Stem Cell Reports
Resource
OPEN ACCESS

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

Maheswara Reddy Emani,[1,4] Elisa Närvä,[1,4] Aki Stubb,[1] Deepankar Chakroborty,[1] Miro Viitala,[1] Anne Rokka,[1] Nelly Rahkonen,[1] Robert Moulder,[1] Konstantin Denessiouk,[1] Ras Trokovic,[2] Riikka Lund,[1] Laura L. Elo,[1,3] and Riitta Lahesmaa[1,*]

[1]Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, 20521 Turku, Finland
[2]Research Programs Unit, Molecular Neurology and Biomedicum Stem Cell Center, University of Helsinki, 00290 Helsinki, Finland
[3]Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland
[4]Co-first author
*Correspondence: riitta.lahesmaa@btk.fi
http://dx.doi.org/10.1016/j.stemcr.2015.01.014

SUMMARY

The RNA-binding protein L1TD1 is one of the most specific and abundant proteins in pluripotent stem cells and is essential for the maintenance of pluripotency in human cells. Here, we identify the protein interaction network of L1TD1 in human embryonic stem cells (hESCs) and provide insights into the interactome network constructed in human pluripotent cells. Our data reveal that L1TD1 has an important role in RNA splicing, translation, protein traffic, and degradation. L1TD1 interacts with multiple stem-cell-specific proteins, many of which are still uncharacterized in the context of development. Further, we show that L1TD1 is a part of the pluripotency interactome network of OCT4, SOX2, and NANOG, bridging nuclear and cytoplasmic regulation and highlighting the importance of RNA biology in pluripotency.

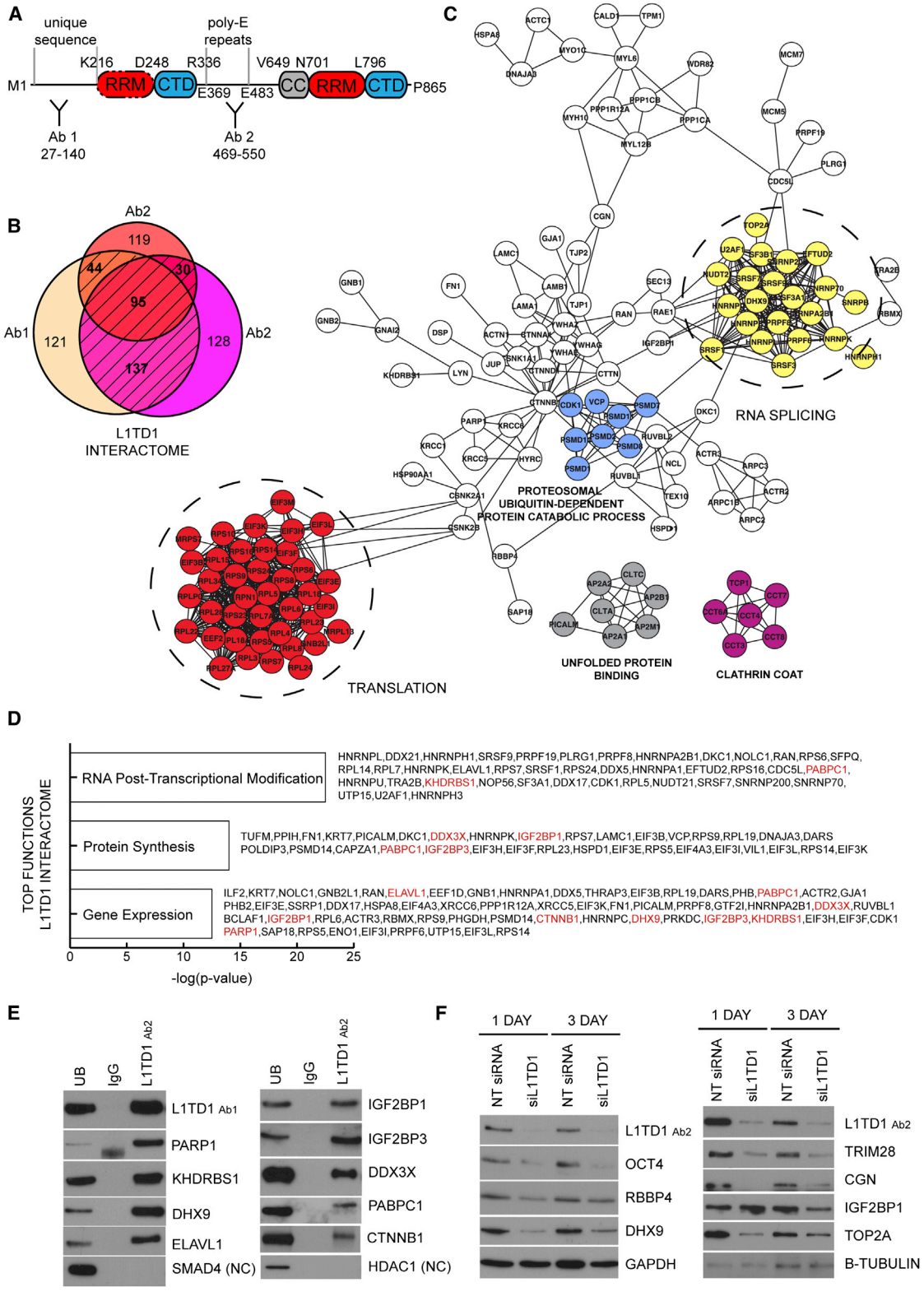
INTRODUCTION

Pluripotent human embryonic stem cells (hESCs) can differentiate into all somatic cells of the human body and hold tremendous potential for developmental biology, drug screening, and regenerative medicine (Thomson et al., 1998). The importance of the core pluripotency transcription factors OCT4, NANOG, and SOX2 in the maintenance and induction of pluripotency has been well documented (Boyer et al., 2005; Takahashi and Yamanaka, 2006; Yu et al., 2007). However, a variety of post-transcriptional processes control and alter the original message after transcription. In addition, protein degradation by proteasomes has been suggested to have a vital role maintaining pluripotency (Buckley et al., 2012; Vilchez et al., 2012). The role of post-transcriptional regulators and protein networks in the maintenance of pluripotency is still largely unknown, especially in human cells.

L1TD1 is highly and specifically expressed in pluripotent cells (Iwabuchi et al., 2011; Mitsui et al., 2003; Närvä et al., 2012; Wong et al., 2011) under the control of OCT4, NANOG, and SOX2 (Boyer et al., 2005; Närvä et al., 2012; Wong et al., 2011). We previously reported that L1TD1 is essential for maintaining the pluripotent state in hESCs (Närvä et al., 2012). Like LIN28, which is one of the well-characterized post-translational regulators of pluripotency (Huang, 2012), L1TD1 is an RNA-binding protein (RBP). RBPs have a fundamental role in a wide variety of cellular processes, including RNA transcription, splicing, processing, localization, stability, and translation. Moreover, each RBP has unique and specific roles (Glisovic et al., 2008). Given its highly specific expression and vitality in pluripotent cells, we expect that L1TD1 regulates the maintenance of the pluripotent state.

