Chapter 6
Deciphering Protein Complexes and Protein Interaction Networks for Stem Cell Pluripotency

Jianlong Wang

Abstract Embryonic stem cells (ESCs) hold great promise in regenerative medicine owing to their unique property of unlimited self-renewal while retaining multilineage differentiation capacities. Stem cell biology has been advanced by high throughput genomics and proteomics approaches toward identifying a fuller repertoire of genetic and epigenetic regulatory factors and understanding how they function individually and/or combinatorially in regulating self-renewal and maintaining pluripotency. Proteins function as members of protein complexes and form a myriad of protein-protein interactions in governing proper transcriptional output and cellular identity. Construction of protein-protein interaction networks together with other large datasets such as expression profiles and target gene occupancy is essential in facilitating a comprehensive understanding of the mechanisms of ESC self-renewal and pluripotency. This chapter will summarize current efforts and ongoing progresses in dissecting the protein complexes and mapping the protein interaction networks associated with the major pluripotency factors Nanog, Oct4 and Sox2, and provide guidance for refining the current methodologies and developing new tools for high throughput data generation to further our understanding of stem cell pluripotency.

Keywords ESCs • Self-renewal • Affinity purification • Protein-protein interaction network • Interactome

J. Wang (✉)
Department of Developmental and Regenerative Biology, Black Family Stem Cell Institute, Mount Sinai School of Medicine, Atran Building, AB7-10D 1428 Madison Ave, New York, NY 10029, USA
e-mail: jianlong.wang@mssm.edu

A. Ma'ayan and Ben D. MacArthur (eds.), *New Frontiers of Network Analysis in Systems Biology*, DOI 10.1007/978-94-007-4330-4_6,

6.1 Introduction

Embryonic stem cells (ESCs), derived from pre-implantation blastocyst stage embryos, are endowed with unlimited self-renewal capacity while maintaining multilineage differentiation potential, a property often referred to as pluripotency. The self-renewal and pluripotency characteristics of ESCs make these cells uniquely attractive for cell-based therapies as they will provide unlimited material supply for nearly all types of cells through differentiation, and thus offering great hope in regenerative medicine.

Several transcription factors, notably the homeobox proteins Oct4 [62] and Nanog [13, 58], as well as the HMG box containing Sox2 [3], play fundamental roles in early development and stem cell pluripotency. These key factors act in combination to sustain pluripotency by activating ESC critical factors (including themselves) and repressing differentiation-promoting genes. ESCs are sensitive to the dosage of Nanog [34], Oct4 [64] and Sox2 [8, 43]: enforced expression of Nanog relieves ESCs from the LIF requirement for stem cell maintenance [13], promotes transfer of pluripotency after cell fusion [76], and ensures direct reprogramming of somatic cells to the so-called induced pluripotent stem cells (iPSCs) [77]. In contrast, overexpression of Oct4 drives primitive endoderm differentiation [64], possibly due to direct repression of the Nanog promoter by excessive Oct4 [66]. In addition, a small increase of Sox2 triggers the differentiation of mouse ESCs toward mesoderm and ectoderm lineages [43], presumably through perturbation of Oct4 expression [57]. Dosage sensitivity suggests that the ESC state reflects a balance of multiple transcriptional inputs that are likely exerted through association and dissociation of multiprotein complexes. In specifying lineages, Oct4 and Cdx2 counteract each other's functions to shift the balance between trophectoderm and inner cell mass (ICM) fates [65], and Nanog and Gata6 antagonize each other to define epiblast and primitive endoderm lineages [15]. The ESC state, therefore, is likely to be maintained by the continuous and direct interplay of multiple nuclear factors, acting in cooperative and antagonistic modes.

Recent efforts employing high throughput and genomewide approaches such as microarray [37, 73], chromatin immunoprecipitation [9, 16, 41, 50] and RNAi studies [18, 21, 26, 35] have led to identification of an array of self-renewal regulators and pluripotency factors in ESCs. These studies have enhanced our understanding of how stem cells maintain the unique state of pluripotency and how stem cell-like characteristics can be imposed on somatic cells *via* fusion-based or factor-based reprogramming processes (see review [38]). In a post-genomic era, however, it becomes obvious that the pluripotency machinery of a stem cell is far more complicated than simply the collection of specific transcripts, proteins and target loci of known pluripotency factors. Through protein complex formation, translational and post-translational modifications, and degradation, the functional output of these systems is difficult to predict based solely upon gene/protein expression and/or genomic occupancy. It is clear that consideration of the transcriptome alone offers an incomplete and biased interpretation of the underlying cell biology [33, 52].

In addition, it has been demonstrated that many target loci bound by certain factors in chromatin immunoprecipitations may not have direct biological significance [6, 78, 103]. Vital cellular functions require the coordinated action of a large number of proteins that assemble into an array of multi-protein complexes of distinct composition and structure to regulate the transcription of target genes [16, 41]. A key towards understanding the molecular basis of self-renewal and pluripotency of ESCs lies in mapping the intricate protein interaction networks encompassing many pluripotency factors. Current efforts in identifying the protein constituents of stem cell protein complexes and characterizing the pluripotency interaction networks (interactomes) surrounding several critical pluripotency factors have uncovered new factors in self-renewal signaling pathways and provided a wealth of valuable information on stem cell pluripotency [7, 67, 88, 91]. For this book chapter, I will review the current status and ongoing efforts in deciphering the pluripotency protein complexes and construction of protein interaction networks in ESCs.

6.2 Overview of the Approaches for Affinity Purification of Protein Complexes in ESCs

Mammalian protein complexes have been studied by combining protein affinity purification (AP) with mass spectrometry (MS) and bioinformatics. AP makes use of specific binding interactions between molecules and generally involves three basic steps: (1) incubation and binding reaction; (2) washing to remove nonspecific binding; (3) dissociation and recovery of the bound material. Various AP methods have been developed with modifications to one or few of these three steps to optimize purification of the protein of interest [4, 19].

AP strategies can be broadly classified into two main approaches according to the nature of the target molecule (Fig. 6.1): if the target molecule is the native endogenous protein, then the antibody-based affinity purification is applied; if however, the target molecule is tagged with an epitope such as FLAG [23], which can be introduced into cells either by a knockin strategy or as an ectopic overexpression vector, then affinity purification will be based on the affinity tag. Two (or more) different affinity tags [47] can be used for tandem (sequential) affinity purification, which can increase specificity and reduce the background [72, 97].

There are several advantages associated with the AP-MS method: first, AP-MS can be performed under physiological conditions, in the native organism or cell type; second, it does not typically perturb relevant post-translational modifications (PTMs), which are often crucial for the organization and/or activity of complexes and can also be identified by MS; third, it can be used to probe dynamic changes in the composition of protein complexes when used in combination with quantitative proteomics techniques such as *i*sobaric *t*ag for *r*elative and *a*bsolute *q*uantitation (iTRAQ) [104] and *s*table *i*sotope *l*abeling with *a*mino acids in *c*ell culture (SILAC)[31] (see more in Sect. 6.4.3 and Chap. 5 in this volume).

