BioMol Concepts, Vol. 2 (2011), pp. 13–25 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BMC.2011.008

Review

Embryonic stem cells: protein interaction networks*

Patricia Miang-Lon Ng and Thomas Lufkin**

Stem Cell and Developmental Biology, Genome Institute of Singapore, 60 Biopolis Street, 138672 Singapore

**Corresponding author
e-mail: lufkin@gis.a-star.edu.sg

Abstract

Embryonic stem cells have the ability to differentiate into nearly all cell types. However, the molecular mechanism of its pluripotency is still unclear. Oct3/4, Sox2 and Nanog are important factors of pluripotency. Oct3/4 (hereafter referred to as Oct4), in particular, has been an irreplaceable factor in the induction of pluripotency in adult cells. Proteins interacting with Oct4 and Nanog have been identified via affinity purification and mass spectrometry. These data, together with iterative purifications of interacting proteins allowed a protein interaction network to be constructed. The network currently includes 77 transcription factors, all of which are interconnected in one network. In-depth studies of some of these transcription factors show that they all recruit the NuRD complex. Hence, transcription factor clustering and chromosomal remodeling are key mechanism used by embryonic stem cells. Studies using RNA interference suggest that more pluripotency genes are yet to be discovered via protein-protein interactions. More work is required to complete and curate the embryonic stem cell protein interaction network. Analysis of a saturated protein interaction network by system biology tools can greatly aid in the understanding of the embryonic stem cell pluripotency network.

Keywords: embryonic stem cells; Oct3/4; pluripotency; protein interaction networks.

Introduction

Embryonic stem (ES) cells were successfully isolated from the mouse in 1981. Two landmark papers opened the doors to this new source of cells that was to become as important as, if not more important than, HeLa cells to medical science. ES cells are derived from the inner cell mass of the blastocyst stage embryo (1, 2). Under the appropriate conditions, these cells replicate indefinitely. Yet unlike other immortalized cell culture, ES cells show a normal karyotype. In addition to their ability to replicate indefinitely, these cells demonstrate pluripotency. Pluripotency is the ability to differentiate into almost all cell types (including the trophectoderm which is sometimes excluded in definitions), without the ability to organize into a whole organism (3–5).

The main thrust for ES cell research comes from its prospects in biomedical research (6), namely, the promises of tissue replacement and regeneration, also referred to as regenerative medicine or regenerative therapy. There are different approaches towards this goal. The two most direct approaches are: (i) to use human ES cells to generate clinically relevant cell populations; and (ii) to use molecular factors to induce pluripotency in adult cells. The product is an induced pluripotent stem (iPS) cell that is then used to generate desired tissues via differentiation. In these approaches, understanding the molecular basis of pluripotency is fundamental. This review will address the protein determinants of pluripotency in ES cells. Recent efforts on the construction of ES cell protein interaction networks and conclusions derived from such data on the molecular mechanism of pluripotency are also covered.

The protein determinants of pluripotency

Since the isolation of ES cells, the focus has advanced to looking for protein determinants of the pluripotent state. Transcription factors play key roles in setting up the embryonic cells for pluripotency because they control gene expression. Three transcriptional factors, Oct4, Sox2 and Nanog, have been identified as key factors in the regulation of pluripotency (4, 7).

Oct4 is considered an important protein for pluripotency because it is an irreplaceable factor in the reprogramming of differentiated cells into iPS cells (8). It was found as a DNA-binding protein that is exclusively expressed during the earliest stages of embryonic development (9–14). Needless to say, Oct4 is expressed in ES cells. *Oct4* null mouse embryos reach the blastocyst stage but the inner cell mass is not pluripotent (15), instead these cells become restricted to the trophoblast lineage.

Sox2 was discovered as a transcription factor that often bound next to the Oct4 motif (16). *Sox2* null mouse embryos have an inner cell mass but with the depletion of maternal Sox2, these embryos fail to maintain the epiblast (16). The importance of the discovery of Sox2 is its interaction with Oct4. Sox2 collaborates with Oct4 to activate *Fgf4*, a gene that is expressed in the inner cell mass and later in distinct embryonic tissues (17, 18). Direct protein-protein interaction between the two transcription factors was shown using a bacterially expressed Oct4-GST fusion protein and *in vitro*-

*Electronic supplementary material to this article with the DOI 10.1515/BMC.2011.008SUP is available from the journal's online content site at www.reference-global.com/toc/bmc/2/1-2.

translated Sox2 (19). The requirement of both Oct4 and Sox2 in the activation of the *Fgf4* gene suggests that protein-protein interaction is a mechanism controlling gene expression in ES cells (20). Subsequently, Oct4 and Sox2 collaboration was also found to regulate expression of *Utf1* (21), *Fbx15* (22) and *Nanog* (23). In addition, the enhancer elements of Oct4 and Sox2 were also found to contain the Oct4-Sox2 binding elements, suggesting that protein-protein interaction is also a mechanism for autoregulation (24–26). In addition to Oct4 and Sox2, the finding that other transcription factors also show clustering at ES cell-specific genes (27, 28) further support the potential of protein-protein interaction as a code for transcriptional activation.

Systematic high-throughput methods further propelled the search for pluripotency factors. Nanog was discovered by two such approaches. The first approach used digital differential display of expressed sequenced tags in mouse ES cells versus somatic tissue (29). The second approach screened cDNA library-transfected ES cells for colonies that remained undifferentiated in the absence of LIF (30). *Nanog* null mouse embryos have the inner cell mass at the blastocyst stage but it fails to become the epiblast and instead differentiates into parietal endoderm-like cells (29).

The strong evidence for the involvement of Oct4, Sox2 and Nanog in pluripotency makes them good starting points (nodes) to study the protein interaction network of pluripotency. In addition, *de novo* discovery of genes with functional association to pluripotency comes from RNA interference (RNAi) studies. Several studies including two genome-wide screens led to the identification of a total of 167 pluripotency-associated genes (Table 1) including Oct4, Sox2 and Nanog (31–34). Out of these 167 proteins, only 15 (Table 1) are currently connected to the Oct4-centered protein interaction network (shaded in grey).

It is envisaged that all of these proteins, particularly factors that have been validated will be nodes in the pluripotency protein interaction network.

Building the ES cell protein interaction network

As more molecular determinants of pluripotency become defined, the next challenge is to integrate them into meaningful mechanisms. Network formulation is useful for the management and understanding of complex mechanisms (35). One type of network is the protein interaction network. A protein interaction network comprises proteins as nodes and undirected edges as the occurrence of binding. The datasets that are used to build the ES cell protein interaction network is generated via affinity purification-mass spectrometry methods of experimentation, and the datasets are mostly Oct4-centric because of its importance in ES and iPS cells. Currently, there are four studies using Oct4 as the 'bait' to find pluripotency-associated proteins (36–39). Other proteins that have been used as baits include Nanog, Sall4, Tcfcp2l1, Dax1, Esrrb, Rex1, Nac1 and Zfp281, all of which also showed interaction with Oct4. Integrating these studies gives

Table 1 Pluripotency-associated genes found via RNAi screens.

