Title

Proteome Profile of Endogenous Retrotransposon-Associated Complexes in Human Embryonic Stem Cells

Journal

Proteomics, 19(15)

ISSN

1615-9853

Authors

Vuong, Linh M
Pan, Songqin
Donovan, Peter J

Publication Date

2019-08-01

DOI

10.1002/pmic.201900169

Peer reviewed

HHS Public Access

Author manuscript

Proteomics. Author manuscript; available in PMC 2021 April 19.

Published in final edited form as:

Proteomics. 2019 August ; 19(15): e1900169. doi:10.1002/pmic.201900169.

Proteome Profile of Endogenous Retrotransposon-Associated Complexes in Human Embryonic Stem Cells

Linh M. Vuong[1,2], **Songqin Pan**[3], **Peter J. Donovan**[1,2,*]

[1]Departments of Developmental and Cell Biology

[2]Departments of Biological Chemistry, UCI

[3]W.M. Keck Proteomics Laboratory, Institute of Integrated Genome Biology, Department of Botany and Plant Sciences, UCR

Abstract

Long INterspersed Element-1 (LINE-1 or L1) are transposable elements similar to retroviruses that have existed in the genome of primates for millions of years. They encode two Open Reading Frame (ORF) proteins (ORF1p and ORF2p) that bind L1 RNA to form a ribonucleoprotein (RNP) complex and are required for L1 integration into the host genome. Humans have evolved with L1 and found ways to limit L1 activity. To identify partners of the L1 RNP, previous studies used ectopic expression of L1 ORF1/2p or RNA in various cancer cells, which express low levels of the ORF proteins. Whether naturally-occurring L1 RNP interacts with the same proteins in non-cancer cells is unknown. Here, we sought to examine the natural assembly of endogenous L1 RNPs in normal human cells. L1 elements are expressed in human embryonic stem cells (hESCs), derived from pre-implantation embryos. Therefore, we used these cells to immunoprecipitate ORF1p followed by MS to identify proteins that associate with the naturally-occurring L1 ORF1p. We found some of the same proteins as well as unique proteins interacting with the endogenous L1 ORF1p complexes. Our analysis of ORF1p-associated proteins in hESCs could help address important questions in both retrotransposon biology and the biology of hESCs.

The human LINE-1 (L1) transposable elements are retrotransposons that can move (transpose) and insert their 6 kb transcript into the host genome. Over the span of human evolution, L1 activity has resulted in accumulation of many L1 copies such that they now are estimated to occupy approximately 17% of the genome [1]. However, because of truncation, internal rearrangement and mutation, only a small number of L1 elements can now actually transpose [2] [3]. Nevertheless, new inserts created through transposition can generate genetic diversity but can also cause disease [4]. The L1 element encodes two open reading frame proteins, ORF1p and ORF2p, that interact with the L1 transcript to form a ribonucleoprotein (RNP) complex that is required for transposing its' own RNA copy back into the host genome [5] [6] [7] [8] [9]. We have yet to fully understand the detailed mechanisms by which L1 RNPs and host cellular proteins co-exist. There have been a number of reports

*Corresponding authors: peter.donovan@uci.edu.

Conflict of interest statement

The authors have declared no conflict of interest

characterizing the L1 RNP complex and its' cellular partners [10] [11] [12] [13] [14] [15]. These studies were carried out using synthetic L1 constructs to overexpress ORF proteins or RNAs exogenously in cancer cell lines. In these studies, proteins of L1 RNP complex were captured either by affinity enrichment of tagged ORF protein or L1 RNA. While these reports address L1 RNP activity in transformed cells, little is known about L1 RNPs in non-transformed cells. In this regard human embryonic stem cells (hESCs) are an interesting cell type which can be maintained in vitro with a normal karyotype and retain developmental pluripotency [16] [17]. Moran and colleagues (2007) showed that the endogenous L1-encoded ORF1p is expressed in hESCs using a synthetic L1 construct tagged with a retrotransposition indicator cassette to show that the L1 element can transpose in these cells [18]. Here we identify cellular proteins that interact with the endogenous L1 ORF1p in hESCs (H9 line) by IP and MS analysis.

