Title

Multi-omics approach reveals posttranscriptionally regulated genes are essential for human pluripotent stem cells

Journal

iScience, 25(5)

ISSN

2589-0042

Authors

Iwasaki, Mio
Kawahara, Yuka
Okubo, Chikako
et al.

Publication Date

2022-05-01

DOI

10.1016/j.isci.2022.104289

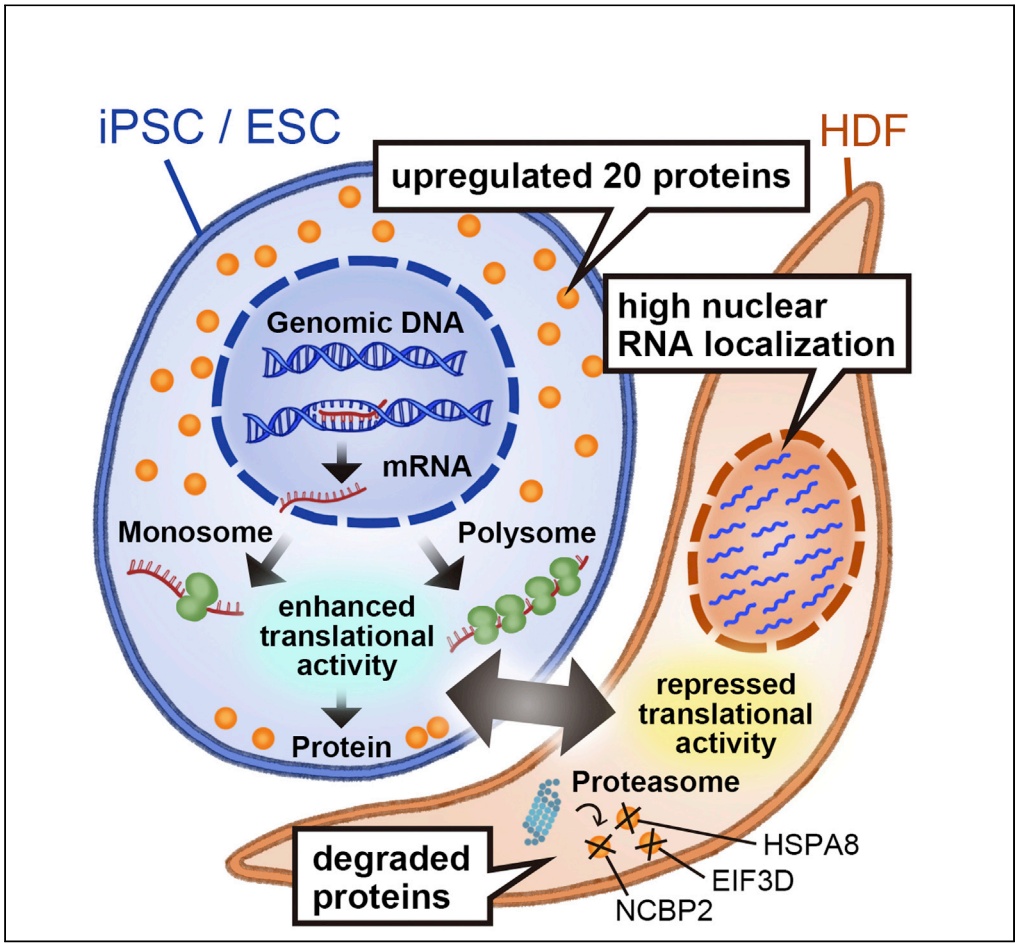
Peer reviewed

iScience

CellPress
OPEN ACCESS

Article

Multi-omics approach reveals posttranscriptionally regulated genes are essential for human pluripotent stem cells



Mio Iwasaki, Yuka Kawahara, Chikako Okubo, ..., Masato Nakagawa, Shinya Yamanaka, Kazutoshi Takahashi

omio@cira.kyoto-u.ac.jp (M.I.)
kazu@cira.kyoto-u.ac.jp (K.T.)

Highlights

The posttranscriptionally regulated 20 genes are necessary for iPSC survival

The proteins of HSPA8, EIF3D, and NCBP2 are quickly degraded in HDFs

mRNA localization affects the protein amounts in most of the 20 genes

Translation is repressed in HDFs despite mRNA binding to ribosomes

Iwasaki et al., iScience *25*, 104289
May 20, 2022
https://doi.org/10.1016/j.isci.2022.104289

iScience

CellPress
OPEN ACCESS

Article

Multi-omics approach reveals posttranscriptionally regulated genes are essential for human pluripotent stem cells

Mio Iwasaki,[1,7,*] Yuka Kawahara,[1] Chikako Okubo,[1] Tatsuya Yamakawa,[1] Michiko Nakamura,[1] Tsuyoshi Tabata,[1] Yohei Nishi,[1,2] Megumi Narita,[1] Akira Ohta,[1] Hirohide Saito,[1] Takuya Yamamoto,[1,3,4] Masato Nakagawa,[1] Shinya Yamanaka,[1,5,6] and Kazutoshi Takahashi[1,*]

SUMMARY

The effects of transcription factors on the maintenance and differentiation of human-induced or embryonic pluripotent stem cells (iPSCs/ESCs) have been well studied. However, the importance of posttranscriptional regulatory mechanisms, which cause the quantitative dissociation of mRNA and protein expression, has not been explored in detail. Here, by combining transcriptome and proteome profiling, we identified 228 posttranscriptionally regulated genes with strict upregulation of the protein level in iPSCs/ESCs. Among them, we found 84 genes were vital for the survival of iPSCs and HDFs, including 20 genes that were specifically necessary for iPSC survival. These 20 proteins were upregulated only in iPSCs/ESCs and not in differentiated cells derived from the three germ layers. Although there are still unknown mechanisms that downregulate protein levels in HDFs, these results reveal that posttranscriptionally regulated genes have a crucial role in iPSC survival.

INTRODUCTION

The processing (Bentley, 2014), export (Hentze et al., 2018), turnover (Dendooven et al., 2020), and accurate decoding of mRNA (Teixeira and Lehmann, 2019), ribosome biogenesis (Pelletier et al., 2018), and protein degradation (Pohl and Dikic, 2019) are critical steps in the posttranscriptional regulation of gene expressions. Among the several hundred human cell types, induced and embryonic pluripotent stem cells (iPSCs/ESCs) have unique posttranscriptional mechanisms (Chen and Hu, 2017), such as specific mRNA processing events, that establish and maintain the pluripotent state (Gabut et al., 2011; Han et al., 2013; Ohta et al., 2013; Salomonis et al., 2010). Interestingly, for many genes, differences in protein levels in ESCs are not accompanied by differences in corresponding mRNA levels (Lu et al., 2009), suggesting posttranscriptionally regulated genes exist in strict regulation at the protein level. One example is the CLOCK gene, in which the amount of mRNA is constant before and after differentiation, whereas the protein expression depends on the cell type and is absent in iPSCs/ESCs (Umemura et al., 2017). Considering that protein levels are more conserved than mRNA levels in primates (Khan et al., 2013), the strict regulation at the protein level with constant mRNA quantities might have important effects on pluripotency and differentiation during development.

To systematically identify genes differently regulated at the mRNA and protein level, transcriptome analysis by RNA sequencing or microarray and proteome analysis by mass spectrometry (MS) has been performed on various cell types (Buccitelli and Selbach, 2020; Matsumoto et al., 2017; Roumeliotis et al., 2017; Wang et al., 2019). Typically, a high correlation in the levels is shown for core metabolic pathway–related genes, but a low correlation is observed for ribosomal and spliceosome genes. However, whether posttranscriptionally regulated genes with tightly controlled protein levels are essential for the maintenance of the cell remains an open question.

Here, we compared global mRNA and protein levels between iPSCs/ESCs and differentiated cells to identify essential posttranscriptionally regulated genes. We found 228 posttranscriptionally regulated genes exclusively in iPSCs/ESCs that showed specific biological functions for RNA and nucleic acid binding.

[1]Department of Life Science Frontiers, Center for iPS Cell Research and Application, Kyoto University, Kyoto 606-8507, Japan

[2]iPSC-Based Drug Discovery and Development Team, RIKEN Bio Resource Research Center (BRC), Kyoto, Japan

[3]Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto 606-8501, Japan

[4]Medical-risk Avoidance based on iPS Cells Team, RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan

[5]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA

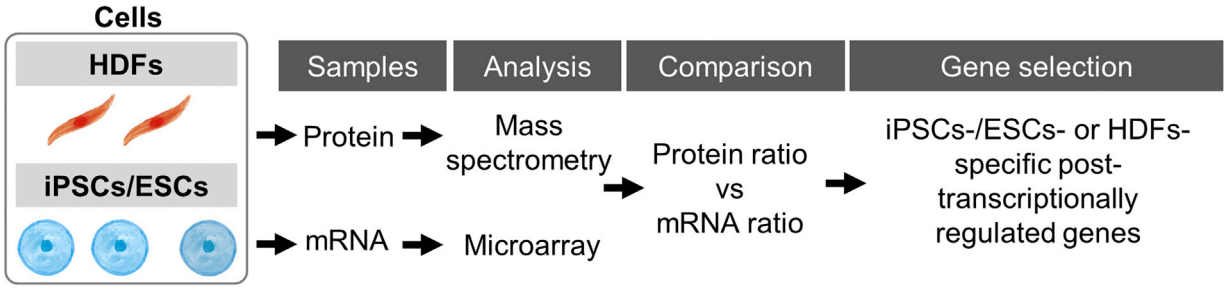
[6]Department of Anatomy, University of California, San Francisco, CA 94143, USA

[7]Lead contact

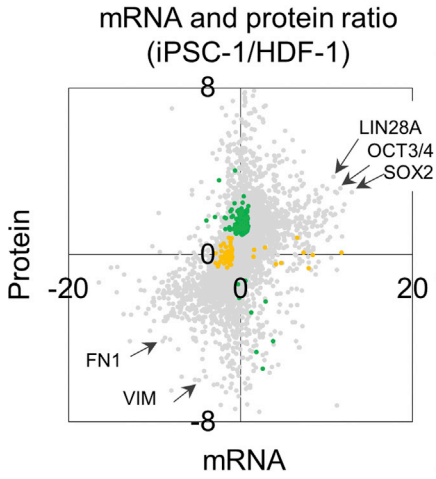
*Correspondence: omio@cira.kyoto-u.ac.jp (M.I.), kazu@cira.kyoto-u.ac.jp (K.T.)

https://doi.org/10.1016/j.isci.2022.104289

A



B



C

	Cell	# Genes	GO (molecular function)	p-value
upregulated genes only at the protein level (p<0.05) (uPRA genes)	iPSCs/ESCs	228	RNA binding	7.12E-97
			nucleic acid binding	6.63E-55
			structural constituent of ribosome	2.96E-50
	HDFs	8	*no enrichment*	--
upregulated genes only at the mRNA level (p<0.05)	iPSCs/ESCs	10	transmembrane transporter activity	0.03
	HDFs	56	cadherin binding	0.01
			cell adhesion molecule binding	0.01

D

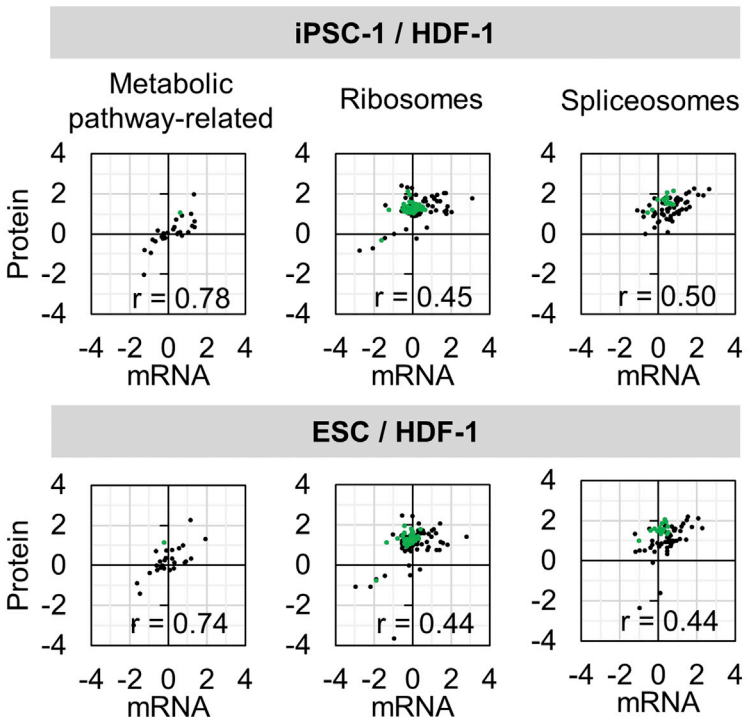


Figure 1. Two hundred twenty-eight and eight genes were respectively identified as posttranscriptionally regulated genes with independent protein upregulation in iPSCs/ESCs and HDFs

(A) Workflow for identifying posttranscriptionally regulated genes using mass spectrometry (MS) for proteins and microarray for mRNA. Two iPSC lines (iPSC-1, 201B7 and iPSC-2, 1418E1), one ESC line (H9), and two HDF lines (HDF-1, HDF1388 and HDF-2, Tig120) were used.

(B) The mRNA-to-protein expression ratio in iPSC-1 and HDF-1 for each gene (Gray). Yellow and green indicate commonly varied genes in iPSCs/ESCs and HDFs, those with only upregulated mRNA, and those with only upregulated protein levels, respectively ($p < 0.05$, two-sample unpaired t test, biological triplicate). Log_2 scale. Comparisons of the mRNA-and-protein ratios between all pairs of iPSCs/ESCs and HDF lines are shown in Figures S1B and S1C. The total number of genes quantified was 6404. All yellow and green plotted data are shown in Table S1. See S1A and "Trans-omics data analysis" in STAR Methods for details.

(C) The number of genes and the GO analysis of molecular functions for genes that were significantly upregulated only at the protein or mRNA level in B.

