

1 **Post-transcriptionally regulated genes are essential for pluripotent stem cell**
2 **survival**

3

4 Mio Iwasaki^{1,7*}, Yuka Kawahara¹, Tsuyoshi Tabata¹, Yohei Nishi^{1,2}, Megumi Narita¹,
5 Akira Ohta¹, Hirohide Saito¹, Takuya Yamamoto^{1,3,4}, Masato Nakagawa¹, Shinya
6 Yamanaka^{1,5,6}, Kazutoshi Takahashi^{1*}

7

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

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

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

14 ⁴Medical-risk Avoidance based on iPS Cells Team, RIKEN Center for
15 Advanced Intelligence Project (AIP), Kyoto 606-8507, Japan

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

17 ⁶Department of Anatomy, University of California, San Francisco, San Francisco, CA
18 94143, USA

19 ⁷Lead Contact

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

21

22 **Keywords**

23 Pluripotent stem cells, post-transcriptional regulation, proteomics, transcriptomics

24

25 **Summary**

26 The effects of transcription factors on the maintenance and differentiation of
27 pluripotent stem cells (PSCs) have been well studied. However, the importance of
28 post-transcriptional regulatory mechanisms, which cause the quantitative
29 dissociation of mRNA and protein expression, has not been explored in detail. Here,

1 by combining transcriptome and proteome profiling, we identified 228 post-
2 transcriptionally regulated genes with strict upregulation of the protein level in PSCs.
3 Among them, we found that 84 genes were vital for the survival of PSCs and HDFs,
4 including 20 genes that were specifically necessary for the survival of PSCs. These
5 20 proteins were upregulated only in PSCs and not in differentiated cells derived
6 from the three germ layers. Subcellular fractionation of the mRNA showed that the
7 expression of most of the 20 proteins was regulated at the mRNA localization stage
8 from the nucleus to the cytoplasm, but their translation efficiency was constant.
9 Together, these results revealed that post-transcriptionally regulated genes have a
10 crucial role in PSC survival.

11

12 **INTRODUCTION**

13 The processing (Bentley, 2014), export (Hentze et al., 2018), turnover (Dendooven
14 et al., 2020), and accurate decoding of mRNA (Teixeira and Lehmann, 2019),
15 ribosome biogenesis (Pelletier et al., 2018), and protein degradation (Pohl and Dikic,
16 2019) are critical steps in the post-transcriptional regulation of gene expressions.
17 Among the several hundred human cell types, pluripotent stem cells (PSCs) have
18 unique post-transcriptional mechanisms (Chen and Hu, 2017), such as specific
19 mRNA processing events, that establish and maintain the pluripotent state (Gabut et
20 al., 2011; Han et al., 2013; Ohta et al., 2013; Salomonis et al., 2010). Interestingly,
21 for many genes, differences in protein levels in PSCs are not accompanied by
22 differences in the corresponding mRNA levels (Lu et al., 2009), suggesting post-
23 transcriptionally regulated genes exist in strict regulation at the protein level. One
24 example is the CLOCK gene, in which the amount of mRNA is constant before and
25 after differentiation, whereas the protein expression depends on the cell type and is
26 absent in PSCs (Umemura et al., 2017). Considering that protein levels are more
27 conserved than mRNA levels in primates (Khan et al., 2013), the strict regulation at
28 the protein level with constant mRNA quantities might have important effects on
29 pluripotency and differentiation during development.

1 To systematically identify genes differently regulated at the mRNA and protein level,
2 transcriptome analysis by RNA sequencing or microarray and proteome analysis by
3 mass spectrometry (MS) have been performed on various cell types (Buccitelli and
4 Selbach, 2020; Matsumoto et al., 2017; Roumeliotis et al., 2017; Wang et al., 2019).
5 Typically, a high correlation in the levels is shown for core metabolic pathway-related
6 genes, but a low correlation is observed for ribosomal and spliceosome genes.
7 However, whether post-transcriptionally regulated genes with tightly controlled
8 protein levels are essential for the maintenance of the cell remains an open question.
9 Here, we compared global mRNA and protein levels between PSCs and
10 differentiated cells to identify essential post-transcriptionally regulated genes. We
11 found 228 post-transcriptionally regulated genes exclusively in PSCs that showed
12 specific biological functions for RNA and nucleic acid binding. siRNA screening
13 revealed that 54% (84 genes) were necessary for cell survival and 20 of them were
14 identified as PSC-specific. Finally, we showed that these 20 genes are mainly
15 regulated at the mRNA localization stage rather than the translation stage,
16 suggesting an importance of cell type-specific post-transcriptional regulation in
17 PSCs.

18

19 **RESULTS**

20 **228 upregulated proteins in PSCs showed constant mRNA expression**

21 First, to identify upregulated genes and post-transcriptionally regulated genes, we
22 analyzed the mRNA and protein levels by microarray and MS, respectively (Figure
23 1A). We used two human dermal fibroblast (HDF) lines (HDF-1, HDF1388; and HDF-
24 2, Tig120), two human induced PSC (iPSC) lines (iPSC-1, 201B7; and iPSC-2,
25 1418E1), and one human embryonic stem cell line (ESC, H9). We compared the
26 ratio of the mRNA and protein levels for 6404 genes for all combinations of PSCs
27 (iPSC-1, -2, ESC) and HDFs (HDF-1, -2) (Figure S1). In the six pairs of different cell
28 types, commonly upregulated genes were defined as having at least a 1.9-fold
29 change ($\log_2(0.9)$). We found that 1275 genes had upregulated mRNA levels in all

1 PSCs, and 1468 genes had upregulated protein levels (Figure 1B). Among these
2 genes, 509 showed upregulated levels for both mRNA and protein, including PSC-
3 specific genes, such as LIN28A, OCT3/4, and SOX2 (Figure 1B, C). Regarding
4 HDFs, the numbers were 1519 and 461, respectively, and the number with common
5 upregulation was 387, including HDF-specific genes, such as FN1 and VIM (Figure
6 1B, C). Overall, each cell-specific gene showed a highly correlated mRNA to protein
7 ratio (Figure 1C and Table S1).

8 Next, we examined genes that were post-transcriptionally regulated by selecting
9 those that were upregulated only at the mRNA or protein level by $\log_2(0.9)$ (Figure
10 1C; see Trans-omics data analysis in Methods). From this analysis, we found 10 and
11 56 genes in PSCs and HDFs, respectively, upregulated only at the mRNA level, and
12 228 and 8 genes upregulated only at the protein level. To investigate the function of
13 these four gene sets, we conducted a Gene Ontology (GO) analysis (Fig. 1D).
14 Especially among the 228 genes in PSCs that showed only upregulated protein
15 levels, we found specific molecular functions, such as RNA binding and nucleic acid
16 binding. Although the 8 genes for HDFs showed no enrichment in molecular
17 functions, we named these 228 and 8 genes “upregulated protein levels independent
18 of mRNA levels (uPRA)” for further analysis and found that 228 PSC-uPRA genes
19 have specific nucleic acid binding functions.

20

21 **20 uPRA genes are essential for PSC cell survival**

22 To investigate which uPRA genes are significant for cell maintenance, we performed
23 siRNA knockdown screening in iPSC-1, -2, and HDF-2 (Figure 2A, 2B, Table S2, S3).
24 We picked siRNAs to target 156 (Table S2) of the 228 PSC-uPRA and 8 HDF-uPRA
25 genes. The siRNA knockdown of OCT3/4 and LaminB2 was respectively used as a
26 pluripotency marker for PSCs and a universal control for both cell types (Figure 2B).
27 The knockdown of none of the 156 uPRA genes induced differentiation (Table S3).
28 On the other hand, the knockdown of 84 of the 156 uPRA genes, such as SRRT and
29 RSL1D1, significantly decreased the number of PSCs (Figure 2B). Among them, the

1 knockdown of 52 uPRA genes decreased the number of HDFs too (Figure 2C). This
2 result indicates that uPRA genes have important effects on the cellular maintenance
3 of HDFs as well. Additionally, we identified 20 PSC-uPRA genes and 1 HDF-uPRA
4 gene that are each specifically necessary for the survival of the respective cells
5 (Figure 2C; images of the knockdown of these 21 genes are shown in Figure S2). A
6 GO analysis showed that the 20 PSC-uPRA genes code for components of
7 ribonucleoprotein and protein-containing complexes and have RNA- and nucleic
8 acid-binding properties (Figure 2D, E). These data suggest that the 20 PSC-uPRA
9 genes are essential for the survival of PSCs and synergistically maintain PSCs via
10 heterocyclic compound binding properties.

11

12 **Upregulated protein levels of the 20 essential uPRA genes are PSC-specific**

13 We examined if the 20 PSC-uPRA genes are regulated post-transcriptionally only in
14 PSCs. We selected another seven differentiated primary cell lines representing the
15 three germ layers to compare with the PSCs, which are undifferentiated. In total,
16 mesoderm-derived cells included adipose tissue-derived mesenchymal stem cells
17 (HAdMSC) and two HDF lines (HDF-1 and 2), endoderm-derived cells included
18 normal human bronchial epithelial cells (NHBEK) and human prostate epithelial cells
19 (PrEC), and ectoderm-derived cells included normal human epidermal keratinocytes
20 (NHEK) and human neural progenitor cells derived from H9 (NPC H9). Quantitative
21 reverse transcription PCR (qRT-PCR) and immunoblotting were performed on the
22 mRNA and proteins, respectively, and the gene expressions relative to HDF-1 were
23 compared (all blotting images and graphs are shown in Figure S3A, C). RSL1D1,
24 one of the 20 essential PSC-uPRA genes, showed a more than 50-fold increase in
25 its protein expression only in PSCs, whereas its mRNA expression was equal among
26 all cell types (Figure 3A, B). Overall, we confirmed that the protein expressions of 18
27 PSC-uPRA genes (RBM22 and SF3B3 were the exceptions) were increased more
28 than two-fold compared with their mRNA expressions in PSCs (Figure 3C). On the
29 other hand, the fold-change between protein and mRNA expressions was around

1 one or less in differentiated cells for 17 PSC-uRPO genes (IMP4, NCBP2, and
2 BUD31 were the exceptions), suggesting the post-transcriptional regulation
3 mechanism is different between PSCs and differentiated cells (Figure 3D). RSL1D1,
4 SF3B4, RBM22, and SF3B3 especially showed protein to mRNA expression ratios
5 less than one in differentiated cells, indicating that the protein expressions of these
6 genes are usually suppressed in cell types other than PSCs. As for AFP, the one
7 HDF-uPRA gene, we confirmed a specific increase in the protein expression in
8 mesoderm-derived cells (Figure S3B). Following these observations, we concluded
9 that the 20 PSC-uPRA genes are post-transcriptionally regulated specifically in
10 PSCs.

