A (RPA) complex (RPA1, RPA2 and RPA3), and the Oct4-interactome included two of the three subunits of this complex (RPA1 and RPA3). Importantly, the homologous complex in yeast (Rpa) has been shown recently to participate not only in DNA replication and repair, but also in transcriptional elongation (32). Thus, we predict that RPA is recruited to Sox2/Oct4 target genes, in particular to MTLs that bind Sox2 and Oct4, and facilitates transcription of these genes. Interestingly, the Sox2-interactome also includes two subunits (Rad23 and Ctn2) of the trimeric Xeroderma Pigmentosum group C (XPC) nucleotide excision repair complex (16). Importantly, other studies have shown that XPC associates with both Sox2 and Oct4, and this complex functions as a co-activator able to mediate the transcriptional activation of the Nanog promoter by Sox2 and Oct4 (33). Furthermore, XPC is recruited to a high percentage of gene regulatory regions co-occupied by Sox2 and Oct4. Aside from the roles of DNA repair/replication machinery participating in gene transcription, the proteomic screens conducted in ESC raise a very intriguing question. Do Sox2 and Oct4 work together to help maintain genome stability? A recent study suggests that this is likely to be the case. After radiation-induced DNA damage, Oct4 (GFP-Oct4) was found to accumulate at regions of UV-damage (34). Thus, it appears that pluripotency-associated transcription factors work together to integrate three of the most critical processes in ESC - DNA replication, maintenance of genome stability and transcriptional regulation.

Changes in the Protein-Protein Interaction Network when ESC Differentiate

Understanding the changes that occur in genome-wide binding of transcription factors and in the protein-protein interaction network during the initial stages of differentiation will help provide a clearer picture of the molecular events that accompany the loss of stem cell self-renewal. With the exception of Sox2, little is known about changes in the interactomes of key transcription factors during the initial stages of differentiation. Comparison of the Sox2-interactomes in ESC and ESC beginning to differentiate indicates that the Sox2-interactome changes dramatically within 24 hours. Although the Sox2-interactome in ESC and ESC initiating differentiation both comprise >60 Sox2-associated proteins, only 18 Sox2-associated proteins are present in both interactomes (15,16). Importantly, the comparison of the Sox2-interactomes was conducted using ESC that expressed similar levels of Sox2. Moreover, Sox2 complexes in these two studies were isolated by the same methodology and Sox2-associated proteins were characterized by the same mass spectrometry platform.

Currently, it is unclear why the Sox2-interactome changes so rapidly. Changes in expression of the proteins involved does not appear to be the main contributor (16). However, changes in the post-translational modifications of interacting proteins are attractive candidates. For example, one hour after human ESC initiate differentiation, their phosphoproteome changes by ~50% (35). Interestingly, the phosphorylation status of many of the proteins present in the Sox2-interactome has been shown to change when human ESC initiate differentiation (16). Furthermore, transcription factors, such as Sox2 and Oct4, are modified in ESC by the whole gamut of post-translational modifications, including phosphorylation, acetylation, poly(ADP-ribosyl)ation, methylation, sumoylation and glycosylation (36-42), which are too numerous to discuss here. Thus, it will be important to determine the extent to which changes in post-translational modifications during differentiation affect both protein-protein interactions and the functions of these proteins. It will also be important to determine how the interactomes of other essential transcription factors change during the initial stages of differentiation.

Conclusions and Future Perspectives

It is evident from the studies discussed here that a high degree of interdependence exists between transcription factors that function as master regulators in ESC. These studies provide a more holistic understanding of the mechanisms that control the self-renewal and pluripotency of ESC. In addition to the three levels of integration discussed in this review, there is a fourth level (4-6), not reviewed here, that connects signal transduction and the expression of key transcription factors. Importantly, the significance of an integrated network goes beyond ESC. The overall principles discussed are undoubtedly applicable to cells in general.

There is little question that significant progress has been made in deciphering the networks that control the fate of pluripotent stem cells. However, many fundamental questions remain to be addressed. For example, it will be important to define the changes that occur during the transition from the “naive” ground state, typified by mouse ESC (43), to the “primed” epiblast stem cell (EpiSC) stage, a stage similar to human ESC. Thus far, important changes in signal transduction that occur during this transition have been identified (4). However, relatively little is known about the changes that occur in transcriptional circuitry and in the interactomes of key transcription factors when naive ESC transition to primed EpiSC. Moreover, the interactomes of key transcription factors in human ESC have not been examined. Equally important, it will be highly beneficial to delve more deeply into the roles of reprogramming transcription factors during the generation of iPSC. Although progress has been made in delineating some of the important changes in the epigenome during reprogramming (44,45), examination of the interactomes of key transcription factors during reprogramming is very limited. Preliminary studies in the author’s laboratory indicate that the Sox2-interactome during the early stages of reprogramming is very different from the Sox2-interactome in ESC.