No.	Gene	Reference
1	Ehmt1	Hu
2	Esrrb	Ivanova
3	Hira	Hu
4	Mbd3	Hu
5	Mga	Hu
6	Nanog	Ivanova
7	Ncoa3	Hu
8	Oct3/4	Hu; Ivanova
9	Pcgf6	Hu
10	Rif1	Hu
11	Rnf2	Ding; Hu
12	Smc1a	Hu
13	Sox2	Hu; Ivanova
14	Yy1	Hu
15	Zfp219	Hu
16	1700067P10Rik	Hu
17	3110070M22Rik	Hu
18	5430407P10Rik	Hu
19	Acadsb	Ding
20	Acoxl	Hu
21	Adk	Hu
22	Aldoa	Hu
23	Amot	Hu
24	Apc	Ding; Hu
25	Ash2l	Zhang
26	Atg3	Hu
27	Atox1	Hu
28	BC018507 (MKIAA0947)	Ding
29	Bcl2l12	Hu
30	Bcorl1	Ding
31	Cbx1	Hu
32	Ccnb1ip1 (Mm343880)	Ivanova
33	Ccrn4l	Hu
34	Cdk9	Hu
35	Cdkn2aip	Hu
36	Cnih3	Hu
37	Cnot1	Ding
38	Cnot3	Hu
39	Coq3	Hu
40	Cpsf1	Hu
41	Cpsf2	Hu
42	Cpsf3	Ding; Hu
43	Ctr9	Ding; Hu
44	Cul3	Hu
45	Cxcl9	Hu
46	Cxxc1	Ding
47	D630039A03Rik	Hu
48	Dab2ip	Hu
49	Dazap1	Hu
50	Dppa4	Ivanova
51	Dppa5a (Dppa5)	Zhang
52	Ear11	Hu
53	Ecel1	Hu
54	Efr3b (KIAA0953)	Ding
55	Eif2s3x	Hu
56	Eif4a1	Hu
57	Eif4g2	Hu
58	Elof1	Hu
59	Eny2	Hu
60	Ep300	Hu

Table 1 (*Continued*)

No.	Gene	Reference
61	Epdr1	Hu
62	Eya1	Hu
63	Eya2	Hu
64	Fbxl8	Hu
65	Fip1l1	Ding; Hu
66	Fry	Hu
67	Gale	Hu
68	Ggh	Hu
69	Golga7	Hu
70	Grk6	Hu
71	Hao1	Hu
72	Hist1h3i	Hu
73	Hnrpul1	Hu
74	Hoxa7	Hu
75	Htatip2	Hu
76	Ift46 (1500035H01Rik)	Hu
77	Il20	Hu
78	Il6st	Hu
79	Ing5	Hu
80	Ino80e (Ccdc95)(AI225782)	Hu
81	Iws1	Ding; Hu
82	Krtap16-7 (Krtap21-1)	Hu
83	Mapk14	Hu
84	Mcrs1	Ding; Hu
85	Med10 (D13Wsu50e)	Hu
86	Metap2	Hu
87	Ms4a6b	Hu
88	Mtch2	Hu
89	MusD elements	Zhang
90	Ncapg2	Hu
91	Ncaph2 (D15Ertd785e)	Hu
92	Ncl	Ding
93	Nedd8	Hu
94	Nfya	Ding
95	Nipbl	Hu
96	Nts	Hu
97	Nup188	Hu
98	Olfr114	Hu
99	Ostf1	Hu
100	P4ha3	Hu
101	Paf1	Hu
102	Pax7	Hu
103	Pcbp1	Hu
104	Pcid2	Hu
105	Pcna	Hu; Zhang
106	Peci	Hu
107	Piwil4	Hu
108	Plac1	Hu
109	Pole4	Hu
110	Ppp4c	Hu
111	Ptbp1	Ding
112	Rad21	Hu
113	Rbx1	Hu
114	Rexo1	Hu
115	Rfwd2	Hu
116	Rnf146	Hu
117	Rprd1b (2610304G08Rik)	Hu
118	Rtf1	Ding
119	Rutbc3	Hu
120	Samd11	Hu
121	Samd5	Hu

Table 1 (*Continued*)

No.	Gene	Reference
122	Sema4a	Hu
123	Setd1b	Hu
124	Sgsm3	Hu
125	Sh2bp1	Hu
126	Shfdg1	Ding
127	Slc16a11	Hu
128	Slc19a3	Hu
129	Smc1l1	Hu
130	Spesp1	Hu
131	Spire1	Hu
132	Sprr2i	Hu
133	Ssu72	Hu
134	Stambpl1	Hu
135	Syngr1	Hu
136	Syt13	Hu
137	Tbx3	Ivanova
138	Tcl1	Ivanova
139	Tekt1	Hu
140	Tgfb1	Hu
141	Thoc2	Ding
142	Thoc5	Hu
143	Thoc5 (Fmip)	Hu
144	Tle4	Zhang
145	Triap1	Hu
146	Trim16	Hu
147	Trim28	Hu
148	Trmt6	Hu
149	Tubd1	Hu
150	Uba1	Hu
151	Ube1x	Hu
152	Ube2m	Ding
153	Uble1b	Zhang
154	Uncx	Hu
155	Uqcr10 (1110020P15Rik)	Hu
156	Vamp2	Hu
157	Wdr61	Ding; Hu
158	Wdr77	Zhang
159	Xpo7	Hu
160	Zadh2	Hu
161	Zfp13	Hu
162	Zfp42 (Rex1)	Zhang
163	Zfp628	Hu
164	Zfp759	Hu
165	Zfp771	Hu
166	Zfp786	Hu
167	Znhit4	Hu

a network comprising 240 proteins (Table 2). Of these, 131 proteins (Table 2) were associated with Oct4. Building the network brings new questions on the completeness and the accuracy of the data. How much of the interactions are we missing? How many false positives are included?

The concern on 'missing interactions' is most strikingly illustrated by the absence of Nanog and Rex1 when Oct4 was the bait (Table 2). Particularly, there are several studies that show association of Oct4 and Nanog (38–40). One reason for the non-reciprocal results could be the different protein levels between Nanog and Oct4 in ES cells. Nanog exists at lower levels than Oct4 making it harder to detect

Table 2 Pluripotency-associated genes found via protein-protein interactions.