To identify and isolate ORF1p, we used a commercial anti-L1 ORF1p antibody (Millipore, clone 4H1) that recognizes the endogenous ORF1p [10]. We performed immunofluorescence on the cells and found that the L1 ORF1p localizes to the cytoplasm as distinct foci in proximity to the nucleus (Fig.1A), consistent with published data on its' localization in human cancer cells and embryonal carcinoma (EC) lines [19] [20] [21]. We used NT2D1, an EC line, as a positive control because these cells are known to express L1 elements [22]. We also performed immunoblot (IB) and MS analysis on hESC extracts. Using IB we identified a 40 kDa band that migrates at the same molecular weight as the band in an NT2D1 EC cell extract (Fig.1B, Fig.S1). We also carried out IP followed by MS analysis (see below) and found that ORF1p was indeed precipitated by the ORF1p antibody and not by a mouse IgG control (Fig.1C). Further, our MS analysis revealed that ORF1p is phosphorylated at residues serine S18 and threonine T203, residues (Fig.S2), known to be important for ORF1p function [23]. Together these data confirm previous studies that L1 elements are active in hESCs [18], suggest that the 4H1 antibody indeed recognizes an endogenous ORF1p in these cells, and that the ORF1p has posttranslational modifications consistent with ORF1p activity.

To profile endogenous L1 ORF1p partners in hESCs, we carried out IP using the anti-ORF1p antibody followed by MS analysis on hESCs in which proteins were cross-linked with formaldehyde (Fig.2A, *see* Supplemental Methods). Three biological replicates of each IP (anti-L1 or control IgG) were analyzed. We identified a total of 2190 proteins, which fall under the false discovery rate (FDR) of <1% ($q \leq 0.01$), that are quantifiable by the PD2.2 software (Table.S1). The majority of the proteins have a $\log_{10}$(P-value) ≤ 2.0 (P-value≤0.010) (Fig.2B). From these 2190 proteins, 219 proteins (179 of which have official gene symbols) have an abundance ratio (AR) of at least 2.0 ($\log2(AR) \geq 1$) (Fig.2B, green spots, Table.S2). Gene Ontology (GO) analysis of the 179 proteins with gene symbols revealed that many of these proteins bind nucleic acids (Table.S3). The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database identified about 10 groups of interconnected proteins (Fig.S3). Some of the functional enrichments in the network include RNA binding complex and RNA metabolism. Twenty-eight proteins from our analysis overlap with the other proteomic studies (Table.S4) [10] [11] [12] [13] [14] [19] [24] [25]. We also compared our list with the list of L1 activator and suppressor genes identified by Liu et al (2018) using genome-wide CRISPR/Cas9 screens in human chronic myeloid leukemia

(K562) and found nine overlapping proteins (Table.S4) [15]. When we applied a higher stringency cutoff on the 219 proteins (by eliminating proteins where peaks are found in at least two of three ORF1p samples, as well as proteins where peaks are not found in at least two of three IgG samples) the list of proteins was reduced to sixty-five. Of these sixty-five proteins, forty-eight have gene symbols, while the rest are protein fragments and globin. (Table 1, *complete list in* Table.S5). These 48 proteins include RNA binding proteins, DNA-damage response proteins and factors involved in chromosome segregation (Fig.2C, Table.S6). Using STRING analysis, we found that eight of the 41 proteins are clustered into three networks (Fig.2D), and include seven proteins (ACT, S100A8, RPL26L1, PURA, ARG1, ZFR, and FANCI) shared with the other proteomic studies (Table 1).

Two proteins that are common in the majority of other proteomic studies (including ours) are MOV10 and PURA as described by Pizarro and Cristofari [26]. MOV10 is an RNA helicase that prevents infection of cells by viruses by interacting with mRNA decay complexes [24] [27] [28] [29], while PURA regulates DNA and RNA molecules for transcription and replication [30] [31] [32] [33].

Proteins not shared with the other studies could represent novel protein-protein interactors of the L1 RNP in individual cell types. Alternatively, the differences seen between the different studies could be the result of different methods of isolating and preparing the ORF1p complexes. Further studies will be required to address this issue.

One protein-protein interaction identified in the STRING analysis from the 41 proteins includes SLIRP and RPL26L1. SLIRP is required for mRNA to associate with the 60S ribosomal protein, RPL26L1, and to maintain mRNA stability [34]. ORF1p could be utilizing these proteins to protect the L1 RNA from degradation. Another network identified in our STRING analysis is between DCP1B and LSM7. These two proteins are involved in deadenylation activity and are required for removing the 5’ cap of mRNA for degradation [35] [36]. DCP1B and LSM7 proteins are found to localize in cytoplasmic foci and, therefore, could potentially co-localize with L1 ORF1p [37] [38]. L1 ORF1p might sequester these proteins to prevent degradation of its’ own RNA or these cellular proteins might be observed in these complexes because they are acting as defense factors to limit L1 activity by regulating L1 RNA.