Finally, the recognition that the transcriptional circuitry of ESC is controlled by a highly interdependent network of transcription factors suggests that the typical image of signal transduction warrants updating. Signal transduction is often referred to as signaling pathways that are linear and reasonably well insulated. Crosstalk is thought to be more the exception than the rule. Given the high degree of integration of the transcriptional circuitry that control the fate of ESC, and undoubtedly cells in general, a more realistic view of signal transduction is that of a three-dimensional signaling network with significant crosstalk. This would allow cells to integrate and efficiently coordinate responses to complex environmental cues that affect their fate.

Acknowledgments

Dr. Jesse Cox is thanked for construction of Figure 1. The author also thanks Drs. Timothy McKeithan and Kaustubh Datta for reading this review and providing helpful comments. The author also thanks members of his laboratory (Dr. Phillip Wilder, Erin Wuebben, Briana Ormsbee and Michelle Desler) for reading this manuscript and Heather Rizzino for editorial assistance. The author also acknowledges support from the National Institute of General Medical Sciences (GM 080751).

References

1. Smith AG. Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol.* 2001; 17:435–462. [PubMed: 11687496]
2. Wilder PJ, Rizzino A. Mouse genetics in the 21st century: using gene targeting to create a cornucopia of mouse mutants possessing precise genetic modifications. *Cytotechnology.* 1993; 11:79–99. [PubMed: 7763692]
3. Cox JL, Rizzino A. Induced pluripotent stem cells: what lies beyond the paradigm shift. *Exp Biol Med* (Maywood). 2010; 235:148–158. [PubMed: 20404029]

4. Hirai H, Karian P, Kikyo N. Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. *Biochem J.* 2011; 438:11–23. [PubMed: 21793804]
5. Xue K, Ng J, Hg H. Mapping the networks for pluripotency. *Phil Trans R Soc B.* 2011; 366:2238–2246. [PubMed: 21727129]
6. Singh AM, Dalton S. The Cell Cycle and Myc Intersect with Mechanisms that Regulate Pluripotency and Reprogramming. *Cell Stem Cell.* 2009; 5:141–149. [PubMed: 19664987]
7. Yuan H, Corbi N, Basilico C, et al. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev.* 1995; 9:2635–2645. [PubMed: 7590241]
8. Feldman B, Poueymirou W, Papaioannou VE, et al. Requirement of FGF-4 for postimplantation mouse development. *Science.* 1995; 267:246–249. [PubMed: 7809630]
9. Nichols J, Zevnik B, Anastasiadis K, et al. Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell.* 1998; 95:379–391. [PubMed: 9814708]
10. Avilion AA, Nicolis SK, Pevny LH, et al. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* 2003; 17:126–140. [PubMed: 12514105]
11. Wilder PJ, Kelly D, Brigman K, et al. Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Dev Biol.* 1997; 192:614–629. [PubMed: 9441693]
12. Niwa H, Toyooka Y, Shimosato D, et al. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell.* 2005; 123:917–929. [PubMed: 16325584]
13. Chew JL, Loh YH, Zhang W, et al. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol.* 2005; 25:6031–6046. [PubMed: 15988017]
14. Kopp JL, Ormsbee BD, Desler M, et al. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells.* 2008; 26:903–911. [PubMed: 18238855]
15. Mallanna SK, Ormsbee BD, Iacovino M, et al. Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. *Stem Cells.* 2010; 28:1715–1727. [PubMed: 20687156]
16. Gao Z, Cox JL, Gilmore JM, et al. Determination of Protein Interactome of Transcription Factor Sox2 in Embryonic Stem Cells Engineered for Inducible Expression of Four Reprogramming Factors. *J Biol Chem.* 2012; 287:11384–11397. [PubMed: 22334693]
17. Papapetrou EP, Tomishima MJ, Chambers SM, et al. Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc Natl Acad Sci USA.* 2009; 106:12759–12764. [PubMed: 19549847]
18. Yamaguchi S, Hirano K, Nagata S, et al. Sox2 expression effects on direct reprogramming efficiency as determined by alternative somatic cell fate. *Stem Cell Res.* 2011; 6:177–186. [PubMed: 21130722]
19. Carey BW, Markoulaki S, Hanna JH, et al. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell.* 2011; 9:588–598. [PubMed: 22136932]
20. Cox JL, Mallanna SK, Luo X, et al. Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. *PLoS One.* 2010; 5:e15486. [PubMed: 21103394]
21. Wang J, Rao S, Chu J, et al. A protein interaction network for pluripotency of embryonic stem cells. *Nature.* 2006; 444:364–368. [PubMed: 17093407]
22. van den Berg DL, Snoek T, Mullin NP, et al. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell.* 2010; 6:369–381. [PubMed: 20362541]
23. Pardo M, Lang B, Yu L, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell.* 2010; 6:382–395. [PubMed: 20362542]
24. Ding J, Xu H, Faiola F, et al. Oct4 links multiple epigenetic pathways to the pluripotency network. *Cell Res.* 2012; 22:155–167. [PubMed: 22083510]
25. Ho L, Ronan JL, Staahl BT, et al. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci USA.* 2009; 106:5181–5186. [PubMed: 19279220]