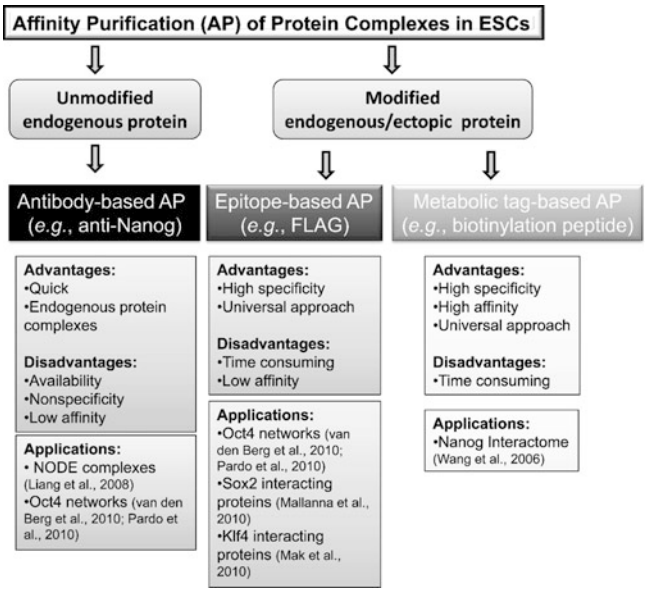

Fig. 6.1 Summary of the strategies for affinity purification of protein complexes in ESCs

6.2.1 Antibody-Based Affinity Purification of Endogenous Protein Complexes

Antibody pull down, also known as immunoprecipitation, is a technique used to isolate a particular protein (and its associated proteins) from solution by means of precipitation. The protein precipitate is formed by coupling the solution with an antibody that specifically binds to the target protein. By isolating a known protein from a complex, other proteins that closely interact with the desired protein may also be pulled out of the complex. Therefore, antibody pull down reveals potential endogenous protein-protein associations. The proteins in a complex may be further separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE) to allow for easy antibody detection of certain known candidate proteins by Western blotting or direct identification of unknown candidate proteins by MS.

The advantages of antibody-based affinity purification are: (1) no transgenic lines need to be established, so the experiment can be done in a speedy manner; and (2) the endogenous protein complexes can be purified, which signifies the biological and functional relevance of the identified interacting proteins. The disadvantages associated with this approach are: (1) most antibodies suffer from non-specific reactivity such that spurious protein complexes will co-purify with the *bona fide* protein complexes; and (2) the affinity between antibody and the target protein is low, and some interactions may be weak and lost during purification procedure. Such a method has been applied to purify endogenous protein complexes of Nanog and Oct4 in mouse ESCs, which yielded a small number of Nanog [48] and Oct4 associated protein

complexes [48, 67, 88] and provided limited and yet still valuable information on Nanog and Oct4 function in stem cell pluripotency (see details in Sect. 6.3.3).

6.2.2 *Epitope Tagging for Affinity Purification of Protein Complexes in ESCs*

Among many epitope tags (see review [94]) that are used for AP studies, the FLAG peptides DYKDDDDK and MDYKDDDDK are the most widely used affinity tags for both immunodetection and AP. The FLAG tag can be multiplied (*e.g.*, 3×FLAG) to increase affinity and specificity and can be placed at either the amino-terminus or carboxyl-terminus and in association with other tags for tandem affinity purification. The FLAG tagged protein and its associated protein complexes can be isolated with anti-FLAG antibody (either in a free form or cross linked with gel matrices such as Protein G-Dynal beads) and eluted with FLAG peptides by competition. The 3×FLAG tagging strategy has been employed by the two recent studies to construct an expanded Oct4 interaction network in mouse ESCs [67, 88] (see Sect. 6.3.2). The main advantage of using the epitope tagging for AP is that it makes AP possible for almost all the proteins of interest, particularly for those that antibodies are not available. This makes high throughput analysis of multiple protein complexes possible. In addition, a relatively higher affinity and specificity of the epitope than that of the endogenous antibody also makes it a favorable choice for AP. However, such advantages of affinity and specificity seem to be dependent on cellular context. Even though the FLAG-based AP has been successfully applied in HeLa [12] and HEK293 [1] cells, its application in ESCs still suffers from high background or non-specific binding, presumably due to the presence of proteins in ESC extracts that are reactive nonspecifically to the FLAG and/or relatively low affinity of the epitope tag and the FLAG antibody. This is manifested by the presence of multiple nonspecific species in Western blotting of ESC lysates or nuclear extracts (our unpublished observation) and presence of Oct4 peptides in control samples of the published Oct4 affinity purification study [88]. Therefore, a refined AP strategy employing FLAG in tandem with a second tag is more often a preferred method. Alternatively, a new metabolic biotin tagging strategy has been developed to complement and improve the AP studies in ESCs [40] (see below).

6.2.3 *Metabolic Biotin Tagging for Affinity Purification of Protein Complexes*

In dissecting the pluripotent state, we have employed *in vivo* biotinylation of critical transcription factors including Nanog and Oct4 in mouse ESCs for affinity purification of protein complexes [89] and ChIP-on-chip [40] for target identification. From these data we have constructed protein-protein [91] and protein-DNA [41]

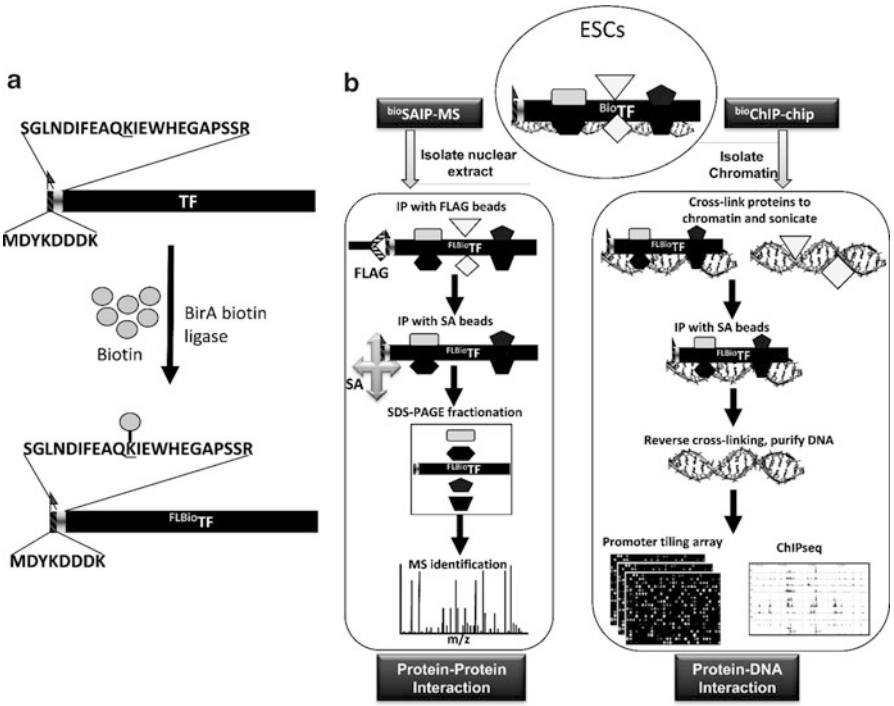

Fig. 6.2 In vivo biotinylation-based strategies for mapping protein-protein interactions in ESCs. (**a**) Modification of a transcription factor (*TF*) with tandem tags (FLAG and Biotinylation Peptide). *E.Coli* BirA biotin ligase catalyzes the addition of biotin to the lysine residue in the biotinylation peptide. (**b**) Affinity methods to capture protein partners (^{bio}SAIP-MS) associated with the biotinylated TF. The ESCs expressing BirA alone will be processed simultaneously as control (not shown). *Triangles* and *diamonds* denote nonspecific binding proteins; other *shapes* denote specific binding proteins