11

12 **HSPA8, EIF3D, and NCBP2 protein expressions are controlled at the ubiquitin-** 13 **dependent degradation stage**

14 Next, we analyzed individual post-transcriptional regulation stages to identify the
15 regulation mechanism of the protein expressions of the 20 essential PSC-uPRA
16 genes (Figure 4A). After transcription, mRNAs are transported from the nucleus to
17 the cytoplasm, where ribosomes bind to translate them. Eventually, the proteins are
18 degraded by the ubiquitin-proteasome system and/or lysosomes. At first, we
19 analyzed the protein degradation efficiency using MG-132 (proteasome inhibitor)
20 and Bafilomycin A1 and Wortmannin (lysosome inhibitors) as controls for the assays
21 (Figure S4A-D). These inhibitors should recover the PSC-uPRA protein expression
22 in HDFs if the proteins are downregulated by fast degradation. We found that the
23 protein expressions of HSPA8, EIF3D, NCBP2, and IMP4 were steadily increased
24 more than two-fold after exposure to proteasome inhibitor for up to 8 hours (Figure
25 4B, S4E). We did the same assay for up to 24 hours and confirmed that the protein
26 expressions of HSPA8, EIF3D, and NCBP2 were controlled by fast degradation via
27 the proteasome system in HDFs (Figure 4C). The lysosome inhibitor assay (Figure
28 4D, S4F) showed high protein expressions of NCBP2 and RSL1D1 for treatment up
29 to 8 hours but not 24 hours (Figure 4E), suggesting no PSC-uPRA genes were

1 regulated by lysosome-dependent degradation. Overall, these data indicate that the
2 protein expressions of three PSC-uPRA genes, HSPA8, EIF3D and NCBP2, are
3 regulated at the ubiquitin dependent degradation stage.

4
5 **The cytosolic localization but not the translation efficiency of uPRA mRNA is**
6 **increased in PSCs**

7 Next, we analyzed the subcellular localization of the mRNAs of the 20 essential PSC-
8 uPRA genes to examine their intracellular abundance between different cell types.
9 We used 18S rRNA and MALAT1 lncRNA, respectively, as cytosolic and nuclear
10 RNA controls in the subcellular fractionation experiment (Figure 5A). For cytosolic
11 RNA, 94.3% and 73.3% of 18S rRNA were observed in PSCs and HDFs, respectively,
12 and for nuclear RNA, 99.5% and 99.8% of MALAT1 lncRNA were observed,
13 validating the subcellular fractionation experiment. Then we analyzed the cytosolic
14 RNA percentage of the 20 essential PSC-uPRA genes by qRT-PCR (Figure 5B). The
15 cytosolic RNA percentage of 18 PSC-uPRA genes (IMP4 and CPSF6 were the
16 exceptions) were increased at least more than two-fold in PSCs. Conversely, the
17 nuclear RNA percentage of PSC-uPRA genes was increased in HDFs (Figure 5C).
18 These results indicated that the mRNAs of the 18 PSC-uPRA genes were retained
19 in the nucleus in HDFs.

20 RNA splicing is known to promote mRNA-export to the cytoplasm, because export
21 factors can be recruited by the spliceosome to mature RNA (Palazzo and Lee, 2018).
22 However, PSCs and differentiated cells do not express different spliced isoforms of
23 the 20 PSC-uPRA genes (Han *et al.*, 2013). We then checked the protein
24 expressions of RNA export factors. We found more than a four-fold increase in the
25 protein expressions of NXT1 and NXF1, two mRNA export receptor heterodimers, in
26 PSCs and NPC 1418E1 compared to the other differentiated cells (Figure 5D, S5A,
27 B). These data indicated that most PSC-uPRA mRNA was retained in the nucleus of
28 HDFs, but the increased NXT1 and NXF1 protein expressions may accelerate mRNA
29 transport to the cytoplasm in PSCs.

1 Next, to investigate the translation efficiency, we extracted mRNAs in the monosome,
2 light polysome, and heavy polysome fractions of PSCs and HDFs using sucrose
3 fractionation. We quantified and compared the PSC-uPRA mRNA in each fraction
4 between PSCs and HDFs (Figure 5E). Except for NIFK and RSL1D1, no PSC-uPRA
5 mRNA was enriched in the heavy polysome fraction of PSCs (Figure 5F). This result
6 suggests that the translation efficiency of most PSC-uPRA mRNAs cannot explain
7 the increased protein levels in PSCs.
8 Among PSC-uPRA mRNAs, SRRT mRNA was the most associated with heavy
9 polysomes in HDFs despite the inhibited protein expression (Figure 5F). Since a
10 lower cytosolic RNA percentage of endogenous SRRT was observed in HDFs
11 (around 10%; Figure 5B), we next analyzed whether its translation occurs if the
12 cytoplasmic RNA content is increased. We constructed full-length FLAG-tagged
13 SRRT mRNA and one day after the mRNA transfection confirmed the translation of
14 FLAG-SRRT protein from the transfected mRNA by immunoblotting (Figure 5G). In
15 summary, these results indicate that there exists an inhibitory mechanism in HDFs
16 for transporting PSC-uPRA mRNAs to the cytoplasm and for the translation of PSC-
17 uPRA genes even after mRNA binds to ribosomes.

18

19 **DISCUSSION**

20 Despite their ubiquitous effects, there is comparatively little understanding about
21 post-transcriptional regulatory mechanisms compared with transcriptional regulatory
22 mechanisms in PSCs. Here, we revealed that 20 post-transcriptionally regulated
23 genes (i.e. uPRA genes) are essential for the maintenance of PSCs. These genes
24 had upregulated protein levels without any major positive upregulation in mRNA
25 levels in PSCs (Figure 3C). The regulatory processes of these 20 essential PSC-
26 uPRA genes are illustrated in Figure 6. In HDFs, 18 PSC-uPRA mRNAs were highly
27 retained in the nucleus (Figure 5C). Among them, the translation efficiency of
28 RSL1D1 and NIFK was upregulated in PSCs (Figure 5F), and the proteins of HSPA8,
29 EIF3D, and NCBP2 were quickly degraded in HDFs at the proteasome degradation

1 stage (Figure 4B, C). Also, we found that the translation efficiency of 18 uPRA genes
2 was comparable between PSCs and HDFs by polysome profiling (Figure 5F),
3 suggesting several inhibitory effects for uPRA mRNA transportation to the cytoplasm
4 and for translation after ribosome binding in differentiated cells.
5 To identify the inhibitory effect, we picked up one uPRA gene, SRRT, for further
6 analysis. The percentage of cytoplasmic SRRT mRNA was 3-fold less in HDFs
7 (Figure 5B), and the mRNA itself was more associated with heavy polysomes in
8 HDFs than in PSCs (Figure 5F). Furthermore, exogenous cytosolic SRRT mRNA
9 was translated in HDFs (Figure 5G), suggesting the cytoplasmic mRNA level is key
10 for proper protein expression. Another possible mechanism for translation inhibition
11 in differentiated cells is ribosome pausing, in which the mRNA binds to ribosomes
12 that are not active or select non-canonical start codons (Chandrasekaran et al.,
13 2019; Darnell et al., 2018; Tresenrider et al., 2021). A third possibility is that aberrant
14 proteins were degraded immediately after the production by a quality control
15 pathway of the mRNA and ribosome (D'Orazio and Green, 2021). More study is
16 needed to clarify which mechanism applies to which gene.
17 All identified 20 PSC-uPRA proteins in the present study are involved in a wide range
18 of cellular functions, from transcription to post-translation. For transcription-related
19 functions, 10 PSC-uPRA genes have polymerase-, splicing- and mRNA maturation-
20 related functions: POLD2 is a component of the DNA polymerase complex; SF3B3,
21 SF3B4, and RBM22 are components of the splicing factor; and PUF60, BUD31
22 SAP18, CPSF6, U2AF2, and SRRT are required for the splicing of pre-mRNA, with
23 SRRT especially known as “a molecular guardian of the pluripotent cell state” by
24 facilitating proper splicing in PSCs (Kainov and Makeyev, 2020). For translation-
25 related functions, 8 PSC-uPRA genes are involved in functions of the ribosome
26 complex: NCBP2 and EIF3D are mRNA cap binding proteins; EIF3D is necessary
27 for specialized translation initiation factors (Lee et al., 2016); RPL7L1 is a putative
28 component of the ribosome complex; and RSL1D1, NIFK, NOL10, IMP4, and TSR1
29 are required for rRNA processing. For post-translational functions, CCT5 and HSPA8
30 are components of the chaperone complex, which helps newly synthesized proteins
31 properly fold. Our observation that these genes are post-transcriptionally regulated

1 is particularly interesting, because they all function as regulators post-transcription.
2 Thus, they might use feedback in their regulation to maintain pluripotency.
3 In conclusion, our study revealed that the protein expressions of 228 genes are post-
4 transcriptionally regulated in PSCs. The protein levels of 20 of these genes are
5 specifically increased in PSCs, and the proteins themselves are essential for PSC
6 maintenance. Finally, different mechanisms regulate the gene expressions between
7 cell types.

8

9 **Limitation of the study**

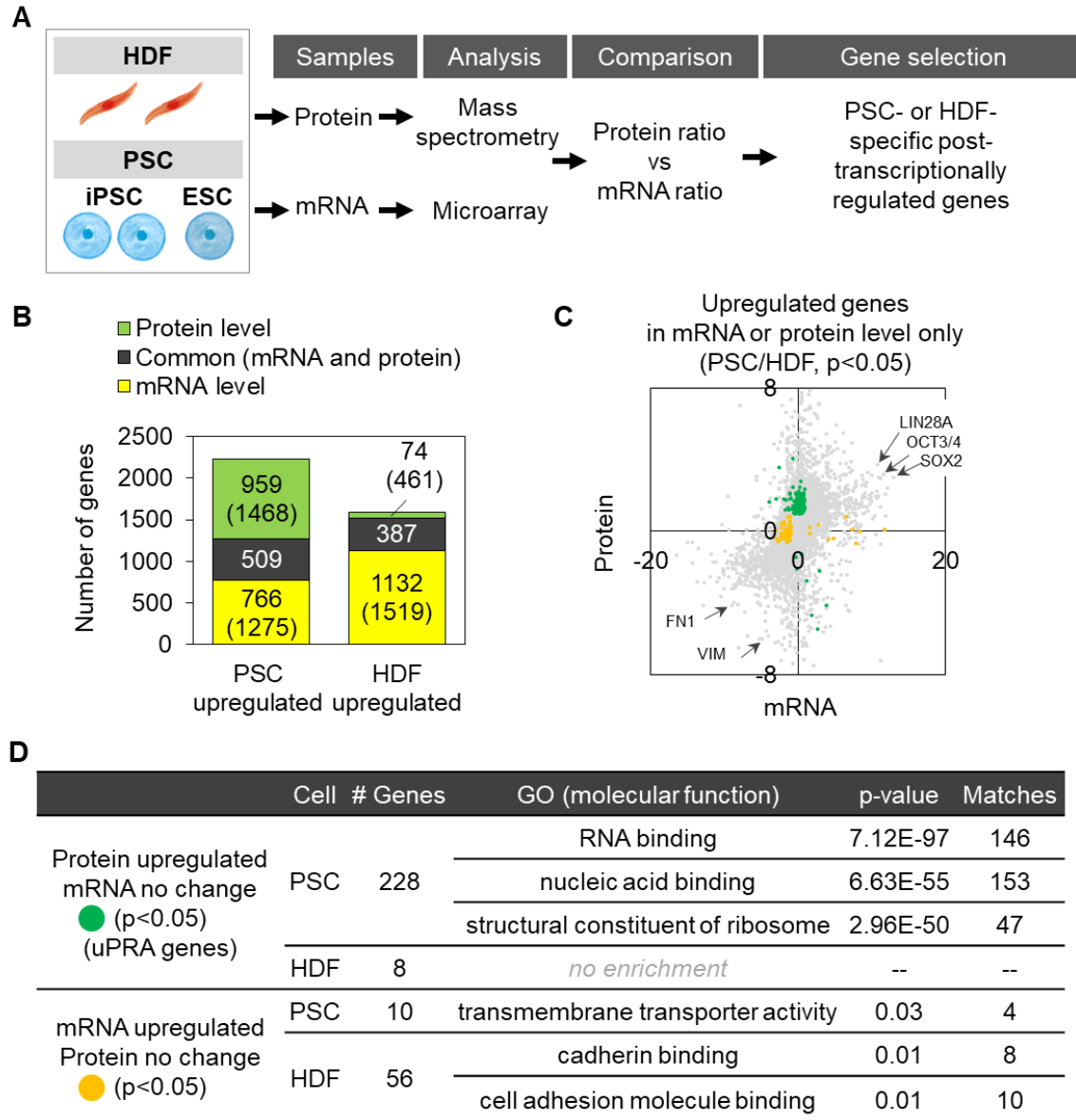
10 We focused on post-transcriptionally regulated genes essential for the maintenance
11 of PSCs. Although we showed critical regulatory steps for controlling uPRA protein
12 levels, a wide variety of regulatory mechanisms seem to be involved. Recent studies
13 have shown that the secondary structure of mRNA is important for regulating protein
14 expression by changing the interaction with the expansion segment of ribosomal
15 RNA (Leppek et al., 2020) and changing the mRNA half-life (Mauger et al., 2019).
16 Future studies are needed to determine if secondary structures affect the translation
17 of the 20 PSC-uPRA genes.