26. Mallanna SK, Rizzino A. Systems Biology Provides New Insights into the Molecular Mechanisms that Control the Fate of Embryonic Stem Cells. *J Cell Physiol.* 2011; 227:27–34. [PubMed: 21412766]
27. Boyer LA, Lee TI, Cole MF, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* 2005; 122:947–956. [PubMed: 16153702]
28. Kim J, Chu J, Shen X, et al. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell.* 2008; 132:1049–1061. [PubMed: 18358816]
29. Chen X, Xu H, Yuan P, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell.* 2008; 133:1106–1117. [PubMed: 18555785]
30. Martello G, Sugimoto T, Diamanti E, et al. Esrrb Is a Pivotal Target of the Gsk3/Tcf3 Axis Regulating Embryonic Stem Cell Self-Renewal. *Cell Stem Cell.* 2012; 11:491–504. [PubMed: 23040478]
31. Wang J, Zhuang J, Iyer S, et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Research.* 2012; 22:1798–1812. [PubMed: 22955990]
32. Sikorski TW, Ficarro SB, Holik J, et al. Subl and RPA associate with RNA polymerase II at different stages of transcription. *Mol Cell.* 2011; 44:397–409. [PubMed: 22055186]
33. Fong YW, Inouye C, Yamaguchi T, et al. A DNA Repair Complex Functions as an Oct4/Sox2 Coactivator in Embryonic Stem Cells. *Cell.* 2011; 147:120–131. [PubMed: 21962512]
34. Bartova E, Sustackova G, Stixova L, et al. Recruitment of Oct4 Protein to UV-Damaged Chromatin in Embryonic Stem Cells. *PLoS One.* 2011; 6:e27281. [PubMed: 22164208]
35. Van Hoof D, Munoz J, Braam SR, et al. Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell.* 2009; 5:214–226. [PubMed: 19664995]
36. Moretto-Zita M, Jin H, Shen Z, et al. Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. *Proc Natl Acad Sci U S A.* 2010; 107:13312–13317. [PubMed: 20622153]
37. Jeong CH, Cho YY, Kim MO, et al. Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. *Stem Cells.* 2010; 28:2141–2150. [PubMed: 20945330]
38. Baltus GA, Kowalski MP, Zhai H, et al. Acetylation of sox2 induces its nuclear export in embryonic stem cells. *Stem Cells.* 2009; 27:2175–2184. [PubMed: 19591226]
39. Gao F, Kwon SW, Zhao Y, et al. PARP1 Poly(ADP-ribosyl)ates Sox2 to Control Sox2 Protein Levels and *FGF4* Expression during Embryonic Stem Cell Differentiation. *J Biol Chem.* 2009; 284:22263–22273. [PubMed: 19531481]
40. Zhao HY, Zhang YJ, Dai H, et al. CARM1 mediates modulation of Sox2. *PLoS One.* 2011; 6:e27026. [PubMed: 22046437]
41. Khidekel N, Ficarro SB, Peters EC, et al. Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. *Proc Natl Acad Sci USA.* 2004; 101:13132–13137. [PubMed: 15340146]
42. Jang H, Kim TW, Yoon S, Choi, et al. O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell.* 2012; 11:62–74. [PubMed: 22608532]
43. Ying QL, Wray J, Nichols J, Battle, et al. The ground state of embryonic stem cell self-renewal. *Nature.* 2008; 453:519–523. [PubMed: 18497825]
44. Sridharan R, Tchieu J, Mason M, et al. Role of the Murine Reprogramming Factors in the Induction of Pluripotency. *Cell.* 2009; 136:364–377. [PubMed: 19167336]
45. Schmidt R, Plath K. The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome Biol.* 2012; 13:251. [PubMed: 23088445]

Figure 1.

The protein-protein interaction network of ESC, which includes Sox2 and Oct4, is part of a highly interdependent signaling and transcription circuitry that integrates all major cellular functions.

Figure 2.

A virtual protein-protein interaction landscape of essential transcription factors in ESC. The interactomes of Sox2, Oct4, Nanog, Sall4, Esrrb, Tcfcp2l1, Zfp281, Dax1, Rex1 and Nac1 in ESC (16,21-23) were integrated into a virtual protein-protein interaction landscape using Cytoscape, an open source platform for integrating and visualizing networks. A more complete protein-protein interaction landscape is provided in Gao and coworkers (16).