Transcriptomic studies have provided a large amount of data concerning genes that are specifically expressed in the pluripotent state. However, this information can only be used to predict the behavior of proteins that are important for the maintenance of pluripotency. Combinations of mass spectrometry (MS) and bioinformatics studies have increased our understanding of the proteins and pathways that are active in ESCs (Jadaliha et al., 2012; Van Hoof et al., 2006, 2009). Only recently, affinity purifications combined with MS have shed light on the protein-protein interaction networks of individual pluripotency factors and the molecular networks that regulate pluripotency. The interactome for Oct4, Sox2, and Nanog, along with other individual transcription factors, has been reported in mouse ESCs (mESCs) (Ding et al., 2012; Gao et al., 2012; Liang et al., 2008; Pardo et al., 2010; van den Berg et al., 2010; Wang et al., 2006). However, the functional protein-protein interactions in hESCs have remained unexplored. Notably, the protein networks that are distinct from transcriptional-control networks in pluripotent cells remain unknown.

To elucidate interacting and functional components involved in the regulation of human pluripotency, we characterized proteins associated with L1TD1. Furthermore, we validated selected interactions in hESCs and human induced pluripotent stem cells (hiPSCs). We show that the pluripotency network of L1TD1 is shared with the



(legend on next page)

OCT4, NANOG, and SOX2 interactomes, and contains uncharacterized protein players that are important for maintenance of pluripotency. Our data significantly expand the current knowledge about the biology of RBPs and provide an important resource of protein interactions in human pluripotent cells.

RESULTS

L1TD1 Interactome

The biological function of a protein is determined by its structure. L1TD1 consists of a C-terminal coiled coil (CC) domain, an RNA recognition motif (RRM) domain, a C-terminal carboxyl tail domain (CTD) domain, and an N-terminal CTD domain (Figure 1A). The N-terminal RRM domain is exceptional in that only 30 of its C-terminal amino acids (K216–D248) are typical of RRM domains. This "unique sequence" possibly contains the CC domain and the rest of the RRM domain, but it clearly differs from the sequences found in typical CC and RRM domains. Based on these structural domains, we conclude that L1TD1 is an RBP. However, the middle region of L1TD1 contains an exceptional poly-E repeat region. Interestingly, nucleolin, whose structure has an unusual combination of Poly-E and RRM domains, has been shown to specifically recruit proteins that are important for its unique function, with the help of the acidic poly-E region (Edwards et al., 2000; Erard et al., 1988; Lapeyre et al., 1987). On the basis of these similarities, we speculated that by determining the proteins that interact with L1TD1, we could identify protein complexes that are important for the maintenance of pluripotency.

To identify the L1TD1 protein network and to exclude secondary partners interacting through RNA, we carried out immunoprecipitation (IP) in the presence of RNase with two specific anti-L1TD1 antibodies (Figure S1) that recognize distinct regions of the L1TD1 protein (Figure 1A). The availability of highly specific antibodies against L1TD1 enabled IP of the endogenous protein, excluding possible overexpression effects. The strategy used to overexpress tagged fusion protein was unsuccessful, as the overexpressed fusion protein of L1TD1 was rapidly cleaved in living cells (data not shown). Liquid chromatography-tandem MS (LC-MS/MS) analysis of three biological replicates resulted in the detection of 306 proteins putatively belonging to the L1TD1 interactome (Figure 1B; Table S1). Immunoglobulin G (IgG) control IPs were used as a negative control. In addition, we carried out a CRAPome analysis (Mellacheruvu et al., 2013) to identify the specificity of the interactions (Table S2).

In order to identify biological processes through which L1TD1 mediates its function, we analyzed the L1TD1 interactome with several of the currently available Gene Ontology and interaction analysis tools. First, we investigated known interactions within the L1TD1 interactome based on the STRING interaction database (Szklarczyk et al., 2011). This revealed highly connected components of proteins involved in RNA-processing biology. In particular, RNA translation (21.6%) and splicing (10%) were identified as key functions (Figure 1C; Table S3). Among the other notable associated functions were intracellular transport (9.6%), protein folding (5.6%), and the proteosomal ubiquitin-dependent protein catabolic process (3%). Ingenuity Pathway Analysis (IPA) further confirmed that the top three functions of L1TD1-interacting proteins were gene expression, post-transcriptional modification, and protein synthesis (Figure 1D). Collectively, these results suggest that L1TD1 is involved in multiple steps ranging from RNA production to the production of a functional protein.

To establish the validity of the detected interaction network, we confirmed randomly selected individual interactions with specific antibodies (Figure 1E). In addition, silencing of L1TD1 resulted in reduced expression of most of the interacting partners (Figure 1F). Taken together

Figure 1. L1TD1-Interacting Proteins Are Involved in RNA Processing

(A) Schematic diagram of the human L1TD1 protein structure: coiled coil (CC) domain, RNA recognition motif (RRM) domain, and carboxyl tail domain (CTD). Regions targeted by antibodies 1 and 2 used in the IPs are shown. Protein domain borders are based on Khazina et al. (2011).

(B) Venn diagram of the proteins identified in the L1TD1 interactome (shaded area) based on MS analysis of three biological replicates.

(C) Clusters of the most enriched biological processes in the L1TD1 interactome. Data from the STRING interaction database (Szklarczyk et al., 2011) were visualized using Cytoscape 2.8.1 (Shannon et al., 2003).

(D) The three most significant functions of L1TD1 interactome proteins based on Ingenuity Pathway Analysis (IPA). Proteins validated experimentally are shown in red.

(E) Western blot (WB) validations of individual protein-protein interactions in L1TD1, IgG control, and unbound (UB) IP reactions. NC, negative control. hESC line HS360. CTNNB1 separate IP reaction, DDX3X and PABPC1 separate IP reaction, hESC line H9.

(F) Protein expression of the interacting proteins after L1TD1 silencing. NT, non-targeting. hESC line HS360. Quantifications can be found in the Supplemental Results.

See also Figure S1 and Tables S1, S2, and S3.

