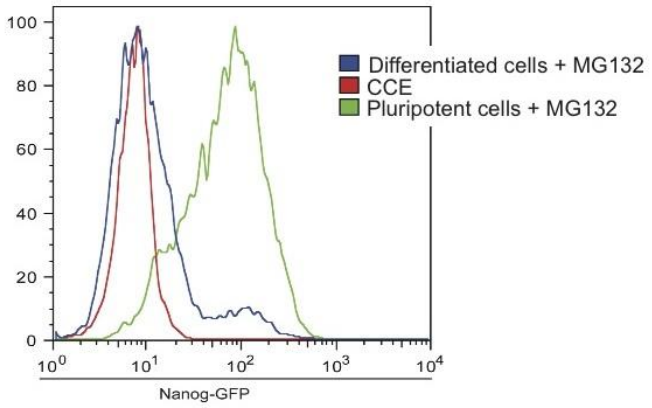
Cell Stem Cell, Volume *11*

Supplemental Information

Regulation of Pluripotency and Cellular Reprogramming by the Ubiquitin Proteasome System

Shannon M. Buckley, Beatriz Aranda-Orgilles, Alexandros Strikoudis, Effie Apostolou, Evangelia Loizou, Kelly Moran-Crusio, Charles L. Farnsworth, Antonius A. Koller, Ramanuj Dasgupta, Jeffrey C. Silva, Matthias Stadtfeld, Konrad Hochedlinger, Emily I. Chen, and Iannis Aifantis

A



B

	Identified peptides
Nanog	112-TVFSQAQLCALK*DR
	156-WQK*NQWLK
	141-YLSLQQMQELSSILNLSYK^QVK*
Oct4	137-ALQK*ELEQFAK
	144-LRPLLEK*WVEEADNNENLQEICK
	199-WVEEADNNENLQEICK*SETLVQAR
	215-WSLETM#FLKCPK*PSLQQITHIANQLGLEK
Dppa5	103-GVLK*LEESMK
	109-LEESMK*TLELGQCIE
	16-K^DIPPWVK*VPEDLK^DPEVFQVQSLVLK^
	16,22-K^DIPPWVK*VPEDLK*DPEVFQVQSLVLK^
	22-DIPPWVK^VPEDLK*DPEVFQVQSLVLK^
	35-VPEDLK^DPEVFQVQSLVLK*YLFGPQGSR^
	79-NLESPEELIEVFIYGSQNNK*IR
	83-AK*WMLQSMAER
	9-K*DIPPWVKVPEDLKDPEVFQVQSLVLK
Jarid2	694-LAK*LQEAYCQYLLSYDSLSPEEHR
	1121-YGSHDGNSTVADGK*K
	879-LVEEK^DCHVAVHCGK^VDTNTHGSGFPVGK*SEPFSR^
Zfp42	132-QVLEASSLLESSLEYMTK^GTK^QEK*
	178-VGASSLLAGGPAEK^PEGGVYCGVLSMLECPQAGCK*K^
	236-HFLVHTGEK^PYQCTFEGCGK*R^
	129-GTK*QEKTEVTQETPPLR
	266-FVCPFDGCEK*SFIQSNNQK
	33-LTLK*QDEARPVQNTR
DNMT3b	668-K*GLYEGTGR
	577-LQDFFTTDPDLEEFEPPKLYPAIPAAK*R
Dax1	251-LITLK*DPQVVCEAASAGLLK
	400-GTVLFNPDLPGLQCVK*YIEGLQWR
Dnmt3l	199-HQLK*AFHDEGAGPMEIYK
	255-SLGFLESGSGSGGGTLK*YVEDVTNVVR
Fbxo15	125-ATSVEETATSLSLLSVWDKEDGYWKK*
	130-EYITK*QISSVK
	147-AALTNSLSPVK*R
	21-GSSVTK*QHAWR
	240-WLSLIEK*YDLSNLR
	155-TSLPSK*TK
	248-K*SAMIGCDR
Setdb1	614-GK*NPLLVPLLYDFR
Myc	149-LVSEK*LASYQAAR
Nacc1	167-VK*TEQELDSVQCTPMAK
	182-VKTEQELDSVQCTPMAK*R
	190-RLWDSSQK*EAGGSGGNNGSR
	203-K*MAKFSTPDLALNR
	206-K^MAK*FSTPDLALNR^
	42-LQGLYCDVSVVVK^GHAFK*
	427-VLHAVK*YYCQNFAPNFK
	438-YYCQNFAPNFK*ESEMNAIAADMCTNAR
Trim28	262-LLASLVK*R
	320-VLVNDAQK*VTEGQQER
	750-LQEK*LSPPYSSPQEFAQDVGR
	774-QFNK*LTEDKADVQSIIGLQR

C

	K-GG enrichment	Published ubiquitination	Stem cell funcion
Dax1	yes	In Cos-7 cells (Ehrlund et al., 2009; Kim et al., 2008)	Regulates Oct4 (Kim et al., 2008; Sun et al., 2009)
Dnmt3a	yes	Not published	Important epigenetic regulator. It is essential role inimprinting. (Guryanova and Levine, 2012; Kaneda et al., 2004)
Dppa5	yes	Not published	Involved in the maintenance of pluripotency (Kim et al., 2005)
Fbxo15	yes	Not published	Target of Oct4 (Tokuzawa et al., 2003)
Jarid2	yes	Not published	Regulates transcriptional priming of bivalent genes in pluripotent ES cells (Landeira et al., 2010)
Jmjd1a	yes	Not published	Regulates self-renewal in ES cells (Loh et al., 2007)
Nanog	yes	In ES cells (Ramakrishna et al., 2011)	Key pluripotency factor (Young, 2011)
Ogt	yes	Not published	Pluripotency regulation (Jang et al., 2012)
Phf17	yes	Not published	Belongs to Oct4 network (Jung et al., 2010)
Oct4	yes	In ES cells (Liao and Jin, 2010)	Key pluripotency factor (Young, 2011)
Rex1	yes	In ES cells (Gontan et al., 2012)	Inhibition of cell differentiation (Scotland et al., 2009)
Setdb1	yes	Not published	Regulates maintencance of pluripotency and represses differentiation factors (Bilodeau et al., 2009)

Figure S1. Identification of Ubiquitin-Modified Proteins Using the diGly-Lysine Affinity Enrichment Method

A) ES cells expressing the Nanog-GFP reporter (knock-in). GFP expression was used as an indicator of ES cell pluripotency. B) Identified peptides modified by ubiquitination in self-renewing ES cells. These peptides correspond to known pluripotency regulators. C) Table indicating the function of

selected pluripotency-related proteins for which ubiquitin modifications were found and the information in the literature about their ubiquitination. (Bilodeau et al., 2009; Ehrlund et al., 2009; Gontan et al., 2012; Guryanova and Levine, 2012; Jang et al., 2012; Jung et al., 2010; Kaneda et al., 2004; Kim et al., 2008; Kim et al., 2005; Landeira et al., 2010; Liao and Jin, 2010; Loh et al., 2007; Ramakrishna et al., 2011; Scotland et al., 2009; Sun et al., 2009; Tokuzawa et al., 2003; Young, 2011).

A

Terms (pluripotent conditions)	Count	PValue	Genes
GO:0006412~**translation**	17	3.23E-07	RPL14, RPS15A, RPS4X, RPS5, RPS18, RPL7, RPS16, RPL13A, EIF3E, RPL21, EIF2S1, RPL7L1, RPL10, EIF1, RPS20, RPL4, RPS24
GO:0045934~**negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process**	17	5.65E-06	COPS2, MSH6, NANOG, GTPBP4, JARID2, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, RAD18, DDX20, DNMT3B, DNAJB6, HELLS
GO:0010605~**negative regulation of macromolecule metabolic process**	19	8.06E-06	MSH6, COPS2, NANOG, GTPBP4, JARID2, SKP2, CD276, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, RAD18, DDX20, DNMT3B, DNAJB6, HELLS
GO:0010558~**negative regulation of macromolecule biosynthetic process**	16	4.36E-05	COPS2, NANOG, GTPBP4, JARID2, CD276, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, DDX20, DNMT3B, HELLS, DNAJB6
GO:0016568~**chromatin modification**	12	4.80E-05	KAT2A, PHF17, CHD8, HDAC1, UTY, RNF168, SETD8, KDM3A, DNMT3B, DAPK3, KDM5B, HELLS
GO:0031327~**negative regulation of cellular biosynthetic process**	16	6.01E-05	COPS2, NANOG, GTPBP4, JARID2, CD276, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, DDX20, DNMT3B, HELLS, DNAJB6
GO:0009057~**macromolecule catabolic process**	20	7.35E-05	RAD23A, SKP2, CDC20, UBE2C, PSMA2, SUMO3, SUMO2, MIB1, CUL5, PSMC4, APOE, EIF3E, PSMB3, UBA2, KLHL13, RAD18, RNF168, FBXO15, CUL4B, TRIP12
GO:0010629~**negative regulation of gene expression**	15	1.34E-04	COPS2, NANOG, JARID2, SKP2, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, DDX20, DNMT3B, HELLS, DNAJB6
GO:0016481~**negative regulation of transcription**	14	1.87E-04	COPS2, NANOG, JARID2, NR0B1, DAXX, TCFCP2L1, PA2G4, HDAC1, POU5F1, ZFP281, DDX20, DNMT3B, HELLS, DNAJB6
GO:0051603~**proteolysis involved in cellular protein catabolic process**	17	2.00E-04	RAD23A, SKP2, CDC20, UBE2C, PSMA2, SUMO3, SUMO2, MIB1, CUL5, PSMB3, UBA2, KLHL13, RAD18, RNF168, FBXO15, CUL4B, TRIP12