No.	Gene	Reference
1	0610010K14Rik	van den Berg
2	2810474O19Rik	van den Berg
3	Acin1	Pardo
4	Actl6a	van den Berg; Pardo
5	Aft2	Pardo
6	Akap8	van den Berg
7	Amotl2	Pardo
8	Arid3b	van den Berg; Pardo; Wang
9	Asf1a	Pardo
10	Brwd1	Pardo
11	Cabin1	van den Berg; Pardo
12	Cad	Pardo
13	Cdk1	Wang
14	Chd1	Pardo
15	Chd3	Pardo
16	Chd4	van den Berg; Pardo
17	Chd5	Pardo
18	Creb1	Pardo
19	Ctbp1	Pardo
20	Ctbp2	van den Berg; Pardo
21	Cubn	Pardo
22	Cul4b	Pardo
23	Dax1	van den Berg; Wang
24	Ddb1	Pardo
25	Dhx9	Pardo
26	Dnaja1	Pardo
27	Dnmt3a	Pardo
28	Dnmt3l	Pardo
29	Emd	Pardo
30	Emsy	van den Berg
31	Ep400	van den Berg
32	Esrrb	van den Berg; Liang; Wang
33	Ewsr1	van den Berg; Wang
34	Foxp4	van den Berg
35	Frg1	van den Berg
36	Gatad2a	van den Berg; Pardo; Liang
37	Gatad2b	van den Berg; Pardo; Liang; Wang
38	Hcfc1	van den Berg; Pardo
39	Hdac1	van den Berg; Pardo; Liang
40	Hdac2	van den Berg; Liang; Wang
41	Hells	van den Berg; Pardo
42	Hira	Pardo
43	Hist1h3e	Pardo
44	Hist1h4b	Pardo
45	Hist3h2bb	Pardo
46	Hnrnpab	van den Berg; Pardo
47	Hnrnpl	Pardo
48	Hnrnpu	Pardo
49	Ifi202b	Pardo
50	Ilf2 (Nf45)	Wang
51	Ino80	Pardo
52	Klf4	Pardo
53	Klf5	van den Berg
54	Kpna2	Pardo
55	Kpna3	Pardo
56	L1td1	van den Berg
57	Lig3	van den Berg; Pardo
58	Lsd1	van den Berg; Pardo; Liang
59	Matr3	Pardo
60	Mbd3	van den Berg; Pardo
61	Mga	van den Berg
62	Mitf	Pardo
63	Msh2	van den Berg
64	Msh6	van den Berg; Pardo
65	Mta1	van den Berg; Pardo; Liang
66	Mta2	van den Berg; Pardo; Liang
67	Mta3	van den Berg; Pardo
68	Myst2	Pardo
69	Nac1	van den Berg; Wang
70	Nfrkb	Pardo
71	Nfyc	Pardo
72	Nudc	Pardo
73	Ogt	van den Berg; Pardo
74	P4ha1	Pardo
75	Parp1	Pardo
76	Phc1	van den Berg
77	Phf17	Pardo
78	Pml	van den Berg; Liang
79	Ppp2r1a	Pardo
80	Psmb6	Pardo
81	Rbbp7	van den Berg
82	Rbm14	van den Berg
83	Rbpj	van den Berg
84	Rcor2	van den Berg; Pardo
85	Requiem	van den Berg; Wang
86	Rfx2	Pardo
87	Rif1	van den Berg; Liang; Wang
88	Rnf2	van den Berg; Wang
89	Rpa1	van den Berg; Pardo
90	Rpa3	Pardo
91	Rybp	van den Berg
92	Sall1	van den Berg; Pardo; Wang
93	Sall3	van den Berg; Pardo
94	Sall4	van den Berg; Pardo; Liang; Wang
95	Smarca4	van den Berg; Pardo; Liang
96	Smarca5	van den Berg; Pardo
97	Smarcc1	van den Berg; Pardo; Wang
98	Smc1a	van den Berg
99	Sox2	van den Berg
100	Sp1	Pardo; Wang
101	Ssrp1	Pardo
102	Supt16h	van den Berg; Pardo
103	Tcfcp2l1	van den Berg
104	Tcfe3	Pardo
105	Tcfeb	Pardo
106	Top2a	Pardo
107	Trim24	Pardo
108	Trim33	van den Berg; Pardo
109	Trrap	van den Berg
110	Ttf2	Pardo
111	Ubn2	Pardo
112	Ubp1	van den Berg
113	Wdr5	van den Berg
114	Xrcc1	van den Berg; Pardo
115	Xrcc5	van den Berg; Pardo
116	Xrcc6	van den Berg; Pardo
117	Zbtb10	Pardo
118	Zbtb2	van den Berg; Pardo
119	Zbtb43	Pardo
120	Zcchc8	van den Berg
121	Zfhx3	Pardo

Table 2 (*Continued*)

No.	Gene	Reference
122	Zfp143	van den Berg
123	Zfp217	Pardo
124	Zfp219	van den Berg; Pardo; Wang
125	Zfp462	van den Berg
126	Zfp513	Pardo
127	Zic2	Pardo
128	Zmym2	van den Berg
129	Zscan4b	Pardo
130	*Nanog	Liang; Wang
131	*Zfp42 (Rex1)	Wang
132	1600027Rik	van den Berg
133	2310057J16Rik	van den Berg
134	4632411B12Rik	van den Berg
135	7420416P09Rik	van den Berg
136	Adnp	van den Berg
137	Arid1a	van den Berg
138	Arid3a	Wang
139	Ashl2	van den Berg
140	Bend3	van den Berg
141	Bptf	van den Berg
142	Brd8	van den Berg
143	Btbd14a	Wang
144	C130039O16Rik	van den Berg
145	Cdc2a	van den Berg
146	Cdk8	van den Berg
147	Cncc	van den Berg
148	Cxxc5	van den Berg
149	Dmap1	van den Berg
150	Ehmt1	van den Berg
151	Elys	Wang
152	Esrra	van den Berg
153	Etl1	Wang
154	Fkbp15	van den Berg
155	Grhl2	van den Berg
156	Ing3	van den Berg
157	Jmjd1c	van den Berg
158	Kap1	Liang
159	L3mbtl2	van den Berg
160	Mbd2	van den Berg
161	Med1	van den Berg
162	Med12	van den Berg
163	Med13	van den Berg
164	Med13l	van den Berg
165	Med14	van den Berg
166	Med15	van den Berg
167	Med16	van den Berg
168	Med17	van den Berg
169	Med18	van den Berg
170	Med19	van den Berg
171	Med23	van den Berg
172	Med24	van den Berg
173	Med25	van den Berg
174	Med26	van den Berg
175	Med27	van den Berg
176	Med29	van den Berg
177	Med30	van den Berg
178	Med4	van den Berg
179	Med6	van den Berg
180	Med7	van den Berg
181	Med8	van den Berg
182	Mll2	van den Berg
183	Mll3	van den Berg

Table 2 (*Continued*)

No.	Gene	Reference
184	Mybbp	Wang
185	Mybl2	van den Berg
186	Myst1	van den Berg
187	Ncoa3	van den Berg
188	Nrip1	van den Berg
189	Oct3/4	van den Berg; Pardo; Liang; Wang
190	Pbrm1	van den Berg; Liang
191	Pcgf6	van den Berg
192	Peg10	van den Berg
193	Pelo	Wang
194	Pnkp	van den Berg
195	Pogz	van den Berg
196	Polb	van den Berg
197	Polr2a	van den Berg
198	Polr2b	van den Berg
199	Polr2c	van den Berg
200	Polr2g	van den Berg
201	Prkdc	van den Berg
202	Prmt1	van den Berg; Wang
203	Rai14	Wang
204	Rbbp4	van den Berg
205	Rbbp5	van den Berg
206	Rest	Wang
207	Ruvbl1	van den Berg
208	Ruvbl2	van den Berg
209	Rypb	Wang
210	Sall2	van den Berg
211	Satb2	van den Berg
212	Scmarca4	van den Berg
213	Set	van den Berg
214	Sin3a	van den Berg; Liang
215	Smarca2	Liang
216	Smarcb1	van den Berg
217	Smarcc2	van den Berg
218	Smarcd1	van den Berg
219	Smarcd2	van den Berg
220	Smarce1	van den Berg
221	Snw1	van den Berg
222	Taf4a	van den Berg
223	Taf6	van den Berg
224	Taf9	van den Berg
225	Tcfcp2	van den Berg
226	Tif1b	Wang
227	Usp9x	van den Berg
228	Vps72	van den Berg
229	Wapl	Wang
230	Wdr18	Wang
231	Wiz	van den Berg
232	Yeats2	van den Berg
233	Yeats4	van den Berg
234	Yy1	Wang
235	Zbtb9	van den Berg
236	Zfp198	Wang
237	Zfp281	Wang
238	Zfp609	Wang
239	Zfp828	van den Berg
240	Zmym4	van den Berg

The bait protein used includes Oct4, Nanog, Sall4, Tcfcp2l1, Dax1, Esrrb, Rex1, Nac1 and Zfp281. Proteins found when Oct4 was the bait are shaded grey. *Nanog and *Zfp42 (Rex1) interacts with Oct4 when they are used as the bait.