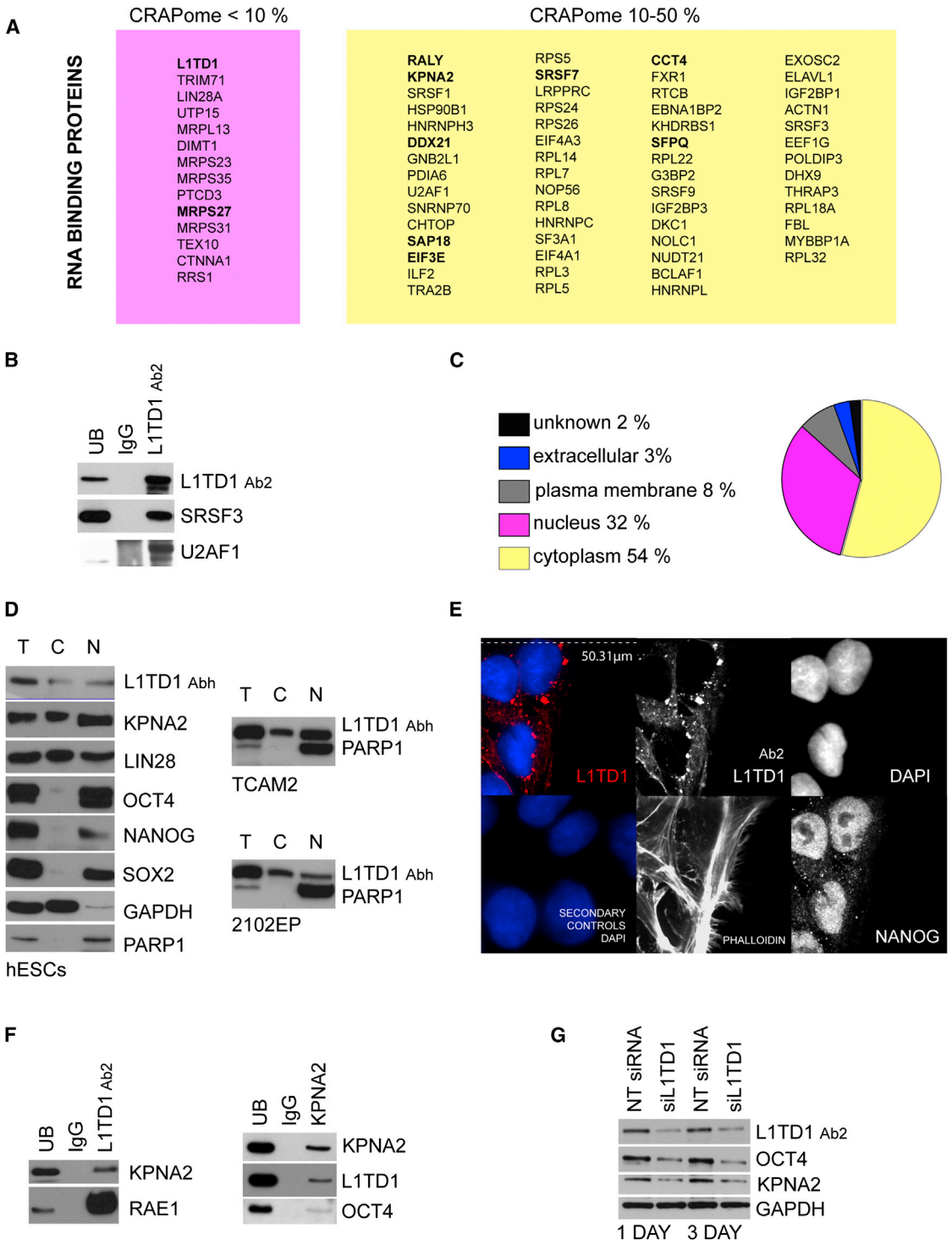


Figure 2. The L1TD1 Interactome Uncovers the Importance of RBPs and the Potential Role of L1TD1 in Nuclear Pore Traffic

(A) RBPs that were both identified here in the L1TD1 interactome in hESCs and previously reported in mESCs (Kwon et al., 2013). Proteins were grouped according to the CRAPome percentages (Workflow 1). Proteins with 0.8-1 SAINT probability are shown in bold (Workflow 3).

(B) WB validation of SRSF3 and U2AF1 interactions in L1TD1, IgG control, and UB IP reactions. hIPSC line HEL 11.4.

(C) Subcellular location of L1TD1-interacting proteins based on IPA software.

(legend continued on next page)

with the previous observation that the silencing of L1TD1 has a rapid effect on pluripotency (Närvä et al., 2012), our results suggest that interacting proteins that respond to L1TD1 silencing participate in the maintenance of pluripotency along with L1TD1.

L1TD1 Interacts with Multiple RBPs

To gain further insight into the network revealed by these measurements, we classified the L1TD1 interactome on the basis of the structural domains of the interaction partners (Table S4). Based on this, 58 proteins (19%) were classified as ribosomal proteins, providing further evidence that L1TD1 is involved in translation. Interestingly, another major class (10%, 31 proteins) identified was proteins with the RNA-binding domain RRM or KH (K homology), suggesting that L1TD1 performs its function in a complex with other RBPs.

Recently, Kwon et al. (2013) reported an RNA-binding repertoire of mESCs including 283 RBPs. A comparison of these RBPs with the L1TD1 interactome resulted in the identification of 89 shared RBPs, 72 of which had high significance based on the CRAPome (Figure 2A). Based on this comparison, 29% of the L1TD1-interacting proteins bind RNA. Moreover, two RBPs that interact with L1TD1, U2af1, and Srsf3 (Figure 2B) were recently shown to be vital for somatic cell reprogramming in mESCs (Ohta et al., 2013).

L1TD1 Traffics between the Nucleus and Cytoplasm

Most of the L1TD1-interacting proteins identified were annotated as cytoplasmic in location (Figure 2C), although more than 30% were nuclear proteins. The presence of nuclear proteins in the L1TD1 interactome led us to reanalyze the cellular location of L1TD1. In a previous study, we localized L1TD1 in the cytoplasmic processing bodies (P-bodies) based on its co-location with AGO2 and LIN28 (Närvä et al., 2012). However, both of these proteins are able to traffic between the cytoplasm and nucleus (Balzer and Moss, 2007; Robb et al., 2005). In addition, a recent MS analysis identified L1TD1 in the nuclear fraction (Sarkar et al., 2012). To test whether L1TD1 is found in the nucleus, we carried out cell fractionation experiments. The results showed the presence of L1TD1 and LIN28 in both the nuclear and cytoplasmic fractions, whereas OCT4, NANOG, and SOX2 showed extensive nuclear expression (Figure 2D). The location was also validated in cancer cell lines expressing endogenous L1TD1 (Figure 2D).

Based on the immunostainings, we concluded that most of the L1TD1 complexes localize adjacent to the nuclear membrane (Figure 2E). This led us to analyze the L1TD1 interactome for the presence of possible binding partners associated with the nuclear membrane and nuclear pore traffic. In addition to identifying multiple proteins involved in cellular transport, we identified key factors in nuclear pore traffic, such as KPNA2, RAN, and RAE1. Interestingly, karyopherin KPNA2 is enriched in undifferentiated ESCs (Van Hoof et al., 2006) and has been shown to be responsible for the nuclear import of OCT4 (Young et al., 2011). Furthermore, RAE1 is a component of the nuclear pore complex involved in RNA export (Murphy et al., 1996). We validated the cytoplasmic and nuclear location of KPNA2 (Figure 2D), as well as interactions of KPNA2 and RAE1 with L1TD1 (Figure 2F). Further, by reciprocal pull-down of KPNA2, we were able to validate the interaction with L1TD1 and OCT4 (Figure 2F). In addition, we demonstrated that silencing of L1TD1 results in decreased KPNA2 and OCT4 levels (Figure 2G). These findings strongly suggest that L1TD1 is able to traffic between the nucleus and cytoplasm and participate in nuclear pore traffic.

L1TD1 Levels Are Rapidly Decreased by Proteasome Inhibitor

The LC-MS/MS analysis resulted in the identification of multiple members of the ubiquitin-dependent proteasome, such as PSMD1, PSMD2, PSMD7, PSMD8, PSMD11, and PSMD14. Of particular interest, PSMD11 was recently reported to be responsible for the increased proteasome activity in hESCs, which is important for the maintenance of pluripotency (Vilchez et al., 2012). The interaction between L1TD1 and PSMD11 was validated with immunoblotting (Figure 3A). Moreover, silencing of L1TD1 resulted in decreased levels of PSMD11 (Figure 3B).

