Stem Cell Reports

Supplemental Information

The L1TD1 Protein Interactome Reveals the Importance of Post-transcriptional Regulation in Human Pluripotency

Maheswara Reddy Emani, Elisa Närvä, Aki Stubb, Deepankar Chakroborty, Miro Viitala, Anne Rokka, Nelly Rahkonen, Robert Moulder, Konstantin Denessiouk, Ras Trokovic, Riikka Lund, Laura L. Elo, and Riitta Lahesmaa

Supplemental Information

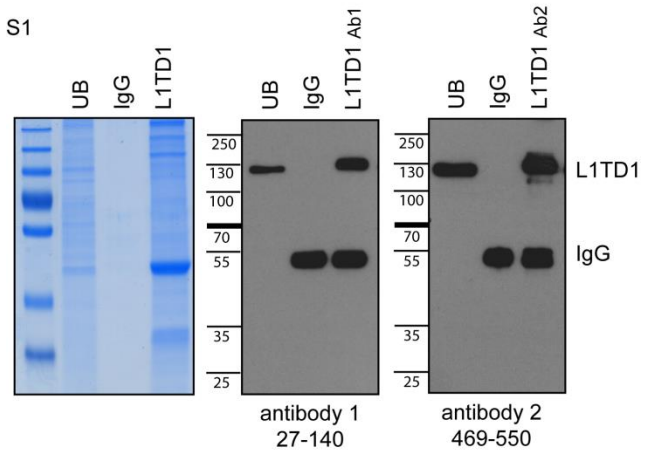


Figure S1, Related to Figure 1. IP with L1TD1 antibodies. Coomassie stained SDS-gel and WB detection of L1TD1 in IP reactions with IgG, L1TD1 HPA030064 (Ab1) and HPA028501 (Ab2), hESC line HS360, (UB) unbound fraction.