Nanog in Oct4 purifications. Conversely, Oct4 exists at higher levels than Nanog and is therefore more easily detected in Nanog purifications. This example suggests that important interactions could be missed for proteins expressed at low levels, as are many transcription factors. Further evidence that a large part of the network remains to be uncovered comes from the low overlap between the components identified from protein-protein interaction and from genome-wide RNA interference studies (Figure 1). The incomplete overlap between the different groups that all study protein-protein interactions (Figure 2) also supports this belief. Alternatively, only the intersection represents true Oct4 interacting proteins (41). However, the observation of interactors such as Sox2 outside the intersection (Figure 2) supports the former opinion rather than the latter.

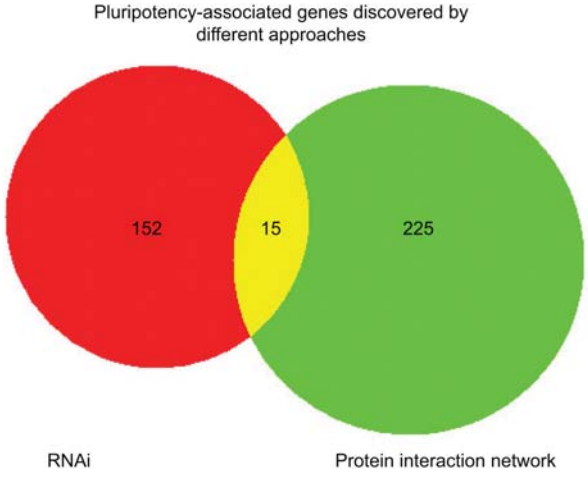

Figure 1 Venn diagram showing the number of pluripotency-associated genes discovered by different approaches.
A total of 167 genes were found in four separate RNAi studies. A total of 240 proteins were found via protein-protein interaction with Oct4, Nanog, Sall4, Tcfcp2l1, Dax1, Esrrb, Rex1, Nac1 and/or Zfp281. Between the two approaches, only 15 genes/proteins are in common.

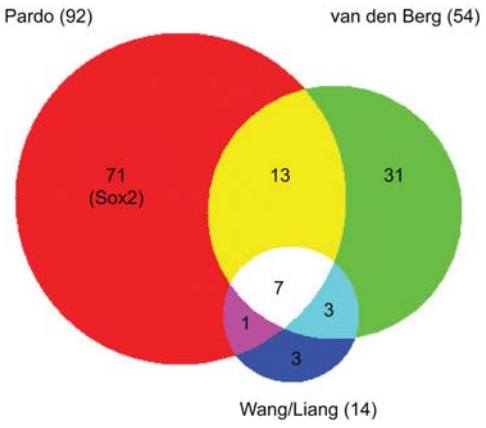

Figure 2 Venn diagram showing the number of proteins identified by protein-protein interaction with Oct4 as the bait.
Proteins from the two smaller datasets by Wang et al. (39) and by Liang et al. (38) are merged into one group.

Certainly, the network is not free of inaccuracies. The weak-yet-important interactions make the distinction of 'false positive' an even greater challenge than it already is. The main challenge comes from the low throughput nature of available validation methods and the shortcomings of each of them.

The most direct method of validating a protein-protein interaction is via reciprocal co-precipitation. This is frequently done by overexpressing the two proteins in a cell culture system. However, some proteins interact indirectly via a common protein, which if not present in the cell, would yield negative results in a co-precipitation analysis.

Furthermore, after direct or indirect association has been verified, it is important to examine the functional significance of proteins in the network. Not all physical association has functional significance. For example, both Oct1 and Oct4 can interact with Sox2, but only the Oct4-Sox2 complex can activate Fgf4 expression (18). Hence, multiple validations are important. Validations that have been employed are as follows: (i) evidence for presence of the interacting protein in ES cells; (ii) evidence that interacting proteins coexist in a common subcellular location; (iii) indication that the level of abundance of the interacting protein changes upon differentiation; (iv) indication that the interacting protein regulates genes of known ES cell transcription factors or vice versa; (v) gain or loss of pluripotency of ES cells when the gene of the interacting protein is knocked-out, suppressed by RNAi or overexpressed. Pluripotency can be monitored by alkaline phosphatase staining, ES cell morphology, transcript levels of Oct4 or Nanog, profiling of lineage markers, and the levels of stage-specific embryonic antigen 1, 3 and 4; and (vi) loss-of-function phenotypes in mice when the gene of the interacting protein is knocked-out, suppressed by RNAi or overexpressed. Given that gene redundancy or functional redundancy is a phenomenon of pluripotency (42), validations that show no effect with a single gene knock-out could be further evaluated via double or triple knock-outs.

Certainly, efforts to extend the boundaries of the ES cell protein interaction network via iteration (36, 39) would help to complete the protein interaction network. However, care should be taken not to go off-tangent in this approach, particularly if protein interaction networks are not truly separable modules in the cell. Yeast 2-hybrid is an alternative approach. However, this approach appears to yield a significantly lower number of Oct4-interacting proteins (38) compared to tandem affinity-mass spectrometry. This would suggest that Oct4 does not show strong binary interactions and rather could be relying on DNA-enhanced associations or complex-mediated indirect associations. Although more research would be required to confirm this, this postulation is corroborated by the observation of its weak interactions with Nanog and Sox2.

Eventually, stricter definitions will be required to trim the ES cell protein interaction network to reveal the core mechanism of pluripotency. This could entail the distinction between genes that control pluripotency and genes that regulate differentiation. Loss of a 'differentiation' gene could

appear as a loss of pluripotency because the ES cell would no longer show the ability to differentiate into its normal repertoire of cell types. However, such defects can be corrected with the reexpression of the gene, suggesting that the pluripotent state is there all the time (7). As proteins tend to demonstrate multifunctionality, it would also be necessary to validate the role of specific interactions rather than components in pluripotency.

Mining the network

There are different levels of analysis in a protein interaction network. A basic analysis is the identification of novel components. Protein-protein interactions added another 225 pluripotency-associated components to those found via RNAi (Figure 1). Ideally, via an iterative approach of protein-protein interaction, all the pluripotency-associated genes identified by RNAi should be rediscovered.