To further study the role of L1TD1 in the proteasome complex, we inhibited proteasome activity with MG-132, which is known to decrease the expression of *SOX2* (Vilchez et al., 2012). Treatment with MG-132 resulted in an immediate reduction of L1TD1 and SOX2, but did not influence the levels of OCT4 or LIN28 at the early time points

(D) Expression of L1TD1, KPNA2, LIN28, OCT4, NANOG, SOX2, GAPDH (cytoplasmic control), and PARP1 (nuclear control) in total (T), cytoplasmic (C), and nuclear (N) cell fractions of hESCs (H9), 2102EP embryoteratocarcinoma, and TCAM2 seminoma cell lines.
(E) Immunostaining of L1TD1 (red), DNA (DAPI staining, blue), NANOG, Phalloidin (actin), and secondary controls in hIPSCs (HEL 11.4).
(F) WB validation of KPNA2, L1TD1, and RAE1 interactions in L1TD1, KPNA2, IgG control, and UB IP reactions. Additional detections from the experiment shown in Figure 1E. KPNA2 IP hIPSC line HEL24.3.
(G) Effect of L1TD1 silencing on the expression of KPNA2 and OCT4. hESC line HS360. Quantifications can be found in the Supplemental Results.
See also Table S2.

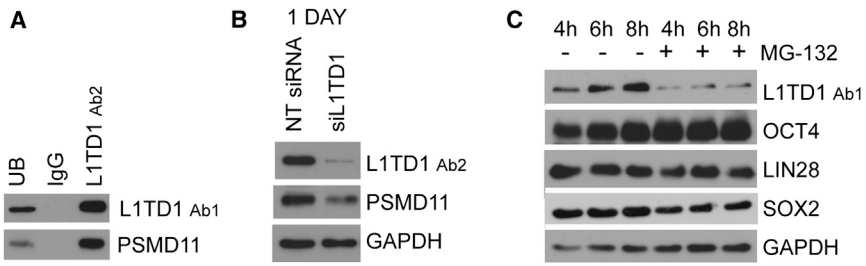


Figure 3. L1TD1 Participates in the High Proteasome Activity of Pluripotency

(A) WB detection of PSMD11 in L1TD1, IgG control, and UB IP reactions. hESC line HS360.

(B) Protein expression of PSMD11 after L1TD1 silencing. hESC line HS360. NT, non-targeting. Quantifications can be found in the Supplemental Results.

(C) Protein expression of L1TD1, OCT4, LIN28, SOX2, and GAPDH after inhibition of proteosomal activity with MG-132 inhibitor (10 μM) versus DMSO control. hESC line H9.

studied (Figure 3C). These results indicate that the protein levels of L1TD1 correlate with proteasome activity and that reduction of L1TD1 expression is an immediate indicator of a decrease in hESC proteasome activity and pluripotency.

L1TD1 Is a Member of the Pluripotency Network

To understand the role of L1TD1 in regulating pluripotency, we compared the L1TD1 interactome with lists of proteins (Jadaliha et al., 2012; Van Hoof et al., 2006) and genes (Assou et al., 2007) that are enriched in hESCs compared with their differentiated counterparts. Based on this analysis, 30% of the L1TD1 interaction partners are associated with pluripotency (Figures 4A and S2A). Nineteen of the interacting proteins had previously been shown to be enriched in hESCs at both the gene and protein levels. Interestingly, the panel contains several factors, such as DPPA4 (developmental pluripotency associated 4), RFC5 (replication factor C subunit 5), and TOP2A (topoisomerase-IIα), whose role in the regulation of pluripotency remains to be investigated. The interaction of L1TD1 with TOP2A was validated (Figure S2B). In addition, silencing of L1TD1 led to decreased TOP2A levels (Figure 1F).

L1TD1 belongs to the interactomes of Nanog (Nitzsche et al., 2011) and Oct4 (Ding et al., 2012; van den Berg et al., 2010) in mESCs. However, NANOG and OCT4 were not detected as L1TD1-interacting proteins in our MS data. It has proved to be challenging to detect these pluripotency factors with MS (Ding et al., 2012; Liang et al., 2008; Pardo et al., 2010; van den Berg et al., 2010). Therefore, we tested the presence of these proteins by immunoblotting of the L1TD1 IPs. This approach led to the detection of OCT4, NANOG, and SOX2 as L1TD1-interacting proteins (Figure 4B). Successful optimization of SOX2-IP further verified L1TD1 interaction with SOX2 (Figure S2C). In addition to these transcription factors, we demonstrated that L1TD1 interacts with the pluripotency-related proteins DNMT3B and TRIM28 (Figure S2D). These interactions were initially identified only in individual MS replicates. In addition, silencing of L1TD1 led to decreased TRIM28 levels (Figure 1F).

Finally, we sought to determine which components of the L1TD1 interactome were also present in the mESC interactomes of OCT4, NANOG, and SOX2 (Figure 4C) (Ding et al., 2012; Nitzsche et al., 2011; van den Berg et al., 2010). We applied IPA and identified 45, 25, and 8 proteins that interact with OCT4, NANOG, and SOX2, respectively. We then analyzed the gene expression of these factors in human undifferentiated, early differentiated (SSEA3−), and differentiated (embryoid bodies [EBs], fibroblasts, and keratinocytes) cells (Figure 4D). This analysis resulted in the identification of proteins that form the key core of pluripotency at the protein level. In addition to NANOG, L1TD1, LIN28A, SOX2, and OCT4, this analysis suggested that MCM5 (minichromosome maintenance complex component 5), PARP1 (poly ADP-ribose polymerase 1), CDK1 (cyclin-dependent kinase 1), RFC4, and TOP2A could have important roles in the regulation of pluripotency. To evaluate the importance of these factors in terms of pluripotency, we selected one of them to determine the influence of its silencing. Depletion of RFC4 led to an immediate reduction of OCT4 and L1TD1 levels (Figure 4E), demonstrating that this factor is required for the maintenance of pluripotency.

To confirm our results, we repeated selected L1TD1 protein interactions in hiPSCs (Figure S2E) and tested the specificity of the L1TD1-bead-antibody complex by performing IP in differentiated HeLa control cells compared with pluripotent cells, including the RFC4 interaction (Figure S2F).

DISCUSSION

RBPs are engaged in every aspect of RNA biology and are essential for carrying specific RNA products from the translation machinery to pre-mRNA splicing, transport, localization, translation, and turnover (Glisovic et al., 2008). Our study indicates that L1TD1 is a functional RBP in human pluripotent cells that is associated with proteins that participate in cellular RNA processes, thus shedding light on the RNA regulation of pluripotent cells.