B

Term (differentiation conditions)	Count	PValue	Genes
GO:0007049~**cell cycle**	22	1.28E-06	FAM33A, SUV39H1, BRCA2, 1700017B05RIK, MYH9, CCNG1, CDC25A, MFN2, REC8, CDKN2A, PLK2, STRA8, KRT7, ZWINT, B230120H23RIK, CYP26B1, PKD2, MAPRE2, HAUS7, PDCD6IP, STAG1, MYH10
GO:0001701~**in utero embryonic development**	13	2.18E-05	UBR3, BRCA2, KEAP1, MYH9, AKT1, MFN2, RDH10, KRT19, ZMIZ1, KRT8, PKD2, MYH10, CYR61
GO:0051301~**cell division**	13	3.60E-05	FAM33A, BRCA2, 1700017B05RIK, CCNG1, MYH9, CDC25A, CDKN2A, ZWINT, HAUS7, MAPRE2, PDCD6IP, MYH10, STAG1
GO:0043009~**chordate embryonic development**	15	1.22E-04	HSPG2, UBR3, BRCA2, KEAP1, MYH9, MFN2, AKT1, RDH10, KRT19, ZMIZ1, KRT8, PKD2, PBX1, MYH10, CYR61
GO:0006720~**isoprenoid metabolic process**	6	1.48E-04	RDH10, HMGCR, CRABP2, CYP26B1, CYP26A1, FDFT1
GO:0000279~**M phase**	12	1.76E-04	FAM33A, REC8, STRA8, ZWINT, CYP26B1, BRCA2, HAUS7, MAPRE2, MYH9, CCNG1, CDC25A, STAG1
GO:0001707~**mesoderm formation**	5	5.06E-04	TWSG1, PRKAR1A, TXNRD1, PRKACA, BMP7
GO:0042573~**retinoic acid metabolic process**	4	6.69E-04	RDH10, CRABP2, CYP26B1, CYP26A1
GO:0001704~**formation of primary germ layer**	5	7.60E-04	TWSG1, PRKAR1A, TXNRD1, PRKACA, BMP7
GO:0010608~**posttranscriptional regulation of gene expression**	8	8.84E-04	AKT1, EIF4A3, EIF2C2, EIF2C1, CDKN2A, KRT7, IGF2BP2, FLNA

C

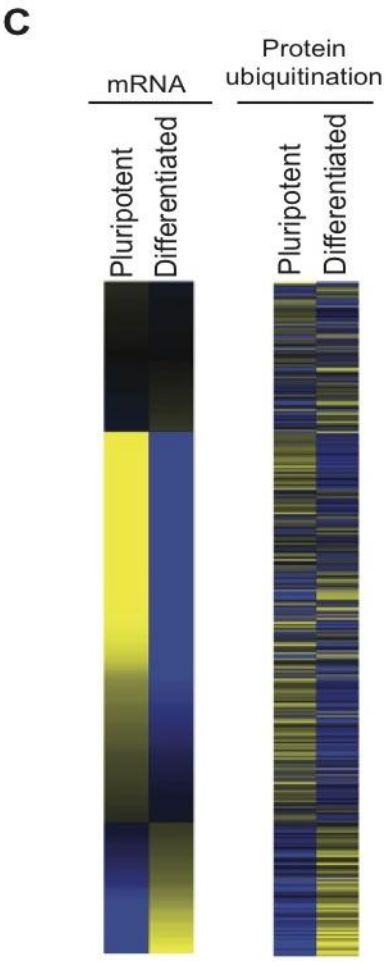


Figure S2. **Ubiquitinated Proteins in Pluripotent and Differentiated ESCs**
A) DAVID Analysis Tool Software gene ontology classification of the proteins showing a fold change cutoff of 2 of the pluripotent population over the differentiated population. B) DAVID Analysis Tool Software gene ontology classification of the proteins showing a fold change cutoff of 2 of the differentiated population over the pluripotent population. C) Heat maps showing mRNA expression in pluripotent mES cells and 4 days differentiated mES cells with retinoic acid (microarray data meta-analysis from Ivanova et al) (left) and quantified ubiquitinated peptides (right).

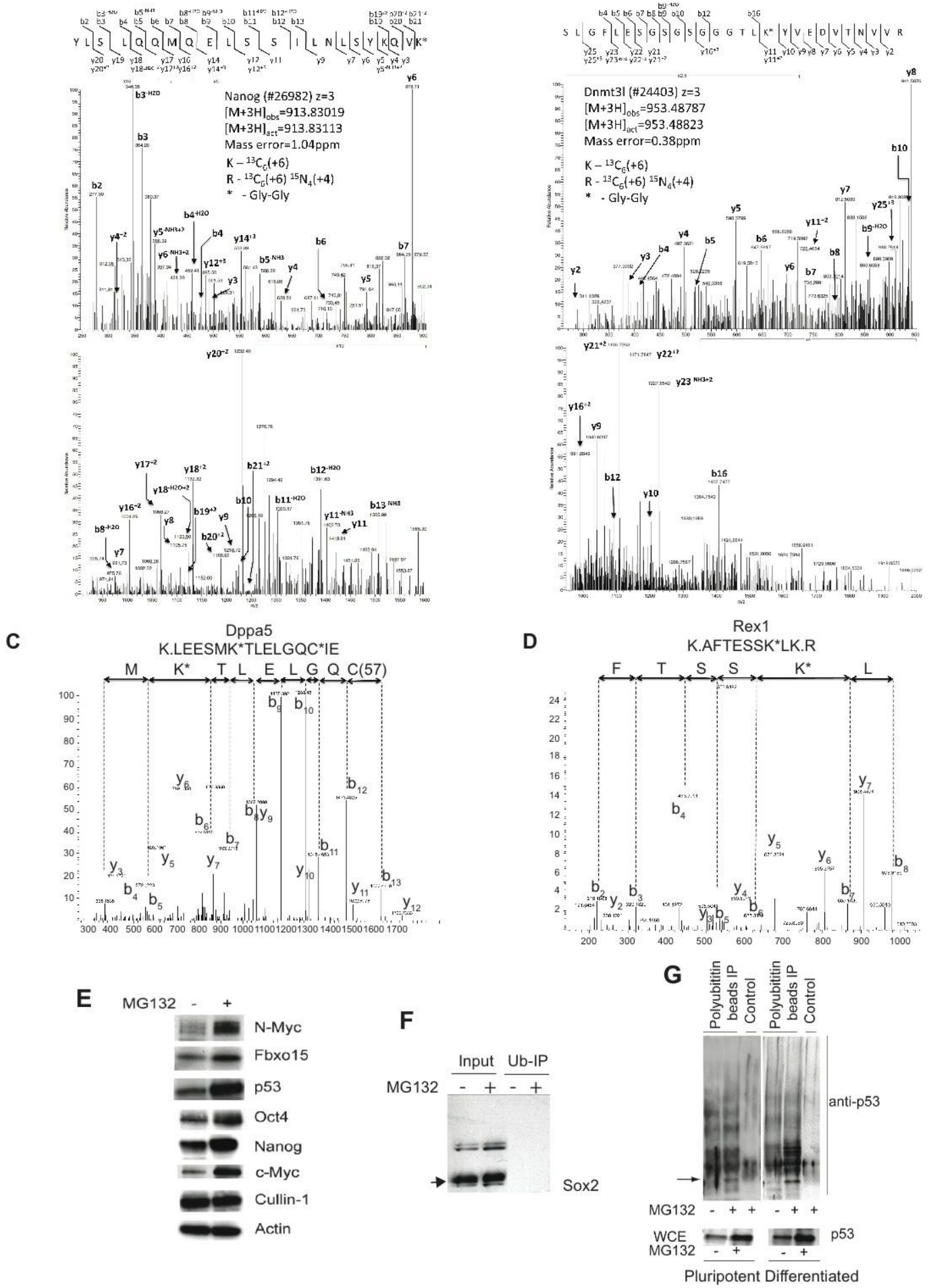

A
Heavy labelled Nanog peptide
Nanog (#26982) z=3
[M+3H]obs=913.83019
[M+3H]act=913.83113
Mass error=1.04ppm
K – 13C6(+6)
R - 13C6(+6) 15N4(+4)
* - Gly-Gly
B
Heavy labelled Dnmt3l peptide
Dnmt3l (#24403) z=3
[M+3H]obs=953.48787
[M+3H]act=953.48823
Mass error=0.38ppm
K – 13C6(+6)
R - 13C6(+6) 15N4(+4)
* - Gly-Gly
C
Dppa5
K.LEESMK*TLELGQC*IE
D
Rex1
K.AFTESSK*LK.R
E
MG132 - +
N-Myc
Fbxo15
p53
Oct4
Nanog
c-Myc
Cullin-1
Actin
F
Input
Ub-IP
MG132 - + - +
Sox2
G
Polyubitin beads IP
Control
anti-p53
MG132 - + + - + +
WCE
MG132 - + - +
p53
Pluripotent Differentiated

Figure S3. Validation of Identified Ubiquitin-Modified Peptides

A-D) Representative annotated MS/MS spectra for identified remnant–containing peptides obtained by immunoprecipitation in pluripotent ES cells. A) heavy labeled Nanog B) heavy labeled Dnmt3l C) Dppa5. D) Rex1 The sequence of the ubiquitinated peptide and the diglycine-modified lysine (K*) are indicatd. E) Immunoblot of pluripotent ES cell extracts after inhibition of the proteasome with MG132 in self-renewal conditions compared to mock control with DMSO. F) *In vivo* ubiquitination of Sox2 showing lack of such modification. G) *In vivo* ubiquitination assay using poly-ubiquitin or control beads of self-renewing and differentiated ES cells (4 days) treated with DMSO, as mock control, or MG132. Western blot was incubated with p53 antibody.