To understand the molecular mechanism of pluripotency, different methods have been employed. Firstly, to unearth key controllers, transcription factors are identified using the Gene Ontology annotation GO:0003700, which is proteins with sequence-specific DNA binding transcription factor activity or other closely related terms. Based on the integrated dataset of all four protein-protein interaction studies, and the annotation 'transcription factor' used in these studies, there are a total of 77 transcription factors. Figure 3A shows

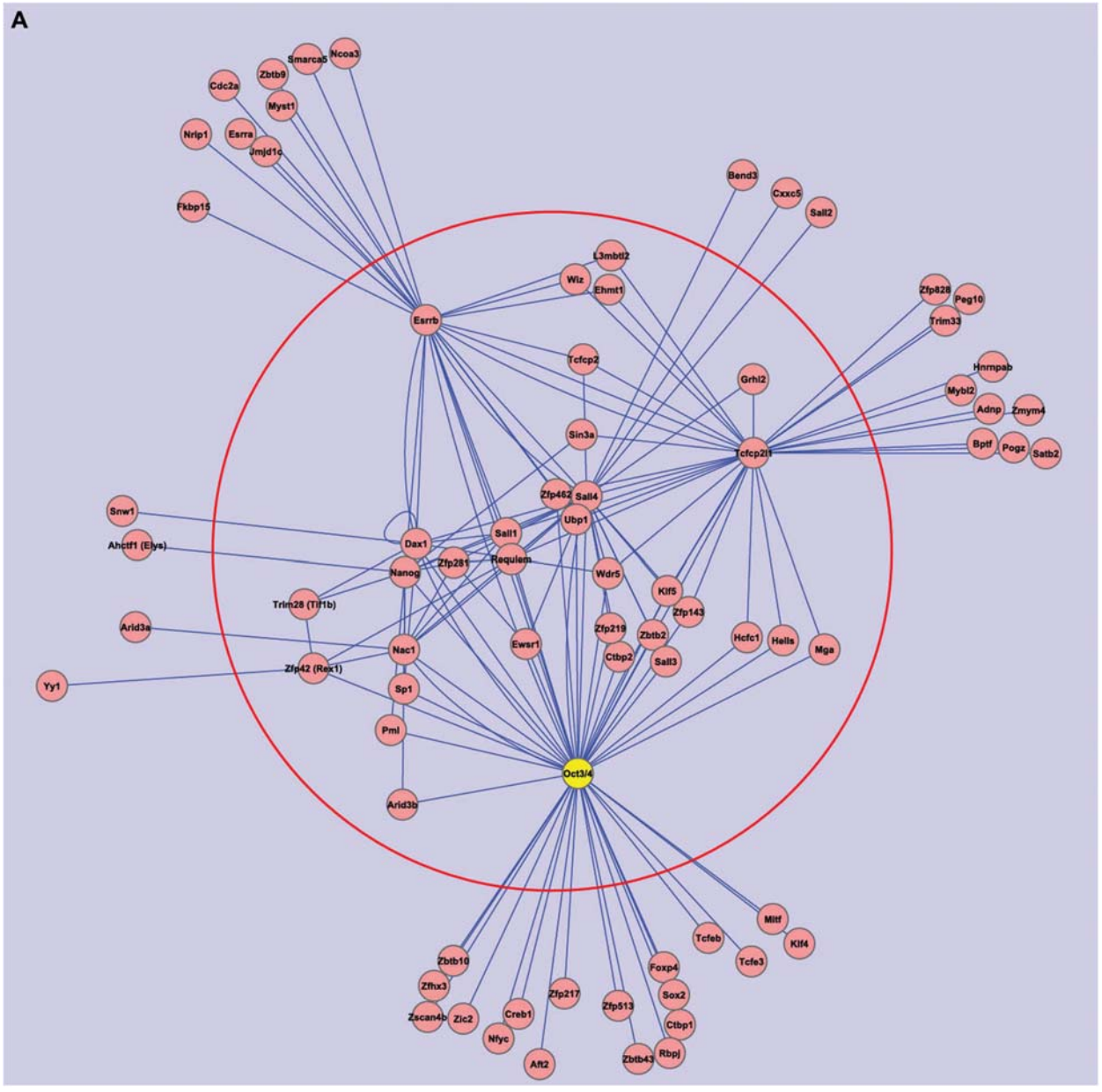

Figure 3 (Continued)