18

19

1 Figure Legends

Iwasaki *et al.*, Figure 1

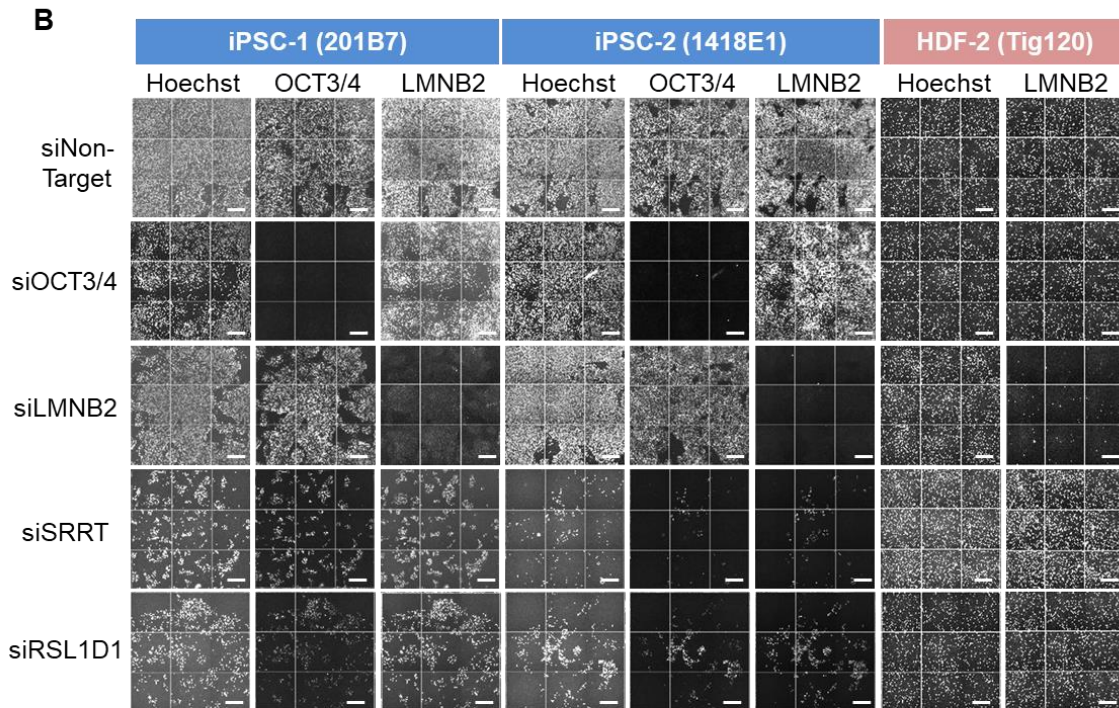
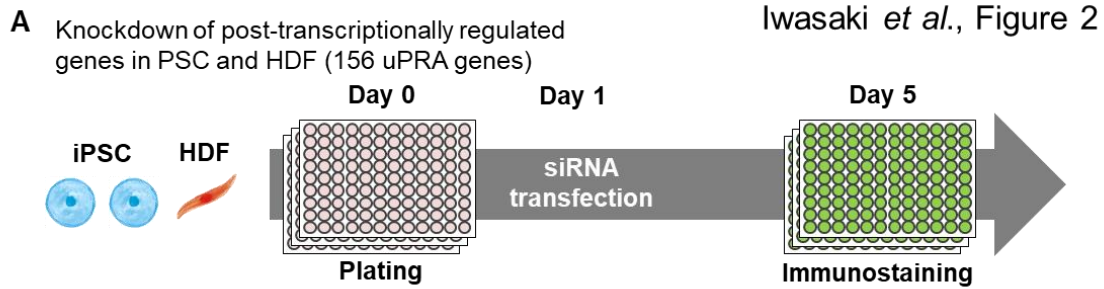


2

3 **Figure 1 | 228 and 8 genes were respectively identified as post-**
 4 **transcriptionally regulated genes with independent protein upregulation in**
 5 **PSCs and HDFs**

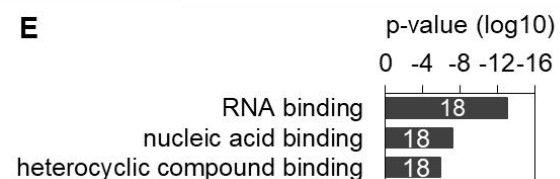
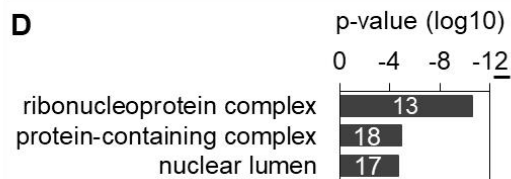
6 A. Workflow for identifying post-transcriptionally regulated genes using mass
 7 spectrometry (MS) for proteins and microarray for mRNA. Two iPSC lines (iPSC-

- 1 1, 201B7; and iPSC-2, 1418E1), one ESC line (H9), and two HDF lines (HDF-1,
2 HDF1388; and HDF-2, Tig120) were used.
- 3 B. Number of commonly upregulated genes by protein levels only, mRNA levels
4 only, and both levels in PSCs and HDFs. Numbers in parentheses show the total
5 numbers.
- 6 C. The mRNA to protein expression ratio in PSCs and HDFs for each gene. Gray,
7 yellow, and green indicate all genes, those with only upregulated mRNA, and
8 those with only upregulated protein levels, respectively ($p < 0.05$, two-sample
9 unpaired t test). Log₂ scale. Comparisons of the mRNA and protein ratios
10 between all PSC and HDF lines are shown in Figure S1. The total number of
11 genes quantified was 6404. All plotted data are shown in Table S1. See “Trans-
12 omics data analysis” in Methods for details.
- 13 D. Number of genes and the GO analysis of molecular functions for genes that were
14 significantly upregulated only at the protein or mRNA level in C.
- 15



C

| Number of uPRA genes | iPSC-1 | iPSC-2 | HDF-2 |
|----------------------------------|--------|--------|-------|
| Cell number decreased | 84 | 80 | 66 |
| Cell number decreased (common) | 52 | | |
| Cell number decreased (specific) | 20 | | 1 |

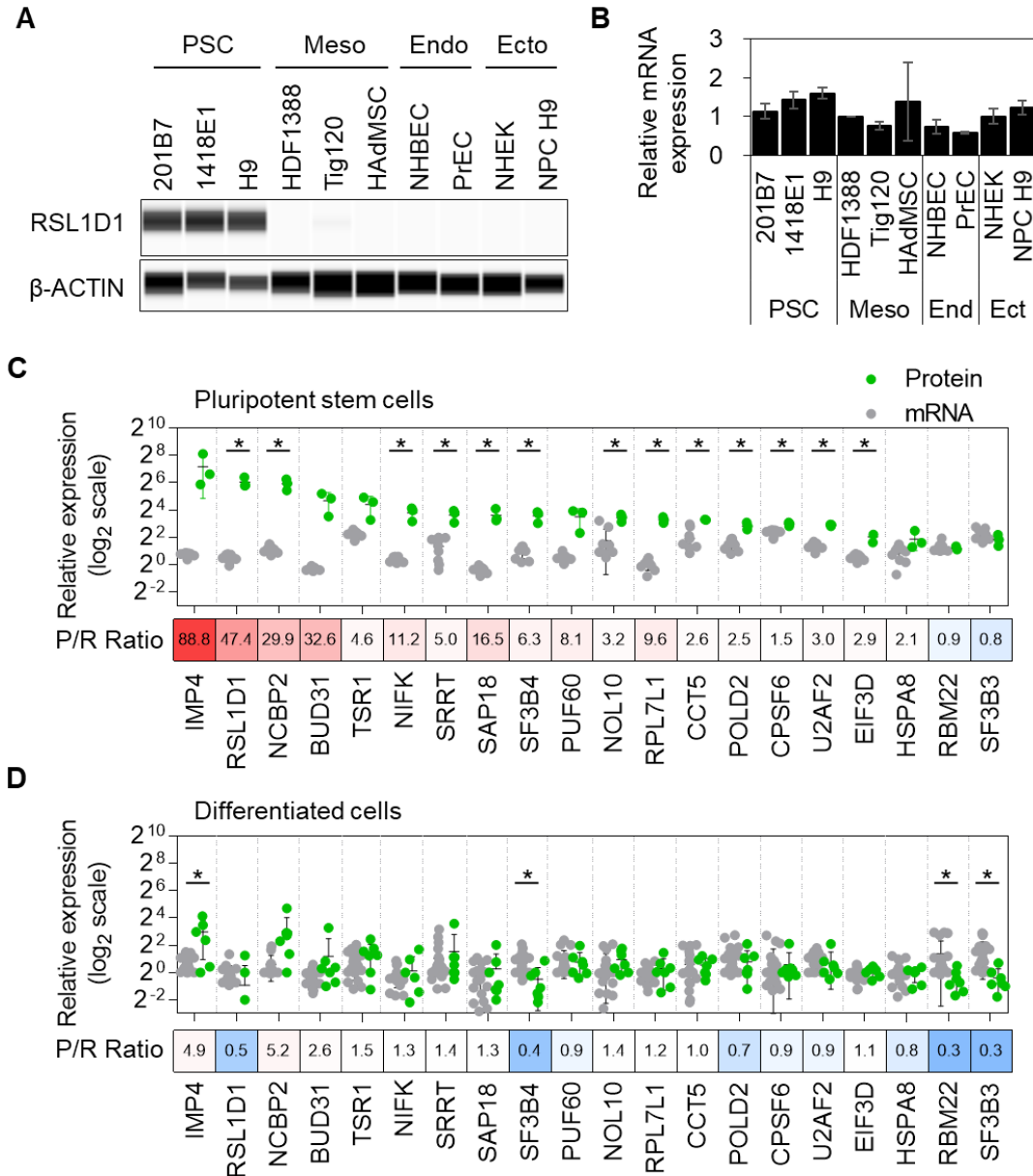


1
2 **Figure 2 | 20 uPRA genes in PSCs and 1 uPRA gene in HDFs were related to**
3 **cell survival**