(D) The mRNA-to-protein expression ratio in iPSC-1 and HDF-1 and in ESC and HDF-1 for metabolic pathway-, ribosome-, and spliceosome-related genes. The correlation coefficients (r) are shown. Black indicates total genes, and green indicates commonly varied genes in iPSCs/ESCs and HDFs that were only upregulated in protein levels in B and C (uPRA genes). Comparisons of the mRNA-and-protein ratios between all pairs of iPSC/ESC and HDF lines are shown in Figure S3.

siRNA screening revealed that 54% (84 genes) was necessary for cell survival and 20 of these genes were identified as iPSC specific. Finally, we show that most of the 20 genes are mainly regulated at the mRNA localization stage, three genes were downregulated at the protein degradation stage in human dermal fibroblast (HDFs), and seven genes were upregulated at the translational stage in iPSCs. Overall, we suggest the importance of cell-type-specific posttranscriptional regulation in iPSCs.

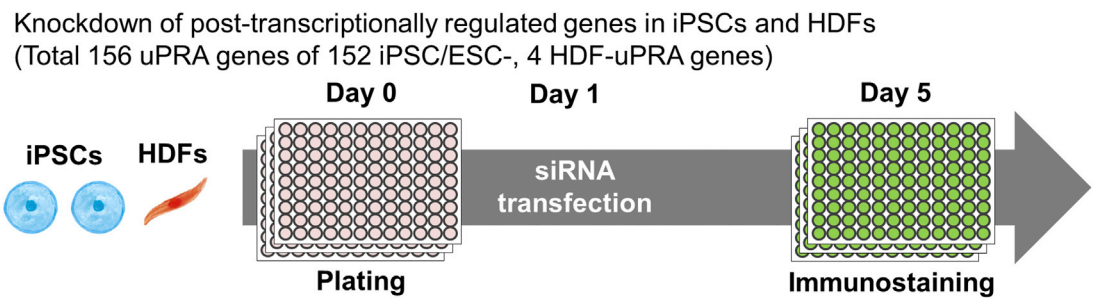
RESULTS

Two hundred twenty-eight upregulated proteins in iPSCs/ESC showed constant mRNA expression

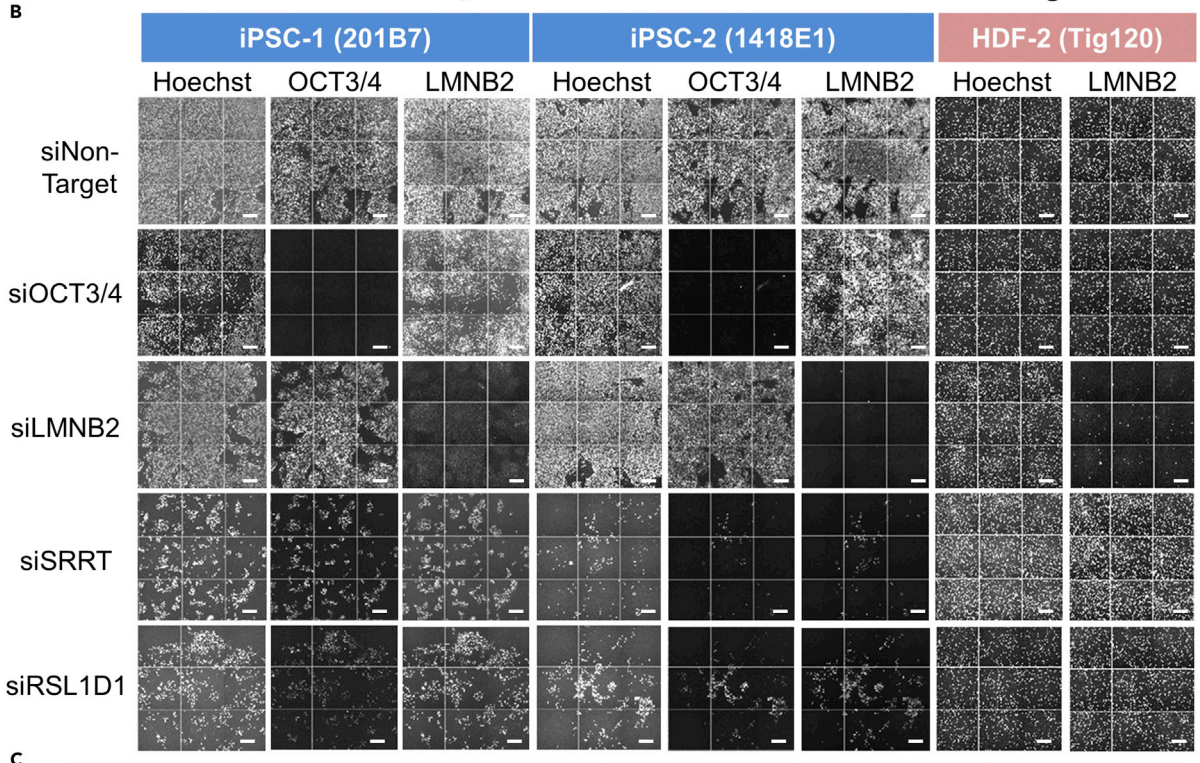
First, to identify upregulated and posttranscriptionally regulated genes, we analyzed the mRNA and protein levels by microarray and MS, respectively (Figures 1A and S1A). We used two HDF lines (HDF-1, HDF1388 and HDF-2, Tig120), two human iPSC lines (iPSC-1, 201B7 and iPSC-2, 1418E1), and one human ESC line (ESC, H9). We compared the ratio of the mRNA and protein levels for 6,404 genes for all combinations of iPSCs (iPSC-1, -2), ESCs, and HDFs (HDF-1, -2) (Figures S1B and S1C). All cell-type-specific genes, including LIN28A, OCT3/4, and SOX2 for iPSCs/ESCs and FN1 and VIM for HDFs, showed a highly correlated mRNA-to-protein ratio (Figure 1B). Then we examined genes that were posttranscriptionally regulated by selecting those that were upregulated only at the mRNA or protein level by log_2 (0.9) (Figure S1B, yellow and green; see Trans-omics data analysis in STAR Methods). We selected and showed commonly varied genes in the six compared pairs of iPSCs/ESCs and HDFs on a representative example of a comparison of iPSC-1 and HDF-1 (Figure 1B, Table S1). From this analysis, we found 10 and 56 genes whose expressions were increased only at the mRNA level in iPSCs/ESCs and HDF, respectively, and 228 and 8 genes whose expressions were increased only at the protein level. A Gene Ontology (GO) analysis was performed to investigate the function of these four gene sets (Figure 1C). In particular, among the 228 genes in iPSCs/ESCs that showed only elevated protein levels, we found specific molecular functions, such as RNA binding and nucleic acid binding. Although the eight genes for HDFs showed no enrichment in molecular functions, we named these 228 and 8 genes "upregulated protein levels independent of mRNA levels (uPRA)" for further analysis.

Ribosome- and spliceosome-related genes, which are known to have low correlation between mRNA and protein levels, are capable of RNA binding. We therefore calculated their correlation coefficients by the normalized quantitative values of mRNAs and proteins for each cell line (Figure S2, Table S2). As previously reported (Buccitelli and Selbach, 2020; Matsumoto et al., 2017), the correlation coefficients for metabolic-pathway-related genes were found to be 0.74 or higher, whereas those for ribosome- and spliceosome-related genes were 0.57 or lower. Correlation coefficients for the mRNA-to-protein ratios were also calculated; except for the HDF1-2 comparison, the correlation coefficients for metabolic-pathway-related factors among iPSCs, ESCs, and HDFs were 0.69 or higher (Figures 1D, S3 and Table S3). This result suggested that the same variation was observed in the mRNA-and-protein ratios regardless of the cell type for metabolic-pathway-related genes. For ribosome- and spliceosome-related genes, the correlation coefficients were less than or equal to 0.57, and comparisons between iPSCs/ESCs and HDFs showed a biased distribution toward higher protein expression levels in iPSCs/ESCs, with some categorized as uPRA genes (Figures 1D and S3). These results suggest that some RNA-binding proteins are posttranscriptionally regulated in iPSCs/ESCs only at the protein level (Table S4).

A



B



C

Cell survival-related genes		iPSC-1	iPSC-2	HDF-2
Total	iPSC-1 ∪ iPSC-2 ∪ HDF-2	84 (83)	80 (79)	66 (**1**)
Common	iPSC-1 ∩ iPSC-2 ∩ HDF-2	52 (52 iPSC/ESC-uPRA genes)		
	iPSC-1 ∩ iPSC-2	21 (**20**)		
	iPSC-1 ∩ HDF-2	4 (4)		4 (0)
	iPSC-2 ∩ HDF-2		2 (2 iPSC/ESC-uPRA genes)	
Unique	Each cell type specific	7 (7)	5 (5)	8 (**1**)

D

E

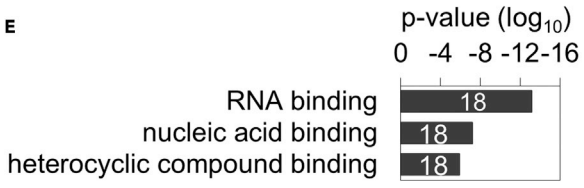


Figure 2. Twenty uPRA genes in iPSCs/ESCs and one uPRA gene in HDFs were related to cell survival

(A) Workflow of the knockdown experiment by siRNA for uPRA genes in iPSCs/ESCs and HDFs (156 uPRA genes in total). iPSC-1 (201B7), iPSC-2 (1418E1), and HDF-2 (Tig120) were used for the assay. See "siRNA screening" in STAR Methods for details. The knockdown targets are shown in Table S5, and the immunoassay cell percentage and observed number of cells are shown in Table S6.

(B) Representative images of the siRNA knockdown experiment. Nuclei were visualized by Hoechst 33342, and iPSCs were visualized by OCT3/4 expression. The knockdown efficiency was visualized by OCT3/4 and LMNB2 expression. siSRRT and siRSL1D1 are representative images for siRNAs experiments in which the cell number decreased compared with controls (siNontarget, siOCT4, and siLMNB2). Cropped image from 96 well, and bars indicate 200 μm.

(C) The number of posttranscriptionally regulated genes (uPRA genes) whose knockdown caused a substantial decrease in cell number. Numbers in parentheses indicate the number of iPSC/ESC or HDF-uPRA genes observed in Figure 1C. Images for the siRNAs of the control and transcriptionally regulated genes, and the total 21 cell-specific uPRA genes are shown in Figures S4A and S4B, respectively.

(D) GO analysis of cellular component for the 20 iPSC/ESC-uPRA genes in C.

(E) GO analysis of molecular function for the 20 iPSC/ESC-uPRA genes in C.

Twenty uPRA genes are essential for iPSC survival

To investigate which uPRA genes are significant for cell maintenance, we performed siRNA knockdown screening in iPSC-1, -2, and HDF-2 (Figures 2A and 2B). We picked siRNAs to target 156 (Table S5) of the 228 iPSC/ESC-uPRA and 8 HDF-uPRA genes. We also selected siRNA to target transcriptionally regulated genes, such as LIN28A, LIN28B, SALL4, SOX2, OCT3/4, and NANOG for iPSCs and VIM, LMNB1, LMNA, ANPEP, and FN1 for HDFs (Figure S4A). In addition, an siRNA that targets LMNB2 was used as a universal transfection control for both cell types (Figures 2B and S4A). Knockdown of none of the transcriptionally regulated genes was associated with cellular viability except for siLMNB1 on HDFs. siOCT3/4 induced iPSC differentiation, but knockdown of any of the 156 uPRA genes did not induce differentiation (Table S6). On the other hand, the knockdown of 84 of the 156 uPRA genes, such as SRRT and RSL1D1, significantly decreased the number of iPSCs compared with control siRNA (siNonTarget, siOCT4, and siLMNB2) (Figures 2B and 2C). The knockdown of 52 of these genes also decreased the number of HDFs (Figure 2C). This result indicates that uPRA genes have essential effects on the cellular maintenance of HDFs as well. In addition, we identified 20 iPSC/ESC-uPRA genes and 1 HDF-uPRA gene that are each specifically necessary for the survival of the respective cells (Figure 2C; numbers in parentheses indicate the number of iPSC/ESC- or HDF-uPRA genes; images of the knockdown of these 21 genes are shown in Figure S4B). We performed *in silico* analysis of motifs in cDNA sequences of the 20 iPSC/ESC-uPRA genes using RBPmap (Paz et al., 2014). We found that the types of motifs are quite diverse (Table S7), and several different combinations of RNA-binding proteins may be responsible for the maintenance of iPSCs. A GO analysis showed that the 20 iPSC/ESC-uPRA genes code for components of ribonucleoprotein- and protein-containing complexes and have RNA- and nucleic-acid-binding properties (Figures 2D and 2E). This result is supported by a previous global RNAi screen analysis (Chia et al., 2010), which found that genes involved in ESC survival on siRNA knockdown are classified in GO categories for nucleic acid binding and ribosomal proteins, mRNA splicing, and processing factors. Overall, these data suggest that the 20 iPSC/ESC-uPRA genes are essential for the survival of iPSCs and synergistically maintain iPSCs via heterocyclic-compound-binding properties.