During human evolution, host and L1 elements have co-existed and they have evolved mechanisms to counteract each others’ activity. We have yet to fully understand their interacting mechanisms in non-cancer cells. Understanding how L1 ORF1p is regulated in non-transformed cells could help us better understand how L1 elements are regulated during normal development and how it is deregulated in cancer cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Drs. Suzanne B. Sandmeyer for advice and encouragement and with G. Wesley Hatfield and Leslie F. Lock for helpful discussions. The authors also thank Terry Nguyen for help with maintaining and plating human ES cells for experiments. This research was funded by a National Institutes of Aging training grant (T32 AG00096) to LMV, and UCI SCRC Seed grant and UCI Institutional funds to PJD. The Orbitrap Fusion mass spectrometer was purchased with funds from a National Institutes of Health shared instrumentation grant (S10 OD010669) to UC Riverside.

Reference

[1]. Smit AF, Current opinion in genetics & development 1996, 6, 743. [PubMed: 8994846]

[2]. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr., Proceedings of the National Academy of Sciences of the United States of America 2003, 100, 5280. [PubMed: 12682288]

[3]. Kulpa DA, Moran JV, Nature structural & molecular biology 2006, 13, 655.

[4]. Chen JM, Stenson PD, Cooper DN, Ferec C, Human genetics 2005, 117, 411. [PubMed: 15983781]

[5]. Dai L, LaCava J, Taylor MS, Boeke JD, Mobile genetic elements 2014, 4, e29319.

[6]. Hohjoh H, Singer MF, The EMBO journal 1996, 15, 630. [PubMed: 8599946]

[7]. Hohjoh H, Singer MF, The EMBO journal 1997, 16, 6034. [PubMed: 9312060]

[8]. Martin SL, Molecular and cellular biology 1991, 11, 4804. [PubMed: 1715025]

[9]. Kulpa DA, Moran JV, Human molecular genetics 2005, 14, 3237. [PubMed: 16183655]

[10]. Taylor MS, LaCava J, Mita P, Molloy KR, Huang CR, Li D, Adney EM, Jiang H, Burns KH, Chait BT, Rout MP, Boeke JD, Dai L, Cell 2013, 155, 1034. [PubMed: 24267889]

[11]. Taylor MS, Altukhov I, Molloy KR, Mita P, Jiang H, Adney EM, Wudzinska A, Badri S, Ischenko D, Eng G, Burns KH, Fenyo D, Chait BT, Alexeev D, Rout MP, Boeke JD, LaCava J, eLife 2018, 7.

[12]. Moldovan JB, Moran JV, PLoS genetics 2015, 11, e1005121.

[13]. Goodier JL, Cheung LE, Kazazian HH Jr., Nucleic acids research 2013, 41, 7401. [PubMed: 23749060]

[14]. Peddigari S, Li PW, Rabe JL, Martin SL, Nucleic acids research 2013, 41, 575. [PubMed: 23161687]

[15]. Liu N, Lee CH, Swigut T, Grow E, Gu B, Bassik MC, Wysocka J, Nature 2018, 553, 228. [PubMed: 29211708]

[16]. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM, Science 1998, 282, 1145. [PubMed: 9804556]

[17]. Kim YE, Park JA, Ha YW, Park SK, Kim HS, Oh SK, Lee Y, Development & reproduction 2012, 16, 353. [PubMed: 25949110]

[18]. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV, Human molecular genetics 2007, 16, 1569. [PubMed: 17468180]

[19]. Goodier JL, Zhang L, Vetter MR, Kazazian HH Jr., Molecular and cellular biology 2007, 27, 6469. [PubMed: 17562864]

[20]. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Kopera HC, Athanikar JN, Hasnaoui M, Bucheton A, Moran JV, Gilbert N, PLoS genetics 2010, 6.

[21]. Skowronski J, Singer MF, Proceedings of the National Academy of Sciences of the United States of America 1985, 82, 6050. [PubMed: 2412228]

[22]. Skowronski J, Fanning TG, Singer MF, Molecular and cellular biology 1988, 8, 1385. [PubMed: 2454389]

[23]. Cook PR, Jones CE, Furano AV, Proceedings of the National Academy of Sciences of the United States of America 2015, 112, 4298. [PubMed: 25831499]

[24]. Goodier JL, Cheung LE, Kazazian HH Jr., PLoS genetics 2012, 8, e1002941.