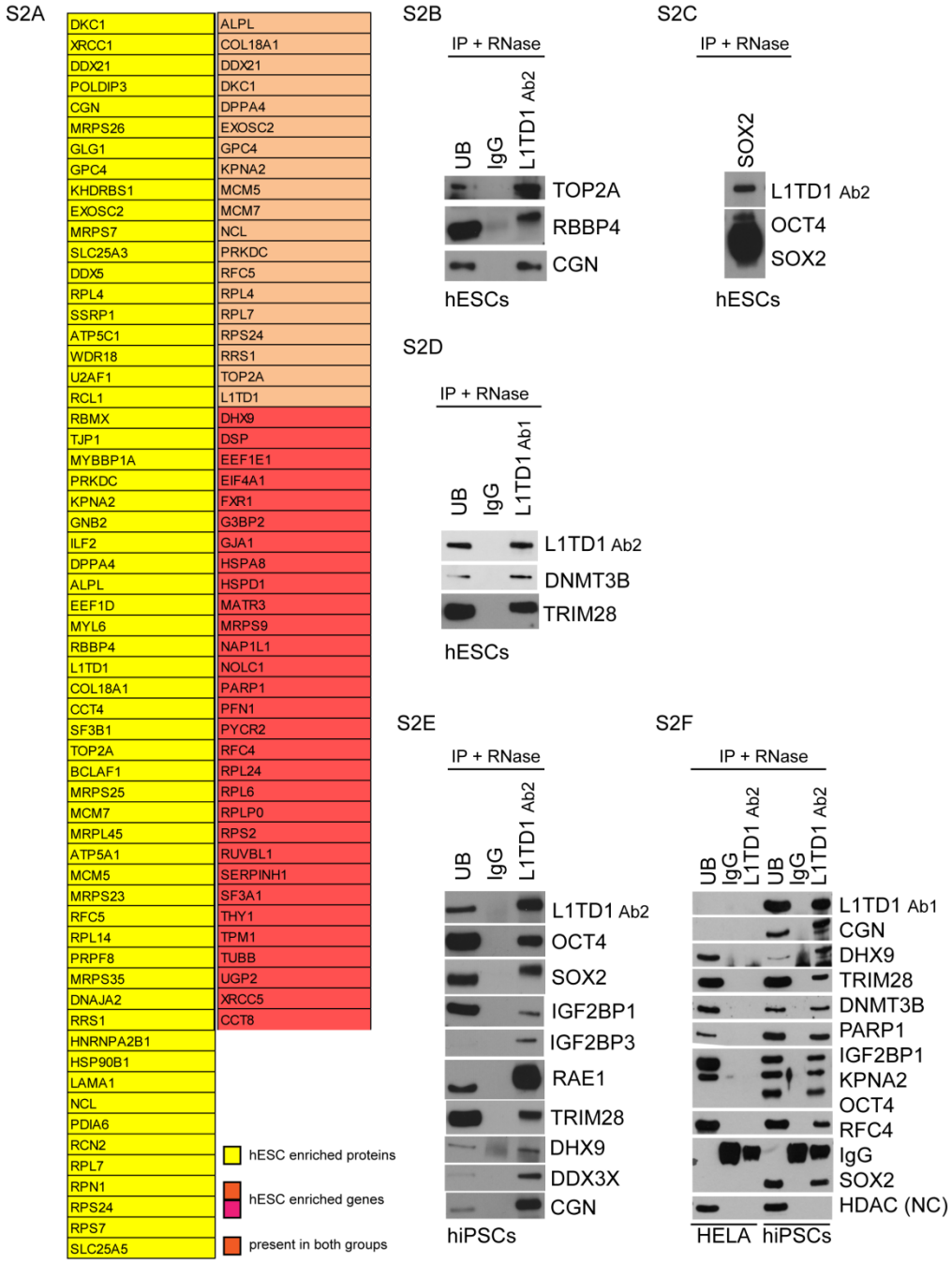


Figure S2, Related to Figure 4. Additional validations. A) hESC enriched genes (Assou et al., 2007) (red) and proteins (yellow) (Jadaliha et al., 2012; Van Hoof et al., 2006) in L1TD1 interactome. Present in both (orange). B) WB detection of TOP2A, RBBP4 and CGN in IP reactions with IgG and L1TD1. Additional detections from the experiment shown in Figure 3A. hESC line HS360. C) WB detection of L1TD1, OCT4 and SOX2

immunoprecipitated with SOX2 antibody. This data is an additional line of experiment shown in main Figure 4B. hESC line HS360. D) WB detection of L1TD1, DNMT3B (hESC line H9) and TRIM28 (hESC line HS360, additional data from the experiment shown in main Figure 1E) in IP reactions with IgG, and L1TD1. E) WB validation of interactions in human induced pluripotent stem cells (hiPSCs), hiPSC line HEL24.3, IP reactions with IgG and L1TD1. F) Validation specificity of key interactions in Hela cells grown on matrigel compared to hIPSCs. Immunoprecipitation reactions with IgG and L1TD1. UB (unbound fraction), NC (negative control).

Table S1, Related to Figure 1, L1TD1 interactome. Final L1TD1 interactome and individual three replicates presented in separate excel sheets. (Excel File)