Upregulated protein levels of the 20 essential uPRA genes are iPSC specific

We examined if the 20 iPSC/ESC-uPRA genes are regulated posttranscriptionally only in iPSCs/ESCs. We selected another seven differentiated primary cell lines representing the three germ layers and compared them with the iPSCs/ESCs, which are undifferentiated. In total, mesoderm-derived cells included adipose-tissue-derived mesenchymal stem cells (HAdMSC) and two HDF lines (HDF-1 and 2), endoderm-derived cells included normal human bronchial epithelial cells (NHBEC) and human prostate epithelial cells (PrEC), and ectoderm-derived cells included normal human epidermal keratinocytes (NHEK) and human neural progenitor cells derived from H9 (NPC H9). Quantitative reverse transcription PCR (qRT-PCR) and antibody-based quantitative protein analysis were performed on the mRNA and proteins, respectively, and the gene expressions relative to HDF-1 were compared (all WES protein images and graphs are shown in Figures S5A and S5C). RSL1D1, one of the 20 essential iPSC/ESC-uPRA genes, showed a more than 50-fold increase in its protein expression only in iPSCs/ESCs, whereas its mRNA expression was equal among all cell types (Figures 3A and 3B). We then compared the mRNA and protein levels of the 20 iPSC/ESC-uPRA genes on iPSCs/ESCs and differentiated cells (Figures 3C and 3D). Overall, we confirmed that the protein expressions of 18 iPSC/ESC-uPRA genes (RBM22 and SF3B3 were the exceptions) were increased more than 2-fold compared with their mRNA expressions in iPSCs/ESCs (Figure 3E). On the other hand, the fold-change between protein and mRNA expressions was around one or less in differentiated cells for 17 iPSC/ESC-uRPA genes (IMP4, NCBP2, and BUD31 were the exceptions), suggesting the

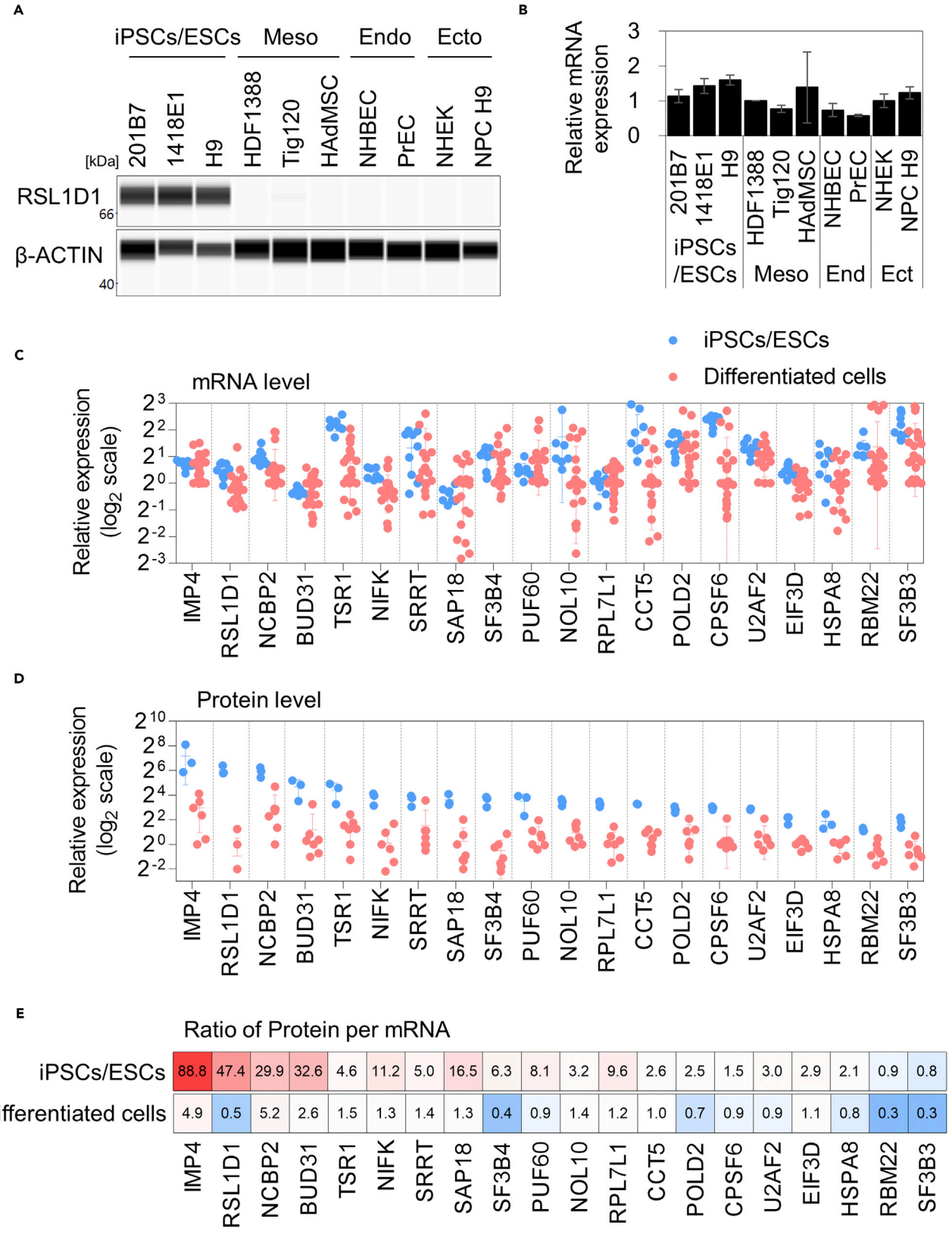

A
iPSCs/ESCs
Meso
Endo
Ecto
201B7
1418E1
H9
HDF1388
Tig120
HAdMSC
NHBEC
PrEC
NHEK
NPC H9
[kDa]
RSL1D1
66
β-ACTIN
40
B
Relative mRNA expression
3
2
1
0
iPSCs /ESCs
Meso
End
Ect
C
iPSCs/ESCs
Differentiated cells
mRNA level
Relative expression (log2 scale)
D
Protein level
IMP4
RSL1D1
NCBP2
BUD31
TSR1
NIFK
SRRT
SAP18
SF3B4
PUF60
NOL10
RPL7L1
CCT5
POLD2
CPSF6
U2AF2
EIF3D
HSPA8
RBM22
SF3B3
E
Ratio of Protein per mRNA
iPSCs/ESCs
88.8 47.4 29.9 32.6 4.6 11.2 5.0 16.5 6.3 8.1 3.2 9.6 2.6 2.5 1.5 3.0 2.9 2.1 0.9 0.8
Differentiated cells
4.9 0.5 5.2 2.6 1.5 1.3 1.4 1.3 0.4 0.9 1.4 1.2 1.0 0.7 0.9 0.9 1.1 0.8 0.3 0.3

Figure 3. Twenty uPRA genes were posttranscriptionally regulated only in iPSCs/ESCs

(A) A representative image of the antibody-based quantitative protein analysis for one iPSC/ESC-uPRA gene, RSL1D1, in various cell types. We used two cell lines for iPSCs, one cell line for ESCs, three cell lines for mesoderm-derived cells, two cell lines for endoderm-derived cells, and two cell lines for ectoderm-derived cells for the gene expression analysis (see text for details). WES protein images of all 20 essential iPSC/ESC-uPRA genes are shown in Figure S5A.

(B) Gene expressions measured by qRT-PCR were normalized by GAPDH. The relative expression ratios were normalized to the result of HDF1388 (HDF-1). Representative results for the RSL1D1 mRNA are shown as the mean ± SD Biological triplicates for the mRNA quantification. Results for all essential iPSC/ESC-uPRA genes are shown in Figure S5C (left panel).

(C) Relative gene expressions of mRNA levels in iPSCs/ESCs and differentiated cells for the 20 essential iPSC/ESC-uPRA genes normalized to HDF1388 (HDF-1). Data are represented as mean ± SD.

(D) Relative gene expressions of protein levels in iPSCs/ESCs and differentiated cells for the 20 essential iPSC/ESC-uPRA genes normalized to HDF1388 (HDF-1). Data are represented as mean ± SD.

(E) A heatmap of the protein-to-mRNA expression ratio (P/R ratio) is shown above each gene name using the average expression value from C and D.

posttranscriptional regulation mechanism is different between iPSCs/ESCs and differentiated cells. RSL1D1, SF3B4, RBM22, and SF3B3 especially showed protein-to-mRNA expression ratios less than one in differentiated cells, indicating that the protein expressions of these genes are usually suppressed in cell types other than iPSCs/ESCs. As for AFP, the one HDF-uPRA gene, we confirmed a specific increase in the protein expression in mesoderm-derived cells (Figure S5B). Following these observations, we concluded that the 20 iPSC/ESC-uPRA genes are posttranscriptionally regulated specifically in iPSCs/ESCs.

HSPA8, EIF3D, and NCBP2 protein expressions are controlled at the ubiquitin-dependent degradation stage in HDFs

Next, we analyzed individual posttranscriptional regulation stages to identify the regulation mechanism of the protein expressions of the 20 essential iPSC/ESC-uPRA genes (Figure 4A). After transcription, mRNAs are transported from the nucleus to the cytoplasm, where ribosomes bind to translate them. Eventually, the proteins are degraded by the ubiquitin-proteasome system and/or lysosomes. At first, we analyzed the protein degradation efficiency using MG-132 (proteasome inhibitor) and Bafilomycin A1 and Wortmannin (lysosome inhibitors) as controls for the assays (Figure S6). These inhibitors should recover the iPSC/ESC-uPRA protein expression in HDFs if the proteins are downregulated by fast degradation. We found that the protein expressions of HSPA8, EIF3D, NCBP2, and IMP4 were steadily increased more than 2-fold after exposure to the proteasome inhibitor for up to 8 h (Figures 4B and S6E). We did the same assay for up to 24 h and confirmed that the protein expressions of HSPA8, EIF3D, and NCBP2 were controlled by fast degradation via the proteasome system in HDFs (Figure 4C). The lysosome inhibitor assay (Figures 4D and S6F) showed high protein expressions of NCBP2 and RSL1D1 for treatment up to 8 h but not 24 h (Figure 4E), suggesting no iPSC/ESC-uPRA genes were regulated by lysosome-dependent degradation. Overall, these data indicate that the protein expressions of three iPSC/ESC-uPRA genes, HSPA8, EIF3D, and NCBP2, are regulated at the ubiquitin-dependent degradation stage.

Cytosolic localization and translation of iPSC/ESC-uPRA mRNA is inhibited in HDFs

Next, we analyzed the subcellular localization of the mRNAs of the 20 essential iPSC/ESC-uPRA genes to examine their intracellular abundance between iPSCs and HDFs. We found that the mRNA percentage in the nucleus is slightly higher in HDFs (Figure S7A). We used 18S rRNA and MALAT1 lncRNA, respectively, as cytosolic and nuclear RNA controls in the subcellular fractionation experiment by qRT-PCR (Figure 5A). For cytosolic RNA, 94.3% and 73.3% of 18S rRNA were observed in iPSCs and HDFs, respectively, and for nuclear RNA, 99.5% and 99.8% of MALAT1 lncRNA were observed, validating the subcellular fractionation experiment. Then we analyzed the cytosolic RNA percentage of the 20 essential iPSC/ESC-uPRA genes by qRT-PCR (Figure 5B). The cytosolic RNA percentage of 18 iPSC/ESC-uPRA genes (IMP4 and CPSF6 were the exceptions) was increased at least more than 2-fold in PSCs. Conversely, the nuclear RNA percentage of iPSC/ESC-uPRA genes was increased in HDFs (Figure 5C). These results indicated that the mRNAs of the 18 iPSC/ESC-uPRA genes were retained in the nucleus in HDFs. We also analyzed global RNA subcellular localization by RNA-seq, as previously reported (Figure S7B) (Carlevaro-Fita and Johnson, 2019; Zuckerman et al., 2020). We compared the nucleus-to-cytoplasm RNA ratios between iPSCs and HDFs (Figure S7C) and found that the overall gene distribution was shifted to the nucleus in HDFs. This observation indicates that the nuclear localization of RNA is not limited to uPRA genes in HDFs.

RNA splicing promotes mRNA-export to the cytoplasm, because the spliceosome can recruit export factors to mature RNA (Palazzo and Lee, 2018). However, ESCs and differentiated cells do not express different spliced isoforms of the 20 iPSC/ESC-uPRA genes (Han et al., 2013). We then checked the protein

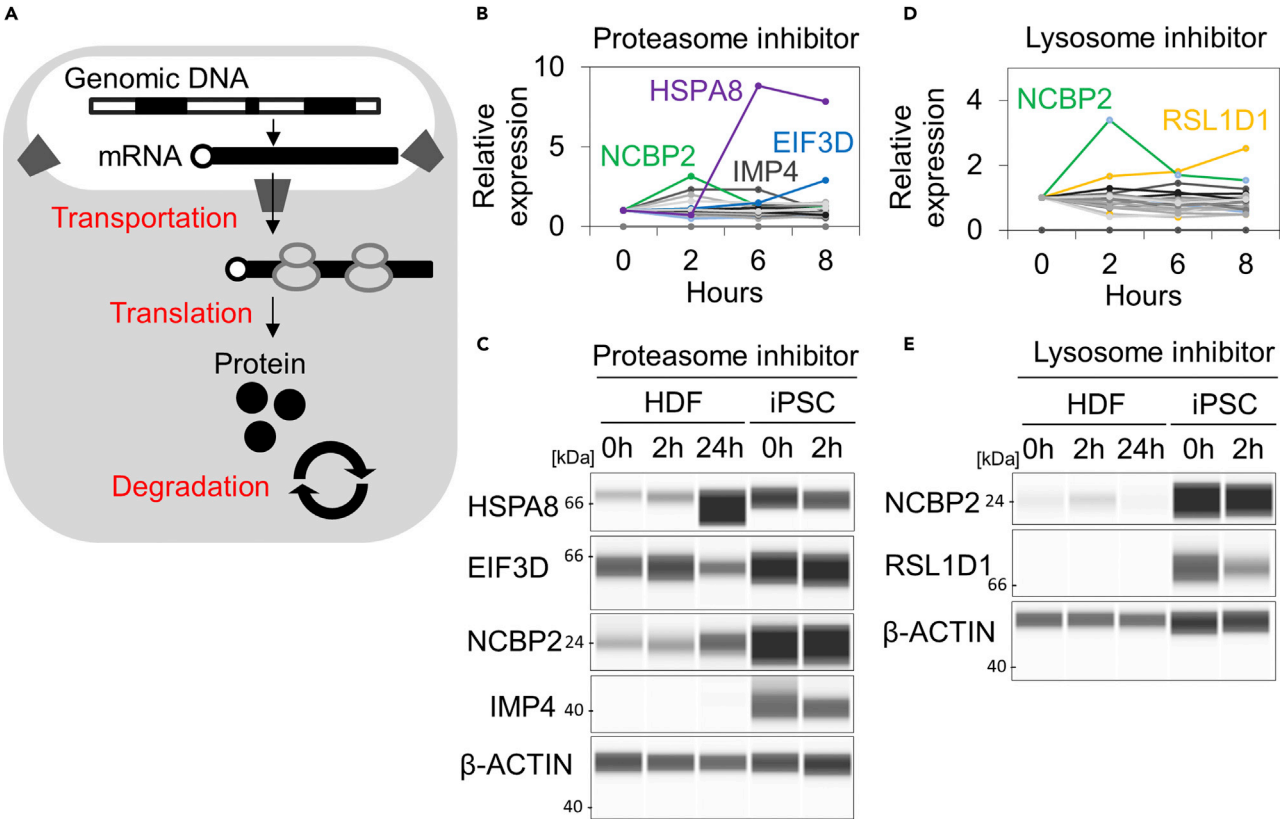


Figure 4. HSPA8, EIF3D, and NCBP2 protein expressions are controlled at the ubiquitin-dependent degradation stage

(A) Overview of the known regulation stages for gene expression after transcription.