[25]. Dai L, Taylor MS, O'Donnell KA, Boeke JD, Molecular and cellular biology 2012, 32, 4323. [PubMed: 22907758]

[26]. Pizarro JG, Cristofari G, Frontiers in cell and developmental biology 2016, 4, 14. [PubMed: 27014690]

[27]. Mooslehner K, Muller U, Karls U, Hamann L, Harbers K, Molecular and cellular biology 1991, 11, 886. [PubMed: 1899287]

[28]. Schnieke A, Stuhlmann H, Harbers K, Chumakov I, Jaenisch R, Journal of virology 1983, 45, 505. [PubMed: 6834466]

[29]. Furtak V, Mulky A, Rawlings SA, Kozhaya L, Lee K, Kewalramani VN, Unutmaz D, PloS one 2010, 5, e9081. [PubMed: 20140200]

[30]. Chang CF, Gallia GL, Muralidharan V, Chen NN, Zoltick P, Johnson E, Khalili K, Journal of virology 1996, 70, 4150. [PubMed: 8648759]

[31]. Jurk M, Weissinger F, Lottspeich F, Schwarz U, Winnacker EL, Nucleic acids research 1996, 24, 2799. [PubMed: 8759014]

[32]. Itoh H, Wortman MJ, Kanovsky M, Uson RR, Gordon RE, Alfano N, Johnson EM, Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research 1998, 9, 651. [PubMed: 9716182]

[33]. Bergemann AD, Ma ZW, Johnson EM, Molecular and cellular biology 1992, 12, 5673. [PubMed: 1448097]

[34]. Lagouge M, Mourier A, Lee HJ, Spahr H, Wai T, Kukat C, Silva Ramos E, Motori E, Busch JD, Siira S, German Mouse Clinic C, Kremmer E, Filipovska A, Larsson NG, PLoS genetics 2015, 11, e1005423.

[35]. Lykke-Andersen J, Molecular and cellular biology 2002, 22, 8114. [PubMed: 12417715]

[36]. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Luhrmann R, The EMBO journal 1999, 18, 5789. [PubMed: 10523320]

[37]. Ingelfinger D, Arndt-Jovin DJ, Luhrmann R, Achsel T, Rna 2002, 8, 1489. [PubMed: 12515382]

[38]. Ariumi Y, Frontiers in chemistry 2016, 4, 28. [PubMed: 27446907]