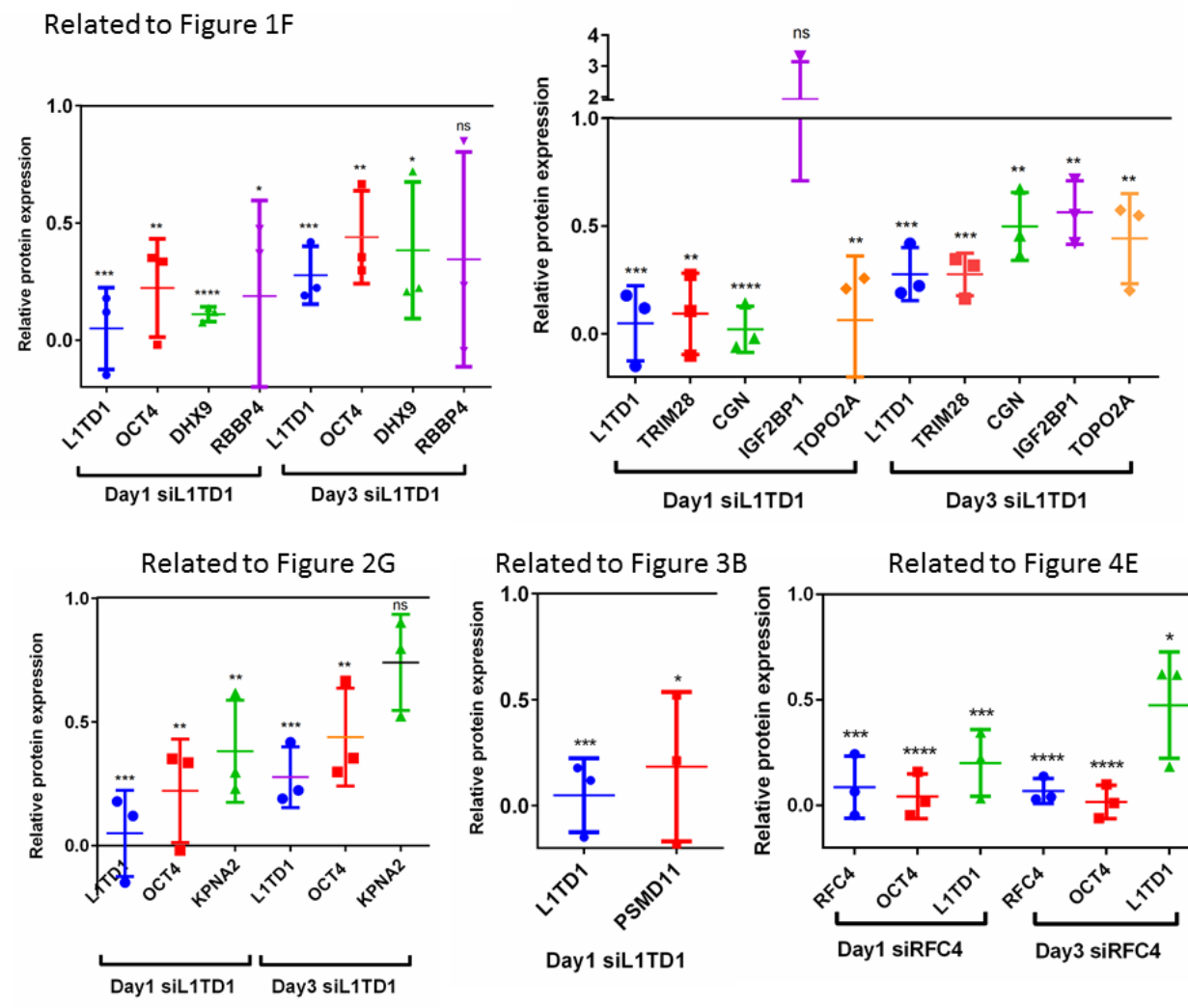
Table S2, Related to Figure 1, CRAPome analysis. Workflow 1 (Percentage found in other studies) and Workflow 3 (Saint Probability) for L1TD1 interactome proteins. (Excel File)

Table S3, Related to Figure 1, STRING interactions. STRING9.0 and DAVID pathway analysis. (Excel File)

Table S4, Related to Figure 2. Protein classifications based on protein domains. (Excel File)

Supplemental Results

Western Blot quantification for siRNA experiments, Related to Figures 1F, 2G, 3B and 4E. Data is expressed as fold changes compared with non-targeting siRNA, normalized to GAPDH and are representative of three biological replicates. (Statistical analyses: The statistical analysis was performed using Graph Pad Prism 6, results were evaluated using unpaired Student's t test. (Statistical significance was accepted at *p<0.05, **p<0.01 and ***p<0.001, **** p<0.0001)



Supplemental Experimental Procedures

Cell Culture

Human ESC line HS360 was obtained from Outi Hovatta (Karolinska Institutet, Sweden) and H9 from WiCell Research Institute (Madison, WI). All experiments were performed in feeder-free culture conditions on Matrigel (BD Biosciences) in mTeSR1 media (Stem Cell Technologies) passaged with dispase (Stem cell technologies). Human iPSC lines used in this study HEL11.4 and HEL24.3 were generated using Sendai viruses (SeV, Cytotune, Lifetechnologies) and were thoroughly characterized as previously described (Mikkola et al., 2013; Toivonen et al., 2013). Hela control cells were grown on matrigel in mTeSR1 media for IP reactions.

Immunoprecipitation

For one IP reaction 3-4 70 % confluent 10 cm plates were washed twice with cold PBS and lysed into NP-40 -buffer (20–50 mM Tris, 150 mM NaCl, 0.5% sodium deoxylate, 0.5% NP-40) containing PhosSTOP and complete EDTA-free inhibitors (Roche). Lysates were preincubated with 10 µg/ml RNaseA (Qiagen) and pre-washed uncoupled 40 µL Dynabeads® Sheep anti-Rabbit Dynabeads® at +4°C for 1 hour to eliminate proteins that bind nonspecifically to the beads. IP was carried out using M-280 Sheep Anti-Rabbit IgG Dynabeads® (Invitrogen – 11203D) according to the manufacturer's protocol. Briefly, 80 µl of µL Dynabeads® Sheep anti-Rabbit Dynabeads® were pre-washed per 2 ug of antibody. IPs were performed with L1TD1 HPA030064 (Ab1), HPA028501 (Ab2) (Sigma), SOX-2 #5024 (Cell signaling) and Normal Rabbit IgG 12-370 (Millipore). 4 ug of antibody was Incubated with gentle tilting and rotation for 2 hours at 4°C with beads. Unbound antibody was washed 8 times with NP-40 buffer. 2 mg of RNAse treated protein lysate was added