Figure S4

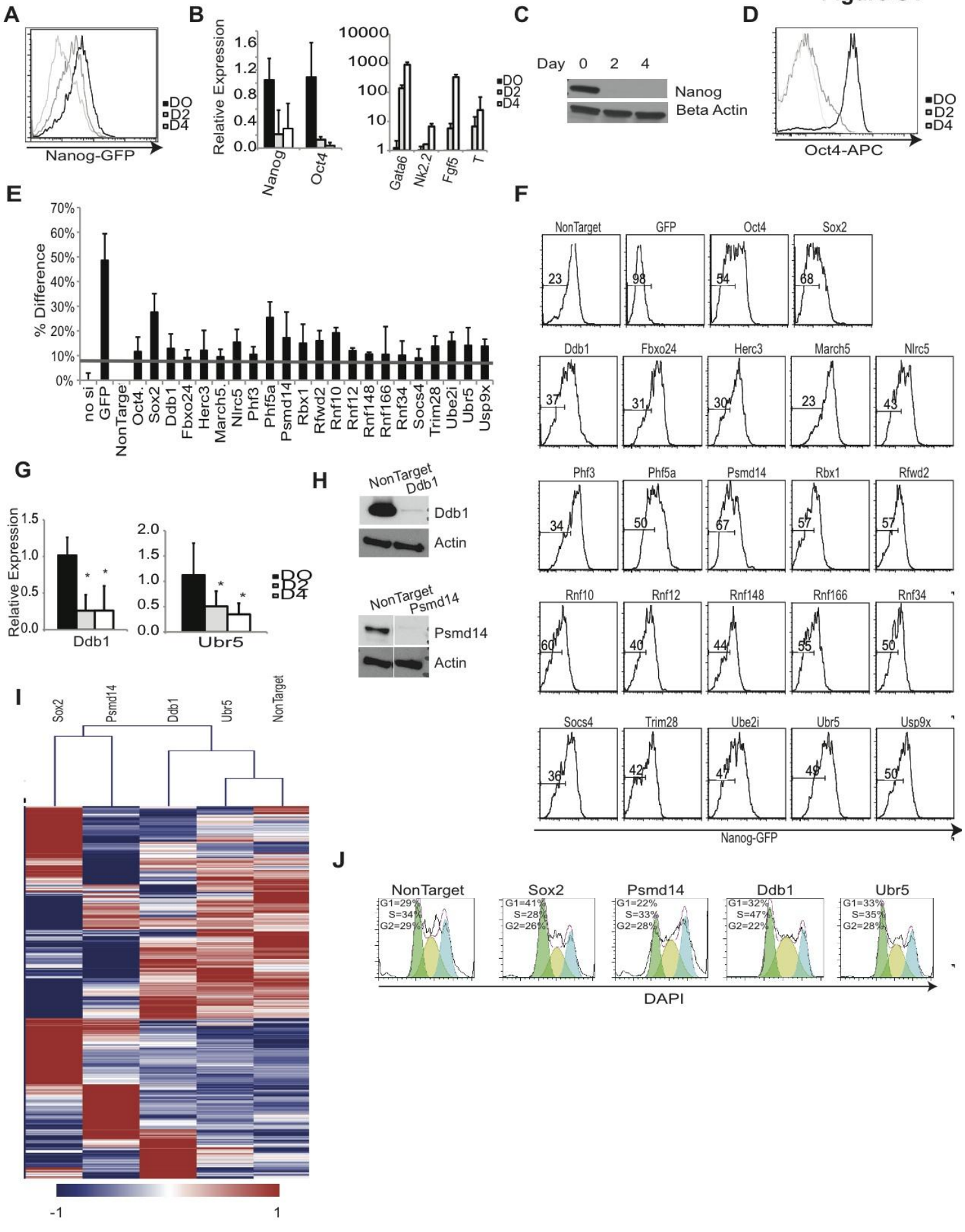
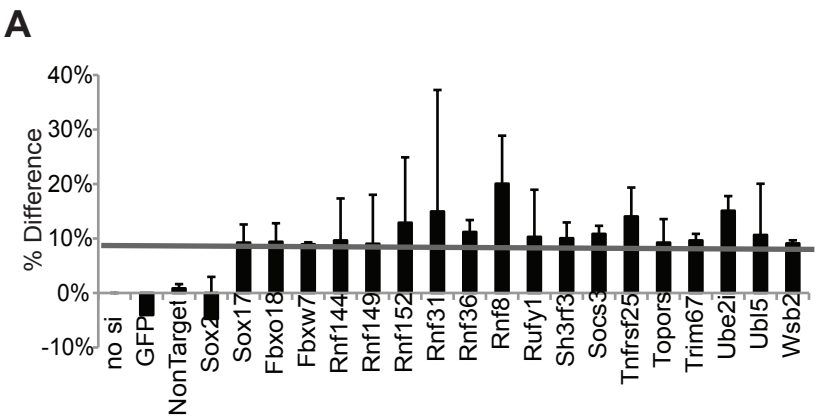
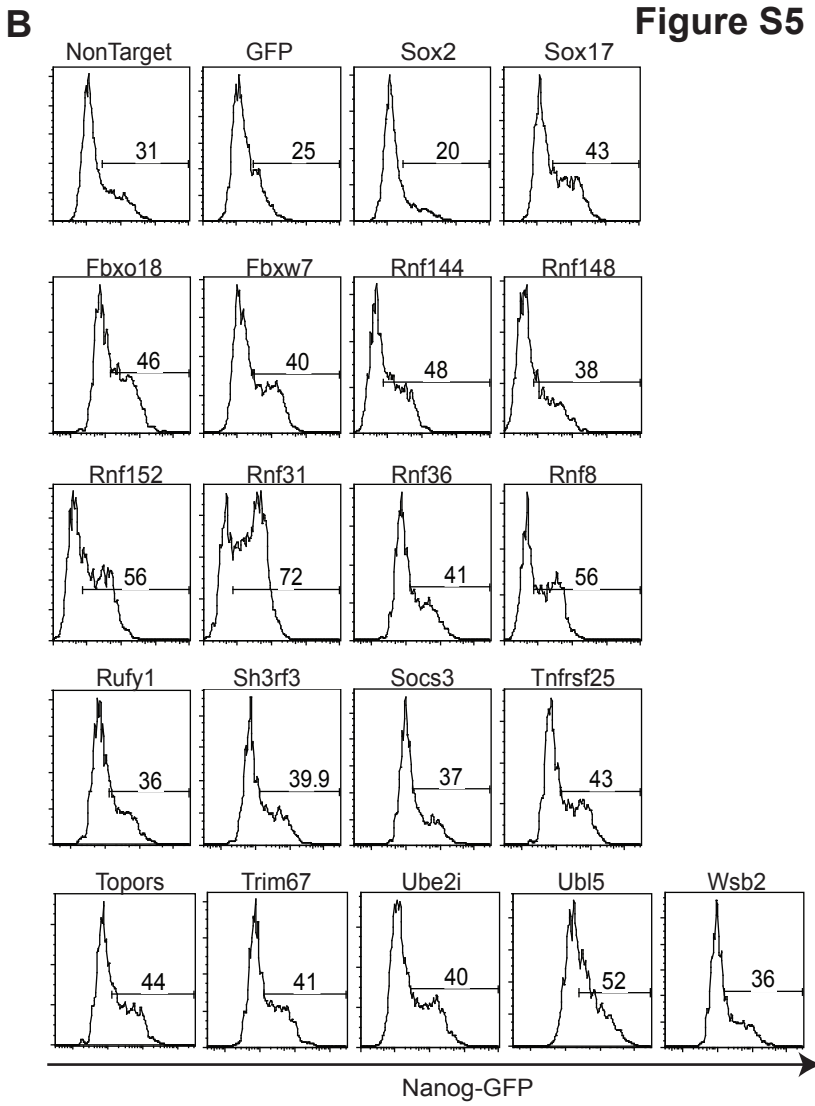

A
Nanog-GFP
DO
D2
D4
B
Relative Expression
Nanog
Oct4
Gata6
Nk2.2
Fgf5
T
C
Day 0 2 4
Nanog
Beta Actin
D
Oct4-APC
E
% Difference
no si
GFP
NonTarge
Oct4.
Sox2
Ddb1
Fbxo24
Herc3
March5
Nlrc5
Phf3
Phf5a
Psmd14
Rbx1
Rfwd2
Rnf10
Rnf12
Rnf148
Rnf166
Rnf34
Socs4
Trim28
Ube2i
Ubr5
Usp9x
F
NonTarget
GFP
Oct4
Sox2
Ddb1
Fbxo24
Herc3
March5
Nlrc5
Phf3
Phf5a
Psmd14
Rbx1
Rfwd2
Rnf10
Rnf12
Rnf148
Rnf166
Rnf34
Socs4
Trim28
Ube2i
Ubr5
Usp9x
Nanog-GFP
G
Relative Expression
Ddb1
Ubr5
H
NonTarget
Ddb1
Ddb1
Actin
NonTarget
Psmd14
Psmd14
Actin
I
Sox2
Psmd14
Ddb1
Ubr5
NonTarget
-1
1
J
NonTarget
Sox2
Psmd14
Ddb1
Ubr5
G1=29%
S=34%
G2=29%
G1=41%
S=28%
G2=26%
G1=22%
S=33%
G2=28%
G1=32%
S=47%
G2=22%
G1=33%
S=35%
G2=28%
DAPI

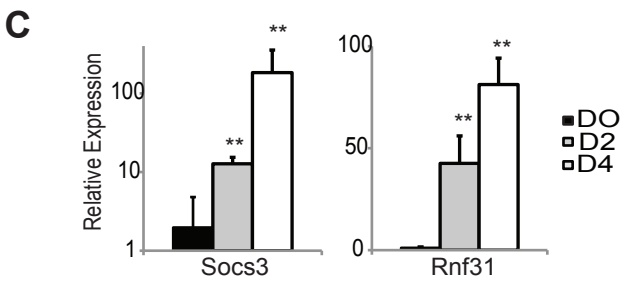
Figure S4. Self-Renewal Hits Identified in siRNA Screen