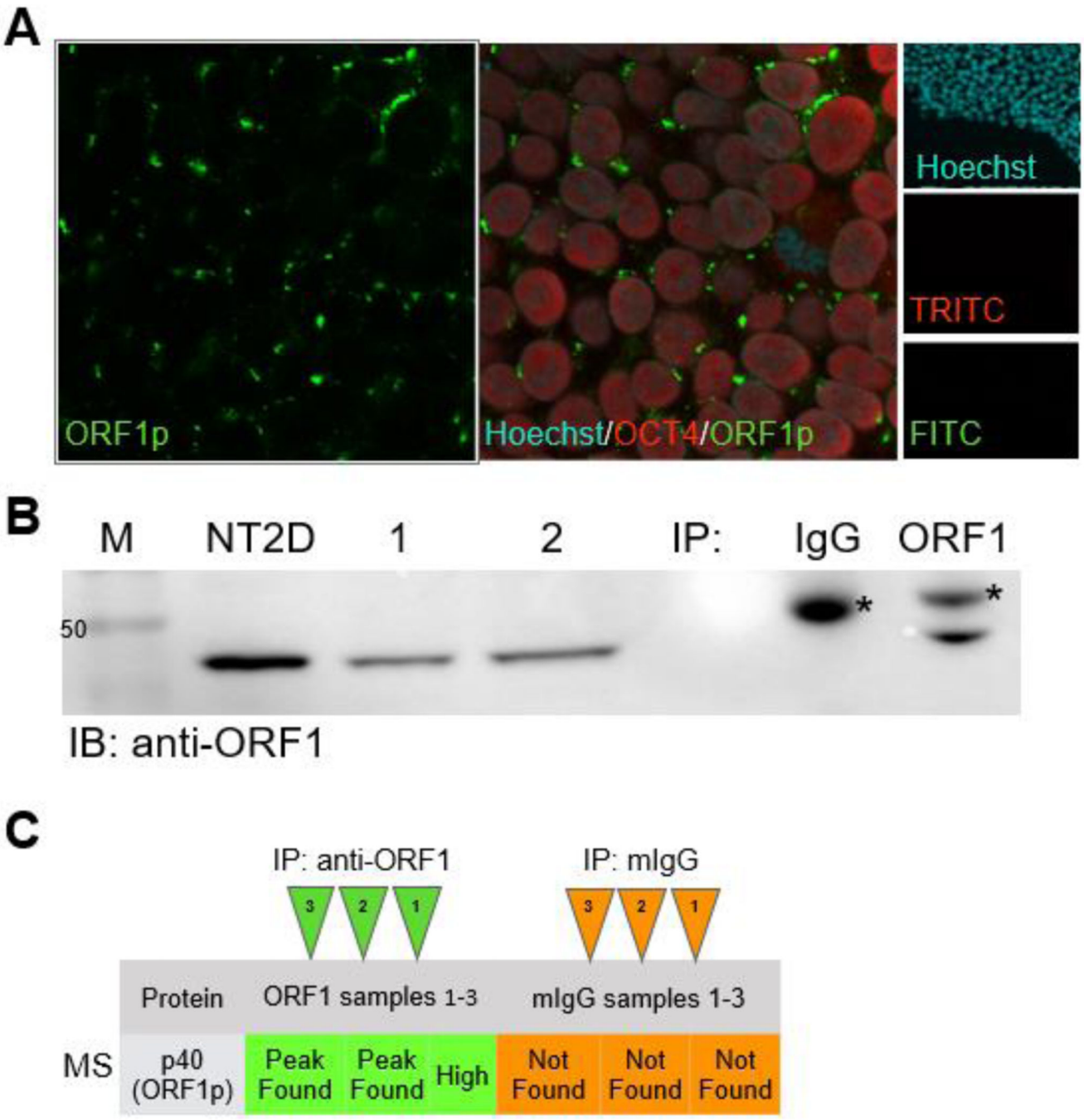

FIG 1.

ORF1p expression in hESCs. (A) IF imaging of ORF1p (left-green) and OCT4 (red) in hESCs, showing single staining of ORF1 (Left) and a merged image of staining for ORF1 (green), OCT4 (nuclei - red) and chromatin (nuclei - blue) (Right). Far right, negative controls for corresponding secondary antibodies. (B) IB analysis of cell extracts to verify expression of ORF1p in hESCs. Lanes 1 and 2 are input at 1 and 2 % from 1.2 mL extracts. The human embryonal carcinoma, NT2D, cell line is a positive control for ORF1p expression. M, molecular markers; ORF1p (40 kDa) band is below the 50 kDa mark. Right,

IP of ORF1p and IgG. Asterisks indicate immunoglobulin heavy chain. (C) Table of mass spectrometry (MS) analysis showing p40 (ORF1) protein in three samples were immunoprecipitated with the anti-ORF1 (green) but not with the mouse IgG control (orange).