regulatory networks controlling stem cell pluripotency. *In vivo* biotinylation is based on a short 'biotinylation peptide' [75] fused to a protein of interest that serves as an *in vivo* substrate mimic for *E. coli* biotin holoenzyme synthetase (BirA), an enzyme that performs highly selective biotinylation of the fusion protein. In mammalian cells, plasmid expression vectors carrying the biotin-tagged transcription factor (^{bio}TF or ^{FLBio}TF with a Flag-biotin dual tag) and BirA (Fig. 6.2a) can be used to obtain high-level production of soluble ^{bio}TF and BirA proteins, and under appropriate culture conditions, the ^{bio}TF protein produced by this system is completely biotinylated. Studies have documented that biotinylation of a tagged TF does not significantly alter protein interactions, DNA-binding properties *in vivo*, or subnuclear distribution [20]. Therefore, it offers a unique methodology to study protein-protein and protein-DNA interactions simultaneously (Fig. 6.2b).

Biotinylation offers a number of advantages over traditional immunoaffinity approaches for protein complex purification. First, the high affinity of biotin for streptavidin (SA) (10^{-15} M kd) allows efficient purification of the biotinylated protein and associated proteins; second, for ChIP applications, the high affinity allows

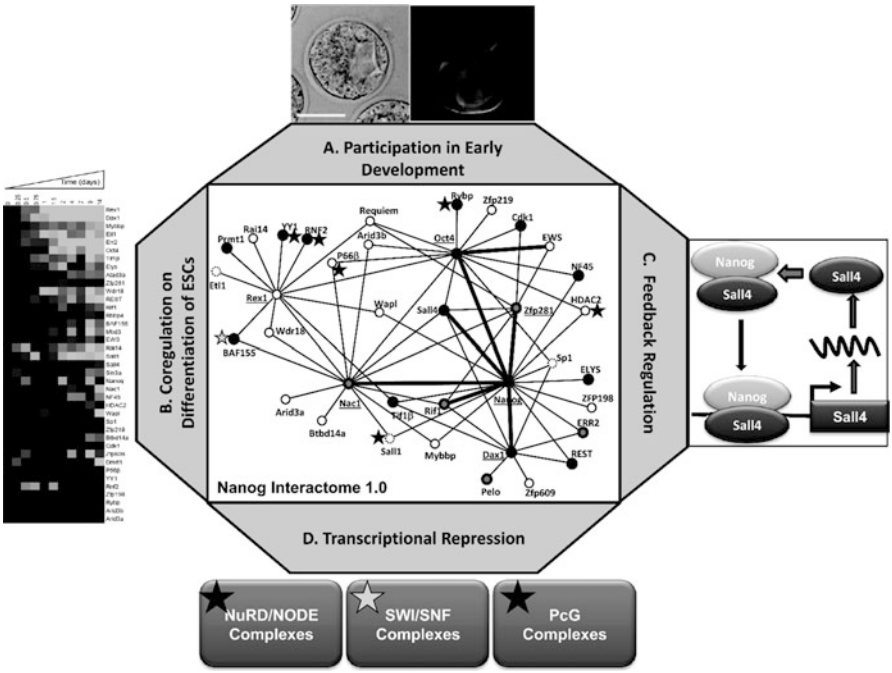

Fig. 6.3 The Nanog interactome for pluripotency of mouse ESCs. Four defining features of the Nanog interactome (labeled **a–d**) for stem cell pluripotency are highlighted. Proteins *underlined* are tagged baits for affinity purification. *Thick lines* indicate interactions confirmed by coIP studies [45, 91, 95]. *Black circles* indicate proteins whose knockout results in defects in proliferation and/or survival of the inner cell mass or other aspects of early development; *gray circles* indicate proteins whose reduction by RNA-mediated interference results in defects in self-renewal and/or differentiation of ESCs; *dotted circles* are proteins whose knockout results in later developmental defects; *white circles* denote proteins for which no loss-of-function data are available

high stringency washing conditions, which reduces background binding that may occur with other affinity tags; third, naturally biotinylated proteins are rare and well-defined [7, 20, 91], and the chance for cross-reaction is quite low; fourth, the approach obviates the need to generate protein-specific antibodies, which often cross-react with other cellular proteins; finally, it is critical that the tagged proteins maintain their functional and structural integrity when expressed. This makes the biotin tag more appealing than other bigger tags (*e.g.*, 3×FLAG) such that structural hindrance from such a small peptide tag is minimal. On the other hand, additional time is required to establish cell lines for *in vivo* biotinylation when compared to direct antibody-based immunoprecipitation, and like epitope tagging, cell lines expressing a controlled level of biotinylated proteins are necessary for analysis [91] since ectopic expression of a protein drastically beyond endogenous levels can result in spurious protein complexes and increases in nonspecific DNA binding. Nevertheless, these concerns are largely outweighed by the superior specificity and highest affinity of the biotin-streptavidin binding which enabled successful construction of "the Nanog interactome" [91] (Fig. 6.3) and an extended transcriptional regulatory network [41] in mouse ESCs.

6.3 Protein Complexes Associated with Nanog, Oct4 and Sox2

6.3.1 The Nanog Interactome

Genetic studies have defined Nanog as a key self renewal regulator that is essential for early development [14, 58] and for ground-state pluripotency of the inner cell mass (ICM) [77] and its *in vitro* derivative ESCs [99]. Nanog is also required for reprogramming of somatic cells to an embryonic pluripotent state [77] and conferring pluripotency to somatic cells upon cell fusion [76]. ESCs lacking Nanog exhibit compromised self-renewal and tend to differentiate toward endodermal lineage. In contrast, enforced expression of Nanog results in enhanced self-renewal at the expense of differentiation propensity [13].

As a divergent homeobox protein, Nanog likely homodimerizes [59, 90] and function in concert with other critical factors such as Oct4 [62] and Sox2 [3]. To further understand the interactive nature of Nanog, we have explored the protein interaction network in which Nanog operates in mouse ESCs. We employed metabolic biotin tagging strategy (Fig. 6.2) for affinity purification of Nanog protein complexes and its associated partner protein complexes followed by MS-based microsequencing [40]. Large-scale purifications were performed with both one-step (streptavidin capture alone) or tandem (FLAG-immunoprecipitation followed by streptavidin capture) [89]. A set of consistent, stringent selection criteria were then applied to each AP-MS experiment to ensure identification of *bona fide* candidate proteins: first, due to the nature of *in vivo* biotinylation, there are background proteins present in both control and tagged samples consisting of mostly naturally biotinylated carboxylases and their associated enzymes as well as some ribosomal proteins. These have been well characterized [20] and thus were removed from the final candidate list; second, some proteins may be identified in both control and tagged samples during the one-step purification. In this case, only candidates with predominantly higher peptide numbers identified by MS in the tagged as compared with the control samples were selected. Alternatively, tandem affinity purification were also performed to confirm such candidates; third, proteins with documented membrane, cytoplasmic, or mitochondrial localization, if present, were excluded; fourth, for proteins specific to tagged samples, only those with ≥ 2 peptides sequenced from at least two independent purifications (either two singles or one single and one tandem) were included in the final candidate list. We have identified a total of 17 proteins of highest confidence that are physically associated with Nanog, either directly or indirectly through other Nanog interacting proteins [91]. In an iterative fashion we then identified partners of several Nanog-associated proteins (including Oct4) and constructed the protein interaction network surrounding Nanog, *i.e.*, the Nanog interactome (Fig. 6.3).