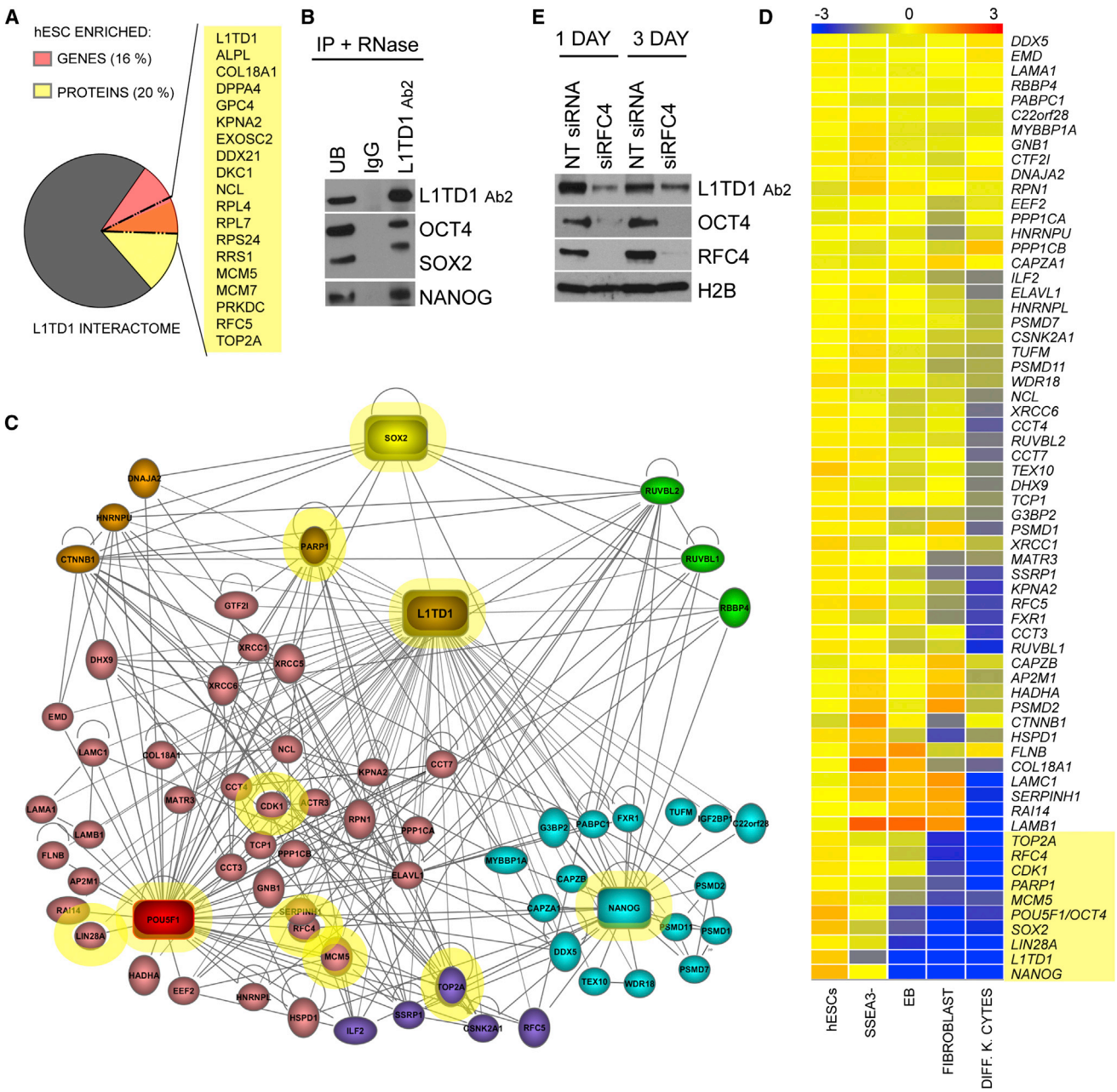


Figure 4. L1TD1 Connects Uncharacterized Factors to Core Pluripotency Regulation

(A) hESC-enriched genes (Assou et al., 2007) and proteins (Jadaliha et al., 2012; Van Hoof et al., 2006) in the L1TD1 interactome. Overlapping hits are listed.

(B) WB detection of OCT4, SOX2, and NANOG in L1TD1, IgG control, and UB IP reactions. hESC line HS360.

(C) Proteins that interact with L1TD1 and OCT4 (n = 45), NANOG (n = 25), or SOX2 (n = 8) (Ding et al., 2012; Nitzsche et al., 2011; van den Berg et al., 2010). The figure was constructed with the use of IPA software.

(D) RNA expression of proteins represented in Figure 4C in hESCs and differentiated cells. EB, embryoid body; DIFF.K.CYTES, differentiated keratinocytes. Microarray data collected from previous studies (Enver et al., 2005; Golan-Mashiach et al., 2005; Skottman et al., 2005). The figure was generated from normalized raw files using GeneSpring GX Software (Agilent Technologies Genomics). The top ten differentiation-responsive genes are highlighted in yellow.

(E) Effect of RFC4 silencing (siRNA1) on the expression of L1TD1 and OCT4. hESC line HS360. Quantifications can be found in the Supplemental Results.

See also Figure S2.

Several observations support the idea that L1TD1 shuttles between the nucleus and cytoplasm. Based on our results and consistent with a previous report on the subcellular proteome of hESCs (Sarkar et al., 2012), we conclude that L1TD1 (like LIN28) can be found in cytoplasmic, nuclear, and membrane fractions. Based on the NetNES 1.1 server, the amino acids L811–L814 could be responsible for the nuclear location of L1TD1. As RBPs are known to participate in RNA traffic through nuclear pores, it is possible that L1TD1 has a role in that process. The interaction of L1TD1 with OCT4, SOX2, NANOG, and the nuclear pore traffic protein karyopherin KPNA2 suggests the intriguing possibility that L1TD1 is responsible for the nuclear import of these transcription factors after their translation to protein in the cytoplasm. Highly stem-cell-specific L1TD1 could be responsible for the nuclear localization, thereby controlling these transcription factors and providing an additional checkpoint for the maintenance of pluripotency.

The network shared among L1TD1, OCT4, SOX2, and NANOG is highly interesting. Our analysis suggests that the nuclear epigenetic regulator PARP1 and CDK1 are key proteins in the core pluripotency network. Recently, PARP1 expression and PARylation activity were shown to be increased in undifferentiated mESCs and to regulate nuclear reprogramming (Chiou et al., 2013; Doege et al., 2012; Lai et al., 2012). In addition, CDK1 was shown to affect the differentiation of mESCs through an interplay with Oct4 (Li et al., 2012). Our results also revealed a panel of proteins that were not previously reported in the context of self-renewal but may be involved in the regulation of pluripotency. Although the role of TOP2A, RCF4, and MCM5 in pluripotency remains uncharacterized, their contribution to this network and function in other cell types suggest they may have a role in regulation of the cell cycle and DNA replication in pluripotent cells. We chose to further investigate the role of RFC4 and demonstrated that it is vital for maintaining pluripotency.

Post-transcriptional regulators and protein turnover determine which factors are active in pluripotency. Recently, the deubiquitinating enzyme Psmd14 was shown to be essential for cellular reprogramming (Buckley et al., 2012), and PSMD11 was shown to regulate increased proteasome activity in hESCs (Vilchez et al., 2012). However, it is still unclear how proteins are selected for ubiquitination in pluripotent cells. The presence of multiple proteasome subunits in the highly pluripotent-specific L1TD1 interactome suggests that L1TD1 could participate in this selection.

One of the central functional categories detected in the L1TD1 interactome was splicing (Figure 1C). Cell-type-specific RBPs are able to regulate tissue-specific splicing (Licatalosi and Darnell, 2010), which has been shown to be unique in mESCs and reorganized during reprogramming by RBPs U2af1 and Srsf3 (Ohta et al., 2013). The interaction between L1TD1 (identified here in hESCs) and these two factors implies that U2AF1 and SRSF3 also have an important role in human pluripotency, and L1TD1 may play a role in stem-cell-specific splicing.