4 A. Workflow of the knockdown experiment by siRNA for uPRA genes in PSCs and

- 1 HDFs (156 uPRA genes in total). iPSC-1 (201B7), iPSC-2 (1418E1), and HDF-
2 2 (Tig120) were used for the assay. See “siRNA screening” in Methods for details.
3 The knockdown targets are shown in Table S2, and the immunoassay cell
4 percentage and observed number of cells are shown in Table S3.
- 5 B. Representative images of the siRNA knockdown experiment. Nuclei were
6 visualized by Hoechst 33342, and PSCs were visualized by OCT3/4 expression.
7 The knockdown efficiency was visualized by OCT3/4 and LMNB2 expression.
8 siSRRT and siRSL1D1 are representative images for siRNAs experiments in
9 which the cell number decreased. Bars indicate 200 μ m.
- 10 C. The number of post-transcriptionally regulated genes (uPRA genes) whose
11 knockdown caused a substantial decrease in cell number. Images for the siRNAs
12 of the total 21 cell-specific uPRA genes are shown in Figure S2.
- 13 D. GO analysis of cellular component for the 20 PSC-uPRA genes in C.
- 14 E. GO analysis of molecular function for the 20 PSC-uPRA genes in C.
- 15

Iwasaki *et al.*, Figure 3



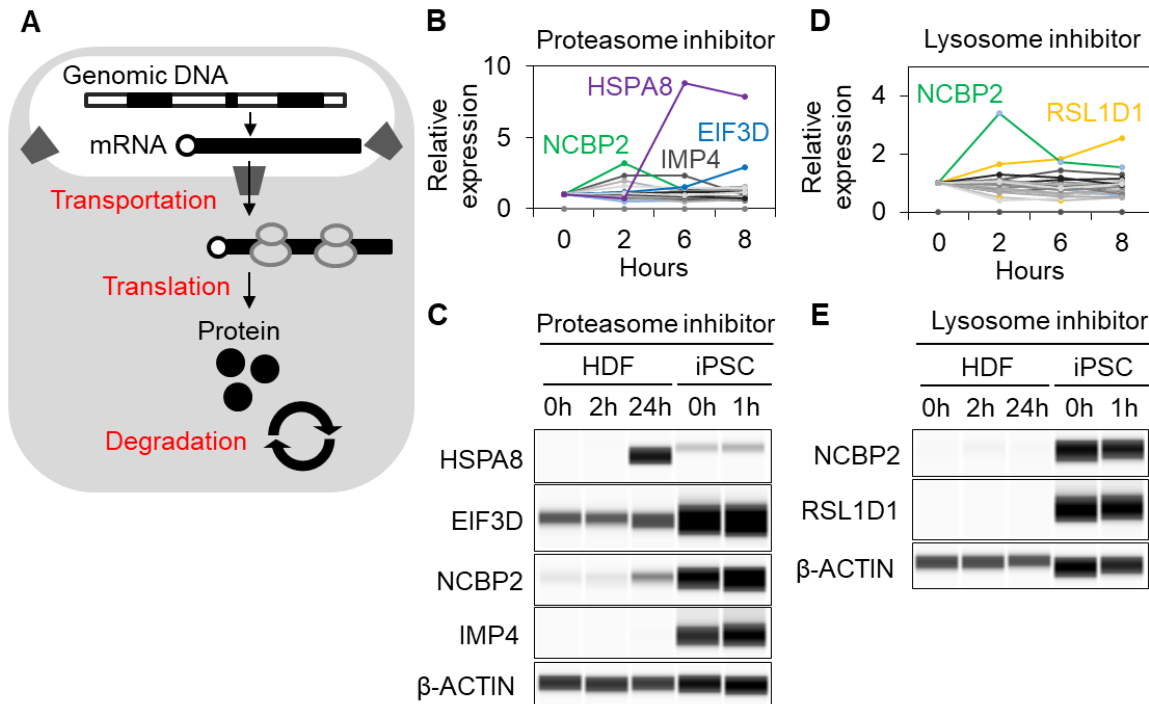
1

2 **Figure 3 | 20 uPRA genes were post-transcriptionally regulated only in PSCs**

3 A. A representative image of the immunoblotting for one PSC-uPRA gene, RSL1D1,
 4 in various cell types. We used three cell lines for PSCs, three cell lines for
 5 mesoderm-derived cells, two cell lines for endoderm-derived cells, and two
 6 lines for ectoderm-derived cells for the gene expression analysis.

- 1 Immunoblotting images of all 20 essential PSC-uPRA genes are shown in Figure
2 S3A.
- 3 B. Gene expression measured by qRT-PCR was normalized by GAPDH. The
4 relative expression ratios were normalized to the result of HDF1388 (HDF-1).
5 Representative results for the RSL1D1 mRNA are shown as the mean \pm S.D.
6 Biological triplicates for the mRNA quantification. Results for all essential PSC-
7 uPRA genes are shown in Figure S3C.
- 8 C. Relative gene expressions of mRNA and protein levels in PSCs for the 20
9 essential PSC-uPRA genes normalized to HDF1388 (HDF-1). A heatmap of the
10 protein to mRNA expression ratio (P/R ratio) is shown above each gene name.
11 * $p < 0.05$, two-sample unpaired t test.
- 12 D. Relative mRNA and protein levels in differentiated cells for the 20 essential PSC-
13 uPRA genes normalized to HDF1388 (HDF-1). A heatmap of the P/R ratio is
14 shown above each gene name. * $p < 0.05$, two-sample unpaired t test.
15

Iwasaki *et al.*, Figure 4



1

2 **Figure 4 | HSPA8, EIF3D, and NCBP2 protein expressions are controlled at the**
 3 **ubiquitin-dependent degradation stage**

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

5 B. HDF-2 (Tig120) was treated with a proteasome inhibitor (20 μ M MG-132) for up
 6 to 8 hours, and the expression levels of the 20 essential PSC-uPRA genes were
 7 measured by immunoblotting. The relative expression of each protein compared
 8 to pre-inhibition is shown. All immunoblotting images are shown in Figure S4A.

9 C. HDF-2 (Tig120) and iPSC-1 (201B7) were treated with 20 μ M MG-132 for up to
 10 24 hours, and the protein expressions for HSPA8, EIF3D, NCBP2, and IMP4
 11 were measured by immunoblotting.

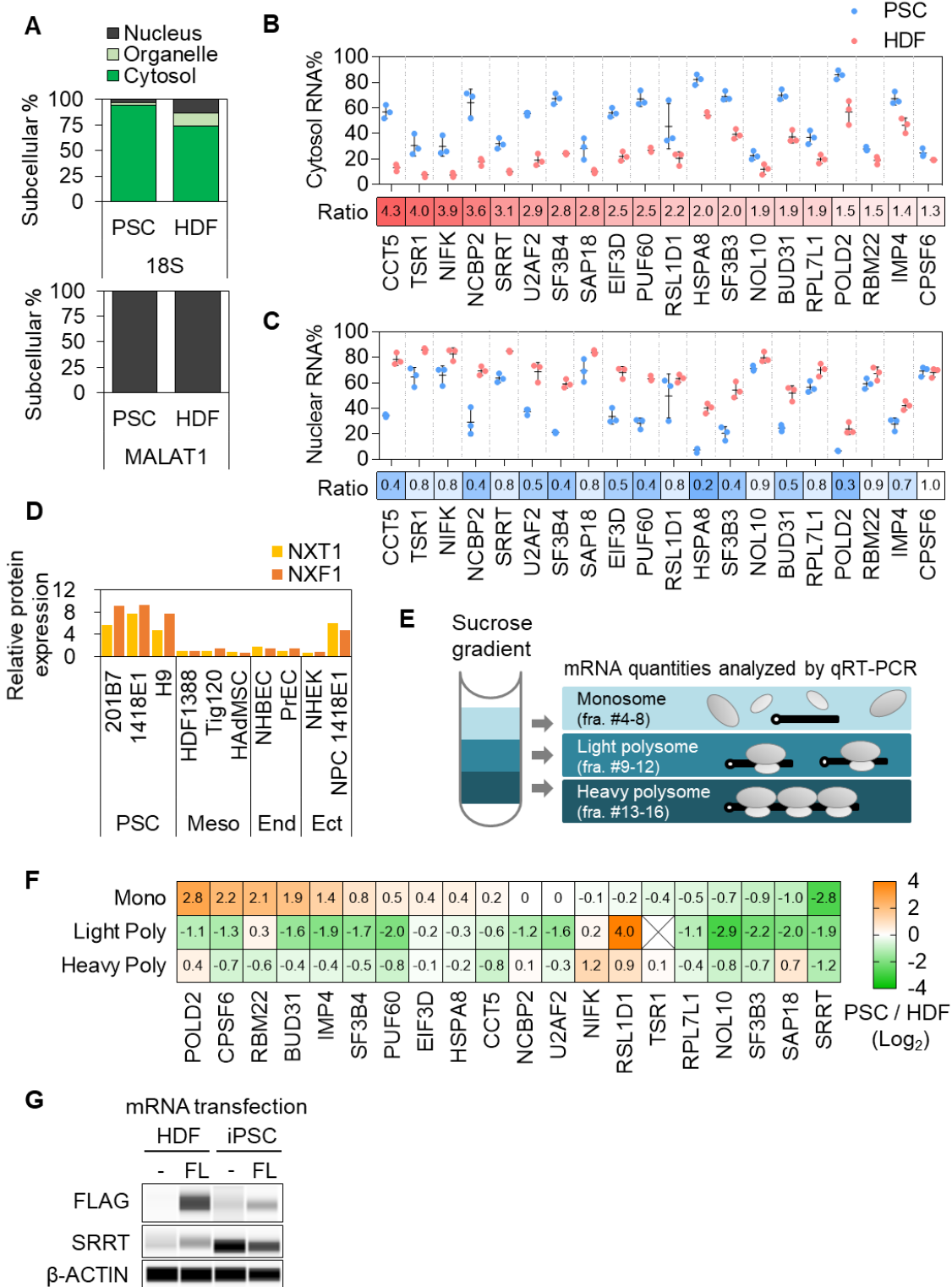
12 D. HDF-2 (Tig120) was treated with lysosome inhibitors (250 nM Bafilomycin A1
 13 and 500 nM Wortmannin) for up to 8 hours, and the protein expression of the 20
 14 essential PSC-uPRA genes were measured by immunoblotting. The relative
 15 expression of each protein to pre-inhibition is shown. All immunoblotting images

1 are shown in Figure S4B.

2 E. HDF-2 (Tig120) and iPSC-1 (201B7) were treated with 250 nM Bafilomycin and
3 500 nM Wortmannin for up to 24 hours, and the protein expressions of NCBP2
4 and RSL1D1 were measured by immunoblotting.