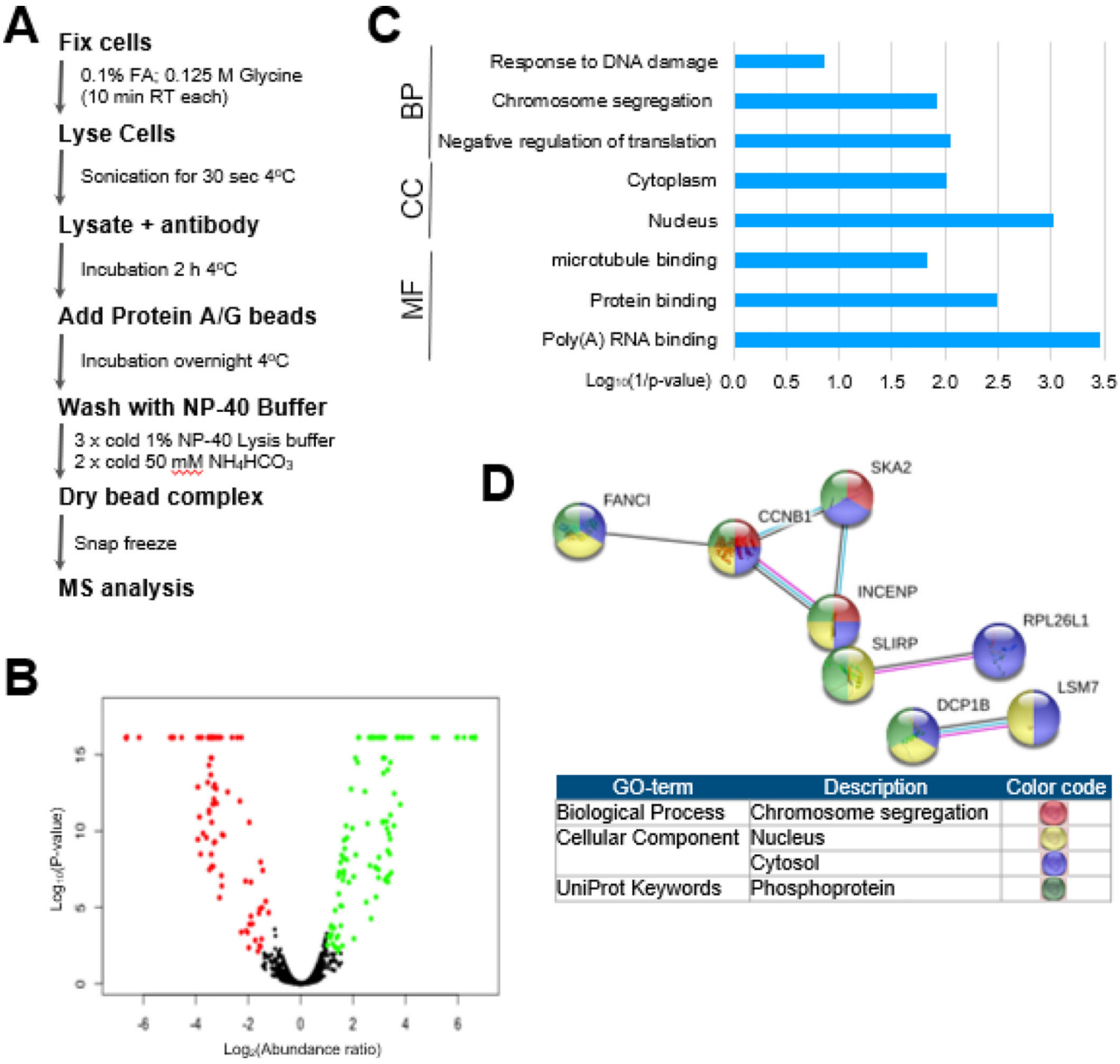

FIG 2.

Proteins identified with ORF1p in hESCs and functional annotation (A) Flowchart for IP. (B) Volcano plot of average abundance ratio (ORF1/IgG) for proteins with an abundance ratio and p-value that are quantifiable by the PD2.2 software. Abundance ratio and p-value are represented as log base 2 and 10, respectively. Green spots represent proteins with an average abundance ratio threshold of ≥2 and p-value of at least 0.010. (C) GO analysis of the 48 out of 65 proteins with annotated gene symbols. BP, Biological Process; CC, Cellular Component, MF, Molecular Function. (D) Networks of protein-protein interactions from 48

proteins. Only eight proteins are closely related. GO analysis of proteins shown in the network. Nodes are color coded as per the GO analysis.

Table 1:

Cellular proteins that interact with the endogenous L1 ORF1p in hESCs (H9)

Gene symbol	Accession #	Description of protein	ORF1 samples			IgG samples		
CAPN5	E7EV01	Calpain-5 OS=Homo sapiens OX=9606 GN=CAPN5 PE=1 SV=2	Peak Found	High	Not Found	Not Found	Not Found	Not Found
RPS19BP1[*]	A0A024R1T1	Ribosomal protein S19 binding protein 1, isoform CRA_a OS=Homo sapiens OX=9606 GN=RPS19BP1 PE=4 SV=1	High	Peak Found	Peak Found	Peak Found	Not Found	Not Found
FAM83D[*]	A0A087WXK8	Chromosome 20 open reading frame 129 OS=Homo sapiens OX=9606 GN=FAM83D PE=1 SV=1	Not Found	High	High	Not Found	High	Not Found
ACT[3]	Q562R8	Actin-like protein (Fragment) OS=Homo sapiens OX=9606 GN=ACT PE=4 SV=1	High	Peak Found	Peak Found	Not Found	Not Found	Not Found
MARF1	Q9Y4F3	Meiosis regulator and mRNA stability factor 1 OS=Homo sapiens OX=9606 GN=MARF1 PE=1 SV=6	High	Peak Found	Peak Found	Not Found	Not Found	Not Found
FMR1[*]	Q06787	Synaptic functional regulator FMR1 OS=Homo sapiens OX=9606 GN=FMR1 PE=1 SV=1	High	High	High	Not Found	Not Found	High
PRB	P02812	Basic salivary proline-rich protein 2 OS=Homo sapiens OX=9606 GN=PRB2 PE=1 SV=3	High	High	Peak Found	Peak Found	Not Found	Not Found
IGHV1OR21-1	A0A087WYE8	Immunoglobulin heavy variable 1/ OR21-1 (non-functional) (Fragment) OS=Homo sapiens OX=9606 GN=IGHV1OR21-1 PE=1 SV=1	Not Found	High	High	Not Found	Not Found	Not Found
S100A8[3]	P05109	Protein S100-A8 OS=Homo sapiens OX=9606 GN=S100A8 PE=1 SV=1	Peak Found	High	Peak Found	Not Found	High	Not Found
GPX4	A0A087WT12	Glutathione peroxidase OS=Homo sapiens OX=9606 GN=GPX4 PE=1 SV=1	Not Found	High	Peak Found	Not Found	Not Found	Not Found
ALG1	Q9BT22	Chitobiosyldiphosphodolichol beta-mannosyltransferase OS=Homo sapiens OX=9606 GN=ALG1 PE=1 SV=2	High	High	Not Found	High	Not Found	Peak Found
PLD2[*]	O14939	Phospholipase D2 OS=Homo sapiens OX=9606 GN=PLD2 PE=1 SV=2	Not Found	Peak Found	High	Not Found	High	Not Found
RBM47	A0AV96	RNA-binding protein 47 OS=Homo sapiens OX=9606 GN=RBM47 PE=1 SV=2	High	High	Peak Found	Not Found	Not Found	Not Found
LIMD2[*]	Q9BT23	LIM domain-containing protein 2 OS=Homo sapiens OX=9606 GN=LIMD2 PE=1 SV=1	Peak Found	High	Not Found	Not Found	Not Found	Not Found
RPL26L1[3]	Q9UNX3	60S ribosomal protein L26-like 1 OS=Homo sapiens OX=9606 GN=RPL26L1 PE=1 SV=1	Peak Found	High	Not Found	Not Found	Not Found	Not Found
MICU2[*]	Q8IYU8	Calcium uptake protein 2, mitochondrial OS=Homo sapiens OX=9606 GN=MICU2 PE=1 SV=2	High	Not Found	High	Not Found	Not Found	Not Found
DMKN[*]	Q6E0U4	Dermokine OS=Homo sapiens OX=9606 GN=DMKN PE=1 SV=3	High	Peak Found	Peak Found	Not Found	Not Found	Peak Found