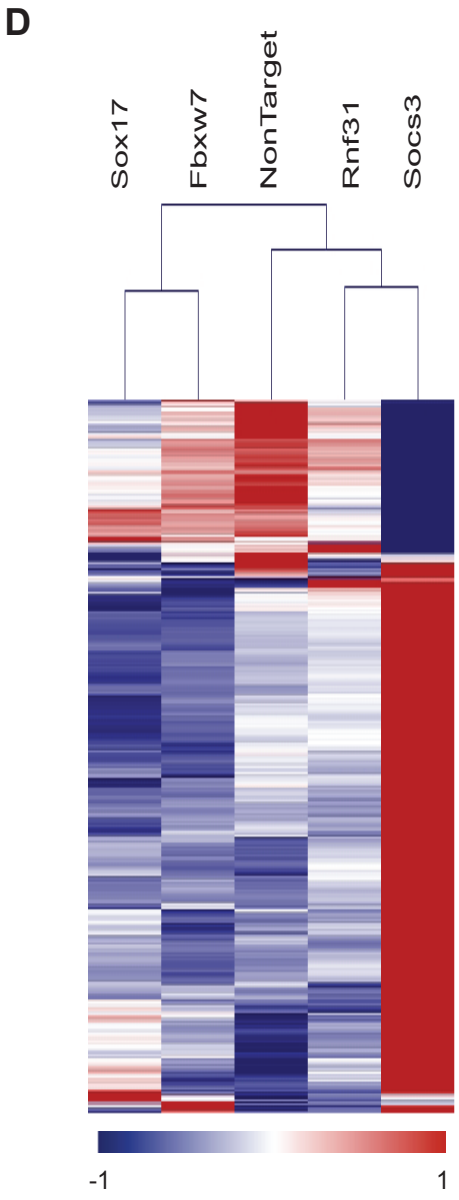
Nanog GFP cells with differentiated for 4 days under treatment with 5uM RA and LIF withdrawl. A) GFP expression was analyze by flow cytometry. B) Relative expression of pluripotency genes (Oct4 and Nanog), early endoderm gene (Gata6) and early ectoderm gene (Nk2.2) by qRT-PCR. C) Western plot during differentiation with anti-Nanog D) Intracellular FACS for Oct4-APC. E) Bar graph representing triplicate experiments. F) Representative FACS plot from one experiment. G) Relative expression of genes during differentiation by qRT-PCR. Cells were treated with 5μM RA for 2 and 4 days. Data represented as +SEM; N=3. *p-value <0.05, **p-value<0.01. H) Knockdown of target genes by western blot. I) Heatmap illustrating Log2 and clustering of all differentially expressed genes (>2-fold) detected using Affymetrix microarray analysis. J) Cell cycle analysis was performed after knockdown. Cells fixed in 4% PFA and stained with 5μg/mL DAPI.

Figure S5


A
40%
30%
20%
10%
0%
-10%
% Difference
no si
GFP
NonTarget
Sox2
Sox17
Fbxo18
Fbxw7
Rnf144
Rnf149
Rnf152
Rnf31
Rnf36
Rnf8
Rufy1
Sh3rf3
Socs3
Tnfrsf25
Topors
Trim67
Ube2i
Ubl5
Wsb2


B
NonTarget
GFP
Sox2
Sox17
31
25
20
43
Fbxo18
Fbxw7
Rnf144
Rnf148
46
40
48
38
Rnf152
Rnf31
Rnf36
Rnf8
56
72
41
56
Rufy1
Sh3rf3
Socs3
Tnfrsf25
36
39.9
37
43
Topors
Trim67
Ube2i
Ubl5
Wsb2
44
41
40
52
36
Nanog-GFP


C
Relative Expression
100
10
1
**
**
Socs3
100
50
0
**
**
Rnf31
DO
D2
D4


D
Sox17
Fbxw7
NonTarget
Rnf31
Socs3
-1
1

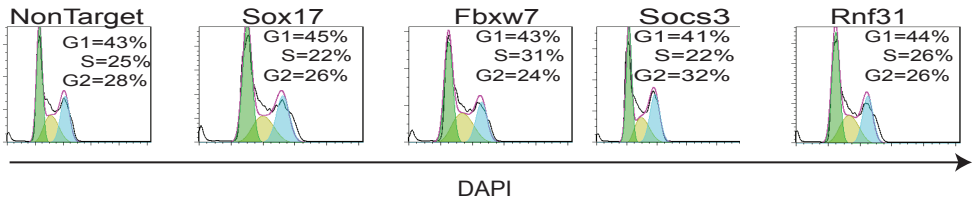

E
NonTarget
G1=43%
S=25%
G2=28%
Sox17
G1=45%
S=22%
G2=26%
Fbxw7
G1=43%
S=31%
G2=24%
Socs3
G1=41%
S=22%
G2=32%
Rnf31
G1=44%
S=26%
G2=26%
DAPI

Figure S5. Differentiation Hits Identified in siRNA Screen

A) Bar graph representing triplicate experiments. B) Representative FACS plots from one experiment. C) Relative expression of genes during differentiation using qRT-PCR. Cells were treated with 5μM RA for 2 and 4 days. Data represented as +SEM; N=3. *p-value <0.05, **p-value<0.01.D) Heatmap illustrating Log2 and clustering of all differentially expressed genes (>2-fold and p<0.05) detected using Affymetrix microarray analysis. E) Cell cycle analysis was performed after knockdown. Cells fixed in 4% PFA and stained with 5μg/mL DAPI.

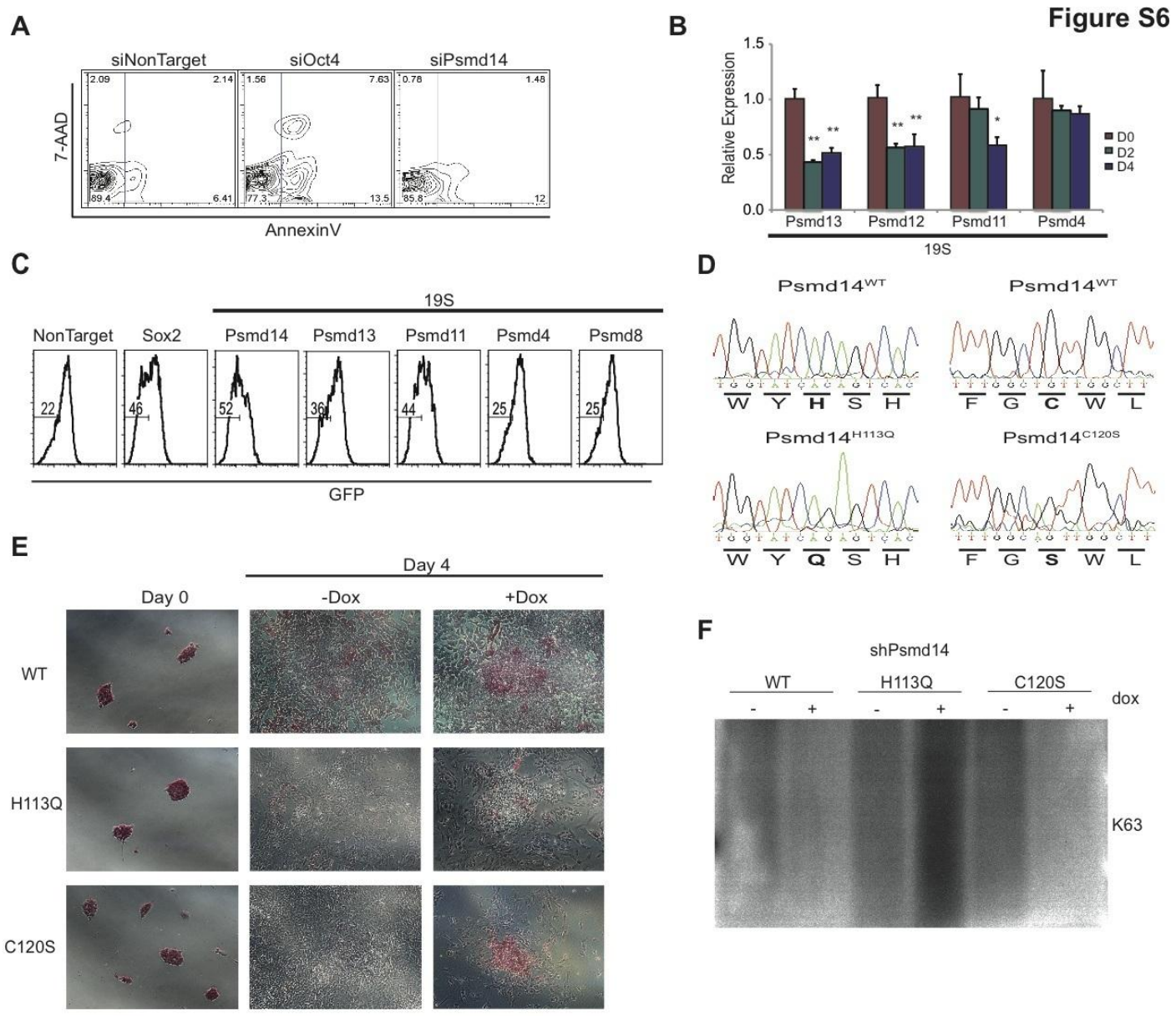
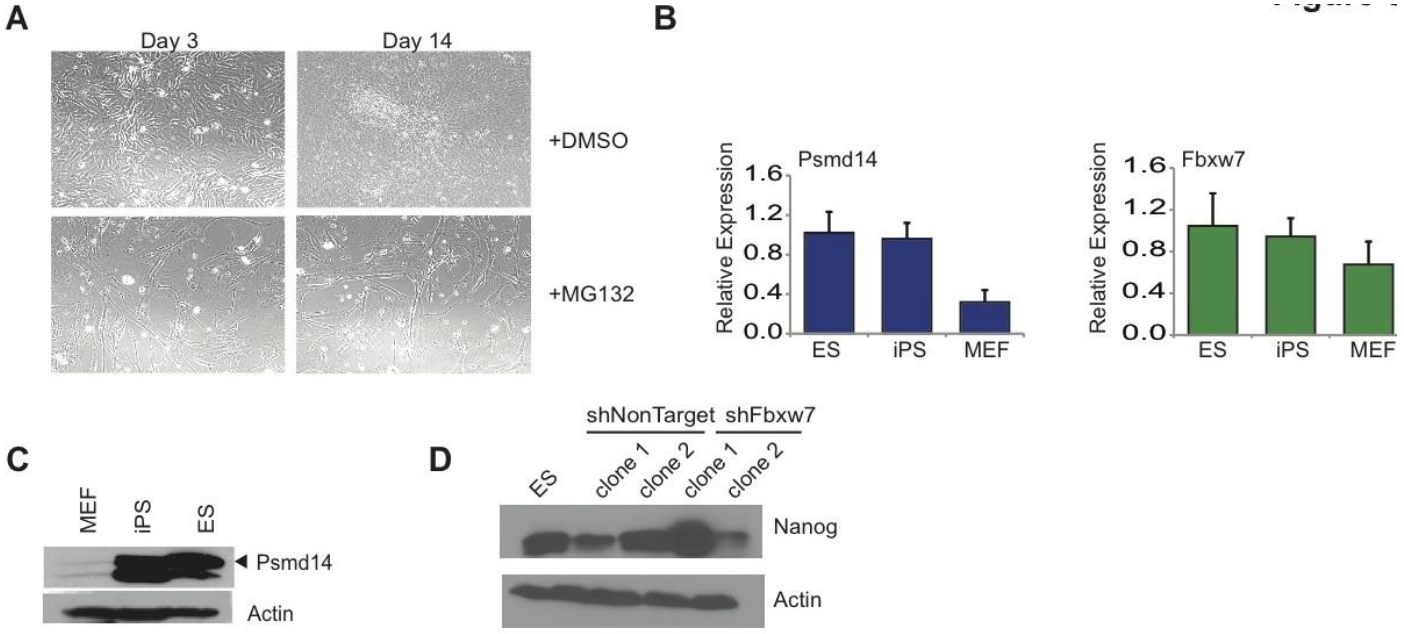