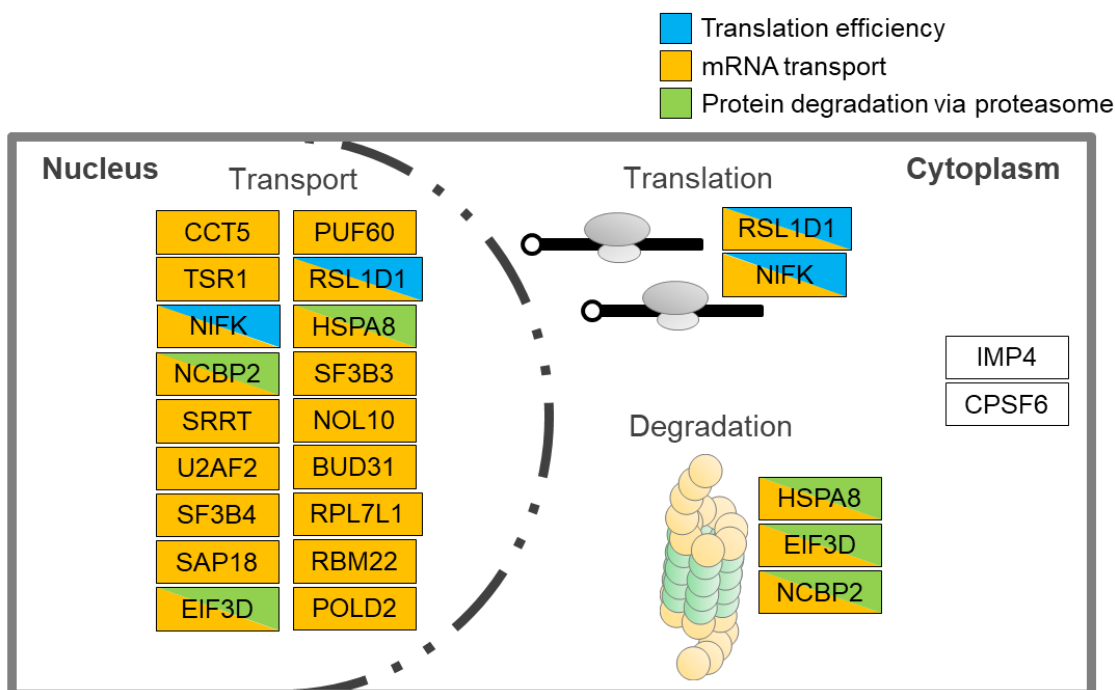
5

Iwasaki *et al.*, Figure 5



- 1 **Figure 5 | PSC-uPRA mRNAs are retained in the nucleus of HDFs without**
2 **changing the translation efficiency**
- 3 A. The percentages of mRNA in the cytoplasm, membrane, and nucleus of iPSC-1
4 (201B7) and HDF-2 (Tig120) are shown for 18S ribosomal RNA and MALAT1
5 lncRNA measured by qRT-PCR.
- 6 B. The percentage of mRNA in the cytoplasm measured by qRT-PCR for the 20
7 essential PSC-uPRA genes. The cytoplasm mRNA expression ratio in PSCs to
8 HDFs is shown above each gene name in the heatmap.
- 9 C. The percentage of mRNA in the nucleus measured by qRT-PCR for the 20
10 essential PSC-uPRA genes. The nucleus mRNA expression ratio in PSCs to
11 HDFs is shown above each gene name in the heatmap.
- 12 D. The protein-fold change between PSCs and other differentiated cells for NXT1
13 and NXF1. Protein expressions measured by immunoblotting were normalized
14 by b-ACTIN. The relative expression ratios were normalized to the result of
15 HDF1388 (HDF-1).
- 16 E. Overview of the analysis of the translation efficiency by the sucrose gradient.
17 The monosome (fraction 4–8), light polysome (fraction 9–12), and heavy
18 polysome (fraction 13–16) were collected to extract RNAs, which were
19 measured by qRT-PCR. See “Monosome and polysome fractionation” in
20 Methods for details.
- 21 F. The percentage of mRNAs in monosome (Mono), light polysome (Light Poly),
22 and heavy polysome (Heavy Poly) were analyzed by qRT-PCR. The average
23 mRNA expression ratio in PSCs to HDFs is shown in the heatmap with log₂ scale
24 (n=3). Values are normalized by spike RNA and compared with the loading
25 sample before the sucrose gradient.
- 26 G. Full-length FLAG-SRRT mRNA was transfected into HDF-2 (Tig120) and iPSC-
27 1 (201B7). Immunoblotting assays for FLAG, SRRT, and b-ACTIN are shown.
28

Iwasaki *et al.*, Figure 6



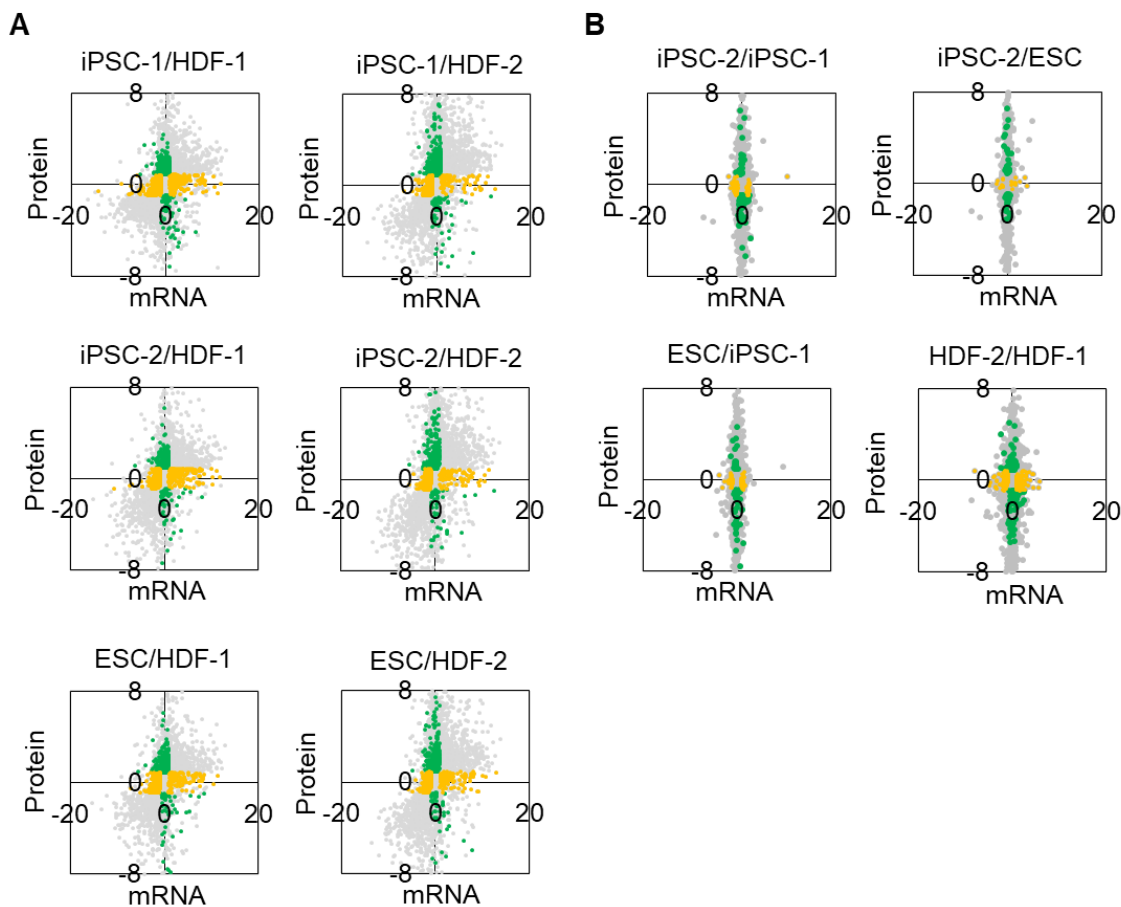
1
2 **Figure 6 | An illustration of how 20 essential PSC-uPRA genes are regulated in**
3 **PSCs and differentiated cells**

4 Blue, yellow, and green indicate the translation efficiency, mRNA transport, and
5 protein degradation via the proteasome of the 20 essential PSC-uPRA genes,
6 respectively. More mRNA of 18 uPRA genes was transported from the nucleus
7 to the cytoplasm in PSCs than in HDFs. The proteins of HSPA8, EIF3D, and
8 NCBP2 were quickly degraded in HDFs at the proteasome degradation stage.
9 The translation efficiency of RSL1D1 and NIFK was higher in PSCs compared
10 to HDFs. No regulation was observed for IMP4 or CPSF6.

11

1 Supporting Information

Iwasaki *et al.*, Supplementary Figure 1



2

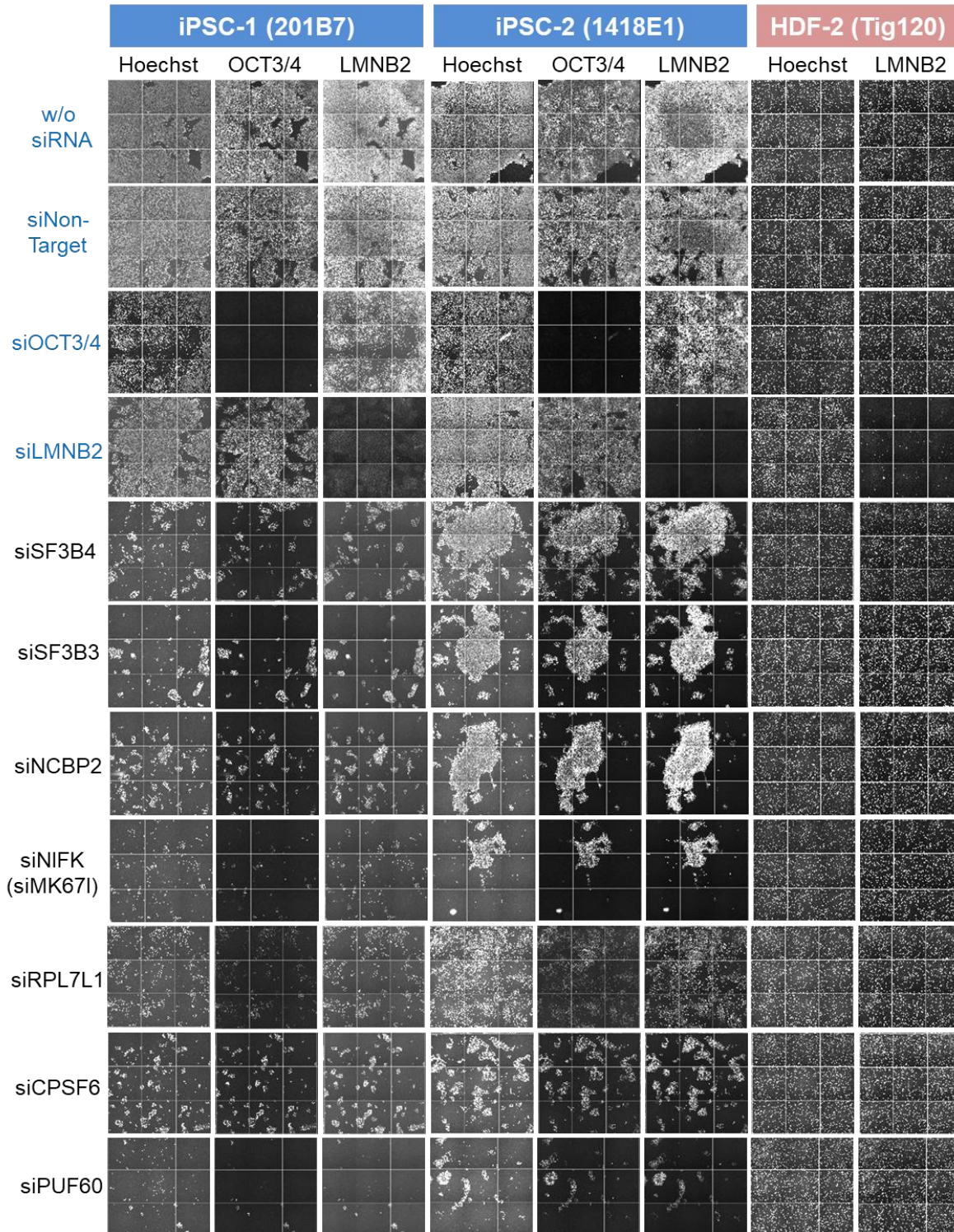
3 Figure S1 | Comparison of the mRNA to protein ratio between cell lines

4 A. The mRNA to protein ratio for each gene between different PSC and HDF
5 lines. Genes with significantly upregulated mRNA and protein levels only are shown
6 in yellow and green, respectively ($p < 0.05$, two sample paired t test).