to the IgG and L1TD1 coupled beads and incubated o/n with gentle tilting and rotation at 4°C. Unbound proteins were washed off 8 times with NP-40 buffer. Elution was done with lower pH PAG Elution buffer 10701 (Admetech).

In case of KPNA2, IP was carried out using Dynabeads® protein G (Invitrogen – 10004D) according to the manufacturer's protocol. 80 µl of Dynabeads® protein G per 4 ug of antibody were pre-washed with NP-40 buffer. IPs were performed with KPNA2 sc-55538 (Sigma) and Normal Mouse IgG 12-371 (Millipore). 4 ug of antibody was Incubated with gentle tilting and rotation for 2hours at 4°C with beads. Unbound antibody was washed 8 times with NP-40 buffer. Beads were then crosslinked to antibodies using DMP (dimethyl pimelimidate dihydrochloride). First bead - antibody complexes were washed with 0.2 M triethanolamine, pH 8.2 (Sigma) and then resuspended to 20 mM DMP (Sigma) in 0.2 M triethanolamine, pH 8.2. Reaction was stopped by resuspending beads in to 50 mM Tris, pH 7.7. After reaction, complexes were washed using PBST (0,01% tween-20). RNAse treated lysate was added on the IgG and KPNA2 coupled beads and incubated o/n with gentle tilting and rotation at +4°C. Unbound proteins were washed off 8 times with NP-40 buffer. Elution was done with lower pH PAG Elution buffer 10701 (Admetech).

Mass Spectrometry

IP proteins were separated on Criterion XT 12% Bis-Tris gel (BioRad) and stained with PageBlue (Thermo Scientific). Lanes were cut to 11-12 pieces and destained with 50% methanol. Proteins were in-gel digested as described earlier (Shevchenko et al., 1996). Tryptic peptides were dissolved in formic acid (0.1%) and submitted for LC-MS/MS analysis.

The LC-MS/MS system consisted of a nanoflow HPLC system (Ultimate 3000, Dionex, Sunnyvale, CA) coupled to a QSTAR Elite mass spectrometer (Applied Biosystems/MDS Sciex, Canada) equipped with a nano-electrospray ionization source (Proxeon, Odense, Denmark). Each sample was injected as one technical replicate. Peptides were first loaded on a trapping column (0.3 x 5 mm PepMap C18, LC Packings) and subsequently separated on a 15 cm C18 column (75 μm x 15 cm, Magic 5 μm 200 Å C_{18}, Michrom BioResources Inc., Sacramento, CA, USA). Peptide separation was achieved using a gradient from 2 to 35% B at a mobile phase in flow rate of 200 nl/min. The mobile phase consisted of water/acetonitrile (98:2 (v/v)) with 0.2% formic acid (solvent A) or acetonitrile/water (95:5 (v/v)) with 0.2% formic acid (solvent B). Two different methods, including either 20 or 45 min gradient, were used depending on the sample complexity.

LC-MS/MS data acquisition and analysis

Data dependent acquisition was performed using Analyst QS 2.0 software (Applied Biosystems/MDS Sciex, Canada). The two or three of the most intensive precursor ions from the survey scan (350 - 1500 *m/z)* were selected for fragmentation with a dynamic exclusion for 60 seconds (for the 20 and 45 minute methods, respectively). MS/MS peak lists were created with Analyst QS 2.0. and database searches were performed using the Mascot search engine (version 2.2.06, Matrix Sciences, London, UK) against the SwissProt database (release 2011_08 containing 20245 sequence entries). The search results from samples cut from the same gel lane were merged. The Mascot search settings included a taxonomy filter 'human', trypsin as an enzyme, one missed trypsin cleavage, precursor-ion mass tolerance of 0.2 Da, fragment-ion mass tolerance of 0.3 Da,

variable modifications of carbamidomethylation of cysteine and methionine oxidation. A significance threshold of $p<0.05$ was used.