Gene symbol	Accession #	Description of protein	ORF1 samples			IgG samples		
BRD4[*]	O60885	Bromodomain-containing protein 4 OS=Homo sapiens OX=9606 GN=BRD4 PE=1 SV=2	Not Found	High	High	Not Found	Not Found	Not Found
AVGR5	Q2A130	Autogenous vein graft remodeling associated protein 5 OS=Homo sapiens OX=9606 GN=AVGR5 PE=2 SV=1	High	High	Peak Found	Not Found	Not Found	Not Found
MYO9A[*]	B2RTY4	Unconventional myosin-IXa OS=Homo sapiens OX=9606 GN=MYO9A PE=1 SV=2	Not Found	Peak Found	High	Not Found	Not Found	High
ZNF268	Q14587	Zinc finger protein 268 OS=Homo sapiens OX=9606 GN=ZNF268 PE=1 SV=2	High	Peak Found	Not Found	Not Found	Not Found	Not Found
ZFR[*5]	B2RNR6	Zinc finger RNA binding protein OS=Homo sapiens OX=9606 GN=ZFR PE=2 SV=1	High	High	Not Found	Not Found	Not Found	Not Found
ANP32E	Q9BTT0	Acidic leucine-rich nuclear phosphoprotein 32 family member E OS=Homo sapiens OX=9606 GN=ANP32E PE=1 SV=1	High	Peak Found	High	Not Found	Not Found	High
SLIRP[*]	G3V2S9	SRA stem-loop-interacting RNA-binding protein, mitochondrial OS=Homo sapiens OX=9606 GN=SLIRP PE=1 SV=1	High	High	High	Not Found	Not Found	Not Found
ZNF624[*]	Q9P2J8	Zinc finger protein 624 OS=Homo sapiens OX=9606 GN=ZNF624 PE=1 SV=3	High	High	Not Found	Not Found	Not Found	Not Found
FHOD1[*]	A0A068F7M9	FH1/FH2 domain-containing protein 1 variant OS=Homo sapiens OX=9606 GN=FHOD1 PE=2 SV=1	Not Found	Peak Found	High	Not Found	Not Found	Not Found
PURA[*1,2,3,4,5]	Q00577	Transcriptional activator protein Pur-alpha OS=Homo sapiens OX=9606 GN=PURA PE=1 SV=2	High	High	Peak Found	Not Found	Not Found	Not Found
CCNB1[*]	P14635	G2/mitotic-specific cyclin-B1 OS=Homo sapiens OX=9606 GN=CCNB1 PE=1 SV=1	Peak Found	High	Not Found	Not Found	Not Found	Not Found
CHD2[*]	O14647	Chromodomain-helicase-DNA-binding protein 2 OS=Homo sapiens OX=9606 GN=CHD2 PE=1 SV=2	High	Peak Found	High	High	Not Found	Not Found
ZC3H7B[*]	Q9UGR2	Zinc finger CCCH domain-containing protein 7B OS=Homo sapiens OX=9606 GN=ZC3H7B PE=1 SV=1	Peak Found	High	High	Not Found	Peak Found	Not Found
IQSEC1[*]	A0A087WWK8	IQ motif and SEC7 domain-containing protein 1 OS=Homo sapiens OX=9606 GN=IQSEC1 PE=1 SV=1	Not Found	High	High	Not Found	Not Found	Not Found
NFX1[*]	Q12986	Transcriptional repressor NF-X1 OS=Homo sapiens OX=9606 GN=NFX1 PE=1 SV=2	Peak Found	High	High	Not Found	Not Found	Peak Found
LSM7[*]	Q9UK45	U6 snRNA-associated Sm-like protein LSm7 OS=Homo sapiens OX=9606 GN=LSM7 PE=1 SV=1	Peak Found	High	High	Not Found	High	Not Found
MMS19[*]	Q96T76	MMS19 nucleotide excision repair protein homolog OS=Homo sapiens OX=9606 GN=MMS19 PE=1 SV=2	Not Found	High	Peak Found	Not Found	High	Not Found