Figure S6. Psmd14-Interacting Proteins and Mutations

A) Annexin V and 7-AAD staining of ES cells transfected with siRNA against non-target control, Oct4 or Psmd14. B) Relative expression by qRT-PCR of selected Psmd14-interacting proteins and components of the 19S regulatory particle of the proteasome during differentiation. C) Histograms depicting Nanog-GFP intensity following knockdown of proteasome lid components. D) Sequencing analysis of the $Psmd14^{WT}$, $Psmd14^{H113Q}$ and $Psmd14^{C120S}$ coding sequence to verify mutagenesis. E) Alkaline phosphatase labeling of pluripotent ES cells and cells differentiated for 4 days with the addition of 5μm RA in the presence or absence of doxycycline. F) Western blot analysis of K63-

linkage specific polyubiquitinated proteins following knockdown of endogenous Psmd14 and induced expression of $Psmd14^{WT}$, $Psmd14^{H113Q}$ or $Psmd14^{C120S}$.



E

Biological processes in IPS and ESC common proteins

Category	Term	Count	%	PValue
GOTERM_BP_FAT	GO:0007049~cell cycle	37	15.4166667	8.81E-14
GOTERM_BP_FAT	GO:0006259~DNA metabolic process	28	11.6666667	2.34E-11
GOTERM_BP_FAT	GO:0051301~cell division	22	9.16666667	2.89E-10
GOTERM_BP_FAT	GO:0006281~DNA repair	18	7.5	1.09E-08
GOTERM_BP_FAT	GO:0007067~mitosis	16	6.66666667	5.52E-08
GOTERM_BP_FAT	GO:0000280~nuclear division	16	6.66666667	5.52E-08
GOTERM_BP_FAT	GO:0010605~negative regulation of macromolecule metabolic process	25	10.4166667	1.19E-07
GOTERM_BP_FAT	GO:0051276~chromosome organization	22	9.16666667	1.73E-07
GOTERM_BP_FAT	GO:0051253~negative regulation of RNA metabolic process	19	7.91666667	2.76E-07
GOTERM_BP_FAT	GO:0006350~transcription	51	21.25	3.56E-07
GOTERM_BP_FAT	GO:0010629~negative regulation of gene expression	21	8.75	9.44E-07
GOTERM_BP_FAT	GO:0016568~chromatin modification	15	6.25	4.83E-06

Biological processes in ESC iPS and MEFs common proteins

Category	Term	Count	%	PValue
GOTERM_BP_FAT	GO:0006412~translation	66	8.979592	5.59E-27
GOTERM_BP_FAT	GO:0030163~protein catabolic process	85	11.56463	3.45E-25
GOTERM_BP_FAT	GO:0051603~proteolysis involved in cellular protein catabolic process	78	10.61224	7.68E-22
GOTERM_BP_FAT	GO:0044257~cellular protein catabolic process	78	10.61224	1.09E-21
GOTERM_BP_FAT	GO:0043632~modification-dependent macromolecule catabolic process	73	9.931973	5.00E-20
GOTERM_BP_FAT	GO:0019941~modification-dependent protein catabolic process	73	9.931973	5.00E-20
GOTERM_BP_FAT	GO:0006511~ubiquitin-dependent protein catabolic process	28	3.809524	3.62E-11
GOTERM_BP_FAT	GO:0006508~proteolysis	87	11.83673	8.76E-10
GOTERM_BP_FAT	GO:0034622~cellular macromolecular complex assembly	32	4.353741	2.77E-09
GOTERM_BP_FAT	GO:0008104~protein localization	67	9.115646	1.36E-08
GOTERM_BP_FAT	GO:0034621~cellular macromolecular complex subunit organization	33	4.489796	1.43E-08
GOTERM_BP_FAT	GO:0051301~cell division	34	4.62585	1.14E-07
GOTERM_BP_FAT	GO:0070647~protein modification by small protein conjugation or removal	19	2.585034	3.40E-07

Figure S7. Members of the UPS and the Proteasome Itself Are Required for Cellular Reprogramming

A) Bright field image of OSKM MEFs treated with Dox and either DMSO or 1μM MG132 for 3 and 14 days. B) Relative expression of UPS genes by qRT-PCR. Data represented as +SEM; N=3. *p-value <0.05, **p-value<0.01. C) Western blot of Psmd14 in ES, iPS, and MEF populations. D) Western blot of Nanog in individual iPS clones derived from OSKM MEFS targetes with either NonTarget shRNA or shRNA against Fbxw7. E) DAVID Analysis Tool Software gene ontology classification of identified ubiquitinated proteins common between ES and iPS cells, and common between ES, iPS cells and MEFs.

The following supplemental tables are available as separate Excel files:

Table S1. Mass-Spectrometry Results of the Ubiproteome Study in Self-Renewing and Differentiated ESCs

Detection and quantification of the peptides and calculations for the generation of the heat map are shown in different sheets. Both SILAC and label-free experiments are shown.

Table S3. Total and Ubiquitin-Modified Protein Expression Profiles of Pluripotent and Differentiated ESCs in the Presence or Absence of Proteasome Inhibitors

Table S7. Mass-Spectrometry Results of the Ubiproteome Study in iPSCs and MEFs

Detection and quantification of the peptides for the two different populations are shown in different sheets.

Table S2. Number of Peptides Identified in the SILAC and Label-free Experiments in the K-GG Enrichment for Selected Proteins of the Pluripotency Network

	Number of peptides identified in the K-GG enrichment	
Gene	***SILAC***	***Label free***
Nanog	2	2
Pou5f1	3	2
Nr0b1	2	2
Zfp42	3	4
Actl6a	1	0
Amotl2	1	2
Chd4	1	0
Dnaja2	2	2
DNMT3b	1	1
Dnmt3l	2	1
Dppa5	9	6
Fbxo15	1	6
Hdac1	2	1
Jarid2	1	2
Las1l	1	2
Med12	1	2
Msh2	2	0
Msh6	2	3
Mybbp1a	1	1
Myc	1	1
Nacc1	8	5
Nup188	2	1
Nup88	3	3
Ogt	2	1
Parp1	1	0
Phf17	5	8
Pml	3	6
Rad23b	3	1
Sall4	0	1
Setdb1	1	1
Tex10	2	3
Top2a	6	1
Tp53	3	4
Trim28	4	3
Utp18	2	2
Znf281	2	1

Table S4. Genes with a *Z* score of >1 and <–1 for the ESC Self-Renewal Screen

Self-renewal Screen			
gene	z-score	gene	z-score
Phf5a	4.9	Birc6	-2.1
Ddb1	4.3	Phf17	-2.0
Psmd14	3.6	Prp19	-1.8
Trim28	3.1	Fbxw5	-1.7
LOC434341	2.6	Fbxo6b	-1.6
Rnf10	2.5	Asb3	-1.6
Ubr5	2.4	Rnf8	-1.5
Herc3	2.2	Mex3a	-1.5
Ube2i	2.0	Usp37	-1.5
Phf3	1.9	Prickle1	-1.4
Usp9x	1.8	LOC244421	-1.4
Rfwd2	1.8	Usp7	-1.3
Rbx1	1.8	Usp25	-1.3
Fbxo24	1.7	Trim33	-1.3
Zfp592	1.7	Wdr59	-1.3
Btrc	1.6	Fbxo5	-1.2
Socs4	1.6	Fbxo47	-1.2
Dcun1d3	1.5	Prickle2	-1.2
Cbll1	1.4	BC059845	-1.2
Fbxo43	1.4	Fbxw7	-1.2
G2e3	1.4	Lmo6	-1.2
March5.	1.4	Ube3e	-1.2
Trim56	1.4	Fbxl18	-1.2
Phf19	1.3	Rnf39	-1.1
Oit3	1.3	Uhrf1	-1.1
Phf15	1.2	Hr	-1.1
Chd5	1.2	Ube3c	-1.1
Rnf12	1.2	Aire	-1.1
Fbxl12	1.2	Cand2	-1.1
Ube2d3	1.2	Neurl2	-1.1
Sytl4	1.2	Pja2	-1.0
Trim7	1.2	Ube2w	-1.0
Cul3	1.2	Usp22	-1.0
LOC380928	1.1	Rnf152	-1.0
Zfand6	1.1	Attp	-1.0
Zfp216	1.0		
Fbxl2	1.0		
Otud7b	1.0		
Rnf148	1.0		
Rnf166	1.0		
Mid2	1.0		