In addition to regulating pluripotency, L1TD1 may have an important function in cancer. We previously showed that L1TD1 is expressed in selected cancer types, such as colorectal carcinoma, ovarian germ cell tumors, and testis seminomas and non-seminomas (Närvä et al., 2012). Interestingly, the L1TD1 interactome reveals many proteins that have an important role or diagnostic value in cancer. For example, KPNA2 has been shown to affect malignant transformation and its elevated expression is correlated with poor diagnosis in multiple forms of cancer (Christiansen and Dyrskjot, 2012).

Of all the factors in the interactome, L1TD1 has the fastest kinetics in response to differentiation. This implies that L1TD1 has a regulatory role in the maintenance of pluripotency. Moreover, interaction with transcriptional, cell cycle, epigenetic, proteosomal, protein traffic, and splicing regulators further supports the regulatory role of L1TD1. The protein interactome constructed here can be used as an important resource for elucidating individual protein interactions in pluripotent stem cells and possibly in cancer stem cells. Moreover, this analysis sheds light on post-translational regulation in pluripotency and the importance of RBPs in pluripotent cells.

EXPERIMENTAL PROCEDURES

All experiments were performed in feeder-free culture conditions on Matrigel (BD Biosciences) in mTeSR1 media (Stem Cell Technologies).

For IP, cells were lysed into NP-40 buffer. Lysates were preincubated with 10 μg/ml RNase A (QIAGEN) at +4°C. IP was carried out using M-280 sheep anti-rabbit IgG Dynabeads (11203D; Invitrogen) with L1TD1 HPA030064 (Ab1), HPA028501 (Ab2) (Sigma), SOX-2 #5024 (Cell Signaling), and normal rabbit IgG 12-370 (Millipore).

IP proteins were gel separated, digested, and submitted for LC-MS/MS analysis. The criteria used for inclusion as an L1TD1-interacting protein were (1) the presence of the protein in two out of three replicates, (2) more than one unique peptide was identified in the L1TD1 IP analysis, and (3) on/off or $\geq$3-fold enrichment of the identified peptides compared with the control IgG IP reaction.

Detailed methods are described in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Results, Supplemental Experimental Procedures, two figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.01.014.

AUTHOR CONTRIBUTIONS

M.R.E. and E.N. conceived and designed the study, performed experimental work, analyzed data, and wrote the manuscript. A.S., M.V., and N.R. performed experimental work. K.D. and D.C. analyzed data. A.R. performed mass spectrometry. R.M. provided scientific support and language editing. R.T. provided supervision for hIPSC cultures. R. Lund and L.L.E. performed data analysis and provided supervision. R. Lahesmaa was the leader and supervisor of the project from conception and design of the study to writing of the manuscript.

ACKNOWLEDGMENTS

We thank Outi Hovatta (Karolinska Institute, Sweden) for the hESC lines and Timo Otonkoski (University of Helsinki) for hiPSC lines. We also thank Marjo Hakkarainen, Päivi Junni, Bogata Fezazi, and Krista Maurinen for excellent technical assistance, and Omid Rasool and Juha-Pekka Pursiheimo for fruitful discussions. We are grateful to the Turku Proteomics Facility, Cell Imaging Core, Turku Centre for Biotechnology (Turku, Finland), and Biomedicum Stem Cell Center (Helsinki, Finland) for their assistance. This study was supported by funding for the Juvenile Diabetes Research Foundation (JDRF), The Academy of Finland (the Centre of Excellence in Molecular Systems Immunology and Physiology Research, 2012-2017, Decision No. 250114 and Postdoctoral Researcher Decision No. 265723), the Finnish Cancer Organizations, the Sigrid Jusélius Foundation, the Turku Graduate School of Biomedical Sciences, the Finnish Cultural Foundation, European Commission Seventh Framework grant EC-FP7-SYBILLA-201106, and the Finnish Cancer Institute.

Received: July 24, 2014
Revised: January 16, 2015
Accepted: January 16, 2015
Published: February 19, 2015

REFERENCES

Assou, S., Le Carrour, T., Tondeur, S., Ström, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Réme, T., Hugnot, J.P., et al. (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells *25*, 961–973.

Balzer, E., and Moss, E.G. (2007). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA Biol. *4*, 16–25.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Buckley, S.M., Aranda-Orgilles, B., Strikoudis, A., Apostolou, E., Loizou, E., Moran-Crusio, K., Farnsworth, C.L., Koller, A.A., Dasgupta, R., Silva, J.C., et al. (2012). Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. Cell Stem Cell *11*, 783–798.

Chiou, S.H., Jiang, B.H., Yu, Y.L., Chou, S.J., Tsai, P.H., Chang, W.C., Chen, L.K., Chen, L.H., Chien, Y., and Chiou, G.Y. (2013). Poly(ADP-ribose) polymerase 1 regulates nuclear reprogramming and promotes iPSC generation without c-Myc. J. Exp. Med. *210*, 85–98.

Christiansen, A., and Dyrskjot, L. (2012). The functional role of the novel biomarker karyopherin α 2 (KPNA2) in cancer. Cancer Lett. *331*, 18–23.

Ding, J., Xu, H., Faiola, F., Ma'ayan, A., and Wang, J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res. *22*, 155–167.

Doege, C.A., Inoue, K., Yamashita, T., Rhee, D.B., Travis, S., Fujita, R., Guarnieri, P., Bhagat, G., Vanti, W.B., Shih, A., et al. (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature *488*, 652–655.

Edwards, T.K., Saleem, A., Shaman, J.A., Dennis, T., Gerigk, C., Oliveros, E., Gartenberg, M.R., and Rubin, E.H. (2000). Role for nucleolin/Nsr1 in the cellular localization of topoisomerase I. J. Biol. Chem. *275*, 36181–36188.

Enver, T., Soneji, S., Joshi, C., Brown, J., Iborra, F., Orntoft, T., Thykjaer, T., Maltby, E., Smith, K., Abu Dawud, R., et al. (2005). Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. Hum. Mol. Genet. *14*, 3129–3140.

Erard, M.S., Belenguer, P., Caizergues-Ferrer, M., Pantaloni, A., and Amalric, F. (1988). A major nucleolar protein, nucleolin, induces chromatin decondensation by binding to histone H1. Eur. J. Biochem. *175*, 525–530.

Gao, Z., Cox, J.L., Gilmore, J.M., Ormsbee, B.D., Mallanna, S.K., Washburn, M.P., and Rizzino, A. (2012). Determination of protein interactome of transcription factor Sox2 in embryonic stem cells engineered for inducible expression of four reprogramming factors. J. Biol. Chem. *287*, 11384–11397.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. *582*, 1977–1986.

Golan-Mashiach, M., Dazard, J.E., Gerecht-Nir, S., Amariglio, N., Fisher, T., Jacob-Hirsch, J., Bielorai, B., Osenberg, S., Barad, O., Getz, G., et al. (2005). Design principle of gene expression used by human stem cells: implication for pluripotency. FASEB J. *19*, 147–149.

Huang, Y. (2012). A mirror of two faces: Lin28 as a master regulator of both miRNA and mRNA. Wiley Interdiscip. Rev. RNA *3*, 483–494.

Iwabuchi, K.A., Yamakawa, T., Sato, Y., Ichisaka, T., Takahashi, K., Okita, K., and Yamanaka, S. (2011). ECAT11/L1td1 is enriched in ESCs and rapidly activated during iPSC generation, but it is dispensable for the maintenance and induction of pluripotency. PLoS ONE *6*, e20461.