Scaffold 3.00.03 (Proteome Software Inc., Portland, OR) was used to combine and control the false discovery rates form MS/MS based peptide and protein identifications. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Peptide identifications were accepted if they could be established at greater than 95.0% probability.

The criteria used for inclusion as a L1TD1-interacting protein were presence in two out of three replicates, more than one unique peptide identified in the L1TD1 IP-analysis and on/off or ≥ 3 fold enrichment of identified peptides compared to control IgG IP-reaction.
To investigate which proteins in the MS data were potential contaminants we used the CRAPome analysis tool (Mellacheruvu et al., 2013). The L1TD1 interactome was uploaded to CRAPome and the analysis was run following the CRAPome Workflow 1. To perform the significance analysis of the interactome SAINT (Choi et al., 2011; Mellacheruvu et al., 2013), the spectral count data for three L1TD1 and three IgG control replicates were uploaded to the CRAPome tool. The analysis was run following the CRAPome Workflow 3 using the default parameters for the scoring functions. As the CRAPome data differed from the IP conditions used in the L1TD1 interactome, the obtained results were not used as an exclusion criterion, but as an estimate of probability and significance for each interacting protein.

Previously reported protein interactions between the L1TD1 interacting proteins were downloaded from the STRING interaction database (Szklarczyk et al., 2011). Only high

confidence (>0.7) interactions derived from experimental data or curated databases were considered. The resulting network was visualized using Cytoscape 2.8.1 (Shannon et al., 2003). The force-directed layout was applied and clusters were identified using the Markov Clustering algorithm (MCL) implemented in the Cytoscape plug-in clusterMaker v. 1.9 with default settings (Morris et al., 2011).

Enriched biological pathways (KEGG), Gene Ontology (GO) terms, and disease associations (OMIM and Genetic Association Database) were identified using the functional annotation tool DAVID 6.7 (Database for Annotation, Visualization and Integrated Discovery)(Huang da et al., 2009a; Huang da et al., 2009b). For the GO analysis, the GO FAT terms were considered.

Western Blotting

In silencing, cell fractioning and proteosomal activity assays cells were lysed in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% TX-100, 5% glycerol, 1% SDS, 1 mM Na_3VO_4, 10 mM NaF, and 1 mM phenylmethanesulfonyl fluoride). Protein concentrations were determined with DC Protein Assay (Bio-Rad) and 6xSDS sample buffer (0.5 M Tris-HCl pH 6.8, 28% glycerol, 9% SDS, 5% 2-mercaptoethanol, 0.01% bromphenol blue) was added. Lysates were electrophoresed on a 10% SDSPAGE gel and transferred to a nitrocellulose membrane. Membranes were incubated overnight at +4 ºC with primary antibodies: L1TD1 HPA028501 (Sigma), L1TD1 HPA030064 (Sigma), OCT-4 (sc-9081) or (sc-5279) (Santa Cruz), NANOG AF1997 (R&D Systems), SOX-2 MAB2018 (R&D Systems), SOX-2 #5024 (Cell signaling), TIF1 beta MA1-2023 (Thermo Scientific), KPNA2 Sc-55538 (Santa Cruz), LIN28 ab46020 (Abcam), RbAp48 sc-8270 (Santa Cruz), IGF2BP-

3 sc-376067 (Santa Cruz), IGF2BP-1 sc-166344 (Santa Cruz), DDX3 sc-81247 (Santa Cruz), KHDRBS1 sc-136062 (Santa Cruz), ELAVL1 sc-56709 (Santa Cruz), Cingulin (CGN) ab117796 (Abcam), DNMT3B ab13604 (Abcam), PSMD11 FZ10R-1150 (EUROPA bioproducts), CTNNB1 sc-59737 (Santa Cruz), TOPO1 sc-10783 (Santa Cruz), TOPO2A sc-365918 (Santa Cruz), PARP1 #9532 (Cell signaling), Mrnp41 (RAE1) sc-374261 (Santa Cruz), PABP sc-32318 (Santa Cruz), DHX9 ab26271 (Abcam), DHX9 ab54593 (Abcam), Smad4 sc-7966 (Santa Cruz), HDAC1 sc-8410 (Santa Cruz), Nodal (H-110) SC-28913 (Santa Cruz), GAPDH 5G4 (HyTest Ltd), β-actin A5441 (Sigma), U2AF1 SAB1402953 (Sigma), SRSF3 (SRp20) sc-73059 (Santa Cruz), anti-rabbit-HRP 554021 (BD Pharmingen), Histone H2B sc-10808 (Santa Cruz), RFC4 (C-9) sc-28301 (Santa Cruz). Secondary antibodies were anti-mouse-HRP sc-2005 (Santa Cruz), anti-goat-HRP sc-2020 (Santa Cruz Biotechnology). Signal was detected with enhanced chemiluminescence reagent (Amersham Biosciences) or Pierce developing solution (Pierce [Thermo Fisher Scientific]).