(B) HDF-2 (Tig120) was treated with a proteasome inhibitor (20 μM MG-132) for up to 8 h, and the expression levels of the 20 essential iPSC/ESC-uPRA genes were measured by an antibody-based quantitative protein analysis. The relative expression of each protein compared with preinhibition is shown. All WES protein images are shown in Figure S6E.

(C). HDF-2 (Tig120) and iPSC-1 (201B7) were treated with 20 μM MG-132 for up to 24 h, and the protein expressions for HSPA8, EIF3D, NCBP2, and IMP4 were measured by an antibody-based quantitative protein analysis.

(D) HDF-2 (Tig120) was treated with lysosome inhibitors (250 nM Bafilomycin A1 and 500 nM Wortmannin) for up to 8 h, and the protein expression of the 20 essential iPSC/ESC-uPRA genes were measured by an antibody-based quantitative protein analysis. The relative expression of each protein to preinhibition is shown. All WES protein images are shown in Figure S6F.

(E) HDF-2 (Tig120) and iPSC-1 (201B7) were treated with 250 nM Bafilomycin and 500 nM Wortmannin for up to 24 h, and the protein expressions of NCBP2 and RSL1D1 were measured by an antibody-based quantitative protein analysis.

expressions of RNA export factors. We found more than a 4-fold increase in the protein expressions of NXT1 and NXF1, two mRNA export receptor heterodimers, in PSCs and NPC 1418E1 compared with the other differentiated cells (Figures S7D–S7F). These data indicated that most iPSC/ESC-uPRA mRNA was retained in the nucleus of HDFs, but the increased NXT1 and NXF1 protein expressions may accelerate mRNA transport to the cytoplasm in iPSCs/ESCs.

Next, to investigate the translation efficiency, we extracted mRNAs in the monosome (single ribosomes), light polysome (multiple ribosomes on mRNA), and heavy polysome fractions (more than four ribosomes on mRNA) of iPSCs and HDFs using sucrose fractionation (Figure 5D). Polysomes are generally known as representatives of translationally active ribosome populations, whereas monosomes are considered inactive ribosome populations (Warner and Knopf, 2002). We quantified and compared the iPSC/ESC-uPRA mRNA percentage (Figure 5E) and ratio (Figure 5F) in each fraction between iPSCs and HDFs. We found that POLD2, CPSF6, RBM22, BUD31, and IMP4 had a relatively high percentage of RNA in the monosome fraction in iPSCs. Recently, synaptic transcripts have been shown to be actively translated in monosome fractions (Biever et al., 2020). Thus, these five identified uPRA genes might be actively translated to achieve increased protein levels in iPSCs. In addition, NIFK and RSL1D1 were enriched in the polysome fraction,

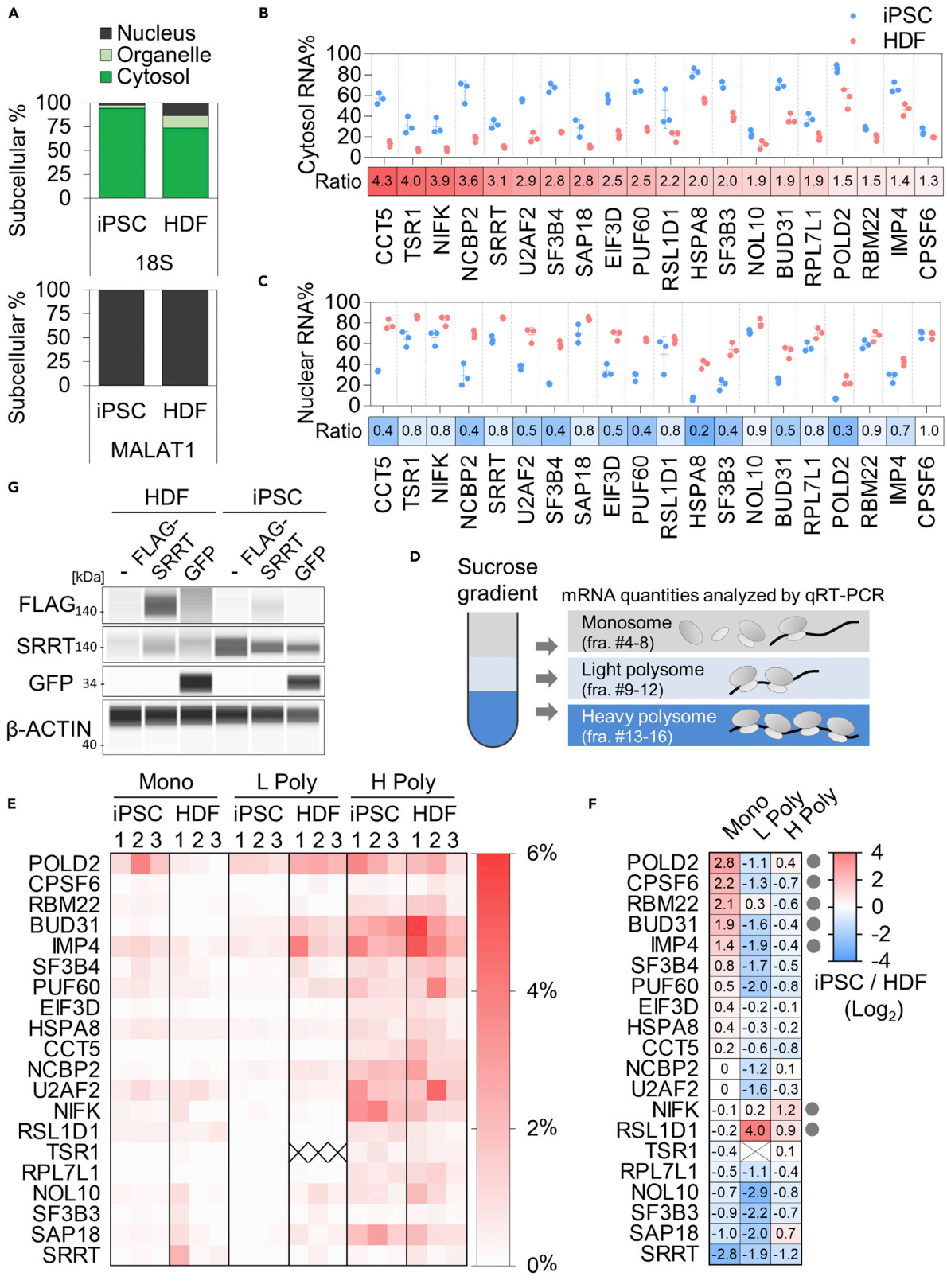

A
Nucleus
Organelle
Cytosol
Subcellular %
100
75
50
25
0
iPSC HDF
18S
MALAT1
B
iPSC
HDF
Cytosol RNA%
Ratio
4.3 4.0 3.9 3.6 3.1 2.9 2.8 2.8 2.5 2.5 2.2 2.0 2.0 1.9 1.9 1.9 1.5 1.5 1.4 1.3
C
Nuclear RNA%
0.4 0.8 0.8 0.4 0.8 0.5 0.4 0.8 0.5 0.4 0.8 0.2 0.4 0.9 0.5 0.8 0.3 0.9 0.7 1.0
CCT5 TSR1 NIFK NCBP2 SRRT U2AF2 SF3B4 SAP18 EIF3D PUF60 RSL1D1 HSPA8 SF3B3 NOL10 BUD31 RPL7L1 POLD2 RBM22 IMP4 CPSF6
G
HDF
iPSC
- FLAG-SRRT GFP
[kDa]
FLAG140
SRRT140
GFP 34
β-ACTIN 40
D
Sucrose gradient
mRNA quantities analyzed by qRT-PCR
Monosome (fra. #4-8)
Light polysome (fra. #9-12)
Heavy polysome (fra. #13-16)
E
Mono
L Poly
H Poly
iPSC HDF
1 2 3
6%
4%
2%
0%
F
Mono L Poly H Poly
POLD2 2.8 -1.1 0.4
CPSF6 2.2 -1.3 -0.7
RBM22 2.1 0.3 -0.6
BUD31 1.9 -1.6 -0.4
IMP4 1.4 -1.9 -0.4
SF3B4 0.8 -1.7 -0.5
PUF60 0.5 -2.0 -0.8
EIF3D 0.4 -0.2 -0.1
HSPA8 0.4 -0.3 -0.2
CCT5 0.2 -0.6 -0.8
NCBP2 0 -1.2 0.1
U2AF2 0 -1.6 -0.3
NIFK -0.1 0.2 1.2
RSL1D1 -0.2 4.0 0.9
TSR1 -0.4 0.1
RPL7L1 -0.5 -1.1 -0.4
NOL10 -0.7 -2.9 -0.8
SF3B3 -0.9 -2.2 -0.7
SAP18 -1.0 -2.0 0.7
SRRT -2.8 -1.9 -1.2
4
2
0
-2
-4
iPSC / HDF
(Log2)

Figure 5. Cytosolic mRNA localization is increased in iPSCs and translation of iPSC/ESC-uPRA mRNA is inhibited in HDFs

(A) The percentages of mRNA in the cytoplasm, organelle, and nucleus of iPSC-1 (201B7) and HDF-2 (Tig120) are shown for 18S rRNA and MALAT1 lncRNA measured by qRT-PCR.

(B) The percentage of mRNA in the cytoplasm measured by qRT-PCR for the 20 essential iPSC/ESC-uPRA genes. The cytoplasm mRNA expression ratio in iPSCs to HDFs is shown above each gene name in the heatmap. Data are represented as mean ± SD.

(C) The percentage of mRNA in the nucleus measured by qRT-PCR for the 20 essential iPSC/ESC-uPRA genes. The nucleus mRNA expression ratio in iPSCs to HDFs is shown above each gene name in the heatmap. Data are represented as mean ± SD.

(D) Overview of the analysis of the translation efficiency by the sucrose gradient. The monosome (fraction 4–8; single ribosome), light polysome (fraction 9–12; multiple ribosomes), and heavy polysome (fraction 13–16; more than four ribosomes) were collected to extract RNAs, which were measured by qRT-PCR. See "Monosome and polysome fractionation" in STAR Methods for details.

(E) The percentage of mRNAs in monosome (Mono), light polysome (Light Poly), and heavy polysome (Heavy Poly) were analyzed by qRT-PCR (n = 3). Values are normalized by spike RNA and compared with the loading sample before the sucrose gradient.

(F) The average mRNA expression ratio in iPSC-1 (201B7) to HDF-2 (201B7) is shown in the heatmap with a $\log_2$ scale using the values in E (n = 3). Gray circles beside the heatmap indicates more than 2-fold enriched uPRA-mRNA in monosomes, light and heavy polysomes in iPSCs.

(G) mRNA transfection experiment. "-" indicates the transfection experiment without mRNA. "FLAG-SRRT" and "GFP" indicate the transfection of full-length FLAG-SRRT mRNA and GFP mRNA as transfection control, respectively. These mRNA were transfected into HDF-2 (Tig120) and iPSC-1 (201B7). Antibody-based quantitative protein analysis assays one day after the transfection of FLAG, SRRT, GFP, and b-ACTIN are shown.

indicating a positive regulation of protein levels at the translational level. However, most uPRA genes were enriched in the light polysome fractions in HDFs. These results indicate that there is a translational repressive effect on HDFs even for mRNA with multiple ribosomes.

To explore the cause of this repressive effect, we picked up one uPRA gene, SRRT, because SRRT mRNA was the most associated with heavy polysomes in HDFs despite the inhibited protein expression among iPSC/ESC-uPRA mRNAs (Figure 5F). Because a lower cytosolic RNA percentage of endogenous SRRT was observed in HDFs (around 10%; Figure 5B), we next analyzed whether its translation occurs if the cytoplasmic RNA content is increased. To this end, we constructed full-length FLAG-tagged SRRT and GFP mRNA and measured the protein levels by an antibody-based quantitative protein analysis one day after the mRNA transfection (Figure 5G and Table S8). We confirmed that the transfection efficiency is similar between iPSCs and HDFs by the GFP protein expression. As a result, FLAG-SRRT protein was observed in HDFs, suggesting a sensitive regulation of mRNA levels in the cytoplasm may contribute to the protein levels. In summary, these results indicate an inhibitory mechanism in HDFs for transporting iPSC/ESC-uPRA mRNAs to the cytoplasm and translating iPSC/ESC-uPRA genes even after mRNA binds to ribosomes.

DISCUSSION

Despite their ubiquitous effects, there is far less understanding about the posttranscriptional regulatory mechanisms compared with transcriptional regulatory mechanisms in iPSCs/ESCs. Here, we revealed that 20 posttranscriptionally regulated genes (i.e., uPRA genes) are essential for the maintenance of iPSCs. These genes had upregulated protein levels without any significant positive upregulation in mRNA levels in iPSCs (Figures 3C–3E). The regulatory processes of these 20 essential iPSC/ESC-uPRA genes are illustrated in Figure 6. The proteins of HSPA8, EIF3D, and NCBP2 were quickly degraded in HDFs at the proteasome degradation stage (Figures 4B and 4C). Eighteen iPSC/ESC-uPRA mRNAs were transported from the nucleus to the cytoplasm more in iPSCs/ESCs than in HDFs (Figure 5B), and high nuclear RNA localization was shown globally in HDFs (Figure S7C). The translation efficiency of POLD2, CPSF6, RBM22, BUD31, and IMP4 was upregulated in monosomes, and the translation of RSL1D1 and NIFK was upregulated in polysomes in iPSCs (Figures 5E and 5F). However, we found that most uPRA genes were enriched in the light polysome fraction in HDFs (Figures 5E and 5F), suggesting several inhibitory effects for uPRA mRNA on translation after ribosome binding in differentiated cells. To identify this inhibitory effect, we picked up one uPRA gene, SRRT, for further analysis. The percentage of cytoplasmic SRRT mRNA was 3-fold less in HDFs (Figure 5B), and the mRNA itself was more associated with monosome, light and heavy polysomes in HDFs than in iPSCs (Figure 5F). Exogenous cytosolic SRRT mRNA was translated in HDFs (Figure 5G), suggesting the cytoplasmic mRNA level might regulate proper protein expression. Another possible mechanism for the translation inhibition in differentiated cells is ribosome pausing, in which mRNA binds to ribosomes that are not active or select noncanonical start codons (Chandrasekaran et al., 2019; Darnell et al., 2018; Tresenrider et al., 2021). A third possibility is that aberrant proteins are degraded immediately after the production by a quality control pathway of the mRNA and ribosome (D'Orazio and Green, 2021). More study is needed to clarify which mechanism applies to which gene.