Table S5. Genes with a *Z* score >1 and <–1 for the ESC Differentiation Screen

Differentiation Screen			
gene	z-score	gene	z-score
Rnf8	3.5	Trim28	-2.0
Ube2i	3.5	Cul3	-2.0
Fbxo41	2.4	Fbxl12	-2.0
Cul4B	2.3	Ube2v2	-1.8
Rnf31	2.1	Fbxo24	-1.8
Wsb2	2.1	Sytl3	-1.8
Ubl5	1.9	Asb8	-1.7
Prp19	1.8	Mll2	-1.6
Ube1x	1.7	LOC434341	-1.6
Rnf152	1.6	Rnf141	-1.5
Socs3	1.5	Ube2e2	-1.4
Rnf149	1.5	Fbxl16	-1.4
LOC244421	1.4	Brpf1	-1.4
Park2	1.4	Cnot4	-1.4
Amfr	1.4	LOC268291	-1.4
Rnf144	1.3	Ube2v1	-1.3
Wdr59	1.3	Zfp592	-1.3
Wsb1	1.3	Bmi1	-1.3
Znrf3	1.2	Phf6	-1.3
Fbxo18	1.2	Rkhd2	-1.2
Fbxl6	1.1	Rnf146	-1.2
Zmynd11	1.1	Dcun1d3	-1.2
Tnfrsf25	1.1	Rnf12	-1.2
Rnf7	1.1	Birc4	-1.2
Fbxw7	1.1	Rkhd1	-1.2
Rnf219	1.1	Phf3	-1.1
Cul2	1.1	Rapsn	-1.1
		Socs6	-1.1
		Ube2z	-1.1
		Rbx1	-1.1
		LOC239618	-1.0
		Pxmp3	-1.0
		Jarid1B	-1.0
		Rc3h1	-1.0
		Rnf2	-1.0
		LOC381534	-1.0
		Wwp1	-1.0
		Ube2e3	-1.0
		Nedd4	-1.0
		Rnf25	-1.0

Table S6. Psmd14-Interacting Proteins

Interacting Proteins as Identified by Mass Spectrometry Analysis of Purified Samples						
			Number of unique peptides			
#	Gene	Accession Number	GFP	PSM14	Molecular Weight	Description
1	MYH7	187956918	1	50	223 kDa	myosin, heavy chain 7, cardiac muscle, beta
2	MYH2	205830428	1	35	223 kDa	myosin, heavy chain 2, skeletal muscle, adult
3	TTN	148695270	0	33	3766 kDa	titin
4	PSMD12	12857304	0	15	53 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12
5	ACTN2	157951643	0	14	104 kDa	actinin, alpha 2
6	PSMD11	134053905	0	13	47 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11
7	ACTA1	4501881	0	12	42 kDa	actin, alpha 1, skeletal muscle
8	MYH4	67189167	0	10	223 kDa	myosin, heavy chain 4, skeletal muscle
9	PSMD13	6755210	0	10	43 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 13
10	TPM2	11875203	0	9	33 kDa	tropomyosin 2 (beta)
11	PSMD14	5031981	0	8	35 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14
12	PSMD7	26326937	0	8	35 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 7
13	TPM1	256000786	0	7	33 kDa	tropomyosin 1 (alpha)
14	FLNC	124487139	0	7	291 kDa	filamin C, gamma
15	MYL3	148677054	0	7	24 kDa	myosin, light chain 3, alkali; ventricular, skeletal, slow
16	PSMD6	46049022	0	6	46 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 6
17	MYL2	153791853	0	6	19 kDa	myosin, light chain 2, regulatory, cardiac, slow
18	PSMD3	19705424	0	6	61 kDa	proteasome (prosome, macropain) 26S subunit, non-ATPase, 3
19	ACTN3	7304855	0	6	103 kDa	actinin, alpha 3
20	DES	148667985	0	6	56 kDa	desmin
21	NEB	148222065	0	5	801 kDa	nebulin
22	MYLPF	7949078	0	5	19 kDa	myosin light chain, phosphorylatable, fast skeletal muscle
23	MYOM2	170763465	0	5	165 kDa	myomesin (M-protein) 2, 165kDa
24	MYL1	164664497	0	5	17 kDa	myosin, light chain 1, alkali; skeletal, fast
25	TNNC1	148692849	0	4	19 kDa	troponin C type 1 (slow)
26	TPM3	20178336	0	4	33 kDa	tropomyosin 3
27	MYH1	187956263	0	3	223 kDa	myosin, heavy chain 1, skeletal muscle, adult
28	ALDOA	293597567	0	3	15 kDa	aldolase A, fructose-bisphosphate
29	TNNT3	123242961	0	3	31 kDa	troponin T type 3 (skeletal, fast)
30	TNNC2	6678371	0	3	18 kDa	troponin C type 2 (fast)
31	CRYAB	6753530	0	3	20 kDa	crystallin, alpha B
32	ATP2A1	148685410	0	3	55 kDa	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1
33	ATP5B	23272966	0	3	57 kDa	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide
34	TNNI1	10946862	0	2	22 kDa	troponin I type 1 (skeletal, slow)
35	MYH6	6754774	0	2	225 kDa	myosin, heavy chain 6, cardiac muscle, alpha
36	ATP5A1	148677500	0	2	52 kDa	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle
37	MYH3	153792649	0	2	224 kDa	myosin, heavy chain 3, skeletal muscle, embryonic
38	MYH8	71143152	0	2	223 kDa	myosin, heavy chain 8, skeletal muscle, perinatal
39	CKM	12845061	0	2	43 kDa	creatine kinase, muscle
40	eGFP	13194618	7	0	27 kDa	enhanced green fluorescent protein

Supplemental Experimental Procedures

Quantification and identification of affinity enriched ubiquitin-modified peptides using the diGly-lysine antibody

SILAC (Stable isotope labeling by amino acids in cell culture) labeling of ES cells.

The pluripotent ES cells were metabolically labeled with $^{13}C_6{}^{15}N_4$-Arginine and $^{13}C_6$-Lysine amino acids using the Mouse Embryonic Stem Cell SILAC Kit (Pierce). The medium was supplemented with GlutaMAX and 20% dialyzed mES cell tested FBS. The SILAC culture media were replaced daily. For the differentiated cells, the the medium was kept with light amino acids For differentiation, 5μM retinoic acid was added and LIF was removed during four days. Medium was replaced daily. 20 μM MG132 was added for 4h to all samples prior to collection. After harvest, cells were snap frozen.

Sample Preparation for ubiquitin-modified peptide enrichment

Lysates were prepared by homogenization in urea lysis buffer (20mM HEPES (pH 8.0), 9M Urea, 1mM sodium orthovanadate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerol-phosphate) followed by sonication at 15W output for 25 seconds and centrifuged 15 minutes at 20,000 x g to remove insoluble material. For the SILAC experiment, the same amount of lysate (17.6mg) of both populations (pluripotent and differentiated) were mixed. For the label free experiments, $2x10^8$ cells for each population (pluripotent, differentiated, iPS cells or MEFs) were used to do duplicates. The resulting "cleared" protein extracts were reduced, carboxamidomethylated, and digested with trypsin. The resulting peptides were separated from non-peptide material by solid-phase extraction with Sep-Pak C18 cartridges and lyophilized. Lyophilized peptides were re-dissolved, and ubiquitin-modified peptides were enriched using the diGly-lysine antibody conjugated to Protein-A agarose. The enriched peptides were eluted from antibody-resin in 100 μL of 0.15% TFA. Eluted peptides were desalted and concentrated with PerfectPure C18 tips immediately prior to LC-MS/MS analysis.

Mass spectrometry analysis for ubiquitin-modified peptides

Peptides were loaded directly onto a 10 cm x 75 μm PicoFrit capillary column packed with Magic C18 AQ reversed-phase resin. Peptides were resolved with a 90-minute linear gradient of acetonitrile in 0.125% formic acid at a flow rate of 280 nL/min. Tandem mass spectra were collected with an LTQ-Orbitrap hybrid mass spectrometer, using a top-ten method, a dynamic exclusion repeat count of 1 and a repeat duration of 30 seconds. The collected MS/MS spectra were evaluated using TurboSequest in the Proteomics Browser package (v.27, rev.13) and searched using the following parameters: peptide ion mass tolerance, 2.5 Da; fragment ion mass tolerance, 1.0 Da; maximum number of differential amino acids per modification, 4; parent ion mass type, monoisotopic; fragment ion mass type, monoisotopic; maximum number of internal cleavage sites, 4; neutral losses of water and ammonia from b and y ions were considered in the correlation analysis and the proteolytic enzyme was specified, trypsin. Searches were performed against the NCBI mouse database, released on October 18, 2006, in both forward- and reversed-sequence directions. Cysteine carboxamidomethylation was specified as a static modification, oxidation of methionine residues was allowed, and the K-e-GG (+114) modification was allowed on lysine. For the SILAC experiments, carbamidomethyl cysteine and the appropriate heavy SILAC residue masses (+6.020 Da) for Lys and (+10.008 Da) for Arg were included as static modifications, and methionine oxidation and diglycine modification of lysine were allowed as dynamic modifications as appropriate.