Cell lines used to validate individual protein interactions:

	HS360	H9	HEL	24.3	EP2102	HEL 11.4
L1TD1	X	X	X		X	X
DHX9	X	X	X		X	
PABPC1	X	X			X	
TRIM28	X	X	X			
OCT4	X	X	X			
SOX2	X		X			
DDX3X	X	X	X			
CGN	X	X	X			
KHDRBS1	X	X				
RBBP4	X	X				
ELAVL1	X					
IGF2BP1	X		X			
IGF2BP3	X		X			
PSMD11	X					
KPNA2	X		X			
RAE1	X		X			
PARP1	X		X			
NANOG	X					
DNMT3B		X	X			
SRSF3						X
U2AF1						X
SMAD4	X					
HDAC1	X					
CTNNB1	X					
TOP2A	X					

RFC4	X

Immunofluorescence

Cells were fixed with 4 % paraformaldehyde for 15 min and permeabilized for 15 min using 0.5% Triton X-100. Cells were stained with primary antibodies for 2 h to o/n: L1TD1 antibody (Ab2) 1:100, Atto 647N-Phalloidin 65906 (Sigma) 1:200 and NANOG AF1997 1:100 (R&D Systems) in 40 % horse serum. Secondary antibody staining was performed with Alexa-488 and 555 (Invitrogen).

Cell fractioning

Cell fractioning was done with commercial kit (Pierce kit 78833). After collecting cytoplasmic fraction excessive washings with 1x PBS was included in the protocol.

RNA Interference and hESC Transfection

RNA interference was performed with siRNA oligonucleotides (Sigma). The sequences of siRNAs are as follows siL1TD1: GAGATGAGTCATGATGAGCATA; NTsiRNA control scr3: CCUACAUCCCGAUCGAUGAUG (Berra et al., 2003); siRFC4 (Arai et al., 2009): RFC4#1 GACGUACCAUGGAGAAGGAGUCGAA, RFC4#2 CAAGGAUCGAGGAGUAGCUGCCAGT. Transfections were performed as previously described (Narva et al., 2012).

Ethical Consideration

Ethics Committee of South-West Finland Hospital District provided the permission to culture the human ESC lines used in this study. Research was carried out following the good scientific practice and guidelines of the National Advisory Board on Research Ethics.

Supplemental References

Arai, M., Kondoh, N., Imazeki, N., Hada, A., Hatsuse, K., Matsubara, O., and Yamamoto, M. (2009). The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells. Liver Int. *29,* 55-62.

Berra, E., Benizri, E., Ginouves, A., Volmat, V., Roux, D., and Pouyssegur, J. (2003). HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1alpha in normoxia. EMBO J. *22,* 4082-4090.

Choi, H., Larsen, B., Lin, Z.Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods *8,* 70-73.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37,* 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4,* 44-57.

Mikkola, M., Toivonen, S., Tamminen, K., Alfthan, K., Tuuri, T., Satomaa, T., Natunen, J., Saarinen, J., Tiittanen, M., Lampinen, M.*, et al.* (2013). Lectin from Erythrina cristagalli supports undifferentiated growth and differentiation of human pluripotent stem cells. Stem Cells Dev. *22,* 707-716.

Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., and Ferrin, T.E. (2011). clusterMaker: a multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics *12,* 436-2105-12-436.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. *68,* 850-858.

Toivonen, S., Ojala, M., Hyysalo, A., Ilmarinen, T., Rajala, K., Pekkanen-Mattila, M., Aanismaa, R., Lundin, K., Palgi, J., Weltner, J.*, et al.* (2013). Comparative analysis of targeted differentiation of human induced pluripotent stem cells (hiPSCs) and human embryonic stem cells reveals variability associated with incomplete transgene silencing in retrovirally derived hiPSC lines. Stem Cells Transl. Med. *2,* 83-93.