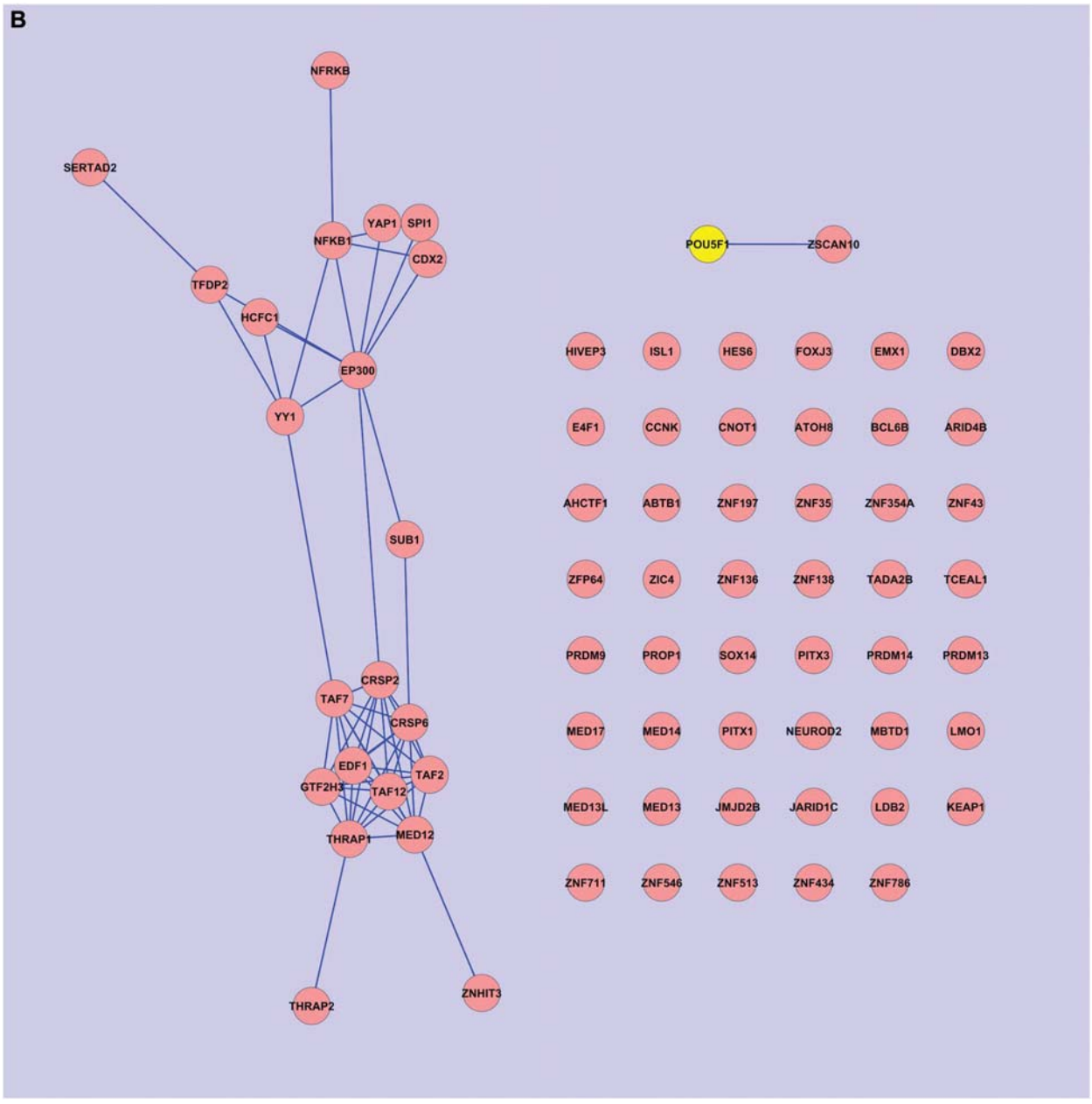

Figure 3 Protein interaction network of transcription factors in the embryonic stem cell of mouse and human.
(A) The mouse network is constructed based on transcription factors interacting with pluripotency-associated factors, Oct4, Nanog, Sall4, Dax1, Esrrb, Tcfcp2l1, Nac1, Zfp281 and Zfp42 (alias Rex1) (36–39). Oct4 is marked yellow. Transcription factors interacting with more than one other transcription factor (inside the circle) could represent the phenomenon of transcription factor clustering for the activation of ES cell genes. For two studies (38, 39), names of proteins were updated to the official one (with the original names used in the publication in parentheses) so that they are consistent across all the studies. Protein functions where not given were also annotated based on gene ontology so that they are consistent across these studies. The compiled list is shown in Supplementary Table S1. From this list, proteins which are transcription factors are selected to generate a protein interaction network. (B) The human network is constructed based on transcription factors that were found via RNAi to have a role in pluripotency (43). The human ortholog of mouse Oct4 is POU5F1 and is marked yellow. A total of 67 transcription factors were uploaded to the online database STRING to search for possible interactions. These interactions are predicted based on experimental as well as homology-based evidences. Both the mouse and human networks are constructed by Cytoscape 2.8.0 (67) and visualized using the force-directed paradigm called *spring embedded* Cytoscape Layout.

a protein-interaction network of these transcription factors using datasets from all four studies. Because clustering of transcription factors on promoters is observed in ES cells, protein-protein interaction between these transcription factors could provide combinatorial codes required for regulation of gene expression for pluripotency. Presumably, transcription factors with two or more interactions (Figure 3A, inside the circle) would be activating more ES cell-spe-

cific genes. Whereas transcription factors with one interaction (Figure 3A, outside the circle) could be activating more general genes. Certainly, there are transcriptions factors that are important to pluripotency but do not cluster into the circle of highly interactive zone because the network is incomplete. For example, the Sox2-interactome has yet to be reported by any lab. The current network therefore serves as a guide for further research.

On this note, this mouse network can also serve as a comparison for data on human embryonic stem cells. Determinants of human embryonic stem cell pluripotency have been identified by a genome-wide RNAi screen (43). The screen identified a total of 566 genes and a protein interaction network based on these has been reported. To compare the transcription factor protein interaction network of mouse and human, we constructed a protein interaction network based solely on the transcription factors, which numbers 67 in the 566 genes. Because the approach of affinity purification-mass spectrometry is yet to be applied to human embryonic stem cells, information regarding possible interactions between any of the 67 transcription factors was obtained via the online database STRING. This results in a network which was reconstructed using Cytoscape (Figure 3B). Clearly, in contrast to the mouse network, most of the transcription factors were unconnected, probably owing to a lack of understanding of these transcription factors. POU5F1, a crucial transcription factor to human ES cell pluripotency, is also highly unexplored with regard to its protein-protein interactions. The only POU5F1 interaction shown in Figure 3B is inferred by studies from mouse ES cells where Oct4 was shown to physically interact with Zscan10 (alias Zfp206) (44). Connections between POU5F1 and SOX2 and between POU5F1 and NANOG cannot be drawn because these genes did not pass the criteria in the RNAi screen for genes that maintain pluripotency in the human ES cells. This suggests that mouse and human pluripotency determinants are highly diverged. It is therefore imperative to investigate the protein interaction network for human ES cells, particularly using POU5F1 as bait.

Secondly, to understand the mechanisms employed by the transcription factors, proteins can be categorized into any of the three gene ontology sections: (i) molecular function; (ii) biological process; and (iii) cellular components (45). Annotations under biological process can help understand the role of a local network of proteins. Annotations under cellular components are extremely useful for the identification of multisubunit enzymes or protein complexes. Presence of all the components of a protein complex is a strong indication that the machinery is assembled for use. It should be noted that proteins can have multiple functions; hence, the assignment of a novel function should be considered if a component is not copurifying with the rest of the complex.

Using the method of gene ontology annotation, it was found that the nucleosome remodeling histone deacetylase (NuRD) complex (46) is the most prominent complex identified in the ES cell protein-interaction network (36–38). All the components of this complex are found in the network and each of the components interacts with one or more of the five transcription factors which have been studied in greater detail (36). These include Nanog, Esrrb, Oct4, Tcfcp2l1 and Sall4 (Figure 4) which are themselves tightly associated with one another. Because some of these transcription factors have been proven to have a direct role in pluripotency, it can be concluded that the ES cell utilizes histone deacetylation mediated by NuRD as a gene repression mechanism to regulate pluripotency. Indeed, case studies have shown that NuRD has specific developmental roles rather than being required for general cellular functions (46–48). In addition to NuRD, other complexes have been reported in the study by Pardo and colleagues (37). Most of these are involved in chromosome remodeling. Confirmation of these findings would surely expand our knowledge of the extent to which each of these complexes contributes to pluripotency. For example, there is evidence that chromosomal remodeling factors such as the polycomb group and polycomb repressive complex are not required for maintenance of pluripotency in ES cells (4, 49–54). Although it is believed that these repressors serve to prevent spontaneous differentiation of the ES cells, the chromatin of the ES cell is deemed, at the same time, to be relatively 'loose' so as to allow free accessibility to the transcription factors. Having the different chromatin modifiers inserted into the protein interaction network can help to clarify their role in pluripotency. In addition to the chromatin modifiers, the basic transcriptional machinery was also found to be recruited to the protein interaction network by Esrrb (36). However, this

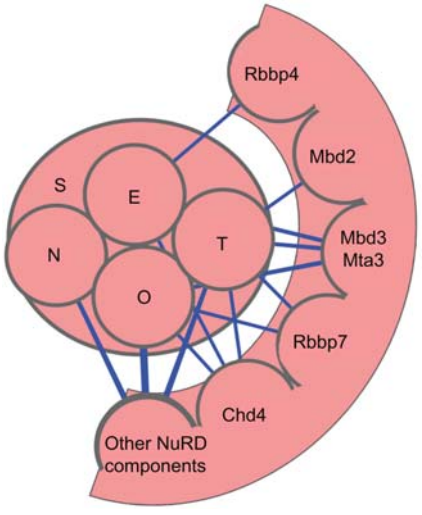

Figure 4 A core set of transcription factors comprising Sall4 (S), Esrrb (E), Nanog (N), Oct4 (O) and Tcfcp2l1 (T) show interaction with one another and with many components of the NuRD complex. The 'other NuRD components' include Gatad2a, Gatad2b, Mta1, Mta2, Hdac1 and Hdac2. This Figure is modified from a protein interaction network constructed with Cytoscape. The Cytoscape input file for this figure is shown in Supplementary Table S1. Edges between the five transcription factors Sall4, Esrrb, Nanog, Oct4 and Tcfcp2l1 are replaced by direct contacts of the nodes to suggest colocalization of these transcription factors. Nodes of proteins belonging to the NuRD complex are merged to suggest their entity as a macromolecule.

mechanism appears not to be utilized by the other transcription factors in the network. It remains unclear if this mechanism is directly related to the regulation of pluripotency.