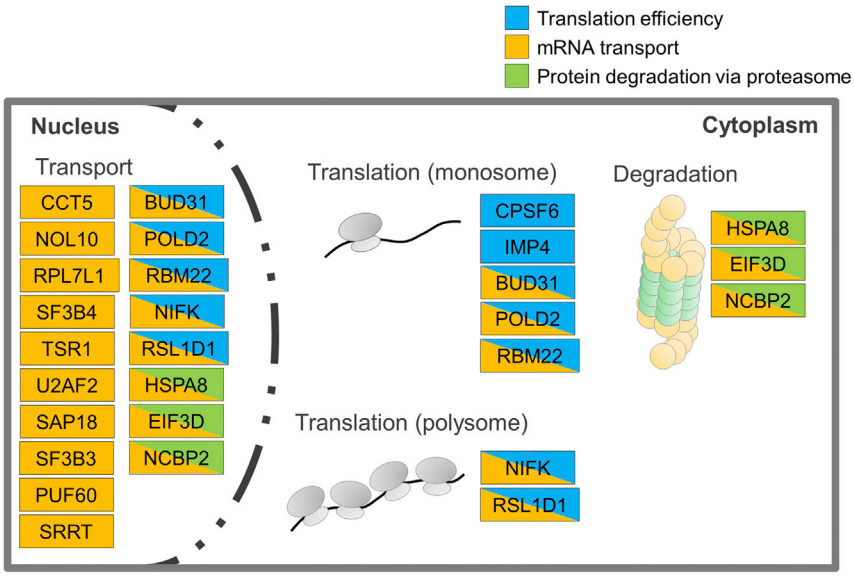


Figure 6. An illustration of how 20 essential iPSC/ESC-uPRA genes are regulated in iPSCs and differentiated cells

Blue, yellow, and green indicate the translation efficiency, mRNA transport, and protein degradation via the proteasome of the 20 essential iPSC/ESC-uPRA genes, respectively. More mRNA of 18 uPRA genes was transported from the nucleus to the cytoplasm in iPSCs than in HDFs. The proteins of HSPA8, EIF3D, and NCBP2 were quickly degraded in HDFs at the proteasome degradation stage. The translation efficiency of CPSF6, IMP4, BUD31, POLD2, and RBM22 was higher in the monosome fraction in iPSCs. The translation efficiency of NIFK and RSL1D1 was higher in the polysome fraction in iPSCs.

All identified 20 iPSC/ESC-uPRA proteins in the present study are involved in a wide range of cellular functions, from transcription to posttranslation. For transcription-related functions, 10 iPSC/ESC-uPRA genes have polymerase-, splicing- and mRNA-maturation-related functions: POLD2 is a component of the DNA polymerase complex; SF3B3, SF3B4, and RBM22 are components of the splicing factor; and PUF60, BUD31 SAP18, CPSF6, U2AF2, and SRRT are required for the splicing of pre-mRNA, with SRRT especially known as "a molecular guardian of the pluripotent cell state" by facilitating proper splicing in iPSCs (Kainov and Makeyev, 2020). For translation-related functions, eight iPSC/ESC-uPRA genes are involved in functions of the ribosome complex: NCBP2 and EIF3D are mRNA cap-binding proteins; EIF3D is necessary for specialized translation initiation factors (Lee et al., 2016); RPL7L1 is a putative component of the ribosome complex; and RSL1D1, NIFK, NOL10, IMP4, and TSR1 are required for rRNA processing. For posttranslational functions, CCT5 and HSPA8 are components of the chaperone complex, which helps newly synthesized proteins properly fold. Our observation that these genes are posttranscriptionally regulated is particularly interesting, because they all function as regulators post-transcription. We found that the motif types of uPRA genes are quite diverse by *in silico* analysis (Table S7), and several different combinations of RNA-binding proteins might be involved in the maintenance of iPSCs. For future screening of the binding motif of individual RNA in each cell type, cross-linking and immunoprecipitation (CLIP) (Ule et al., 2003) or folded RNA element profiling with structure library (FOREST) (Komatsu et al., 2020) should be considered.

In conclusion, our study revealed that the protein expressions of 228 genes are posttranscriptionally regulated in iPSCs. Moreover, the protein levels of 20 of these genes are specifically increased in iPSCs, and the proteins themselves are essential for iPSC maintenance. Finally, although the translational repression mechanisms of uPRA genes in HDFs still need further investigation, we showed that different mechanisms regulate the posttranscriptional gene expressions across cell types.

Limitations of the study

We focused on posttranscriptionally regulated genes essential for the maintenance of human iPSCs. We carefully selected iPSC lines that closely resemble ESCs to classify the two cell types into one group (iPSCs/ESCs). We observed that similar genes were equally variable among iPSCs and ESCs, but the

iPSC origin and lab culture protocols might impact differences between ESCs and iPSCs. Secondly, although we showed critical regulatory steps for controlling uPRA protein levels, a wide variety of regulatory mechanisms seem to be involved. Recent studies have shown that the secondary structure of mRNA is important for regulating protein expression by changing the interaction with the expansion segment of rRNA(Leppek et al., 2020) and changing the mRNA half-life (Mauger et al., 2019). Future studies are needed to determine if secondary structures affect the translation of the 20 iPSC/ESC-uPRA genes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Cell culture
  - Protein extraction
  - Gene expression analysis by antibody-based protein quantification
  - Gene expression analysis by qRT-PCR and microarray
  - Subcellular fractionation of mRNA and proteins
  - Gene expression analysis by RNA-sequencing
  - Gene expression analysis by nano liquid chromatography (nanoLC)-mass spectrometry (MS)
  - *Trans*-omics data analysis
  - siRNA screening
  - Proteasome and lysosome inhibitor assay
  - Monosome and polysome fractionation
  - mRNA transfection
  - *In silico* analysis of motifs in cDNA sequences
  - Data availability
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104289.

ACKNOWLEDGMENTS

We would like to thank Y. Fujita and T. Matsushita for the technical assistance, P. Karagiannis for editorial assistance, and S.Takeshima for administrative support. This work was supported by K-CONNEX from the Japan Science and Technology Agency (JST), Core Center for iPS Cell Research from Japan Agency for Medical Research and Development (AMED; JP21bm0104001), AMED-PRIME (21gm6410003h0001(M.I.)), the Japanese Society for the Promotion of Science KAKENHI Grant (JSPS; 19K16104(M.I.)), and the iPS Cell Research Fund.

AUTHOR CONTRIBUTIONS

M.I. and Y.K. performed most of the experiments. T.T generated the program to quantify the MS data. M.Narita performed the microarray. M.Nakagawa., Y.N., and A.O. helped perform the siRNA screening analysis. T.Y., H.S., C.O., M.N., and T.Y. performed the mRNA-related experiments. M.I., K.T., and S.Y. designed and directed the research. M.I. analyzed the data and wrote the manuscript with editing by all authors.

DECLARATION OF INTERESTS

M.I. is a scientific adviser (without salary) of xFOREST Therapeutics. H.S. is an outside director of aceRNA Technologies, Co., Ltd. K.T. is on the scientific advisory board (without salary) of I Peace, Inc. S.Y. is a scientific advisor (without salary) of iPS Academia Japan. The authors declare no competing interests.

Received: October 8, 2021
Revised: April 1, 2022
Accepted: April 20, 2022
Published: May 20, 2022

REFERENCES

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175. https://doi.org/10.1038/nrg3662.

Biever, A., Glock, C., Tushev, G., Ciirdaeva, E., Dalmay, T., Langer, J.D., and Schuman, E.M. (2020). Monosomes actively translate synaptic mRNAs in neuronal processes. Science *367*, eaay4991. https://doi.org/10.1126/science.aay4991.

Buccitelli, C., and Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene expression control. Nat. Rev. Genet. *21*, 630–644. https://doi.org/10.1038/s41576-020-0258-4.

Carlevaro-Fita, J., and Johnson, R. (2019). Global positioning system: understanding long noncoding RNAs through subcellular localization. Mol. Cell *73*, 869–883. https://doi.org/10.1016/j.molcel.2019.02.008.

Chandrasekaran, V., Juszkiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. Nat. Struct. Mol. Biol. *26*, 1132–1140. https://doi.org/10.1038/s41594-019-0331-x.

Chen, Q., and Hu, G. (2017). Post-transcriptional regulation of the pluripotent state. Curr. Opin. Genet. Dev. *46*, 15–23. https://doi.org/10.1016/j.gde.2017.06.010.

Chen, Y.A., Tripathi, L.P., Fujiwara, T., Kameyama, T., Itoh, M.N., and Mizuguchi, K. (2019). The TargetMine data warehouse: enhancement and updates. Front Genet. *10*, 934. https://doi.org/10.3389/fgene.2019.00934.

Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S., et al. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature *468*, 316–320. https://doi.org/10.1038/nature09531.

D'Orazio, K.N., and Green, R. (2021). Ribosome states signal RNA quality control. Mol. Cell *81*, 1372–1383. https://doi.org/10.1016/j.molcel.2021.02.022.

Darnell, A.M., Subramaniam, A.R., and O'Shea, E.K. (2018). Translational control through differential ribosome pausing during amino acid limitation in mammalian cells. Mol. Cell *71*, 229–243.e11. https://doi.org/10.1016/j.molcel.2018.06.041.

Dendooven, T., Luisi, B.F., and Bandyra, K.J. (2020). RNA lifetime control, from stereochemistry to gene expression. Curr. Opin. Struct. Biol. *61*, 59–70. https://doi.org/10.1016/j.sbi.2019.10.002.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O., et al. (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell *147*, 132–146. https://doi.org/10.1016/j.cell.2011.08.023.

Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., et al. (2013). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature *498*, 241–245. https://doi.org/10.1038/nature12270.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat. Rev. Mol. Cell Biol. *19*, 327–341. https://doi.org/10.1038/nrm.2017.130.

Howe, K.L., Achuthan, P., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., et al. (2021). Ensembl 2021. Nucleic Acids Res. *49*, D884–D891. https://doi.org/10.1093/nar/gkaa942.

Iwasaki, M., Tabata, T., Kawahara, Y., Ishihama, Y., and Nakagawa, M. (2019). Removal of interference MS/MS spectra for accurate quantification in isobaric tag-based proteomics. J. Proteome Res. *18*, 2535–2544. https://doi.org/10.1021/acs.jproteome.9b00078.

Kainov, Y.A., and Makeyev, E.V. (2020). A transcriptome-wide antitermination mechanism sustaining identity of embryonic stem cells. Nat. Commun. *11*, 361. https://doi.org/10.1038/s41467-019-14204-z.

Khan, Z., Ford, M.J., Cusanovich, D.A., Mitrano, A., Pritchard, J.K., and Gilad, Y. (2013). Primate transcript and protein expression levels evolve under compensatory selection pressures. Science *342*, 1100–1104. https://doi.org/10.1126/science.1242379.

Komatsu, K.R., Taya, T., Matsumoto, S., Miyashita, E., Kashida, S., and Saito, H. (2020). RNA structure-wide discovery of functional interactions with multiplexed RNA motif library. Nat. Commun. *11*, 6275. https://doi.org/10.1038/s41467-020-19699-5.

Koyanagi-Aoi, M., Ohnuki, M., Takahashi, K., Okita, K., Noma, H., Sawamura, Y., Teramoto, I., Narita, M., Sato, Y., Ichisaka, T., et al. (2013). Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. Proc. Natl. Acad. Sci. U S A. *110*, 20569–20574. https://doi.org/10.1073/pnas.1319061110.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

Lee, A.S.Y., Kranzusch, P.J., Doudna, J.A., and Cate, J.H.D. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. Nature *536*, 96–99. https://doi.org/10.1038/nature18954.

Leppek, K., Fujii, K., Quade, N., Susanto, T.T., Boehringer, D., Lenarčič, T., Xue, S., Genuth, N.R., Ban, N., and Barna, M. (2020). Gene- and species-specific hox mRNA translation by ribosome expansion segments. Mol. Cell *80*, 980–995.e13. https://doi.org/10.1016/j.molcel.2020.10.023.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Lu, R., Markowetz, F., Unwin, R.D., Leek, J.T., Airoldi, E.M., MacArthur, B.D., Lachmann, A., Rozov, R., Ma'ayan, A., Boyer, L.A., et al. (2009). Systems-level dynamic analyses of fate change in murine embryonic stem cells. Nature *462*, 358–362. https://doi.org/10.1038/nature08575.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet. J. *17*, 10. https://doi.org/10.14806/ej.17.1.200.

Matsumoto, M., Matsuzaki, F., Oshikawa, K., Goshima, N., Mori, M., Kawamura, Y., Ogawa, K., Fukuda, E., Nakatsumi, H., Natsume, T., et al. (2017). A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nat. Methods *14*, 251–258. https://doi.org/10.1038/nmeth.4116.

Mauger, D.M., Cabral, B.J., Presnyak, V., Su, S.V., Reid, D.W., Goodman, B., Link, K., Khatwani, N., Reynders, J., Moore, M.J., and McFadyen, I.J. (2019). mRNA structure regulates protein expression through changes in functional half-life. Proc. Natl. Acad. Sci. U S A. *116*, 24075–24083. https://doi.org/10.1073/pnas.1908052116.

Nakagawa, M., Taniguchi, Y., Senda, S., Takizawa, N., Ichisaka, T., Asano, K., Morizane, A., Doi, D., Takahashi, J., Nishizawa, M., et al. (2014). A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. Sci. Rep. *4*, 3594. https://doi.org/10.1038/srep03594.

Ohta, S., Nishida, E., Yamanaka, S., and Yamamoto, T. (2013). Global splicing pattern reversion during somatic cell reprogramming. Cell Rep. *5*, 357–366. https://doi.org/10.1016/j.celrep.2013.09.016.

Okuda, S., Watanabe, Y., Moriya, Y., Kawano, S., Yamamoto, T., Matsumoto, M., Takami, T., Kobayashi, D., Araki, N., Yoshizawa, A.C., et al. (2017). jPOSTrepo: an international standard data repository for proteomes. Nucleic Acids Res. *45*, D1107–D1111. https://doi.org/10.1093/nar/gkw1080.