There are four outstanding features associated with the Nanog interactome. First, the network is remarkably enriched for proteins that are required individually to control the survival or differentiation of the ICM or aspects of early embryonic development (Fig. 6.3a). Second, most genes encoding the proteins within the network are co-regulated and, specifically, downregulated during ESC differentiation

based on available microarray data from ESC differentiation studies [69] (Fig. 6.3b). Third, when compared with the target genes of Nanog and Oct4 from the ChIP-on-chip [9] or ChIP-PET [50] studies, a notable feedback regulation mode was evident: many (at least 56%) of the genes encoding the proteins of the network [*e.g.*, Sall4 [95]] are putative Nanog and/or Oct4 targets that also serve as "upstream" effectors to control, either positively or negatively, their own transcriptional regulation (Fig. 6.3c). Fourth, a number of factors in the Nanog interactome with both ESC-specific and ubiquitous expression patterns connect to several epigenetic regulatory pathways. These include the histone deacetylase NuRD (P66β and HDAC2), PRC1 (YY1, RNF2/Ring1B and Rybp) and SWI/SNF chromatin remodelling (BAF155) complexes (Fig. 6.3d). The ESC state is marked by open chromatin and hypertranscription [86] such that multiple loci encoding developmental regulators are often associated with bivalent chromatin marks and poised for imminent activation upon differentiation [5]. Therefore, the repressive machinery embedded within the pluripotency Nanog interactome provides a failsafe mechanism to prevent premature expression of key developmental genes under such dynamically open chromatin conformation of ESCs. Taken together, the Nanog interactome illustrates the requirement for both genetic and epigenetic regulatory control of ESC pluripotency [63] and highlights the importance of transcriptional repression for stem cell pluripotency [17].

Functional studies using RNAi by other groups subsequently confirm several proteins within the Nanog interactome for their function in ESC maintenance. For example, Err2 (Esrrb) [36, 50], Rif1 [50], Sall4 [24, 49, 98, 102], Dax1 (Nr0b1) [61, 79] were individually confirmed by other candidate approaches. In addition, genome wide RNAi studies [21, 26, 35] further validate a number of other network proteins for their roles in self-renewal and pluripotency of ESCs. These studies highlight the efficiency and validity of the biotin-mediated AP-MS strategy for studying protein-protein interactions in ESCs. On the other hand, subsequent studies identified additional Nanog interacting proteins that are not present in the Nanog interactome, such as Smad1 [80], NFκB [85] and Med12 [87], suggesting that our AP method is non-saturating. Alternatively, the purification condition, particularly the high salt (350 mM NaCl) used in our initial study may be biased toward purifying stronger interacting proteins. Therefore, future optimization of affinity purification conditions will be needed to maximize AP strategies for proteomic studies in ESCs, and future studies using improved AP conditions or complementary AP approaches may uncover new partners of Nanog and reveal novel functions for Nanog in stem cell pluripotency.

6.3.2 The Oct4 Interactome

Oct4 is vital to the development of an embryo during its early stages of differentiation into somatic cells. Early studies found that the loss of Oct4 expression in mouse embryos causes lack of development of the ICM of the zygote [62]. Oct4 interacts with multiple core pluripotency TFs and connects with several epigenetic regulators in the Nanog interactome [91] (Fig. 6.3). Growing evidence

suggests that Oct4 is the key player for genetic and epigenetic regulation of stem cell pluripotency. First, Oct4 is an essential factor that functions alone or with other ESC TFs for factor-based somatic cell reprogramming [68, 81, 82, 100]. Second, it cooperates with Sox2 and Nanog to repress *Xist* [60] and interacts with CTCF to activate *Tsix* [22] in coupling X inactivation reprogramming to the control of pluripotency during embryogenesis. Third, it interacts with Eset, the histone H3K9 methyltransferase, to restrict extra-embryonic trophoblast lineage potential in ESCs [101]. Fourth, it controls the chromatin architecture of ESCs through direct regulation of downstream targets encoding histone H3K9 demethylases Jmjd1a and Jmjd2c, which function in part by modulating H3K9 methylation of the pluripotency factors Tcl1 and Nanog, respectively [51].

While the genomic loci occupied by Oct4 have been extensively identified in both mouse [10, 41, 50] and human [9] ESCs, the spectrum of Oct4 interaction partners is underexplored. While the Nanog interactome encompasses Oct4 and several Oct4-interacting proteins, the Oct4-centered protein interaction network had not been fully explored until two recent complementary studies [67, 88]. These two studies both employed the 3×FLAG tag for affinity purification. The van den Berg study [88] took advantage of the ZHBTc4 ESCs [64] and established an ESC line that expresses 3×FLAG-Oct4 in the absence of endogenous Oct4 (Fig. 6.4a). The Pardo study [67] employed BAC transgenesis to introduce a tandem tag (3×FLAG and calmodulin binding peptide-CBP) fused to Oct4 under the endogenous Oct4 regulatory elements (Fig. 6.4) at a sub-endogenous level (30% wt) and integrated it in a precise location in the mouse genome. Again, as we pointed out for the Nanog interactome study, the current two Oct4 network studies are also limited: there are 54 and 92 Oct4-interacting proteins identified, respectively, by the two studies with mere 20 common proteins (Fig. 6.4). This is largely due to the different experimental platforms and conditions employed by the two studies. It is unclear at this point whether the *bona fide* Oct4-centered interactome should be constructed as a union or intersection of the two data sets [46]. Therefore, additional complementary AP-MS studies are needed to resolve the issue. Nevertheless, several consistent features are associated with the two Oct4-centered interactomes: first, the Oct4 interactome is also enriched for factors critically important for ESC maintenance and early development; second, the majority of the genes encoding these two Oct4 interactome proteins are co-regulated, and particularly downregulated upon ESC differentiation; third, expression of the majority of Oct4-interacting proteins is controlled by Oct4 and other key ESC transcription factors, suggestive of the intricate linkage between transcriptional regulatory networks and protein interaction networks; fourth, both Oct4 networks connect with epigenetic regulatory complexes such as NuRD and SWI/SNF (Fig. 6.4b), which is consistent with the findings in the Nanog interactome (Fig. 6.3). Overall, these data suggest that the Nanog and Oct4 interactomes are inherently connected. Although the detection of Nanog in the two Oct4 interactome studies is either not present [88] or minimal [67], Oct4 has been confirmed to be one of the interacting proteins of Nanog in the Nanog interactome [91] and in the endogenous Nanog protein complexes [48].