Gene symbol	Accession #	Description of protein	ORF1 samples			IgG samples		
GIGYF1	O75420	GRB10-interacting GYF protein 1 OS=Homo sapiens OX=9606 GN=GIGYF1 PE=1 SV=2	Peak Found	High	High	Not Found	High	Not Found
HBB[*]	C8C504	Beta-globin OS=Homo sapiens OX=9606 GN=HBB PE=3 SV=1	High	Peak Found	Not Found	Peak Found	Not Found	Not Found
EIF4ENIF1[*]	A0A024R1K0	Eukaryotic translation initiation factor 4E nuclear import factor 1, isoform CRA_a OS=Homo sapiens OX=9606 GN=EIF4ENIF1 PE=4 SV=1	Not Found	High	High	Not Found	Not Found	Peak Found
ARG1[3]	P05089	Arginase-1 OS=Homo sapiens OX=9606 GN=ARG1 PE=1 SV=2	High	Not Found	Peak Found	Peak Found	Not Found	Not Found
INCENP[*]	Q9NQS7	Inner centromere protein OS=Homo sapiens OX=9606 GN=INCENP PE=1 SV=3	Peak Found	High	Not Found	Not Found	Peak Found	Not Found
LATS1[*]	O95835	Serine/threonine-protein kinase LATS1 OS=Homo sapiens OX=9606 GN=LATS1 PE=1 SV=1	Not Found	Peak Found	High	Not Found	Not Found	High
DCP1B[*]	Q8IZD4	mRNA-decapping enzyme 1B OS=Homo sapiens OX=9606 GN=DCP1B PE=1 SV=2	Not Found	High	High	Not Found	Peak Found	Not Found
NOSIP[*]	A0A075B6F9	Nitric oxide synthase-interacting protein OS=Homo sapiens OX=9606 GN=NOSIP PE=1 SV=1	Not Found	Peak Found	High	Not Found	High	Not Found
SKA2[*]	Q8WVK7	Spindle and kinetochore-associated protein 2 OS=Homo sapiens OX=9606 GN=SKA2 PE=1 SV=1	High	High	Not Found	Not Found	High	Not Found
IGF2R[*]	P11717	Cation-independent mannose-6-phosphate receptor OS=Homo sapiens OX=9606 GN=IGF2R PE=1 SV=3	High	Not Found	Peak Found	Peak Found	Not Found	Not Found
ITPA	A0A0S2Z3W7	Nucleotide diphosphatase (Fragment) OS=Homo sapiens OX=9606 GN=ITPA PE=2 SV=1	Not Found	Peak Found	High	Not Found	Not Found	Peak Found
FANCI[*5]	B7ZMF2	Fanconi anemia, complementation group I OS=Homo sapiens OX=9606 GN=FANCI PE=2 SV=1	Not Found	High	Peak Found	Not Found	Not Found	Peak Found
SPATA24[*]	D6R9Q3	Spermatogenesis-associated protein 24 OS=Homo sapiens OX=9606 GN=SPATA24 PE=1 SV=2	High	Peak Found	Not Found	Not Found	Peak Found	Not Found
SYNPO2[*]	B9EG60	Synaptopodin 2 OS=Homo sapiens OX=9606 GN=SYNPO2 PE=2 SV=1	Not Found	Peak Found	High	Not Found	Not Found	High

[*] Nucleic acid binding

[1] Goodier JL et al., NAR (2013). 41(15): 7401–19.

[2] Moldovan JB and Moran JV. PLoS Genetics (2015) 11(5): e1005121.

[3] Taylor MS et al., Cell (2013). 155(5): 1034–48.

[4] Taylor MS et al., Elife (2018). 7.

[5] Liu N et al., Nature (2018). 553(7687): 228–232.