Palazzo, A.F., and Lee, E.S. (2018). Sequence determinants for nuclear retention and cytoplasmic export of mRNAs and lncRNAs. Front. Genet. *9*, 440. https://doi.org/10.3389/fgene.2018.00440.

Paz, I., Kosti, I., Ares, M., Cline, M., and Mandel-Gutfreund, Y. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res. *42*, W361–W367. https://doi.org/10.1093/nar/gku406.

Pelletier, J., Thomas, G., and Volarević, S. (2018). Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat. Rev. Cancer *18*, 51–63. https://doi.org/10.1038/nrc.2017.104.

Pohl, C., and Dikic, I. (2019). Cellular quality control by the ubiquitin-proteasome system and autophagy. Science *366*, 818–822. https://doi.org/10.1126/science.aax3769.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem. *75*, 663–670. https://doi.org/10.1021/ac026117i.

Roumeliotis, T.I., Williams, S.P., Gonçalves, E., Alsinet, C., Del Castillo Velasco-Herrera, M., Aben, N., Ghavidel, F.Z., Michaut, M., Schubert, M., Price, S., et al. (2017). Genomic determinants of protein abundance variation in colorectal cancer cells. Cell Rep. *20*, 2201–2214. https://doi.org/10.1016/j.celrep.2017.08.010.

Salomonis, N., Schlieve, C.R., Pereira, L., Wahlquist, C., Colas, A., Zambon, A.C., Vranizan, K., Spindler, M.J., Pico, A.R., Cline, M.S., et al. (2010). Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. Proc. Natl. Acad. Sci. U S A. *107*, 10514–10519. https://doi.org/10.1073/pnas.0912260107.

Takahashi, K., Jeong, D., Wang, S., Narita, M., Jin, X., Iwasaki, M., Perli, S.D., Conklin, B.R., and Yamanaka, S. (2020). Critical roles of translation initiation and RNA uridylation in endogenous retroviral expression and neural differentiation in pluripotent stem cells. Cell Rep. *31*, 107715. https://doi.org/10.1016/j.celrep.2020.107715.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872. https://doi.org/10.1016/j.cell.2007.11.019.

Teixeira, F.K., and Lehmann, R. (2019). Translational control during developmental transitions. Cold Spring Harb Perspect. Biol. *11*, a032987. https://doi.org/10.1101/cshperspect.a032987.

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. https://doi.org/10.1126/science.282.5391.1145.

Tresenrider, A., Morse, K., Jorgensen, V., Chia, M., Liao, H., van Werven, F.J., and Ünal, E. (2021). Integrated genomic analysis reveals key features of long undecoded transcript isoform-based gene repression. Mol. Cell *81*, 2231–2245.e11. https://doi.org/10.1016/j.molcel.2021.03.013.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science *302*, 1212–1215. https://doi.org/10.1126/science.1090095.

Umemura, Y., Koike, N., Ohashi, M., Tsuchiya, Y., Meng, Q.J., Minami, Y., Hara, M., Hisatomi, M., and Yagita, K. (2017). Involvement of posttranscriptional regulation of Clock in the emergence of circadian clock oscillation during mouse development. Proc. Natl. Acad. Sci. U S A. *114*, E7479–E7488. https://doi.org/10.1073/pnas.1703170114.

Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. *15*, e8503. https://doi.org/10.15252/msb.20188503.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28* (16), 2184–2185. https://doi.org/10.1093/bioinformatics/bts356.

Warner, J.R., and Knopf, P.M. (2002). The discovery of polyribosomes. Trends Biochem. Sci. *27*, 376–380. https://doi.org/10.1016/s0968-0004(02)02126-6.

Yamanaka, S. (2012). Induced pluripotent stem cells: past, present, and future. Cell Stem Cell *10*, 678–684. https://doi.org/10.1016/j.stem.2012.05.005.

Zuckerman, B., Ron, M., Mikl, M., Segal, E., and Ulitsky, I. (2020). Gene architecture and sequence composition underpin selective dependency of nuclear export of long RNAs on NXF1 and the TREX complex. Mol. Cell *79*, 251–267.e6. https://doi.org/10.1016/j.molcel.2020.05.013.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies (dilution ratio used in this study)		
Rabbit polyclonal anti-BUD31 (1:100)	Proteintech	Cat# 11798-1-AP RRID:AB_2274894
Rabbit monoclonal anti-CCT5 (TCP1e) (1:500)	Abcam	Cat# ab129016 RRID:AB_11154964
Rabbit polyclonal anti-CPSF6 (1:100)	Proteintech	Cat# 15489-1-AP RRID:AB_10694140
Rabbit polyclonal anti-EIF3D (1:250)	Proteintech	Cat# 10219-1-AP RRID:AB_2096880
Rabbit monoclonal anti-HSPA8 (D12F2) (1:200)	Cell Signaling Technology	Cat# 8444 RRID:AB_10831837
Rabbit polyclonal anti-IMP4 (1:500)	Proteintech	Cat# 16205-1-AP RRID:AB_2239858
Rabbit polyclonal anti-NCBP2 (CBP20) (1:250)	Proteintech	Cat# 11950-1-AP RRID:AB_2150142
Rabbit polyclonal anti-NIFK (MKI67IP) (1:500)	Proteintech	Cat# 12615-1-AP RRID:AB_2142384
Rabbit monoclonal anti-NOL10 (1:500)	Abcam	Cat# ab181161
Rabbit polyclonal anti-POLD2 (1:20)	Proteintech	Cat# 10288-1-AP RRID:AB_2284041
Goat polyclonal anti-PUF60 (1:200)	Abcam	Cat# ab22819 RRID:AB_777559
Rabbit polyclonal anti-RBM22 (1:100)	Abcam	Cat# ab157105
Rabbit polyclonal anti-RPL7L1 (1:100)	Proteintech	Cat# 16707-1-AP RRID:AB_1851609
Rabbit monoclonal anti-RSL1D1 (1:2000)	Abcam	Cat# ab181100
Rabbit polyclonal anti-SAP18 (1:100)	Proteintech	Cat# 13841-1-AP RRID:AB_2301353
Rabbit polyclonal anti-SF3B3 (1:1000)	Proteintech	Cat# 14577-1-AP RRID:AB_2270189
Rabbit polyclonal anti-SF3B4 (1:100)	Proteintech	Cat# 10482-1-AP RRID:AB_2301639
Rabbit polyclonal anti-SRRT (ARS2) (1:100)	Abcam	Cat# ab220991
Rabbit polyclonal anti-TSR1 (1:500)	Proteintech	Cat# 16887-1-AP RRID:AB_2222589
Rabbit polyclonal anti-U2AF65 (U2AF2) (1:50)	Proteintech	Cat# 15624-1-AP RRID:AB_2211330
Rabbit monoclonal anti-GAPDH (14C10) (1:500)	Cell Signaling Technology	Cat# 2118 RRID:AB_561053
Mouse monoclonal anti-beta-actin (1:250)	Sigma-Aldrich	Cat# A5441 RRID:AB_476744
Rabbit polyclonal anti-AFP (1:1000)	Proteintech	Cat# 14550-1-AP RRID:AB_2223933
Rabbit polyclonal anti-MYO1E (1:50)	Proteintech	Cat# 17768-1-AP RRID:AB_2251045

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse monoclonal anti-FLAG (M2) (1:40)	Sigma-Aldrich	Cat# F1804 RRID:AB_262044
Rabbit polyclonal anti-Ubiquitin (1:50)	Abcam	Cat# ab19247 RRID:AB_444805
Rabbit polyclonal anti-LC3B (1:10)	Novus Bio	Cat# NB100-2220SS RRID:AB_791015
Mouse monoclonal anti-GFP (GF200) (1:100)	Nacalai Tesque	Cat# 04363-24
Mouse monoclonal anti-OCT3/4 (1:500)	BD Biosciences	Cat# BD611202
Rabbit monoclonal anti-NXF1 (1:20)	Abcam	Cat# ab129160 RRID:AB_11142853
Mouse monoclonal anti-NXT1 (1:20)	Proteintech	Cat# 67680-1-Ig RRID:AB_2882873
Mouse polyclonal Alexa488 (1:300)	Thermo Fisher Scientific	Cat# A11001 RRID:AB_2534069
Rabbit polyclonal Cyanine3 (1:300)	Thermo Fisher Scientific	Cat# A10520 RRID:AB_2534029
Rabbit monoclonal anti-LaminB2 (D8P3U) (1:500)	Cell Signaling Technology	Cat# 12255 RRID:AB_2797859
Goat anti-rabbit secondary HRP-conjugated (1:1)	Protein Simple	Cat# 042-206 RRID:AB_2860577
Goat anti-mouse secondary HRP-conjugated (1:1)	Protein Simple	Cat# 042-205 RRID:AB_2860576
Donkey anti-goat secondary HRP-conjugated (1:1)	Protein Simple	Cat# 043-491-2
Chemicals, peptides, and recombinant proteins		
SDB-XC Empore disc cartridge	3M	Cat# 2340
Sodium dodecyl sulfate (SDS)	Nacalai Tesque	Cat# 31606-75
Sodium deoxycholate (SDC)	WAKO	Cat# 190-08313
Sodium lauroyl sarcosinate (SLS)	WAKO	Cat# 192-10382
Sucrose	WAKO	Cat# 195-07925
Lys-C, Mass Spec Grade	Promega	Cat# VA1170
Sequencing Grade Modified Trypsin (Lyophilized)	Promega	Cat# V5111
Ethyl acetate	WAKO	Cat# 051-00356
Acetonitrile	WAKO	Cat# 018-19853
Acetic acid	WAKO	Cat# 018-20061
Methanol	WAKO	Cat# 134-14523
Trifluoroacetic acid (TFA)	WAKO	Cat# 204-02743
Dithiothreitol (DTT)	WAKO	Cat# 045-08974
Iodoacetamide (IAA)	WAKO	Cat# 095-02151
Ammonium bicarbonate	WAKO	Cat# 018-21742
Dimethyl sulfoxide (DMSO)	WAKO	Cat# 045-28335
Triethylammonium bicarbonate buffer	Sigma-Aldrich	Cat# T7408-100ML
1M Tris-HCl (pH 9.0)	Nippon gene	Cat# 314-90381
1M Tris-HCl (pH 8.0)	Nippon gene	Cat# 314-90065
1M Tris-HCl (pH 7.5)	Nippon gene	Cat# 316-90221
NaCl (5 M), RNase-free	Thermo Fisher Scientific	Cat# AM9760G
$MgCl_2$ (1 M)	Thermo Fisher Scientific	Cat# AM9530G

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Triton X-100 10%	Teknova	Cat# T1105
Ultrapure water	Kanto Chemical	Cat# 11307-79
Nuclease-Free Water (not DEPC-Treated)	Thermo Fisher Scientific	Cat# AM9937
Y-27632	Sigma-Aldrich	Cat# Y0503
TrypLE select	Thermo Fisher Scientific	Cat# 12563011
Trypsin-EDTA (0.25%), phenol red	Thermo Fisher Scientific	Cat# 25200056
Laminin-511 E8 (iMatrix-511)	Nippi	Cat# 892012
StemFiT AK03N	Ajinomoto	Cat# AK03N
STEMdiff Neural Progenitor Medium	Stem Cell Technologies	Cat# 05833
Dulbecco's Modified Eagle Medium (DMEM)	Nacalai Tesque	Cat# 08459-35
Fetal Bovine Serum, qualified, New Zealand	Thermo Fisher Scientific	Cat# 10091148
BEGM Bronchial Epithelial Cell Growth Medium BulletKit	Lonza	Cat# CC-3170
PrEGM Prostate Epithelial Cell Growth Medium BulletKit	Lonza	Cat# CC-3166
MesenPRO RS Medium	Thermo Fisher Scientific	Cat# 12746012
STEMdiff Neural Progenitor Medium	STEMCELL Technologies	Cat# ST-05833
ReagentPack Subculture Reagents	Lonza	Cat# CC-5034
InSolution MG-132	Calbiochem	Cat# 474791
Bafilomycin A1	InvivoGen	Cat# tlrl-baf1
InSolution WORTMANNIN	Calbiochem	Cat# 681676
TurboDNase (2 U/μL)	Thermo Fisher Scientific	Cat# AM2238
SUPERaseIn	Thermo Fisher Scientific	Cat# AM2694
Cycloheximide solution, 100 mg/mL in DMSO	Sigma-Aldrich	Cat# C4859-1ML
Protease Inhibitor	Sigma-Aldrich	Cat# P8340-1ML
cOmplete Mini, EDTA-free Protease Inhibitor Cocktail	Sigma-Aldrich	Cat# 11836170001
Phosphatase Inhibitor cocktail 2	Sigma-Aldrich	Cat# P5726-1ML
Phosphatase Inhibitor cocktail 3	Sigma-Aldrich	Cat# P0044-1ML
Critical commercial assays		
SurePrint G3 Human GE 8x60K v3	Agilent Technologies	Cat# G4858A
iTRAQ Reagents Multiplex Kit	Sciex	Cat# 4352135
Subcellular Protein Fractionation Kit for Cultured Cells	Thermo Fisher Scientific	Cat# 78840
BCA Protein Assay Kit	Thermo Fisher Scientific	Cat# 23225
Quantitative Fluorometric Peptide Assay	Thermo Fisher Scientific	Cat# 23290
Anti-Rabbit Detection Module for Jess, Wes, Peggy Sue or Sally Sue	Protein Simple	Cat# DM-001
Anti-Mouse Detection Module for Jess, Wes, Peggy Sue or Sally Sue	Protein Simple	Cat# DM-002
12-230 kDa Jess or Wes Separation Module, 8 x 25 capillary cartridges	Protein Simple	Cat# SM-W004
miRNeasy Mini Kit	QIAGEN	Cat# 217004
Cellstain - Hoechst 33342 solution	Dojindo	Cat# 346-07951
QIAzol lysis reagent	QIAGEN	Cat# 79306
Trizol LS reagent	Thermo Fisher Scientific	Cat# 10296028