Jadaliha, M., Lee, H.J., Pakzad, M., Fathi, A., Jeong, S.K., Cho, S.Y., Baharvand, H., Paik, Y.K., and Salekdeh, G.H. (2012). Quantitative proteomic analysis of human embryonic stem cell differentiation by 8-plex iTRAQ labelling. PLoS ONE *7*, e38532.

Khazina, E., Truffault, V., Büttner, R., Schmidt, S., Coles, M., and Weichenrieder, O. (2011). Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. Nat. Struct. Mol. Biol. *18*, 1006–1014.

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013).

The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1122–1130.

Lai, Y.S., Chang, C.W., Pawlik, K.M., Zhou, D., Renfrow, M.B., and Townes, T.M. (2012). SRY (sex determining region Y)-box2 (Sox2)/poly ADP-ribose polymerase 1 (Parp1) complexes regulate pluripotency. Proc. Natl. Acad. Sci. USA *109*, 3772–3777.

Lapeyre, B., Bourbon, H., and Amalric, F. (1987). Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc. Natl. Acad. Sci. USA *84*, 1472–1476.

Li, L., Wang, J., Hou, J., Wu, Z., Zhuang, Y., Lu, M., Zhang, Y., Zhou, X., Li, Z., Xiao, W., and Zhang, W. (2012). Cdk1 interplays with Oct4 to repress differentiation of embryonic stem cells into trophectoderm. FEBS Lett. *586*, 4100–4107.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., and Songyang, Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat. Cell Biol. *10*, 731–739.

Licatalosi, D.D., and Darnell, R.B. (2010). RNA processing and its regulation: global insights into biological networks. Nat. Rev. Genet. *11*, 75–87.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods *10*, 730–736.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Murphy, R., Watkins, J.L., and Wente, S.R. (1996). GLE2, a Saccharomyces cerevisiae homologue of the Schizosaccharomyces pombe export factor RAE1, is required for nuclear pore complex structure and function. Mol. Biol. Cell *7*, 1921–1937.

Närvä, E., Rahkonen, N., Emani, M.R., Lund, R., Pursiheimo, J.P., Nästi, J., Autio, R., Rasool, O., Denessiouk, K., Lähdesmäki, H., et al. (2012). RNA-binding protein L1TD1 interacts with LIN28 via RNA and is required for human embryonic stem cell self-renewal and cancer cell proliferation. Stem Cells *30*, 452–460.

Nitzsche, A., Paszkowski-Rogacz, M., Matarese, F., Janssen-Megens, E.M., Hubner, N.C., Schulz, H., de Vries, I., Ding, L., Huebner, N., Mann, M., et al. (2011). RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. PLoS ONE *6*, e19470.

Ohta, S., Nishida, E., Yamanaka, S., and Yamamoto, T. (2013). Global splicing pattern reversion during somatic cell reprogramming. Cell Rep. *5*, 357–366.

Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M., and Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell *6*, 382–395.

Robb, G.B., Brown, K.M., Khurana, J., and Rana, T.M. (2005). Specific and potent RNAi in the nucleus of human cells. Nat. Struct. Mol. Biol. *12*, 133–137.

Sarkar, P., Collier, T.S., Randall, S.M., Muddiman, D.C., and Rao, B.M. (2012). The subcellular proteome of undifferentiated human embryonic stem cells. Proteomics *12*, 421–430.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Skottman, H., Mikkola, M., Lundin, K., Olsson, C., Strömberg, A.M., Tuuri, T., Otonkoski, T., Hovatta, O., and Lahesmaa, R. (2005). Gene expression signatures of seven individual human embryonic stem cell lines. Stem Cells *23*, 1343–1356.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. *39*, D561–D568.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

van den Berg, D.L., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369–381.

Van Hoof, D., Passier, R., Ward-Van Oostwaard, D., Pinkse, M.W., Heck, A.J., Mummery, C.L., and Krijgsveld, J. (2006). A quest for human and mouse embryonic stem cell-specific proteins. Mol. Cell. Proteomics *5*, 1261–1273.

Van Hoof, D., Muñoz, J., Braam, S.R., Pinkse, M.W., Linding, R., Heck, A.J., Mummery, C.L., and Krijgsveld, J. (2009). Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell *5*, 214–226.

Vilchez, D., Boyer, L., Morantte, I., Lutz, M., Merkwirth, C., Joyce, D., Spencer, B., Page, L., Masliah, E., Berggren, W.T., et al. (2012). Increased proteasome activity in human embryonic stem cells is regulated by PSMD11. Nature *489*, 304–308.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

Wong, R.C., Ibrahim, A., Fong, H., Thompson, N., Lock, L.F., and Donovan, P.J. (2011). L1TD1 is a marker for undifferentiated human embryonic stem cells. PLoS ONE *6*, e19355.

Young, J.C., Major, A.T., Miyamoto, Y., Loveland, K.L., and Jans, D.A. (2011). Distinct effects of importin α2 and α4 on Oct3/4 localization and expression in mouse embryonic stem cells. FASEB J. *25*, 3958–3965.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

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

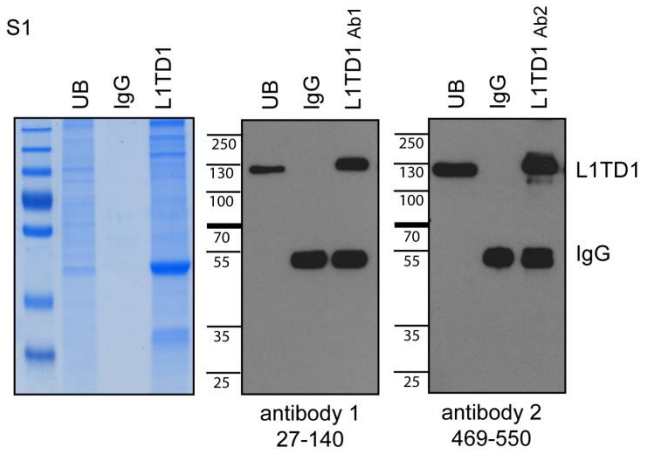


Figure S1, Related to Figure 1. IP with L1TD1 antibodies. Coomassie stained SDS-gel and WB detection of L1TD1 in IP reactions with IgG, L1TD1 HPA030064 (Ab1) and HPA028501 (Ab2), hESC line HS360, (UB) unbound fraction.