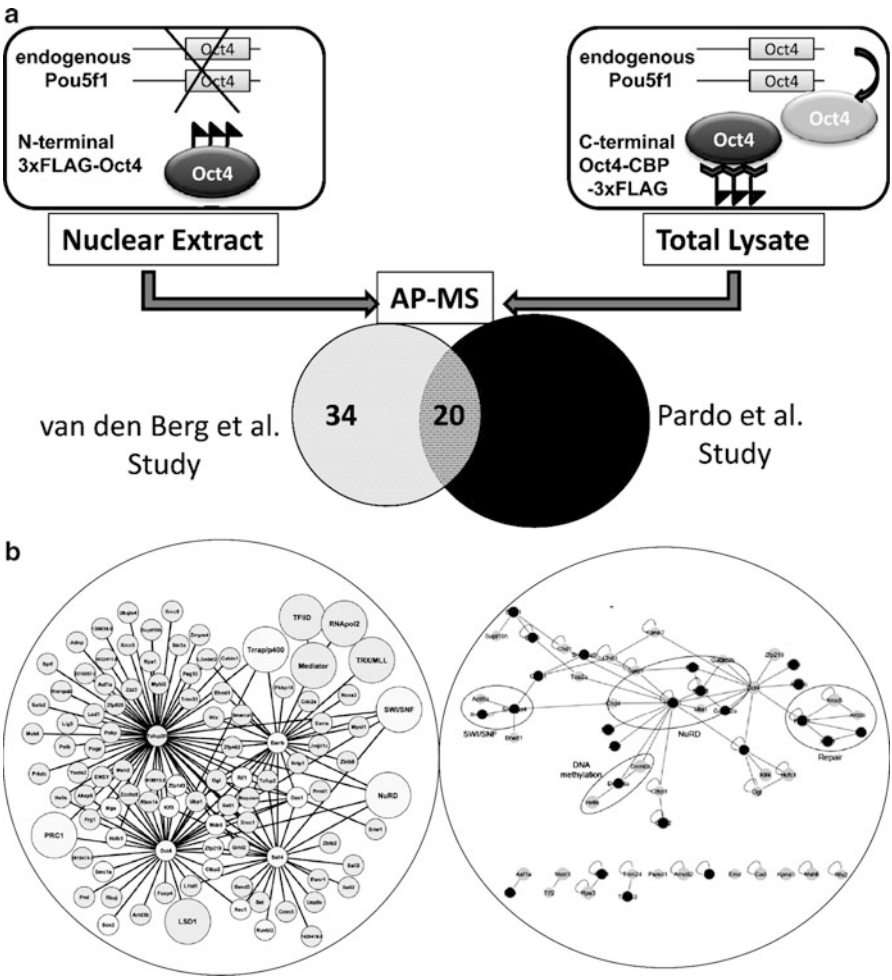

Fig. 6.4 The Oct4 interactome for pluripotency of mouse ESCs. (**a**) Summary of the strategies and results of the two Oct4 network studies [67, 88]. Note the limited overlap between the two studies. (**b**) Network presentation of the two Oct4 interactomes (Reproduced from Cell Stem Cell with permission)

6.3.3 The Endogenous Nanog and Oct4 Protein Complexes

Due to their essential roles in regulating self-renewal and pluripotency of ESCs, the two homeodomain transcription factors Nanog and Oct4 are under intensive study and their endogenous protein complexes have also been sought by antibody-based affinity purification strategies [48, 67, 88]. The study by Liang et al. [48] showed that both Nanog and Oct4 interact with each other and associate with proteins from multiple repression complexes, such as NuRD, Sin3A and Pml complexes.

Surprisingly, of the various core components in the NuRD complex with which Nanog and Oct4 interact, Mta1 was preferred, whereas Mbd3 and Rbbp7 were either absent or present at a sub-stoichiometric levels [48]. The so-called Nanog and Oct4 associated deacetylase (NODE) contains histone deacetylase (HDAC) activity that is comparable to NuRD, and retains its association with Nanog and Oct4 in $Mbd3^{-/-}$ ESCs [48]. However, the NODE complex may simply be the byproduct of the insensitivity of the affinity approach and suboptimal immunoprecipitation condition employed in that study, which was manifested by the limited number of Nanog- and Oct4-interacting proteins identified [48]. The association of Nanog and Oct4 with the conventional NuRD rather than the peculiar NODE repression complex in ESCs was supported by the two independent Oct4 immunoprecipitation studies that yielded a greater number of candidate interacting proteins in the Oct4 complexes including Mbd3 and several other NuRD components with high confidence [67, 88]. Nevertheless, it is clear that Nanog and Oct4 associate with multiple repressor complexes to regulate target gene expression and control ESC fate [48, 91].

The number of candidate interacting proteins identified with such an antibody-based IP approach from all the three studies (Fig. 6.5) is much smaller than that from the epitope tagging strategies (Fig. 6.4), due likely to aforementioned inherent limitation associated with the antibody-based AP, *i.e.*, low affinity and high non-specificity (Fig. 6.1). Therefore, it is important to bear in mind that antibody-based AP should serve only as a complementary approach, but not a gold standard, for identification and validation of potential interacting proteins of a certain protein of interest.

6.3.4 The Sox2 Protein Complexes in ESCs

Although Sox2 is one of the core ESC factors (together with Nanog and Oct4), the mechanism by which Sox2 controls the fate of ESCs is much less well defined. Like Nanog and Oct4, Sox2 expression is also dosage sensitive for stem cell maintenance [43], which highlights the need to decipher Sox2 protein complexes and its interaction network in understanding stem cell pluripotency. Sox2 is well known for its close partnership with Oct4 in target gene regulation [74, 93], however, the evidence for their physical association in protein complexes remains elusive. In one study, Sox2 has been co-purified with Oct4 as part of an expanded Oct4 interaction network [88], however, Sox2 has not been identified in another related Oct4 network study [67] or in the Nanog interactome containing Oct4 [91]. This may reflect the non-saturating nature of affinity purification approaches or the weak interaction between Oct4 and Sox2. Alternatively, there could be only substoichiometric levels of Sox2 in the Oct4 protein complexes and *vice versa*.

Small increase of Sox2 expression level in ESCs promotes their differentiation [43], which compromises the epitope-based AP strategy with ectopic expression of tagged Sox2 in ESCs. Therefore, future endeavor to dissect Sox2 protein complexes in ESCs will require careful manipulation of the ectopic expression of

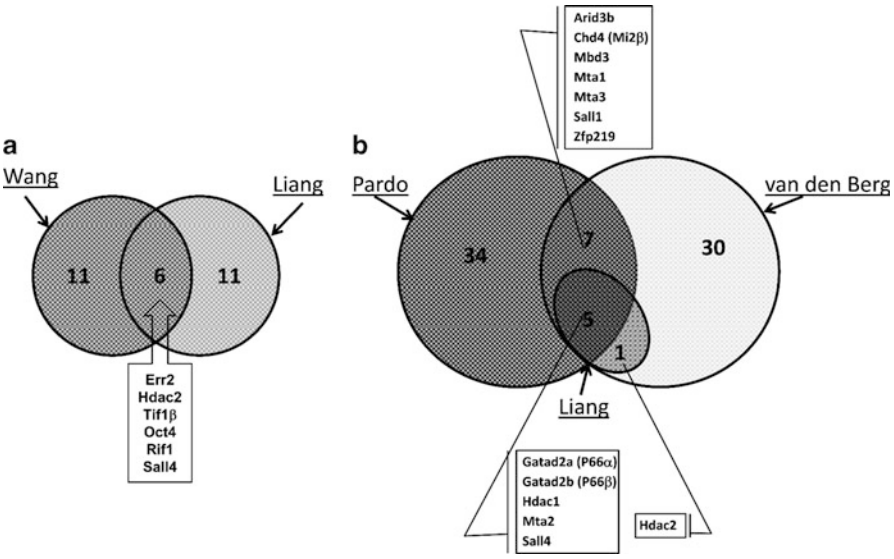

Fig. 6.5 Comparison of the Nanog and Oct4 interacting proteins identified by multiple studies. (**a**) Nanog-interacting proteins; (**b**) Oct4 interacting proteins. Data presented are from four published studies: Wang [91], Liang [48], Pardo [67] and van den Berg [88]

the epitope-tagged Sox2 or adoption of a strategy employing the transgenic rescue of the Sox2 knockout ESCs with the epitope-tagged Sox2. Interestingly, the proteomics studies of Sox2 during early differentiation of ESCs has identified >60 nuclear proteins that associate with Sox2, and a significant number of the identified Sox2-associated proteins also interact with Oct4 and Nanog [56]. While the study may have captured many Sox2-interacting proteins that are required for stem cell maintenance and yet are still active during early stage of differentiation, future studies purifying Sox2 complexes and mapping the Sox2-centered interactome in ESCs are needed to gain a comprehensive understanding of Sox2 function in regulating genes required for self-renewal and pluripotency of ESCs.