A third method that is yet to be fully utilized for the analysis of the protein interaction networks is to employ the tools of system biology. This is because the protein interaction network is currently incomplete. At this stage, the network structure can be strongly skewed by the methods used to generate the network (55). The observation that essential proteins tend to be more highly connected than nonessential proteins could also be a true property or a consequence of them having been more thoroughly studied, or a combination of the two (56). However, as data accumulates, the power of systems biology to catalogue and integrate data will be necessary (35). Concepts from graph theory (35, 57, 58) can provide us with insight into the 'molecular characteristic' or the 'functional characteristic' of the ES cell. Simulations can be used to allow us to predict and explain the outcomes of experimental manipulations.

The future network

A protein interaction network by virtue of the protocols employed is a single snapshot of the protein-protein interactions of the cell at any given time. To understand how ES cells have the ability to differentiate into different cell types, further information will have to be integrated. The final protein interaction network should include information on protein subcellular location and protein concentration. For example, ES cell fate has been shown to be highly sensitive to Oct4 dosage levels (59). All information in the network will change as a function of time as the cell undergoes cell cycling and when the cell undergoes fate changes. A study on the systems level changes across the three mechanistic layers: epigenetic, transcriptional and translational during fate change in mouse ES cell data show that changes in nuclear protein levels are not accompanied by concordant changes in the corresponding mRNA levels, suggesting that translational and post-translational mechanisms, rather than transcriptional regulation, play important roles during lost of pluripotency (60). For full understanding and successful simulation, information from the protein interaction network, the gene regulatory network and microRNA networks of ES cells should be fed back into one another. Ultimately, the goal of using systems biology is to be able to show how the properties of individual components collaborate into a meaningful integrated process, and how the different processes result in the emergent property of pluripotency.

Expert opinion

Ironically, pluripotency is best demonstrated by its loss. A population of cells is pluripotent if it can differentiate into many cell types; but once that happens, pluripotency is lost. In the ES cell, molecules for pluripotency work to balance two opposing features: the readiness to initiate differentiation and the prevention of differentiation. To understand the molecular mechanism of pluripotency, we need to keep in mind this concept of pluripotency.

The current protein interaction network encompasses both of these features of pluripotency. To complicate matters, most proteins are multifunctional and can play different roles in both aspects of pluripotency. In view of this, looking at proteins for the assignment of processes can be more confusing than helpful. Assignment of processes can be more meaningful if it is done to the edges of the network rather than to the nodes. This opinion can be best illustrated with an example. The readiness to differentiate is established by keeping the chromatin in an accessible state. This can be achieved by close cooperation between chromatin modifiers and transcription factors such as Oct4. Hence, the edge between the Oct4 node and the NuRD protein nodes can be assigned with the purpose of 'keeping chromatin relaxed'. For the prevention of differentiation, one mechanism is via protein-protein interaction of transcription factors and again Oct4 can be involved. For example, physical interaction between Oct4 and Cdx2 forms a repressor complex which provides autoregulation of the two genes. Furthermore, physical interaction between Oct4 and Sox2 forms an activation complex for the transcription of genes such as *Fgf4*. Hence, the edges between Oct4 and other transcription factors can be assigned with the purpose of 'auto-repressor' and 'compulsory co-activation', respectively.

Another perspective which should be incorporated when looking at the protein interaction network is the presence of two types of protein-protein interactions. Transient protein-protein interactions occur between transcription factors or between transcription factors and other protein complexes. Static protein-protein interactions occur between protein subunits of a stable protein complex. The first type of interaction usually encodes instructions or messages, whereas the second type of interaction functions mainly to execute the processes as a module. Identifying these interactions allow us to understand how cell fate decisions are made and how these decisions are executed.

In view of the large number of proteins that have been associated with pluripotency. It is possible that there are alternate means of achieving pluripotency. After all, pluripotency is a cellular state rather than a cellular composition. Proteins such as Ronin (61, 62), which show strong associations with pluripotency, can operate via a separate network. Observations that different combinations of factors (8, 63–66) can also induce pluripotency are another sign of the multiple means of achieving this state.

Overall, we envisage great promise in obtaining answers and insights from a mature protein interaction network. However, this will require construction of the network to be closely accompanied with attempts to annotate the purpose and nature of the interaction as discussed above.

Outlook

An example of how knowledge derived from mouse ES cells has contributed towards the goal of regenerative stem cell therapy is the generation of iPS cells. The factors discovered

from ES cells were used to induce pluripotency in adult cells. This removes the need to use human embryos which is highly controversial in stem cell therapy. In the next 10 years, ES cells will continue to be a source of guidance until iPS technology is perfected. During this time, data accumulation should continue until a point where the boundaries of the protein interaction network are felt. At the same time extra efforts will be needed towards looking for interactions among low concentration proteins and towards validation of the network. With a more complete protein interaction network, new hypothesis can be formulated. As more system biology data is generated from other fields, it will become possible to compare between non-pluripotent and pluripotent networks. The ES cell protein interaction network, once ready, will serve as a point of comparison with other stem cells, with differentiating cells and with cancer cells. Such comparisons can potentially bring out unique features of operation in each of these cellular conditions. Finally, in view of the differences between human and mice, the same work will have to be repeated with human ES cells. However, from the challenges encountered in mouse ES cell research, the working knowledge gained will ensue much faster progress with the human ES cell project.

Highlights

- Three transcription factors, Oct4, Sox2 and Nanog, show strong evidence in their role as determinants of pluripotency.
- Another 237 proteins are associated with these determinants by protein-protein interactions.
- Another 152 proteins discovered to have a role in pluripotency by genome-wide RNAi screening are yet to be connected via protein-protein interactions.
- Further protein-protein interaction studies to connect and extend on these proteins are necessary.
- Multiple validations to confirm the involvement of these proteins in pluripotency are necessary.
- Transcription factors show collaboration in the protein interaction network.
- NuRD is frequently recruited by a core of ES cell transcription factors.
- Other chromatin modification machineries are also potentially recruited.
- When the network is reasonably saturated, system biology analysis should be employed to give insight into network properties.
- Assignment of purpose to edges rather than to nodes in the network will drive understanding of the network.
- Inclusion of information on dynamic properties of the protein interaction network would facilitate predictive capabilities.