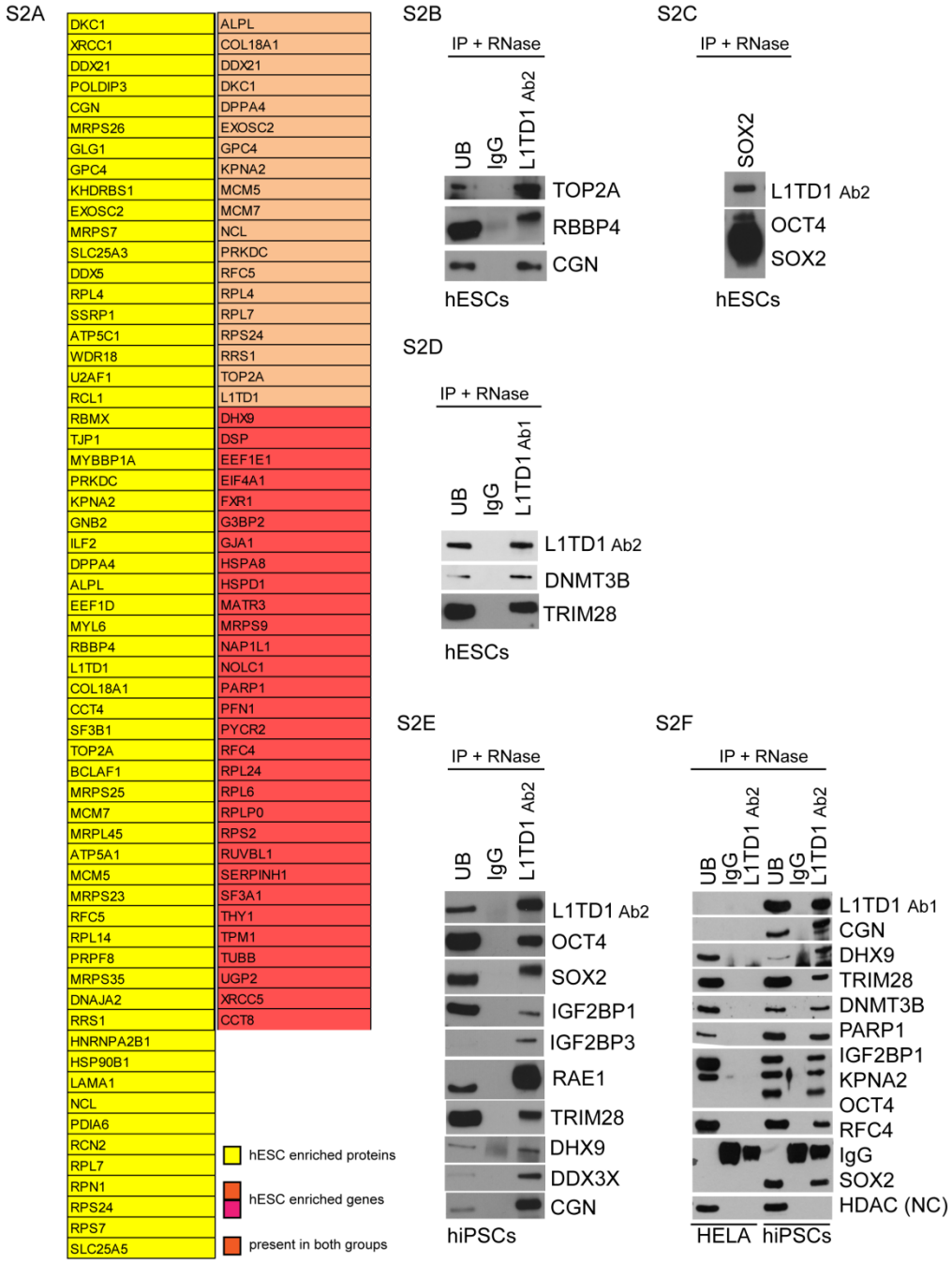


Figure S2, Related to Figure 4. Additional validations. A) hESC enriched genes (Assou et al., 2007) (red) and proteins (yellow) (Jadaliha et al., 2012; Van Hoof et al., 2006) in L1TD1 interactome. Present in both (orange). B) WB detection of TOP2A, RBBP4 and CGN in IP reactions with IgG and L1TD1. Additional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

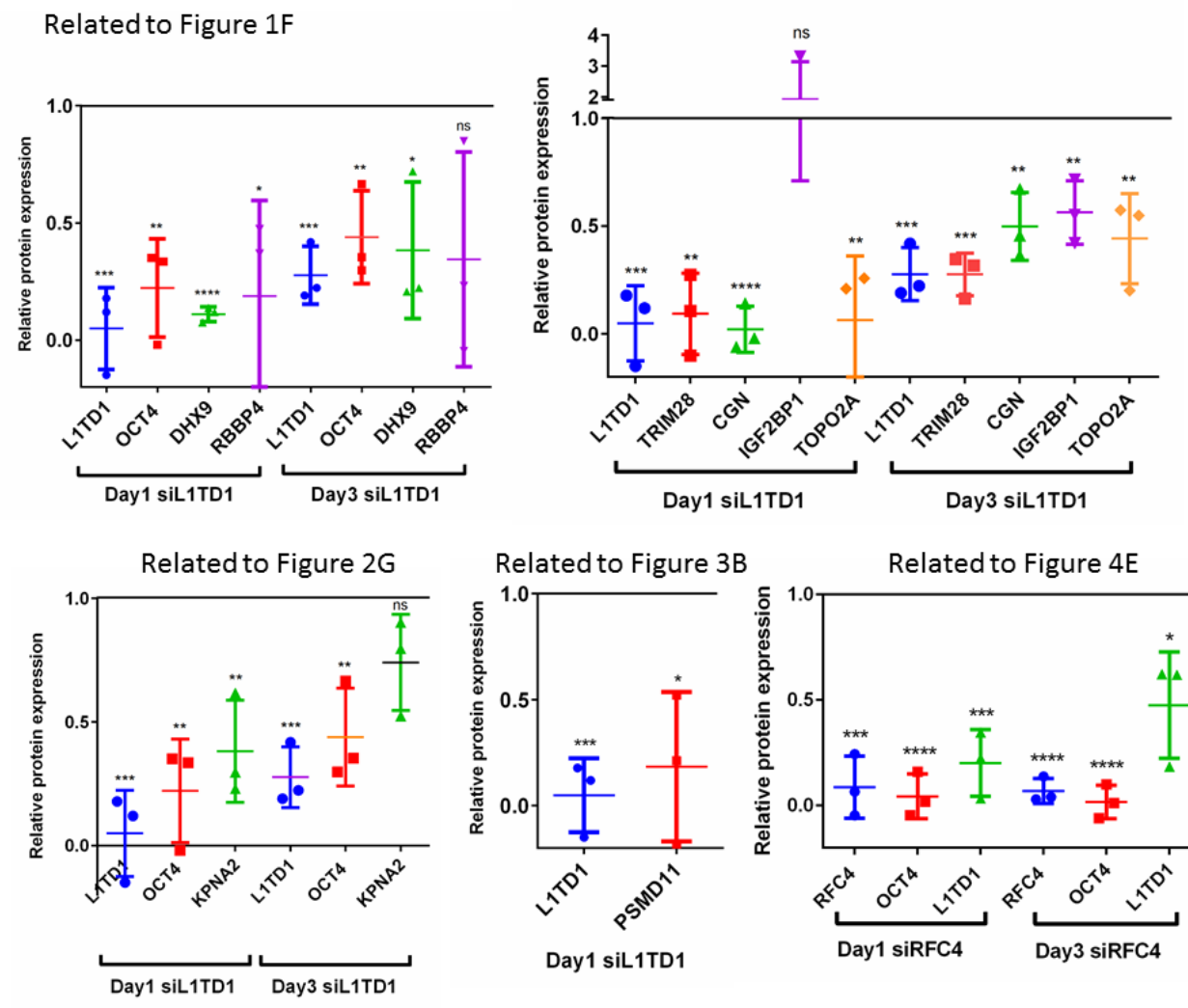
Table S2, Related to Figure 1, CRAPome analysis. Workflow 1 (Percentage found in other studies) and Workflow 3 (Saint Probability) for L1TD1 interactome proteins. (Excel File)

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.