6.4 Future Directions

Stem cell research has evolved in the post-genomics era with increasing application of proteomics approaches [92] and high throughput systems biology approaches to define lists of molecular "parts" and regulatory interactions between the "parts" in both undifferentiated ESCs and their differentiated progenies [53, 54, 96]. The analysis of protein complexes and protein-protein interactions is essential for understanding mechanisms of ESC self-renewal and pluripotency. Deploying MS-based proteomics approaches such as AP-MS to decipher protein complexes and dissect protein interaction networks surrounding key pluripotency transcription factors will likely be one of the major scientific inquires in the near future.

6.4.1 High Throughput Experimental System for Network Reconstruction in ESCs

Proteins are much more chemically and structurally diverse than their nucleic acid counterparts, making them intrinsically less suitable for large-scale high-throughput analyses. Large scale protein-protein interaction studies using a high throughput and systematic AP-MS approach were largely confined to cellular systems that are amenable to efficient transfection or genetic engineering such as *Escherichia coli* [11], *Saccharomyces cerevisiae* [27, 44] and human HEK293 cells [25, 29, 39]. The current AP-MS strategies (one gene-one protein-one ESC line) employed in the ESC research [67, 88, 91] are still quite low throughput and time consuming. Considering the exponential growth of the "part-lists" of stem cell factors identified during the past few years owing to the development of microarrays and deep sequencing as well as genome wide RNAi screens, improved AP-MS strategies will have to be developed to meet the demand of high throughput network reconstruction. Future efforts should be directed to develop an integrated strategy for AP-MS in ESCs with the following features: (1) rapid generation of multiple ESC lines bearing epitope tagged baits of interests. Combined lentiviral technology with Gateway cloning technology will expedite cloning of cDNA compatible with publicly and commercially available cDNA libraries and establish stable ESC lines by lentivirus infection [55]; (2) increased yields in protein complex preparation for affinity purification. Due to the dosage sensitivity of many key stem cell factors, the ectopic expression of epitope tagged baits is controlled in a minimal level to preserve the functional integrity of ESCs and avoid formation of the spurious protein complexes due to overexpression. Therefore the total amount of the protein complexes associated with the tagged baits is also limited. A combined vector system allowing tagged bait cDNA expression with simultaneous knockdown of the endogenous protein will be one of the options for increasing recovery of the protein complexes associated with the bait. The added bonus of this transgenic rescue strategy is that the functionality of the tagged bait is also confirmed; (3) direct liquid chromatography tandem MS (LC-MS/MS) analysis of purified complexes to improve the sensitivity and reproducibility of protein identification. The current AP-MS workflow with SDS-PAGE fractionation before MS analysis (Fig. 6.2b) presents a major experimental bottleneck in large-scale and high throughput studies of the protein complexes in ESCs. Development of a strategy to remove endogenously biotinylated proteins (the major background proteins of high abundance in the *in vivo* biotinylation system) and optimization of the affinity purification condition for chemical compatibility of the sample with subsequent LC-MS/MS will be necessary to reach this goal.

6.4.2 Protein-Protein Interactions and Stem Cell Heterogeneity

Heterogeneity is a hallmark of ESCs that might have evolved as a mechanism that enables stem cells to respond to differentiation-inducing signals while retaining

their self-renewal potential [30]. ESCs under undifferentiated culture condition show fluctuating expression levels of a number of genes including stem cell specific factors such as Nanog, Rex1, Dppa3 (Stella), Pecam1, Zscan4 and genes normally associated with cell differentiation such as Brachyury/T and Twist2 (see review [83]). The current methodology in mapping the interactome only averages interaction profiles over a large quantity of individual cells with heterogeneous or stochastic expression of some network proteins such as Nanog and Rex1 in the Nanog interactome. Therefore, it is quite possible that, although the extensive protein interactions will likely be present in individual cells, not all of the identified interactions operate within any given individual cell. For example, in pluripotent Nanog null ESCs [14] the Nanog interactome may be compensated by other factors and the ESC state is stabilized by other interactomes such as the Oct4 interactome. It is not currently possible to construct interactomes at the single cell level; however, one could interrogate limited protein-protein interactions from the existing interactome using fluorescence resonance energy transferor related techniques combined with imaging to gain insights into protein-protein interactions and functional stem cell heterogeneity in single cells. In this regard, implementation and improvement of single cell analysis platforms [32, 84] (also see Chap. 1 Review by Arai in this book) will be the necessary step toward this goal.

6.4.3 *Quantitative Protein-Protein Interactions for Stem Cell Pluripotency*

ESCs are sensitive to the dosage of Nanog [34], Oct4 [64] and Sox2 [8, 43]. As a key pluripotency factor, Nanog has been demonstrated to form dimers [59, 90], and Nanog dimerization is critical for interaction with multiple pluripotency network proteins [90]. However, current interactome studies cannot distinguish the mode of action of these dosage sensitive TFs relative to other factors, and it is also unknown whether the monomers or dimers of Nanog form different protein complexes in controlling subset of genes important for the pluripotent state. In addition, Oct4 and Sox2 often act together to regulate a subset of target genes, however, the stoichiometric level of each partner in the Oct4-Sox2 protein complexes cannot be defined by current AP-MS approaches. Reprogramming somatic cells to naïve pluripotency is associated with erasure of epigenetic memory [42, 71] concomitant with gradual increase of Nanog, Oct4 and Sox2 expression and consolidation of the core transcriptional regulatory network [71] that is interdependent on the pluripotency interactome. How dosage sensitivity alters composition of protein complexes and transcriptional regulation of genes controlling stem cell pluripotency remains to be explored. The future efforts to refine the biochemical approaches to isolate distinct protein complexes in combination with quantitative MS technology such as iTRAQ for relative quantitation [70, 104] (see Chap. 5 in this volume for more on iTRAQ) and MRM/MS (*m*ultiple *r*eaction *m*onitoring-MS) [2, 28] for absolute

quantitation will be necessary to decipher quantitative relationship of the interacting proteins. In particular, the MRM/MS assay quantifies a specific tryptic peptide that is selected as a stoichiometric representative of the cleaved protein against an internal synthetic stable isotope-labeled peptide, allowing for the absolute measurement of protein concentration [2, 28].