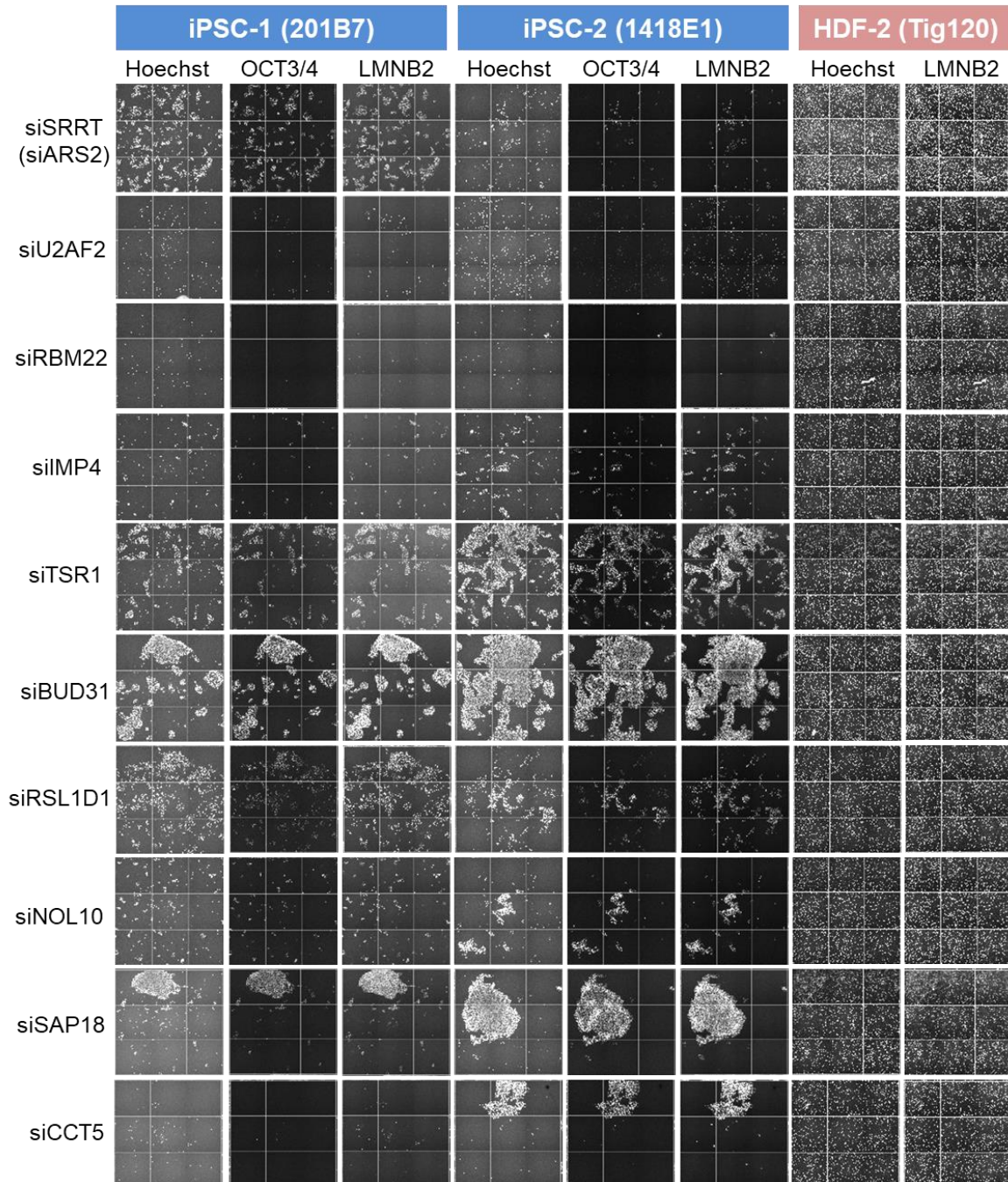
7 B. The mRNA to protein ratio for each gene between the same cell type. Genes
8 with significantly upregulated mRNA and protein levels only are shown in yellow and
9 green, respectively ($p < 0.05$, two sample paired t test).

10

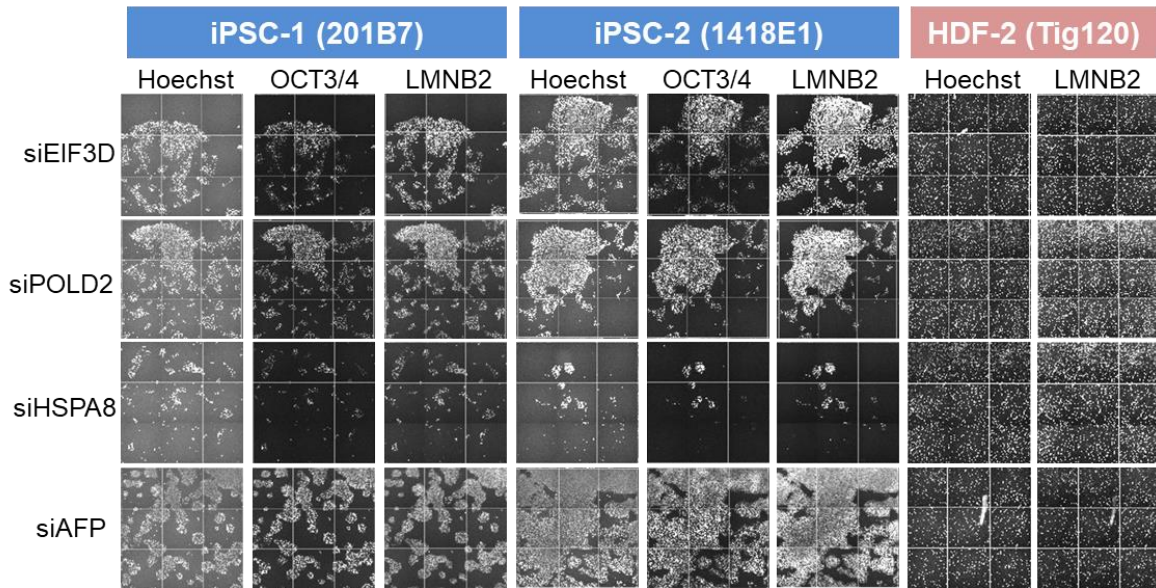
Iwasaki *et al.*, Supplementary Figure 2



Iwasaki *et al.*, Supplementary Figure 2



Iwasaki *et al.*, Supplementary Figure 2



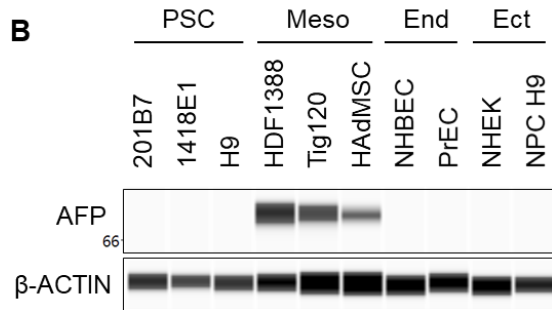
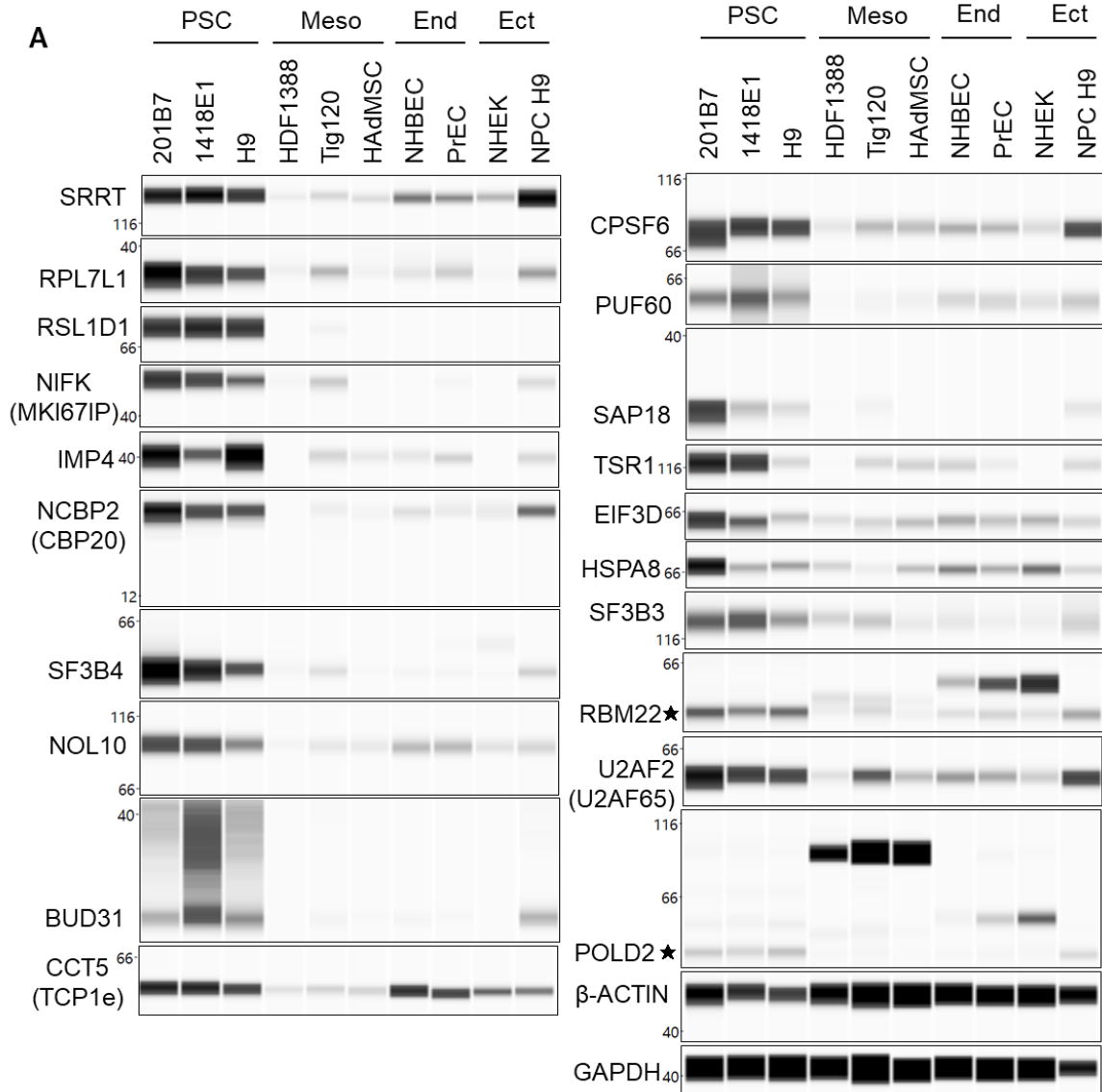
1