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SuperScript III First-Strand Synthesis SuperMix for qRT-PCR	Thermo Fisher Scientific	Cat# 11752050
ReverTraAce	TOYOBO	Cat# TRT-101
In-Fusion HD Cloning Kit	Clontech Laboratories	Cat# 639648
TaqMan Gene Expression Master Mix	Thermo Fisher Scientific	Cat# 4369016
PowerUp SYBR Green Master Mix	Thermo Fisher Scientific	Cat# A25742
Amicon Ultra centrifugal filters (10K)	Merck Millipore	Cat# UFC501096
Lipofectamine Stem Reagent	Thermo Fisher Scientific	Cat# STEM00015
Lipofectamine MessengerMAX	Thermo Fisher Scientific	Cat# LMRNA001
Opti-MEM Reduced Serum Medium	Thermo Fisher Scientific	Cat# 31985062
Stemfect RNA Transfection Kit	STEMGENT	Cat# 00-0069
Agilent RNA6000 Pico Kit	Agilent	Cat# 5067-1513
Agilent High-Sensitivity DNA Kit	Agilent	Cat# 5067-4626
Qubit™ RNA High Sensitivity (HS) Kit	Thermo Fisher Scientific	Cat# Q32855
Qubit™ dsDNA High Sensitivity (HS) Kit	Thermo Fisher Scientific	Cat# Q32853
IDT for Illumina RNA UD Indexes Set A, Ligation	Illumina	Cat# 20040553
Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero Plus	illumina	Cat# 20040529
NextSeq 500/550 High Output v2 Kit	illumina	Cat# FC-404-2005
Deposited data		
Raw and analyzed data (Human gene expression microarray)	This study	GSE184546
Raw images	Mendeley	https://doi.org/10.17632/y6b3bgng9p.1
Raw and analyzed data (mass spectrometry)	jPOST	JPST001308 (PXD028489)
Raw and analyzed RNA-seq files	This study	GSE199820
Experimental models: Cell lines		
201B7 human iPSC line	(Takahashi, K. et al., 2007)	RRID:CVCL_A324
1418E1 human iPSC line	This study	N/A
HDF1388 Human dermal fibroblast	Purchased from Cell applications, Inc.	N/A
Tig120 Human dermal fibroblast	National Institutes of Biomedical Innovation, Health and Nutrition.	N/A
H9 embryonic stem cell	(Thomson, J. A. et al., 1998)	WA09 RRID:CVCL_9773
Normal human bronchial epithelial cells; NHBEC	Lonza Bioscience	Cat# CC-2541
Human prostate epithelial cells; PrEC	Lonza Bioscience	Cat# CC-2555
Human adipose tissue-derived mesenchymal stem cells; HAdMSC	Life Technologies	Cat# R7788115
Normal human epidermal keratinocytes; NHEK	Lonza	Cat# 00192627
H9 ESC-derived neural progenitor cells	Thermo Fisher Scientific	N7800200, RRID:CVCL_IU37
1418E1 iPSC-derived neural progenitor cells	This study	N/A
Oligonucleotides		
BUD31 (Hs00696974_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
CCT5 (Hs04362335_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
CPSF6 (Hs01101212_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
EIF3D (Hs01044815_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HSPA8 (Hs03044880_gH) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
IMP4 (Hs00369187_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
NCBP2 (Hs01597558_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
NIFK (Hs00757500_s1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
NOL10 (Hs01042161_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
POLD2 (Hs00371757_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
PUF60 (Hs01050525_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
RBM22 (Hs00216159_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
RPL7L1 (Hs02339924_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
RSL1D1 (Hs00378363_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
SAP18 (Hs00705532_s1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
SF3B3 (Hs00418633_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
SF3B4 (Hs00538859_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
SRRT (Hs00210818_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
TSR1 (Hs00250762_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
U2AF2 (Hs00200737_m1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
ACTB (Hs01060665_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
GAPDH (Hs02786624_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
18S (Hs99999901_s1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
MALAT1 (Hs00273907_s1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
GammaTub23C (Dm01841764_g1) TaqMan Assay	Thermo Fisher Scientific	Cat# 4331182
siRNAs, see Table S2	Dharmacon	N/A
DNA sequence for mRNA transfection, see Table S4	This study	N/A
Software and algorithms		
Excel 2016	Microsoft	https://www.office.com/
ProteinPilot v5.0	Sciex	https://sciex.com/products/software/proteinpilot-software
Mascot	Matrix Science	https://www.matrixscience.com/
Compass	Protein Simple	https://www.proteinsimple.com/compass/downloads/
HCS Studio Cell Analysis Software	Thermo Fisher Scientific	N/A
GeneSpring version 14.9.1	https://www.agilent.com/	Agilent Technologies
TargetMine	https://targetmine.mizuguchilab.org/targetmine/	Mizuguchi Laboratory https://doi.org/10.3389/fgene.2019.00934
Removal of interference mixture MS/MS spectra (RiMS) perl script	https://sourceforge.net/projects/rimsprogram/files/	Mio Iwasaki Laboratory (Iwasaki et al., 2019)
STAR Aligner (version 2.5.3a)	https://github.com/alexdobin/STAR	(Dobin et al., 2013)
bowtie 2 (version 2.2.5)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtmll	(Langmead and Salzberg, 2012)
cutadapt-1.12	http://gensoft.pasteur.fr/docs/cutadapt/1.18/index.html	(Martin, 2011)
SAM tools (version 1.10)	https://sourceforge.net/projects/samtools/files/samtools/1.7/	(Li et al., 2009)
RSeQC (version 4.0.0)	http://rseqc.sourceforge.net/	(Wang, L. et al., 2012)
RSEM (version 1.3.3)	https://github.com/deweylab/RSEM	(Li and Dewey, 2011)

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R (version 4.1)	https://www.r-project.org/	N/A
Perl 5 (version 24)	https://www.activestate.com/products/perl/	N/A
Other		
MonoCap C18 HighResolution 4000 (0.1 x 4000 mm)	GL Sciences	Cat# 5020-4000

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents may be directed to and will be fulfilled upon reasonable request by the Lead Contact author, Mio Iwasaki (omio@cira.kyoto-u.ac.jp).

Materials availability

Unique reagents generated in this study are available from the lead contact with a Materials Transfer Agreement.

Data and code availability

- Microarray, RNA-sequencing and proteome data are accessible in the Gene Expression Omnibus database of the National Center for Biotechnology Information website and in the Japan Proteome Standard Repository/Database, respectively. Accession numbers are listed in the key resources table. The WES images that were not shown in the paper have been deposited at Mendeley and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The human pluripotent stem cell (PSC) lines (iPSC-1, 201B7; and iPSC-2, 1418E1), one embryonic stem cell line (ESC; H9), two types of human dermal fibroblasts (HDF-1, HDF1388; and HDF-2, Tig120), endodermal cells (normal human bronchial epithelial cells, NHBEC; and human prostate epithelial cells, PrEC), mesodermal cells (human adipose tissue-derived mesenchymal stem cells, HAdMSC), ectodermal cells (normal human epidermal keratinocytes, NHEK), and human neural progenitor cells derived from H9 (NPC H9) and 1418E1 (NPC 1418E1) were cultured in humidified incubators at 37°C in 5% CO_2 and 20% O_2. All reagents were warmed in a water bath set at 23°C before use. Karyotypes of both iPSCs were verified as normal. We carefully selected the iPSC lines that closely resemble ESCs to classify the two cell types into one group: iPSCs/ESCs. In general, variations in the quality of iPSCs are larger than in ESCs (Yamanaka, 2012), but one of the iPSC clones used in our study (iPSC-1; 201B7) was shown to resemble ESCs (iPSC-1; 201B7, (Koyanagi-Aoi et al., 2013)). The other clone (iPSC-2; 1418A1) also has similar expression patterns to ESCs (Figure S1C). The experiments using H9 ESCs were conducted in conformity with "The Guidelines on the Distribution and Utilization of Human Embryonic Stem Cells" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

METHOD DETAILS

Cell culture

Two human induced pluripotent stem cell lines (iPSC-1, 201B7; and iPSC-2, 1418E1) and one embryonic stem cell line (ESC; H9) were maintained under feeder-free conditions on iMatrix-511 (Nippi) with StemFit AK03N medium (Ajinomoto). A Rock inhibitor (Y-27632, final concentration, 10 μM) was used only on the day of plating (Nakagawa et al., 2014). Two types of human dermal fibroblasts (HDF-1, HDF1388; and HDF-2, Tig120) were cultured in Dulbecco's modified Eagle's medium (DMEM; Nacalai Tesque) with 10% fetal bovine serum (FBS, Thermo Fisher Scientific) and 0.5% penicillin and streptomycin (gibco). Endodermal cells (normal human bronchial epithelial cells, NHBEC; and human prostate epithelial cells, PrEC), mesodermal cells (human adipose tissue-derived mesenchymal stem cells, HAdMSC), and

ectodermal cells (normal human epidermal keratinocytes, NHEK) were cultured according to the manufacturer's harvesting protocol (Lonza). Human neural progenitor cells derived from H9 (NPC H9) and 1418E1 (NPC 1418E1) were cultured in STEMdiff Neural Progenitor Medium according to the manufacturer's harvesting protocol (Veritas).

Protein extraction

The cells were washed once with ice-cold PBS (Nacalai Tesque) and directly lysed with ice-cold lysis buffer (PTS buffer: 12 mM SDC, 12 mM SLS, 100 mM Tris-HCl (pH 9.0)) or SDS buffer (1% SDS, 20 mM Tris-HCl (pH 8.0)), with 1% phosphatase and protease inhibitors (Sigma-Aldrich). Cell lysates were collected by scraping and pipetting on pre-chilled Protein LoBind 2 mL tubes (Eppendorf). After sonication and heat shock at 95°C for 5 min, the protein concentration was determined using a BCA protein assay kit (Thermo Fisher Scientific, #23227).

Gene expression analysis by antibody-based protein quantification

The antibody-based protein quantification was performed using a WES automated capillary electrophoresis system (Protein Simple) with 12-230 kDa Separation Module (Protein Simple) as instructed. The protein concentration was aligned to around 0.5 mg/mL of the sample. Information about the primary and secondary antibodies and the dilutions are provided in Key resources table. The data was analyzed and visualized using Compass for Simple Western software (Protein Simple). The WES protein images were used throughout the manuscript.

Gene expression analysis by qRT-PCR and microarray

The cells were lysed with QIAzol lysis reagent (QIAGEN) for whole cell lysis, and fractionated samples were lysed with TRIZOL LS reagent (Thermo Fisher Scientific). Total RNA was purified using a miRNeasy Mini Kit (QIAGEN). Purified RNA (0.1–1 μg) was used for single strand complementary DNA (cDNA) synthesis using a SuperScript III First-Strand Synthesis SuperMix for qRT-PCR (Thermo Fisher Scientific) or ReverTra Ace (TOYOBO). Quantitative RT-PCR was performed using TaqMan Gene Expression Master Mix (Thermo Fisher Scientific) or PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) on a Quant Studio 3 instrument (Thermo Fisher Scientific). The mRNA levels were normalized to the human GAPDH or drosophila gamma-Tubulin at 23C (gammaTub23C) expression, and then the relative expressions were normalized with the control.

Microarray was performed as described previously (Takahashi et al., 2020) using the purified total RNA evaluated by a 2100 Bioanalyzer (Agilent Technologies). Briefly, two hundred nanograms of total RNA were labeled with Cyanine 3-CTP and used for hybridization with SurePrint G3 Human GE 8x60K ver. 3 (G4851C, Agilent Technologies) with the one-color protocol. The arrays were scanned with a Microarray Scanner System (G2565BA, Agilent Technologies), and extracted signals were analyzed using the GeneSpring version 14.9.1 software program (Agilent Technologies). Gene expression values were normalized by the 75th percentile shifts.

Subcellular fractionation of mRNA and proteins

Subcellular fractionations of the cytoplasm, organelles, and nuclei were performed using the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Fisher Scientific, #78840). Briefly, the cells were washed once with ice-cold PBS (Nacalai Tesque). After completely removing PBS, the cells were lysed with 200 μL (24-well plate), 500 μL (6-well plate), or 1000 μL (9-cm dish) of CEB solution at 4°C for 5 min. Cell lysates were collected by scraping and pipetting on pre-chilled Protein LoBind 2 mL tubes (Eppendorf). The lysates were centrifuged at 500 x g for 5 min, and then the supernatants were collected as cytoplasmic fractions. The cell pellets were dissolved using 200 μL (24-well plate), 500 μL (6-well plate), or 1000 μL (9-cm dish) of MEB solution. After a 5-s vortex, the lysate tubes were inverted for 10 min at 4°C. The supernatants were collected as organelle fractions after centrifugation at 3000 x g for 5 min. The pellets were collected as nucleus fractions. For protein extraction, the cytoplasm and organelle fractions were ultra-filtrated using an AmiconUltra-0.5 (10K) at 14000 x g for 10 min and solubilized by PTS buffer. For the nucleus fractions, 200 μL (24-well plate), 500 μL (6-well plate), or 1000 μL (9-cm dish) of PTS buffer was added to solubilize the pellet. Protein extraction was performed as described above. For RNA extraction, an equal volume of TRIZOL LS reagent (Thermo Fisher Scientific) was added to the cytosol and organelle fractions. For the nucleus fraction, 400 μL (24-well plate) or 1000 μL (6-well plate) of QIAzol

lysis reagent (QIAGEN) was added to solubilize the pellet. RNA extraction was performed as described above.

Gene expression analysis by RNA-sequencing

The subcellular fractionated RNA of iPSC-1 and HDF-2 was purified as described above. Purified RNAs were quantified using the Agilent RNA6000 Pico Kit (Agilent) on a Bioanalyzer 2100 (Agilent) and Qubit RNA HS assay Kit (Thermo Fisher Scientific). Sequencing libraries were prepared using Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero Plus (illumina) following the manufacturer's protocol with some modifications. In brief, purified RNA (500 ng) was applied for ribosomal RNA depletion. The remaining RNA was evaluated using the Qubit™ RNA High Sensitivity (HS) Kit (Thermo Fisher Scientific), and 15 ng of RNA was used for later steps. Indices were used from IDT for Illumina RNA UD Indexes Set A (illumina). The quality and concentration of libraries were evaluated using an Agilent High-Sensitivity DNA Kit (Agilent) and Qubit dsDNA HS assay Kit (Thermo Fisher Scientific). The libraries were sequenced using NextSeq500 (illumina) with NextSeq 500/550 High Output v2 Kit (illumina). The sequenced reads were first trimmed the adapter sequence using cutadapt-1.12 (Martin, 2011), and excluded reads were mapped to ribosomal RNA using bowtie2 (version 2.2.5) and SAM tools (version 1.10) (Langmead and Salzberg, 2012; Li et al., 2009). Then the reads were aligned to the human genome (hg38) using STAR (version 2.5.3a) (Dobin et al., 2013), and RSeQC (version 4.0.0) was used for the quality check. Finally, uniquely mapped reads were used to calculate transcripts per million (TPM) using RSEM (version 1.3.3) with GENCODE annotation file (version 35) (Li and Dewey, 2011). To quantify subcellular localization, we performed previously described methods (Zuckerman et al., 2020) (Carlevaro-Fita and Johnson, 2019) with R (version 4.1). In brief, the genes whose expression levels were between the 50^{th} and 90^{th} percentiles of whole cell samples were extracted for each sample set. Using these genes, the normalization coefficients (Figure S7B) were estimated using a linear regression analysis. Using these coefficients, the TPM values were normalized to obtain localization values. To confirm accuracy of the coefficients, we compared TPM values of the whole sample and predicted values (normalized Cyto+ normalized Nuc) for all expressed genes. For all sample sets, Pearson correlation values were over 0.96.