6.4.4 Interactome Dynamics During Stem Cell Fate Changes

The pluripotency interactome so far provides static "snapshots" of undifferentiated ESC state. The exciting therapeutic and regenerative potential of ESCs will only be realized during their fate change to generate more differentiated progenies. Molecular regulation of stem cell fate entails a complex and coordinated action among multilayered regulatory pathways (DNA, mRNA/miRNA and protein) that eventually converge in dynamic protein expression and intricate protein interactions governing distinct cellular identity. Therefore, the interactomes during stem cell fate change are dynamic, and methodologies need to be developed in the future to measure interactome dynamics during cell fate change so as to maximize generation of specific cell types for therapeutic application. One recent study has provided a powerful experimental system to interrogate the Nanog interactome dynamics after shRNA-mediated downregulation of Nanog in ESCs [52]. In addition, another study employed an inducible system to drive Sox2 overexpression and monitor the composition of ectopic Sox2-associated protein complexes during early stage of ESC differentiation, which identified another Sox family protein, Sox21, as a novel regulator for stem cell fate [56]. However, a direct comparison with Sox2 network in the pluripotent state and its dynamics during cell fate change is not possible due to the lack of the Sox2 interactome in ESCs.

6.5 Concluding Remarks

Recent efforts in systematically profiling gene expression in ESCs and global genome-wide functional RNAi screens [18, 21, 26, 35] have yielded a wealth of high throughput data and provided a long molecular 'parts list' of regulatory factors important for ESC self-renewal and pluripotency. How these ever-growing 'molecular building blocks' are interconnected into functional regulatory networks that are ultimately responsible for self-renewal and differentiation of ESCs is unclear. The proteomic studies deciphering protein complexes and protein-protein interactions in ESCs will continue to be an area of active research that likely will bear fruit in the near future.

Acknowledgement The author thanks the editors of this book chapter and all members in the lab for critical reading of the manuscript. The author's current research is supported by a Seed Fund from the Black Family Stem Cell Institute at Mount Sinai School of Medicine, an IDEA grant (C026420) from the Empire State Stem Cell Fund through New York State Department of Health (NYSTEM), and a grant from the NIH (1R01-GM095942-01A1).

Notes While this book chapter is in the final production stage, a more sophisticated Oct 4 interactome has been published by the author's group (Ding et al., Cell Research 22:155–167, 2012; PMID 22083510).

References

1. Adams DG, Wadzinski BE (2007) Isolation and characterization of PP2A holoenzymes containing FLAG-tagged B subunits. Methods Mol Biol (Clifton, NJ) 365:101–111
2. Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5:573–588
3. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 17:126–140
4. Bauer A, Kuster B (2003) Affinity purification-mass spectrometry. Powerful tools for the characterization of protein complexes. Eur J Biochem 270:570–578
5. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326
6. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE et al (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816
7. Blelloch R, Wang Z, Meissner A, Pollard S, Smith A, Jaenisch R (2006) Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells (Dayton, Ohio) 24:2007–2013
8. Boer B, Kopp J, Mallanna S, Desler M, Chakravarthy H, Wilder PJ, Bernadt C, Rizzino A (2007) Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2:Oct-3/4 target genes. Nucleic Acids Res 35:1773–1786
9. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG et al (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122:947–956
10. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK et al (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349–353
11. Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N et al (2005) Interaction network containing conserved and essential protein complexes in Escherichia coli. Nature 433:531–537
12. Carey MF, Peterson CL, Smale ST (2010) Purification of mediator from HeLa cell lines expressing a flag-tagged mediator subunit. Cold Spring Harb Protoc 2010:pdb prot5451
13. Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113:643–655
14. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A (2007) Nanog safeguards pluripotency and mediates germline development. Nature 450:1230–1234
15. Chazaud C, Yamanaka Y, Pawson T, Rossant J (2006) Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell 10:615–624
16. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J et al (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133:1106–1117

17. Cheong CY, Lufkin T (2010) Transcriptional repression in ES cells. J Cell Biochem 110:288–293
18. Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS et al (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature 468:316–320
19. Collins MO, Choudhary JS (2008) Mapping multiprotein complexes by affinity purification and mass spectrometry. Curr Opin Biotechnol 19:324–330
20. de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, Heck A, Grosveld F, Strouboulis J (2003) Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Proc Natl Acad Sci USA 100:7480–7485
21. Ding L, Paszkowski-Rogacz M, Nitzsche A, Slabicki MM, Heninger AK, de Vries I, Kittler R, Junqueira M, Shevchenko A, Schulz H et al (2009) A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. Cell Stem Cell 4:403–415
22. Donohoe ME, Silva SS, Pinter SF, Xu N, Lee JT (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature 460:128–132
23. Einhauer A, Jungbauer A (2001) The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. J Biochem Biophys Methods 49:455–465
24. Elling U, Klasen C, Eisenberger T, Anlag K, Treier M (2006) Murine inner cell mass-derived lineages depend on Sall4 function. Proc Natl Acad Sci USA 103:16319–16324
25. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M et al (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 3:89
26. Fazzio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell 134:162–174
27. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM et al (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141–147
28. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100:6940–6945
29. Glaser S, Schaft J, Lubitz S, Vintersten K, van der Hoeven F, Tufteland KR, Aasland R, Anastassiadis K, Ang SL, Stewart AF (2006) Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development 133:1423–1432
30. Graf T, Stadtfeld M (2008) Heterogeneity of embryonic and adult stem cells. Cell Stem Cell 3:480–483
31. Gruhler S, Kratchmarova I (2008) Stable isotope labeling by amino acids in cell culture (SILAC). Methods Mol Biol (Clifton, NJ) 424:101–111
32. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell 18:675–685
33. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19:1720–1730
34. Hatano SY, Tada M, Kimura H, Yamaguchi S, Kono T, Nakano T, Suemori H, Nakatsuji N, Tada T (2005) Pluripotential competence of cells associated with Nanog activity. Mech Dev 122:67–79
35. Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 23:837–848
36. Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR (2006) Dissecting self-renewal in stem cells with RNA interference. Nature 442:533–538

37. Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR (2002) A stem cell molecular signature. Science (New York, NY) 298:601–604
38. Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132:567–582
39. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J et al (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol Cell 27:262–274
40. Kim J, Cantor AB, Orkin SH, Wang J (2009) Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. Nat Protoc 4:506–517
41. Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132:1049–1061
42. Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI et al (2010) Epigenetic memory in induced pluripotent stem cells. Nature 467:285–290
43. Kopp JL, Ormsbee BD, Desler M, Rizzino A (2008) Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. Stem Cells (Dayton, Ohio) 26:903–911
44. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP et al (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature 440:637–643
45. Lee J, Rhee BK, Bae GY, Han YM, Kim J (2005) Stimulation of Oct-4 activity by Ewing's sarcoma protein. Stem Cells (Dayton, Ohio) 23:738–751
46. Lemischka IR (2010) Hooking up with Oct4. Cell Stem Cell 6:291–292
47. Li Y (2010) Commonly used tag combinations for tandem affinity purification. Biotechnol Appl Biochem 55:73–83
48. Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D et al (2008) Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol 10:731–739
49. Lim CY, Tam WL, Zhang J, Ang HS, Jia H, Lipovich L, Ng HH, Wei CL, Sung WK, Robson P et al (2008) Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell 3:543–554
50. Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J et al (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38:431–440
51. Loh YH, Zhang W, Chen X, George J, Ng HH (2007) Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev 21:2545–2557
52. Lu R, Markowetz F, Unwin RD, Leek JT, Airoldi EM, MacArthur BD, Lachmann A, Rozov R, Ma'ayan A, Boyer LA et al (2009) Systems-level dynamic analyses of fate change in murine embryonic stem cells. Nature 462:358–362
53. MacArthur BD, Ma'ayan A, Lemischka IR (2008) Toward stem cell systems biology: from molecules to networks and landscapes. Cold Spring Harb Symp Quant Biol 73:211–215
54. Macarthur BD, Ma'ayan A, Lemischka IR (2009) Systems biology of stem cell fate and cellular reprogramming. Nat Rev Mol Cell Biol 10:672–681
55. Mak AB, Ni Z, Hewel JA, Chen GI, Zhong G, Karamboulas K, Blakely K, Smiley S, Marcon E, Roudeva D et al (2010) A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. Mol Cell Proteomics 9:811–823
56. Mallanna SK, Ormsbee BD, Iacovino M, Gilmore JM, Cox JL, Kyba M, Washburn MP, Rizzino A (2010) 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 (Dayton, Ohio) 28:1715–1727

57. Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA et al (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol 9:625–635
58. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113:631–642
59. Mullin NP, Yates A, Rowe AJ, Nijmeijer B, Colby D, Barlow PN, Walkinshaw MD, Chambers I (2008) The pluripotency rheostat Nanog functions as a dimer. Biochem J 411:227–231
60. Navarro P, Chambers I, Karwacki-Neisius V, Chureau C, Morey C, Rougeulle C, Avner P (2008) Molecular coupling of Xist regulation and pluripotency. Science (New York, NY) 321:1693–1695
61. Niakan KK, Davis EC, Clipsham RC, Jiang M, Dehart DB, Sulik KK, McCabe ER (2006) Novel role for the orphan nuclear receptor Dax1 in embryogenesis, different from steroidogenesis. Mol Genet Metab 88:261–271
62. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95:379–391
63. Niwa H (2007) How is pluripotency determined and maintained? Development 134:635–646
64. Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 24:372–376
65. Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J (2005) Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell 123:917–929
66. Pan G, Li J, Zhou Y, Zheng H, Pei D (2006) A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. FASEB J 20:1730–1732
67. Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J (2010) An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell 6:382–395
68. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ (2008) Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451:141–146
69. Perez-Iratxeta C, Palidwor G, Porter CJ, Sanche NA, Huska MR, Suomela BP, Muro EM, Krzyzanowski PM, Hughes E, Campbell PA et al (2005) Study of stem cell function using microarray experiments. FEBS Lett 579:1795–1801
70. Pewsey E, Bruce C, Tonge P, Evans C, Ow SY, Georgiou AS, Wright PC, Andrews PW, Fazeli A (2010) Nuclear proteome dynamics in differentiating embryonic carcinoma (NTERA-2) cells. J Proteome Res 9:3412–3426
71. Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, Apostolou E, Stadtfeld M, Li Y, Shioda T et al (2010) Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol 28:848–855
72. Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods (San Diego, Calif) 24:218–229
73. Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA (2002) "Stemness": transcriptional profiling of embryonic and adult stem cells. Science (New York, NY) 298:597–600
74. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P (2005) Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem 280:24731–24737
75. Schatz PJ (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. Biotechnology (N Y) 11:1138–1143
76. Silva J, Chambers I, Pollard S, Smith A (2006) Nanog promotes transfer of pluripotency after cell fusion. Nature 441:997–1001

77. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A (2009) Nanog is the gateway to the pluripotent ground state. Cell 138:722–737
78. Struhl K (2007) Transcriptional noise and the fidelity of initiation by RNA polymerase II. Nat Struct Mol Biol 14:103–105
79. Sun C, Nakatake Y, Akagi T, Ura H, Matsuda T, Nishiyama A, Koide H, Ko MS, Niwa H, Yokota T (2009) Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. Mol Cell Biol 29:4574–4583
80. Suzuki A, Raya A, Kawakami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodriguez-Esteban C, Izpisua Belmonte JC (2006) Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proc Natl Acad Sci USA 103:10294–10299
81. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872
82. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676
83. Tanaka TS (2009) Transcriptional heterogeneity in mouse embryonic stem cells. Reprod Fertil Dev 21:67–75
84. Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA (2010) Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell 6:468–478
85. Torres J, Watt FM (2008) Nanog maintains pluripotency of mouse embryonic stem cells by inhibiting NFkappaB and cooperating with Stat3. Nat Cell Biol 10:194–201
86. Turner BM (2008) Open chromatin and hypertranscription in embryonic stem cells. Cell Stem Cell 2:408–410
87. Tutter AV, Kowalski MP, Baltus GA, Iourgenko V, Labow M, Li E, Kadam S (2009) Role for Med12 in regulation of Nanog and Nanog target genes. J Biol Chem 284:3709–3718
88. van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA (2010) An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell 6:369–381
89. Wang J, Cantor AB, Orkin SH (2009) Tandem affinity purification of protein complexes in mouse embryonic stem cells using in vivo biotinylation. Curr Protoc Stem Cell Biol Chapter 1:Unit1B.5
90. Wang J, Levasseur DN, Orkin SH (2008) Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. Proc Natl Acad Sci USA 105:6326–6331
91. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH (2006) A protein interaction network for pluripotency of embryonic stem cells. Nature 444:364–368
92. Wang J, Trowbridge JJ, Rao S, and Orkin SH (2008) Proteomic studies of stem cells. Stembook 1.4.1 (edited by Bernstein BE. and Lemischka I.), published online in stembook.org. StemBook [Internet]. Cambridge (MA): Harvard Stem Cell Institute; 2008-. PMID: 20614589.
93. Wang ZX, Teh CH, Kueh JL, Lufkin T, Robson P, Stanton LW (2007) Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. J Biol Chem 282:12822–12830
94. Waugh DS (2005) Making the most of affinity tags. Trends Biotechnol 23:316–320
95. Wu Q, Chen X, Zhang J, Loh YH, Low TY, Zhang W, Zhang W, Sze SK, Lim B, Ng HH (2006) Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J Biol Chem 281:24090–24094
96. Xu H, Schaniel C, Lemischka IR, Ma'ayan A (2010) Toward a complete in silico, multi-layered embryonic stem cell regulatory network. Wiley Interdiscip Rev Syst Biol Med 2:708–733
97. Xu X, Song Y, Li Y, Chang J, Zhang H, An L (2010) The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. Protein Expr Purif 72:149–156

98. Yang J, Chai L, Fowles TC, Alipio Z, Xu D, Fink LM, Ward DC, Ma Y (2008) Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc Natl Acad Sci USA 105:19756–19761
99. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008) The ground state of embryonic stem cell self-renewal. Nature 453:519–523
100. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R et al (2007) Induced pluripotent stem cell lines derived from human somatic cells. Science (New York, NY) 318:1917–1920
101. Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh YH, Yaw LP, Robson P, Lim B, Ng HH (2009) Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev 23:2507–2520
102. Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L et al (2006) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol 8:1114–1123
103. Zhang X, Odom DT, Koo SH, Conkright MD, Canettieri G, Best J, Chen H, Jenner R, Herbolsheimer E, Jacobsen E et al (2005) Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. Proc Natl Acad Sci USA 102:4459–4464
104. Zieske LR (2006) A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J Exp Bot 57:1501–1508