2 Figure S2 | Images of 20 PSC-uPRA genes and 1 HDF-uPRA gene whose knockout
3 caused a decrease in cell survival.

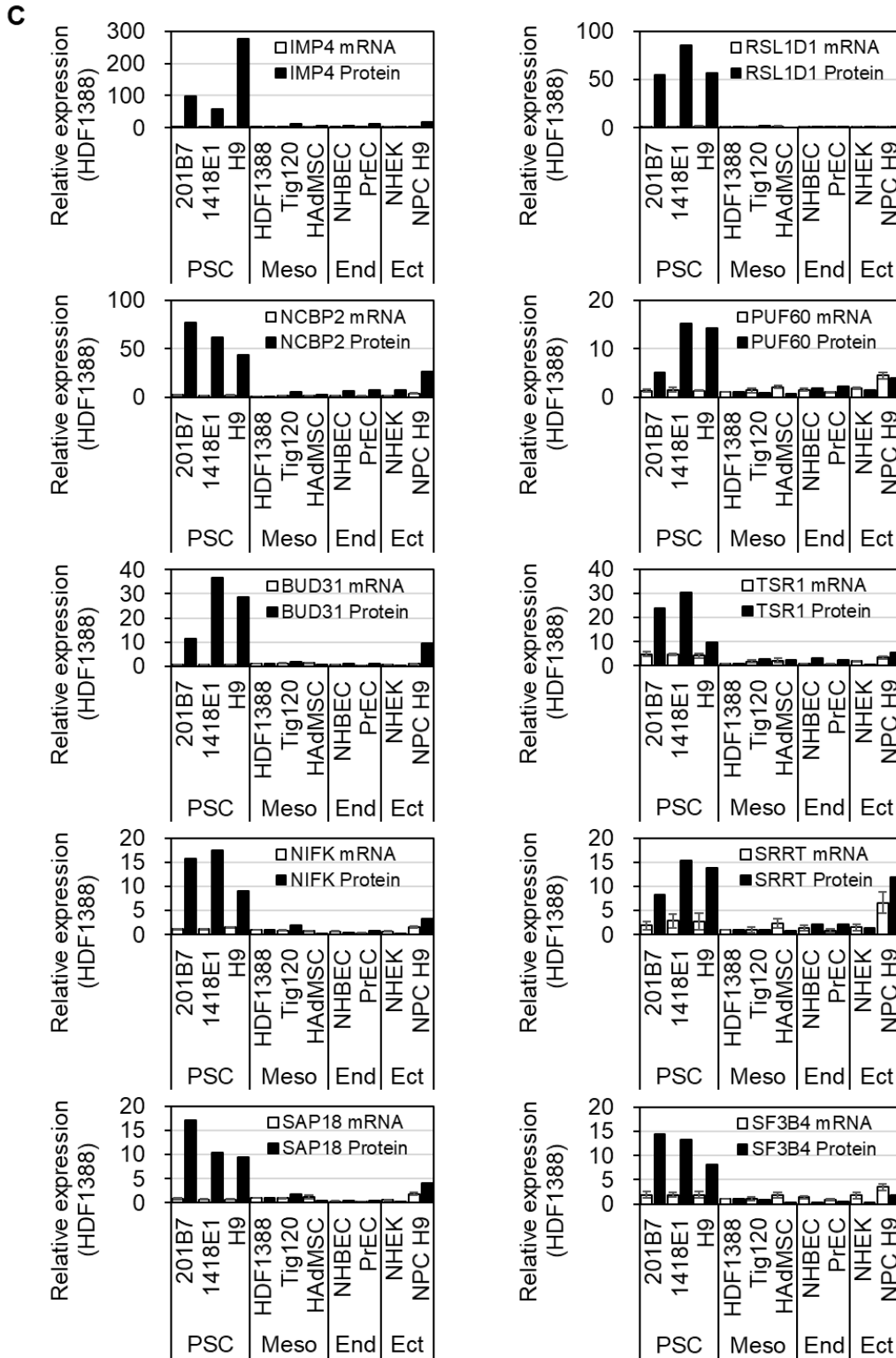
4 Hoechst 33342-stained nuclei, OCT3/4-stained PSCs, and LaminB2 (LMNB2)-
5 stained PSCs and HDFs are shown. Knockdown efficiency was visualized by siRNA
6 for OCT3/4 and LaminB2. The 20 essential PSC-uPRA genes are SF3B4, SF3B3,
7 NCBP2, NIFK, RPL7L1, CPSF6, PUF60, SRRT, U2AF2, RBM22, IMP4, TSR1,
8 BUD31, RSL1D1, NOL10, SAP18, CCT5, EIF3D, POLD2, and HSPA8. The one
9 essential HDF-uPRA gene is AFP.

10

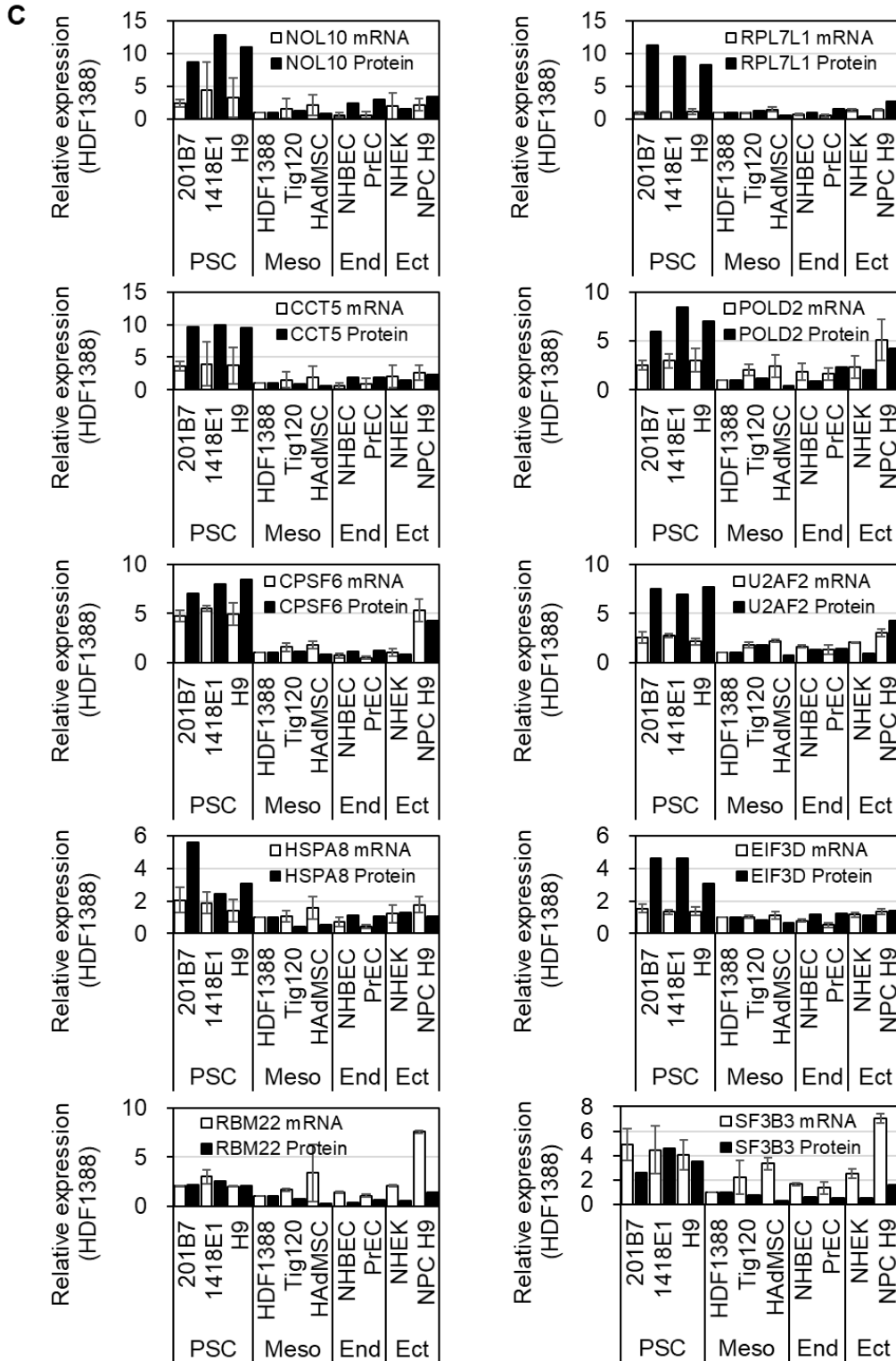
Iwasaki *et al.*, Supplementary Figure 3