The false positive rate was approximated by taking the ratio of the reversed database assignements to the forward database assignments after filtering the initial SEQUEST/Sorcerer search results based on

XCorr (≥ 1.5), mass accuracy (± 10 ppm) and a minimum RSp score of below 50. The final false positive rate was set at 1%. The mass accuracy range was narrowed further based on the XCorr-mass error plot for each experiment. The range in false-positive assignment rates obtained from the SEQUEST/Sorcerer search of all samples for each motif antibody in this study was determined and is shown in each section summary.

MS/MS Quantitation

Each MS/MS spectrum arises from a parent ion observed during a survey MS scan and can be linked to the intensity of that parent ion at its chromatographic apex, essentially measuring the abundance of the peptide in the sample. Parent ion intensities were obtained from each sample's LC-MS data file using XCalibur software 2.0.5 from ThermoFinnigan (San Jose, CA) and are reported in the summary tab of the quantification table (Table S1 for SILAC and Table S1 for Label Free).

Changes in ubiquitin peptide levels were measured by taking the ratio of raw intensities between self-renewal, and differentiated cells, with the self-renewal sample as the reference (denominator) in each case. To generate the Heat Map, raw intensity values were normalized by dividing each value identified for each ubiquitin-modified peptide, by the median of the two samples analyzed in duplicates (Supplementary Tab. S4). The data were converted to log2 for further used and they were represented using MeV v4.7.

A proprietary computer program was used to search for parent ions in the ion chromatogram files based on their chromatographic retention times and their mass-to-charge (m/z) ratios for all identified ubiquitinated peptides. A coarse retention time tolerance of ±4.0 min and an m/z tolerance of ±10 ppm were used to cluster identical peptides across all samples, requiring that the matching peptides migrate within a fine retention time tolerance of ±0.5 min throughout the LC-MS run. The computer program collected the retention time of each MS peak, observed m/z ratio, and ion intensity. Peak intensity quantification showing more than 2.5-fold increase in the treated relative to untreated samples was manually reviewed in the ion chromatogram files. This eliminated the possibility that the automated process selects the wrong chromatographic peak from which to derive the corresponding intensity measurement. When a given peptide was not observed in one of the conditions, a threshold of 30,000 counts was used for the peak detection algorithm for quantification.

Shotgun proteomic analysis of total and ubiquitin-modified protein expression by label-free quantification

Sample Preparation for Mass Spectrometry-based Analysis

Frozen cell pellets were fractionated into nuclear and cytoplasmic fractions using the ProteoExtract Subcellular Proteome Extraction Kit (Millipore, MA). 1mg of proteins were precipitated with the methanol/chloroform method and digested with trypsin (1:20 ratio) in 50mM Ammonium Bicarbonate buffer over night at 37^{o}C. The resulting peptide mixture was fractionated based on pH and analyzed by MudPIT [Ruse CI et al., J Proteome Res, 2008 May; 7(5):2140-50].

Multidimensional chromatography and tandem mass spectrometry

Peptide mixtures were pressure-loaded onto a 250 μm inner diameter (i.d.) fused-silica capillary packed first with 3 cm of 5 μm strong cation exchange material (Partisphere SCX, Whatman), followed by 3 cm of 10 μm C18 reverse phase (RP) particles (Aqua, Phenomenex, CA). Loaded and washed microcapillaries were connected *via* a 2 μm filtered union (UpChurch Scientific) to a 100 μm i.d. column, which had been pulled to a 5 μm i.d. tip using a P-2000 CO_2 laser puller (Sutter Instruments), then packed with 13 cm of 3 μm C18 reverse phase (RP) particles (Aqua, Phenomenex, CA) and equilibrated in 5% acetonitrile, 0.1 % formic acid (Buffer A). This split-column was then installed in-line with a NanoLC Eskigent HPLC pump. The flow rate of channel 2 was set at 300

nl/min for the organic gradient. The flow rate of channel 1 was set to 0.5µl/min for the salt pulse. Fully automated 11-step chromatography runs were carried out. Three different elution buffers were used: 5% acetonitrile, 0.1 % formic acid (Buffer A); 98% acetonitrile, 0.1% formic acid (Buffer B); and 0.5 M ammonium acetate, 5% acetonitrile, 0.1% formic acid (Buffer C). In such sequences of chromatographic events, peptides are sequentially eluted from the SCX resin to the RP resin by increasing salt steps (increase in Buffer C concentration), followed by organic gradients (increase in Buffer B concentration). The last chromatography step consists in a high salt wash with 100% Buffer C followed by acetonitrile gradient. The application of a 2.5 kV distal voltage electrosprayed the eluting peptides directly into a LTQ-Orbitrap XL mass spectrometer equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Full MS spectra were recorded on the peptides over a 400 to 2000 m/z range by the Orbitrap, followed by five tandem mass (MS/MS) events sequentially generated by LTQ in a data-dependent manner on the first, second, third, and fourth most intense ions selected from the full MS spectrum (at 35% collision energy). Mass spectrometer scan functions and HPLC solvent gradients were controlled by the Xcalibur data system (ThermoFinnigan, San Jose, CA).

Database search and interpretation of MS/MS datasets

Tandem mass spectra were extracted from raw files, and a binary classifier - previously trained on a manually validated data set - was used to remove the low quality MS/MS spectra. The remaining spectra were searched against a mouse IPI database (v3.75, September 19, 2010, 56,871 entries) and 124 common contaminant proteins. To calculate confidence levels and false positive rates, we used a decoy database containing the reverse sequences of the protein database appended to the target database (Elias and Gygi, 2007), and the SEQUEST algorithm to find the best matching sequences from the combined database. SEQUEST searches were done through the Integrated Proteomics Pipeline (IP2, Integrated Proteomics Inc., CA) on Intel Xeon X5450 X/3.0 PROC processor clusters running under the Linux operating system. The peptide mass search tolerance was set to 50ppm. No differential modifications were considered. A fully tryptic status was imposed on the database search.
The validity of peptide/spectrum matches was assessed in DTASelect2 (Tabb et al., 2002) using SEQUEST-defined parameters, the cross-correlation score (XCorr) and normalized difference in cross-correlation scores (DeltaCN). The search results were grouped by charge state (+1, +2, and +3) and tryptic status (fully tryptic, half-tryptic, and non-tryptic), resulting in 9 distinct sub-groups. In each one of the sub-groups, the distribution of XCorr and DeltaCN values for (a) direct and (b) decoy database hits was obtained, and the two subsets were separated by quadratic discriminant analysis. Outlier points in the two distributions (for example, matches with very low Xcorr but very high DeltaCN were discarded. Full separation of the direct and decoy subsets is not generally possible; therefore, the discriminant score was set such that a false discovery rate of 1% was determined based on the number of accepted decoy database peptides. This procedure was independently performed on each data subset, resulting in a false positive rate independent of tryptic status or charge state. In addition, a minimum sequence length of 7 amino acid residues was required, and each protein on the final list was supported by at least two independent peptide identifications unless specified. These additional requirements – especially the latter - resulted in the elimination of most decoy database and false positive hits, as these tended to be overwhelmingly present as proteins identified by single peptide matches. After this last filtering step, the false discovery rate was reduced to below 1%. Relative fold difference between samples was derived using the spectral counting method[4,5].

Database search and interpretation of MS/MS datasets (modified proteins)
MS/MS spectra were extracted from the RAW file with Readw.exe (http://sourceforge.net/projects/sashimi). The resulting mzXML file contains all the data for all MS/MS spectra and can be read by the subsequent analysis software. The MS/MS data was searched with Inspect 6* against a the same protein database used to identified the unmodified proteins with a

fixed modification of +57.0214 on Cystein, and optional modifications of +15.9994 on Methionine, +79.9663 on Threonine, Serine and Tyrosine and +114.0429 on Lysine. Only peptides with a p value less than 0.01 and less than 3ppm mass difference between the expected and experimental peptide mass were chosen for the subsequent analysis. Common contaminants (e.g. keratins) were removed from the results. (Carvalho et al., 2008; Elias and Gygi, 2007; Liu et al., 2004; Tanner et al., 2005)

Computational Analysis

The data from the mass spectrometry experiments were analyzed and the networks were generated using the Ingenuity Systems IPA software 9.0. This software uses a proprietary algorithm to generate networks of interconnected molecules. Each connection shown represents known relationships between the molecules, which can be found in the Ingenuity Knowledge Base. The most highly connected molecules in our dataset were consolidated on different networks and we focused on the network containing the pluripotency factors. Subsequently, we expanded this network using our dataset and choosing Nanog and Oct4 as nodes, since they are two key pluripotency factors ubiquitinated in our dataset.

Western blotting

Cells were collected and harvested in RIPA buffer (50mM Tris-HCl, 150mM NaCl, 0.1% SDS, 0.5 % NaDeoxycholate, 1% NP40) supplemented with proteinase inhibitors (Sigma), NaF, PMSF and Na_3VO_4. 30-50μg of each sample were analyzed by Western blot. Primary antibodies used were against Nanog (A300-397A, Bethyl), Oct-3/4 (C-10, Santa Cruz) and 2840S, Cell Signalling), Sox2 (ab97959, Abcam), p53 (Leica Microsystems), N-myc (Calbiochem), Fbxo15 (Genetex), p-c-myc (Cell Signaling), Cullin-1 (71-8700, Invitrogen), Psmd14 (2941-1, Epitomics), Psmd11 (O-23, Santa Cruz), Psmd12 (N-12, Santa Cruz), Psmb4 (β7) (H-104, Santa Cruz), K48-specific (Apu2, Millipore), K63-specific (Apu3, Millipore), Ubiquitin (P4D1, Cell Signalling), Actin (C4, Millipore), Flag (M2, Sigma) and (M2, Agilent Technologies), c-Myc (9402S, Cell Signalling), mTor (2972, Cell Signalling), Notch1 IC (Val1744, Cell Signalling), c-Jun (60A8, Cell Signalling), p100 (4882, Cell Signalling), Mcl1 (S19, Santa Cruz), Cyclin E (ab7959, Abcam), Klf5 (24331, Abcam) and SREPB-1 (K-10, Santa Cruz). Blots incubated with secondary horseradish peroxidase (HRP)-conjugated anti-mouse, anti-rabbit or anti goat IgG antibodies (Amersham, GE Healthcare) in TBS-T according the manufacturer's protocol. Bands were visualized using chemiluminescence (GE Healthcare).