Gene expression analysis by nano liquid chromatography (nanoLC)-mass spectrometry (MS)

Protein samples lysed with PTS buffer were subjected to reduction, alkylation, Lys-C/trypsin digestion (enzyme ratio 1/100), and desalting as previously described (Iwasaki et al., 2019). Briefly, 1000 μL of an organic solvent was added to 1000 μL of the digested protein solution, and the mixture was acidified with 0.5% TFA (final concentration). The mixture was shaken for 5 min, then centrifuged at 15700 x g for 2 min to obtain aqueous and organic phases. The aqueous phase was collected and desalted using SDB-StageTip. The resulting peptides were labeled with isobaric tags for relative and absolute quantification (Sciex). Briefly, 120 μg of desalted peptide samples were dried and dissolved in 10 μL of 500 mM triethyammonium bicarbonate. Approximately 20 μL of iTRAQ reagents (Sciex) was added to 23 μL of ethanol and mixed with the peptide sample. After incubation for 1.5 h at room temperature, 16 μL of 10% TFA and 400 μL of loading buffer (0.5% trifluoroacetic acid and 4% (v/v) acetonitrile) were added to quench the reaction, and the sample mixture was desalted using StageTip (Rappsilber et al., 2003). A total of 8 μg of iTRAQ labelled sample set was subjected to nanoLC-MS/MS using a TripleTOF 5600 System (AB Sciex) equipped with an HTC-PAL autosampler (CTC Analytics). All nanoLC-MS conditions were the same as previously described (Iwasaki et al., 2019) and used the monolithic column (4 m length, 100 μm i.d., GL Science). The mobile phases were composed of 0.5% acetic acid with 5% (v/v) DMSO (solution A) and 0.5% acetic acid in 80% (v/v) acetonitrile with 5% (v/v) DMSO (solution B). A flow rate of 400 μL/min of 5-15% solution B for 205 min, 15-35% solution B for 549 min, 35-40% solution B for 103 min, 40-100% solution B for 5 min, 100% solution B for 118 min and 5% solution B for 100 min was used (total 1,080 min). NanoLC-MS/MS analysis was performed in triplicate, and blank runs were inserted between the samples. The proteome data analysis method was previously reported (Iwasaki et al., 2019). The raw data files were analyzed using ProteinPilot v5.0 (Sciex) with acceptable modifications of N-terminal iTRAQ, iTRAQ of lysine, carbamidomethylation of cysteine, oxidation of methionine, phosphorylation of serine, threonine or tyrosine, deamidation of asparagine or glutamine, the N-terminal pyro-glutamic acid of glutamine or glutamic acid, and protein N-terminal acetylation. Peak lists, which were generated from a ProteinPilot.group file, were analyzed by Mascot v2.5 (Matrix Science) with the carbamidomethylation of cysteine as the fixed modification, and the N-terminal iTRAQ, iTRAQ of lysine, and methionine oxidation as the variable modification. Both database search engines were used against human entries of UniProt/Swiss-Prot release 2016_06 (8-June-2016) with a precursor mass tolerance of 20 ppm, a fragment ion

mass tolerance of 0.1 Da, and strict trypsin and Lys-C specificity, which allowed for up to two missed cleavages. For the peptide identification, peptides were rejected if any of the following conditions were not satisfied: (a) if the same scan was assigned to different peptides between ProteinPilot and Mascot, (b) peptide confidence was below 0.05, (c) the charge state was more than 5, (d) or the peptide length was less than 6 amino acids. For the protein identification, at least two confidently ($p < 0.05$) identified peptides per protein were used. Single peptides with higher confidence ($p < 0.01$) were allowed. Finally, peptides were grouped into protein groups. False discovery rates (FDRs) were estimated by searching against a decoy sequence database (<1%).

For quantification, the accumulated intensity of iTRAQ label spectra was calculated after the application of the RiMS method to remove interference spectra (Iwasaki et al., 2019). Then, the final normalized accumulated intensity of iTRAQ label spectra was calculated for the whole and subcellular fraction samples in technical triplicates to acquire the protein quantification value for each cell type.

Trans-omics data analysis

The mRNA and protein ratio between cell types was calculated using the normalized mRNA and protein quantification values described above for iPSC-1, iPSC-2, ESC, HDF-1, and HDF-2. Then, the mRNA and protein ratios of iPSCs or ESCs to HDFs were compared based on the same gene symbol or Entrez Gene. To determine statistical significance, we conducted unpaired t tests with Excel 2016 (Microsoft) in biological triplicate. The definition of independent mRNA and protein upregulation is as follows. Figure S1A shows how the actual distribution of the mRNA-and-protein ratios corresponds to each cluster. Commonly varied genes were selected in the six compared pairs of iPSCs/ESCs and HDFs for further analysis (Figure 1C). Gene ontology analysis was performed using TargetMine (Chen et al., 2019). To identify upregulation, the following conditions were applied.

Upregulated genes only at the protein level in iPSCs/ESCs (Log_2):

mRNA ratio < 0.9, Protein ratio ≥ 0.9

Upregulated genes only at the protein level in HDFs (Log_2):

mRNA ratio > -0.9, Protein ratio ≤ -0.9

Upregulated genes only at the mRNA level in iPSCs/ESCs (Log_2):

$-0.9 <$ Protein ratio < 0.9, mRNA ratio ≥ 0.9

Upregulated genes only at the mRNA level in HDFs (Log_2):

$-0.9 <$ Protein ratio < 0.9, mRNA ratio ≤ -0.9

siRNA screening

For the siRNA screening, we referred to (Chia et al., 2010). 201B7 and 1418E1 were seeded on iMatrix-511-coated 96-well plates at 4,940 cells/well. Y-27632 (final concentration, 10 μM) was used for the iPSC maintenance for two days after passage and transfection. Tig120 was seeded directly on 96-well plates at 2,470 cells/well. On the next day of the passage (day 1), a manually aliquoted siGENOME SMARTpool siRNA library (0.2 μM, 7.5 μL/well; Horizon Discovery) was mixed with an equal volume of 0.5% of Stemfect RNA Transfection Reagent (STEMGENT). Then, 20 μL of siRNA solution and 80 μL of fresh medium were mixed, and the cell medium was replaced with 80 μL of this solution. On the next day, the medium of the wells with transfection reagent was manually replaced. As controls for the siRNA screening, we used siNontarget, siOCT3/4, and LaminB2 for every 3 wells per plate. Four days after the transfection, the cells were fixed and analyzed by immunostaining. In brief, the cells were washed twice with PBS (Nacalai Tesque) and fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature. Then the fixed cells were treated with PBS containing 0.2% Triton X-100 and 1% bovine serum albumin (BSA, Thermo Fisher Scientific) for 15 min at room temperature after two PBS washes. The cells were incubated with the primary antibody anti-OCT3/4 (B&D) or Lamin B2 (CST) diluted in PBS containing 1% BSA for more than one hour at room temperature with protection from light. After washing twice with PBS, the cells were incubated with 0.01%

Hoechst 33342 (Dojindo) and the secondary antibody Alexa488 or Cy3 diluted in PBS containing 1% BSA for one hour at room temperature with protection from the light. After two PBS washes, PBS was added to the well, and 10x immunofluorescence images were acquired using an ArrayScan VTI 600 Series (Thermo Fisher Scientific). Obtained images were analyzed using HCS Studio (Thermo Fisher Scientific, version 6.5.0, build 7614). The siRNA knockdown of OCT3/4 and LaminB2 was respectively used as a pluripotency marker for iPSCs and a universal control for both iPSCs/ESCs and HDFs. We measured the cell number by Hoechst staining and checked the general knockdown efficiency by immunostaining. The average knockdown efficiencies by siOCT3/4 were 100.0% and 99.7% for 201B7 (iPSC-1) and 1418E1 (iPSC-2), respectively, and by siLaminB2 it was 98.4% for Tig120 (HDF-2). The criteria for a change in cell number by siRNA against OCT3/4, LaminB2, or no target was a value more than ± 2 S.D. the mean cell number and checked manually. Dilution conditions of the primary and secondary antibodies are provided in Key resources table. A list of the siRNA used in this study is provided in Table S5.

Proteasome and lysosome inhibitor assay

One semiconfluent 6-well dish of adherent 201B7 or Tig120 was treated with medium containing 20 μM MG-132 (Calbiochem) or 250 nM Bafilomycin A1 (InvivoGen) and 500 nM Wortmannin (Calbiochem) for the proteasome or lysosome assays, respectively. After 0, 2, 6, 8, and 24 h, the cells were washed once with ice-cold PBS (Nacalai Tesque) and directly lysed with ice-cold PTS buffer with 1% phosphatase and protease inhibitors (Sigma-Aldrich). Cell lysates were collected by scraping and pipetting on pre-chilled Protein LoBind 2 mL tubes (Eppendof). After sonication and heat shock at 95°C for 5 min, the protein concentration of the sample was determined using the BCA protein assay kit.

Monosome and polysome fractionation

One semiconfluent 100-mm dish of adherent 201B7 or Tig120 was treated with medium containing 100 μg/mL Cycloheximide (Sigma-Aldrich) for 5 min at 37°C. The cells were placed on ice and gently washed once with 10 mL ice-cold PBS. Then they were lysed with 0.4 mL ice-cold lysis buffer (20 mM Tris-HCl, pH7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), Complete EDTA-free Protease Inhibitor Cocktail, 100 μg/mL Cycloheximide, 1% Triton X-100, 25 units/mL Turbo DNase (Thermo Fisher Scientific), and 100 units/mL SUPERaseIn (Thermo Fisher Scientific)), scraped, and collected into a 1.5 mL chilled DNA LoBind Tube (Eppendorf). The lysate was incubated on ice for 10 min and triturated through a 25-gauge needle (Terumo) ten times before centrifugation at 20,000 xg for 10 min at 4°C. The supernatant was collected in a new 1.5 mL tube for the sucrose gradient analysis and as the sample for loading. A 10–45% continuous sucrose gradient was prepared in a polyclear tube (Seton) using 10 and 45% sucrose buffers containing 100 μg/mL Cycloheximide and 1 mM DTT in polysome buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl and 15 mM $MgCl_2$) and the Biocomp Gradient Master program (Biocomp). An equal amount of cell lysate as sample (300 μL) was loaded on the prepared gradient solution. Monosomes and polysomes were separated in the sucrose gradient by ultracentrifugation using a SW-41 rotor (Beckman Coulter) at 36,000 rpm for 2.5 h at 4°C. The profile of relative RNA abundances of monosomes and polysomes were visualized at 254-nm wavelength, and equal-volume fractions were collected simultaneously with the Biocomp Piston Gradient Fractionator (Biocomp). We collected the fractions as monosomes using 40S, 60S, and 80S ribosomes. For light polysomes, we collected two and three polysomes. For heavy polysomes, we collected four or more polysomes. For the RNA analysis, an equal sample volume of TRIzol LS reagent (Thermo Fisher Scientific) was immediately added to the fractions and load sample. RNA was purified using an miRNeasy Mini Kit according to manufacturer's instruction. Purified RNAs along with 1 ng of spiked drosophila RNA were used for the cDNA synthesis and the following qRT-PCR. The cycle threshold for detectable gene expression was set as Ct = 45. Ct values were normalized by the spiked drosophila RNA and compared with the loading sample before the sucrose gradient.

mRNA transfection

One day before the transfection, we plated 201B7 and Tig120 at a density of 2×10^5 cells and 0.7×10^5 cells per well on a 24-well plate, respectively. For the iPSCs, we prepared an LN511-coated 24-well plate, and Y-27632 (final concentration, 10 μM) was used for the iPSC maintenance for two days after passage. For the mRNA transfection, N1-methyl-pseudouridine ($1m\Psi$)-modified mRNA was prepared, and 100 ng mRNA was transfected using Lipofectamine MessengerMAX (Thermo Fisher Scientific) on the next day of passage (day 1) according to the manufacturer's protocol. The next day, the cells were washed twice with PBS (Nacalai Tesque) and lysed with 150 μL PTS buffer. We used GFP mRNA as a control for the transfection efficiency.

In silico analysis of motifs in cDNA sequences

We obtained DNA sequences from BioMart in Ensemble database (Howe et al., 2021). We downloaded cDNA sequences from the dataset of Human genes (GRCh38.p13). The cDNA sequences are used for searching motif sequences in RBPmap (Paz et al., 2014) with the setting of database assembly (Dec.2013 (GRCh38/hg38)), genome (Human) and selecting motifs from RBPmap full list (Human/Mouse motifs). The result text file was analyzed by perl 5 version 24, and the motifs with minimum p value was exported.

Data availability

Gene expression microarray and RNA-seq results are accessible in the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information website (accession number: GSE184546, GSE199820, respectively). The mass spectrometry data have been deposited to the ProteomeXchange Consortium via jPOSTrepo (Okuda et al., 2017) (https://repository.jpostdb.org/) with the dataset identifier JPST001308 (PXD028489) for the 2-plex analysis of iPSC-1 (201B7) and HDF-1 (HDF1388), and for the 4-plex analysis of iPSC-1 (201B7), iPSC-2 (1418E1), ESC (H9), and HDF-2 (Tig120). Original immunoblot data have been deposited to Mendeley Data (https://doi.org/10.17632/y6b3bgng9p.1).

QUANTIFICATION AND STATISTICAL ANALYSIS

We performed statistical analysis using two-sample unpaired t-test to calculate p values for the difference between mRNA and protein ratios using Excel (Microsoft). The p values less than 0.05 were considered significant and are indicated by green and yellow in Figures 1B, S1B, and S1C.