References

1. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci USA 1981; 78: 7634–8.
2. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981; 292: 154–6.
3. Smith AG. Embryo-derived stem cells: of mice and men. Annu Rev Cell Dev Biol 2001; 17: 435–62.
4. Niwa H. How is pluripotency determined and maintained? Development 2007; 134: 635–46.
5. Solter D. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. Nat Rev Genet 2006; 7: 319–27.
6. Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell 2008; 132: 661–80.
7. Silva J, Smith A. Capturing pluripotency. Cell 2008; 132: 532–6.
8. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 2008; 26: 101–6.
9. Rosner MH, Vigano MA, Ozato K, Timmons PM, Poirier F, Rigby PW, Staudt LM. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature 1990; 345: 686–92.
10. Scholer HR, Balling R, Hatzopoulos AK, Suzuki N, Gruss P. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. Embo J 1989; 8: 2551–7.
11. Scholer HR, Hatzopoulos AK, Balling R, Suzuki N, Gruss P. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. Embo J 1989; 8: 2543–50.
12. Scholer HR, Ruppert S, Suzuki N, Chowdhury K, Gruss, P. New type of POU domain in germ line-specific protein Oct-4. Nature 1990; 344: 435–9.
13. Scholer HR, Dressler GR, Balling R, Rohdewohld H, Gruss P. Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. Embo J 1990; 9: 2185–95.
14. Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell 1990; 60: 461–72.
15. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 1998; 95: 379–91.
16. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 2003; 17: 126–40.
17. Dailey L, Yuan H, Basilico C. Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. Mol Cell Biol 1994; 14: 7758–69.
18. Yuan H, Corbi N, Basilico C, Dailey L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev 1995; 9: 2635–45.
19. Ambrosetti DC, Basilico C, Dailey L. Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. Mol Cell Biol 1997; 17: 6321–9.
20. Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, Wilmanns M. Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes Dev 2003; 17: 2048–59.
21. Nishimoto M, Fukushima A, Okuda A, Muramatsu M. The gene for the embryonic stem cell coactivator UTF1 carries a regula-

tory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. Mol Cell Biol 1999; 19: 5453–65.
22. Tokuzawa Y, Kaiho E, Maruyama M, Takahashi K, Mitsui K, Maeda M, Niwa H, Yamanaka S. Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. Mol Cell Biol 2003; 23: 2699–708.
23. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem 2005; 280: 24731–7.
24. Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol Cell Biol 2005; 25: 6031–46.
25. Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N, Niwa H, Muramatsu M, Okuda A. Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res 2002; 30: 3202–13.
26. Okumura-Nakanishi S, Saito M, Niwa H, Ishikawa F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. J Biol Chem 2005; 280: 5307–17.
27. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. Cell 2008; 132: 1049–61.
28. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 2008; 133: 1106–17.
29. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 2003; 113: 631–42.
30. Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 2003; 113: 643–55.
31. Ding L, Paszkowski-Rogacz M, Nitzsche A, Slabicki MM, Heninger AK, de Vries I, Kittler R, Junqueira M, Shevchenko A, Schulz H, Hubner N, Doss MX, Sachinidis A, Hescheler J, Iacone R, Anastassiadis K, Stewart AF, Pisabarro MT, Caldarelli A, Poser I, Theis M, Buchholz F. A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. Cell Stem Cell 2009; 4: 403–15.
32. Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 2009; 23: 837–48.
33. Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR. Dissecting self-renewal in stem cells with RNA interference. Nature 2006; 442: 533–8.
34. Zhang JZ, Gao W, Yang HB, Zhang B, Zhu ZY, Xue YF. Screening for genes essential for mouse embryonic stem cell self-renewal using a subtractive RNA interference library. Stem Cells 2006; 24: 2661–8.
35. Pieroni E, de la Fuente van Bentem S, Mancosu G, Capobianco E, Hirt H, de la Fuente A. Protein networking: insights into global functional organization of proteomes. Proteomics 2008; 8: 799–816.
36. van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA. An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell 2010; 6: 369–81.
37. Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell 2010; 6: 382–95.
38. Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D, Songyang Z. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol 2008; 10: 731–9.
39. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. A protein interaction network for pluripotency of embryonic stem cells. Nature 2006; 444: 364–8.
40. Zhang L, Rayner S, Katoku-Kikyo N, Romanova L, Kikyo N. Successful co-immunoprecipitation of Oct4 and Nanog using cross-linking. Biochem Biophys Res Commun 2007; 361: 611–4.
41. Lemischka IR. Hooking up with Oct4. Cell Stem Cell 2010; 6: 291–2.
42. Jiang J, Chan YS, Loh YH, Cai J, Tong GQ, Lim CA, Robson P, Zhong S, Ng HH. A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol 2008; 10: 353–60.
43. Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, Huss M, Soh BS, Kraus P, Li P, Lufkin T, Lim B, Clarke ND, Bard F, Ng HH. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature 2010; 468: 316–20.
44. Yu HB, Kunarso G, Hong FH, Stanton LW. Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. J Biol Chem 2009; 284: 31327–35.
45. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25–9.
46. Ahringer J. NuRD and SIN3 histone deacetylase complexes in development. Trends Genet 2000; 16: 351–6.
47. Ch'ng, Q, Kenyon C. egl-27 generates anteroposterior patterns of cell fusion in C. elegans by regulating Hox gene expression and Hox protein function. Development 1999; 126: 3303–12.
48. Mannervik M, Levine M. The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. Proc Natl Acad Sci USA 1999; 96: 6797–801.
49. de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R, Nesterova TB, Silva J, Otte AP, Vidal M, Koseki H, Brockdorff N. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 2004; 7: 663–76.
50. O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 2001; 21: 4330–6.
51. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. Control of developmental regulators by polycomb in human embryonic stem cells. Cell 2006; 125: 301–13.
52. Montgomery ND, Yee D, Chen A, Kalantry S, Chamberlain SJ, Otte AP, Magnuson T. The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol 2005; 15: 942–7.
53. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 2006; 8: 532–8.

54. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006; 441: 349–53.
55. Futschik ME, Tschaut A, Chaurasia G, Herzel H. Graph-theoretical comparison reveals structural divergence of human protein interaction networks. Genome Inform 2007; 18: 141–51.
56. Hakes L, Pinney JW, Robertson DL, Lovell SC. Protein-protein interaction networks and biology – what's the connection? Nat Biotechnol 2008; 26: 69–72.
57. Dong J, Horvath S. Understanding network concepts in modules. BMC Syst Biol 2007; 1: 24.
58. Mason O, Verwoerd M. Graph theory and networks in Biology. IET Syst Biol 2007; 1: 89–119.
59. Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 2000; 24: 372–6.
60. Lu R, Markowetz F, Unwin RD, Leek JT, Airoldi EM, MacArthur BD, Lachmann A, Rozov R, Ma'ayan A, Boyer LA, Troyanskaya OG, Whetton AD, Lemischka IR. Systems-level dynamic analyses of fate change in murine embryonic stem cells. Nature 2009; 462: 358–62.
61. Zwaka TP. Ronin and caspases in embryonic stem cells: a new perspective on regulation of the pluripotent state. Cold Spring Harb Symp Quant Biol 2008; 73: 163–9.
62. Dejosez M, Krumenacker JS, Zitur LJ, Passeri M, Chu LF, Songyang Z, Thomson JA, Zwaka TP. Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. Cell 2008; 133: 1162–74.
63. Feng B, Jiang J, Kraus P, Ng JH, Heng JC, Chan YS, Yaw LP, Zhang W, Loh YH, Han J, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol 2009; 11: 197–203.
64. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. Nature 2007; 448: 313–7.
65. Yamanaka S. A fresh look at iPS cells. Cell 2009; 137: 13–7.
66. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006; 126: 663–76.
67. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498–504.