Ubiquitin pull down experiments.

Poly-ubiquitin pull down was done with the ubiquitined protein enrichment kit (Calbiochem) following manufacture's instructions. Briefly, cell pellets were disrupted in lysis buffer (50 mM HEPES (pH 7.5), 5mM EDTA, 150 mM NaCl, 1% Triton X-100) supplemented with proteinase inhibitors and 10mM of fresh-prepared N-Ethylmaleimide (NEM) for 20 minutes on ice and centrifuged for 10 min at 14000 rpm. 500μg of protein was incubated with 40μl of polyubiquitin beads or control beads. Incubation was performed overnight at 4^{o}C with constant rotation and next days beads were washed 4 times for 10 minutes. Next, beads were boiled in SDS buffer and a western blot was performed as previously described.

ES cell culture and transfection

ES cell lines were maintained on gelatin coated plates with DMEM media (Cellgro) supplemented with 15% fetal bovine serum (FBS) (ThermoFisher Scientific), 1,000 U/ml LIF (Chemicon), 2-mercaptoethanol, sodium pyruvate, non-essential amino acids, and L-glutamine (Gibco) For the self-renewal screen, Nanog-GFP ES cells (kind gift from Dr. I. Lemischka, Mount Sinai School of Medicine) were reverse transfected by mixing 30nM siRNA (Dharmacon) with 0.3μl Lipofectamine

2000 (Invitrogen) in 50μl DMEM media supplemented with 2-mercaptoethanol, sodium pyruvate, non-essential amino acids, and L-glutamine. siRNA mix was added to a gelatin coated 96-well plate, and ES cells were plated at 1,500 cells/well in 50μl DMEM media (Cellgro) supplemented with 30% fetal bovine serum (FBS), 2,000 U/ml LIF (Chemicon), 2-mercaptoethanol, sodium pyruvate, non-essential amino acids, and L-glutamine. GFP fluorescence was measured using an LSRII (BD Biosciences) 48 hours post-transfection. For differentiation screen reverse transfection was set up as previously with self-renewal screen. 12 hours post-transfection media was changed to differentiation media, which contained DMEM, supplemented with 10% FBS and 5μM retinoic acid (RA). GFP fluorescence was measured using flowcytometry 48 hours after addition of differentiation media.

One-step Purification of tagged proteins and Mass Spectrometry

$5*10^7$ KH2 cells targeted with vectors expressing SF-Psmd14 or SF-eGFP were induced with 2μM doxycycline (Sigma) for 4 days and treated in culture with 10μM MG132 (Peptides International) for 3h at the day of collection. The pellets were resuspended in Lysis Buffer (100mM Tris-HCl pH7.5, 150mM NaCl, 1% Triton-X100, 1mM EDTA, 2mM $MgCl_2$, and supplemented with Complete Mini protease inhibitors (Roche), 10mM N-elthylmaleimide (Sigma), 10mM NaF (Sigma), 1mM Na_3VO_4 (Sigma) and 250 units Benzonase nuclease (Novagen), and passed 8 times through a 25 and 5/8 gauge syringe. Protein purification was performed as described before with modifications (Gloeckner et al., 2009). The supernatant was left to bind for 3hours at 4^oC with 150μL net bead volume StrepTactin macroprep beads (IBA), pre-washed three times with Lysis Buffer. The beads were washed 8 times with Washing Buffer (100mM Tris-HCl pH7.5, 150mM NaCl, 1% Triton-X100, 1mM EDTA, 2mM $MgCl_2$) with 4 column volumes each time and changing 5 tubes in total. Washed beads were transferred in a new tube with 5mL 1X Elution Buffer (Buffer E, IBA: 100mM Tris-HCl pH8.0, 150mM NaCl, 1mM EDTA and 2.5mM Desthiobiotin). After incubating for 15min, the proteins were eluted by spinning through filter columns (CytoSignal). The resulting volume was concentrated with Amicon Ultra centrifugal filter units, 10,000 MW cutoff, (Millipore) to a final volume of 200μL. After trypsin digestion peptides were analysized by LC-MS/MS on Orbitrap Velos MS. The MS/MS spectra were searched against NCBI database using a local MASCOT search engine (V.2.3). At least two peptides were identified for each protein (the false discovery rate <0.01%) with a confidence interval; no less than 95%.

Quantitative real-time PCR.

Total RNA was harvested from cells using the Qiagen RNeasy Kit (Qiagen). RNA was quantified by absorbance at 260 nm and 2μg of total RNA used for cDNA synthesis using Superscript III first strand synthesis kit (Invitrogen) Q-RT-PCR was carried out using SYBR green universal mix PCR reaction buffer (Roche) using an Roche lightcycler 480 II (Roche). Bar graphs represent average of duplicate experiments $\pm$ standard derivation.

Flow cytometry analysis and antibodies

For SSEA1 staining, cells were were collected by trypsinization, washed twice with PBS and were incubated for 30 min with Alexa Fluor 647 anti mouse/human SSEA1 antibody clone MC-480 (BioLegend). For cell cycle profiling, as previously cells were fixed and permeabilzed. Following permeabilzation cells were washed at stained with 2μg/ml DAPI, and 5μg/ml RNaseA for 15min at RT then analyzed by flowcytometry. Apoptosis was detected using Annexiv-V PE-conjugated detection kit (BD Pharmingen) along with 7-AAD following manufacturers protocol.

Supplemental References

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev *23*, 2484-2489.

Carvalho, P.C., Hewel, J., Barbosa, V.C., and Yates, J.R., 3rd (2008). Identifying differences in protein expression levels by spectral counting and feature selection. Genet Mol Res *7*, 342-356.

Ehrlund, A., Anthonisen, E.H., Gustafsson, N., Venteclef, N., Robertson Remen, K., Damdimopoulos, A.E., Galeeva, A., Pelto-Huikko, M., Lalli, E., Steffensen, K.R., *et al.* (2009). E3 ubiquitin ligase RNF31 cooperates with DAX-1 in transcriptional repression of steroidogenesis. Mol Cell Biol *29*, 2230-2242.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods *4*, 207-214.

Gloeckner, C.J., Boldt, K., Schumacher, A., and Ueffing, M. (2009). Tandem affinity purification of protein complexes from mammalian cells by the Strep/FLAG (SF)-TAP tag. Methods Mol Biol *564*, 359-372.

Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van, I.W., Grootegoed, J.A., and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature *485*, 386-390.

Guryanova, O., and Levine, R. (2012). DNMT3A and stem cell function: new insights into old pathways. Haematologica *97*, 324.

Jang, H., Kim, T.W., Yoon, S., Choi, S.Y., Kang, T.W., Kim, S.Y., Kwon, Y.W., Cho, E.J., and Youn, H.D. (2012). O-GlcNAc Regulates Pluripotency and Reprogramming by Directly Acting on Core Components of the Pluripotency Network. Cell Stem Cell.

Jung, M., Peterson, H., Chavez, L., Kahlem, P., Lehrach, H., Vilo, J., and Adjaye, J. (2010). A data integration approach to mapping OCT4 gene regulatory networks operative in embryonic stem cells and embryonal carcinoma cells. PLoS One *5*, e10709.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Kim, S.K., Suh, M.R., Yoon, H.S., Lee, J.B., Oh, S.K., Moon, S.Y., Moon, S.H., Lee, J.Y., Hwang, J.H., Cho, W.J., *et al.* (2005). Identification of developmental pluripotency associated 5 expression in human pluripotent stem cells. Stem Cells *23*, 458-462.

Landeira, D., Sauer, S., Poot, R., Dvorkina, M., Mazzarella, L., Jorgensen, H.F., Pereira, C.F., Leleu, M., Piccolo, F.M., Spivakov, M., *et al.* (2010). Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. Nat Cell Biol *12*, 618-624.

Liao, B., and Jin, Y. (2010). Wwp2 mediates Oct4 ubiquitination and its own auto-ubiquitination in a dosage-dependent manner. Cell research *20*, 332-344.

Liu, H., Sadygov, R.G., and Yates, J.R., 3rd (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem *76*, 4193-4201.

Loh, Y.H., Zhang, W., Chen, X., George, J., and Ng, H.H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev *21*, 2545-2557.

Ramakrishna, S., Suresh, B., Lim, K.H., Cha, B.H., Lee, S.H., Kim, K.S., and Baek, K.H. (2011). PEST Motif Sequence Regulating Human NANOG for Proteasomal Degradation. Stem Cells Dev.

Scotland, K.B., Chen, S., Sylvester, R., and Gudas, L.J. (2009). Analysis of Rex1 (zfp42) function in embryonic stem cell differentiation. Dev Dyn *238*, 1863-1877.

Sun, C., Nakatake, Y., Akagi, T., Ura, H., Matsuda, T., Nishiyama, A., Koide, H., Ko, M.S., Niwa, H., and Yokota, T. (2009). Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. Mol Cell Biol *29*, 4574-4583.

Tabb, D.L., W.H. McDonald, and J.R. Yates, r. (2002). DTASelect and Contrast: Tools for Assembling and Comparing Protein Identification from Shotgun Proteomics. J Proteome Res *1*, 21-26.

Tanner, S., Shu, H., Frank, A., Wang, L.C., Zandi, E., Mumby, M., Pevzner, P.A., and Bafna, V. (2005). InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. Anal Chem *77*, 4626-4639.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. Mol Cell Biol *23*, 2699-2708.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.