Iwasaki *et al.*, Supplementary Figure 3



Iwasaki *et al.*, Supplementary Figure 3



1 Figure S3 | Immunoblotting images and relative expressions of the 20 uPRA genes
2 and one HDF-uPRA gene.

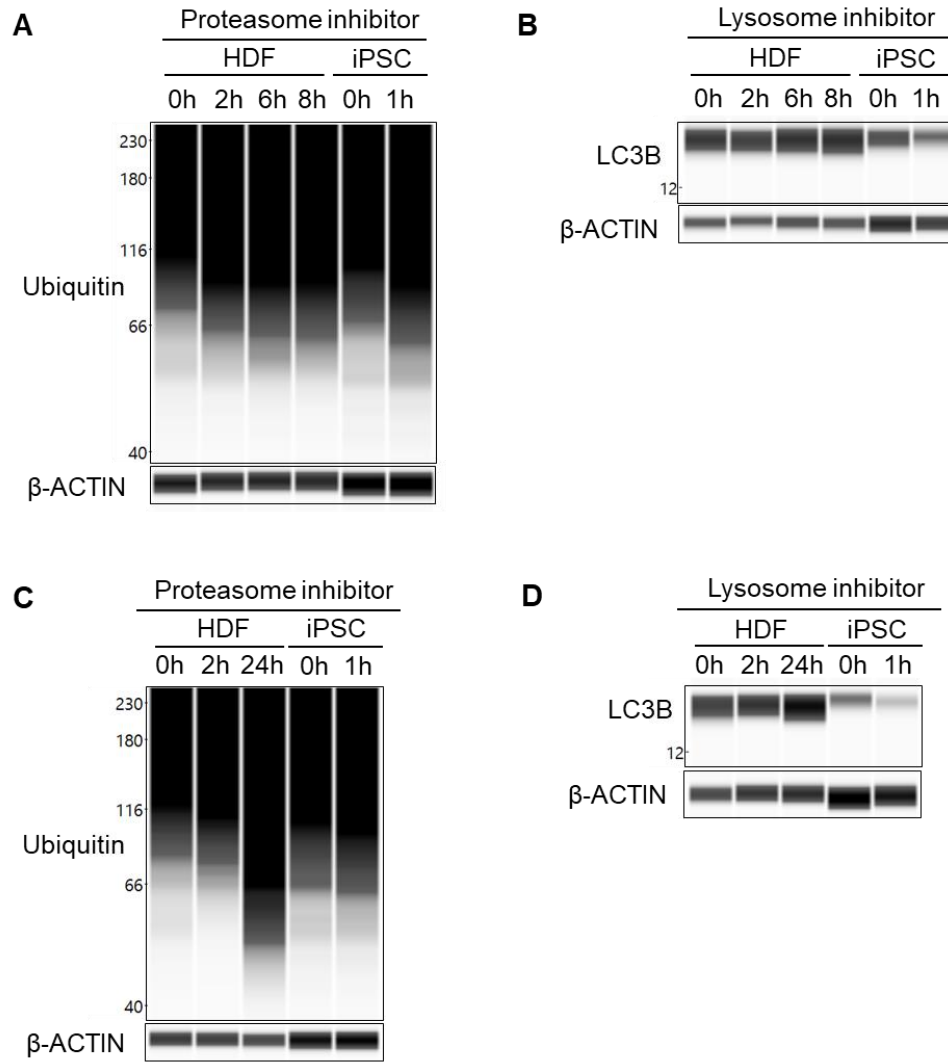
3 A. Immunoblotting images of the 20 essential PSC-uPRA genes. Meso,
4 mesoderm; End, endoderm; Ect, ectoderm. Three cell lines were used for PSCs,
5 three cell lines for mesoderm-derived cells, two cell lines for endoderm-derived cells,
6 and two cell lines for ectoderm-derived cells.

7 B. Immunoblotting image of the 1 essential HDF-uPRA gene (AFP). We could
8 not quantify the mRNA by qRT-PCR because of the low mRNA level.

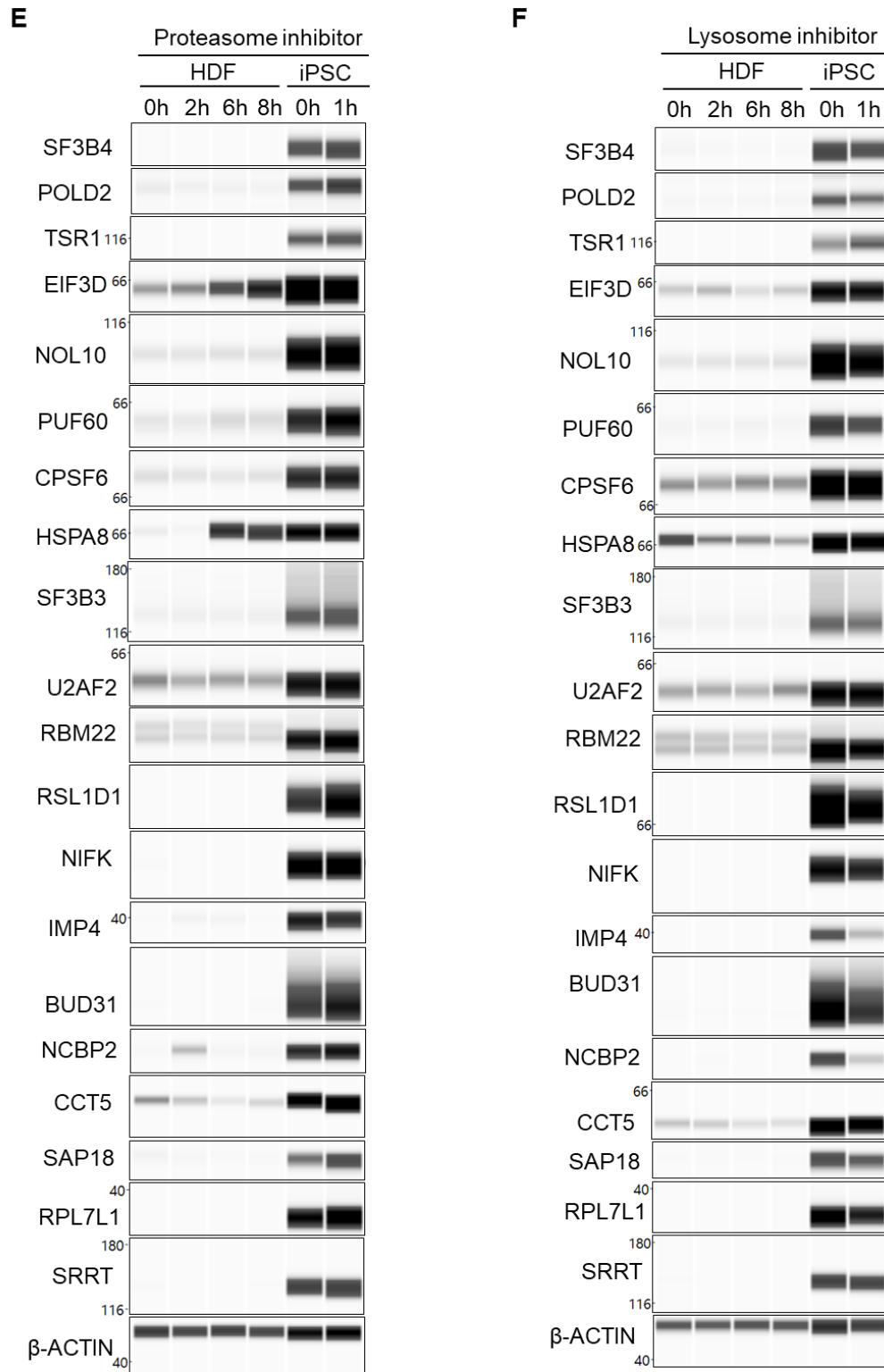
9 C. Gene expressions measured by qRT-PCR and immunoblotting normalized
10 to HDF-1 (HDF1388). For the mRNA quantification, GAPDH was used for the
11 normalization, and bars represent mean \pm S.D., biological triplicate. For the protein
12 quantification, β -ACTIN was used for the normalization.

13

Iwasaki *et al.*, Supplementary Figure 4

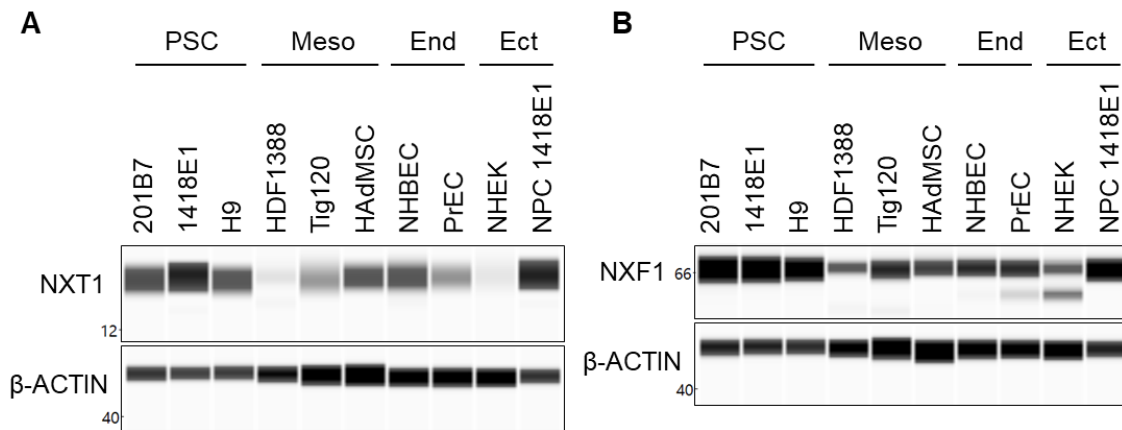


Iwasaki *et al.*, Supplementary Figure 4



- 1 Figure S4 | Immunoblotting images of the proteasome inhibitor and lysosomal
2 inhibitor assays.
- 3 A. Ubiquitin protein levels measured by immunoblotting in iPSC-1 (201B7) and
4 HDF-2 (Tig120) samples treated with a proteasome inhibitor (20 μ M MG-132) for up
5 to 8 hours as a control for the proteasome inhibitor assay.
- 6 B. LC3B protein levels measured by immunoblotting in iPSC-1 (201B7) and
7 HDF-2 (Tig120) samples treated with lysosomal inhibitors (250 nM Bafilomycin A1
8 and 500 nM Wortmannin) for up to 8 hours as a control for the lysosome inhibitor
9 assay.
- 10 C. iPSC-1 (201B7) and HDF-2 (Tig120) were treated with 20 μ M MG-132 for
11 up to 24 hours, and ubiquitin protein levels were measured by immunoblotting.
- 12 D. iPSC-1 (201B7) and HDF-2 (Tig120) were treated with 250 nM Bafilomycin
13 A1 and 500 nM Wortmannin for up to 24 hours, and LC3B protein levels were
14 measured by immunoblotting.
- 15 E. The gene expression levels for the 20 essential PSC-uPRA genes were
16 measured by immunoblotting with the same samples used in A.
- 17 F. The gene expression levels for the 20 essential PSC-uPRA genes were
18 measured by immunoblotting with the same samples used in B.
- 19
20

Iwasaki *et al.*, Supplementary Figure 5



1
2 Figure S5 | Immunoblotting images of the NXT1 and NXF1 protein expressions
3 between different cell lines

4 Immunoblots of NXT1 (A) and NXF1 (B) for the same samples used in Figure S3
5 except for NPC 1418E1. Meso, mesoderm; End, endoderm; Ect, ectoderm.

6
7 Table S1. Quantified gene list of the mRNA and protein expressions in PSCs and
8 HDFs.

9 Table S2. List of siRNA used for Figure 2.

10 Table S3. Number of cells and percentage of immunoassayed proteins following
11 siRNA treatment for 156 targeted genes, related to Figure 2A and B.

12 Table S4. DNA sequence of the transfected full length FLAG-SRRT.

13
14 **ACKNOWLEDGMENTS**

15 We would like to thank Y. Fujita and T. Matsushita for the technical assistance, P.
16 Karagiannis for editorial assistance, and S. Takeshima for administrative support.

17 This work was supported by K-CONNEX from the Japan Science and Technology
18 Agency (JST), Core Center for iPS Cell Research from Japan Agency for Medical
19 Research and Development (AMED; JP21bm0104001), AMED-PRIME
20 (21gm6410003h0001 (M.I.)), and the Japanese Society for the Promotion of

1 Science KAKENHI Grant (JSPS; 19K16104 (M.I.)).

2

3 **AUTHOR CONTRIBUTIONS**

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

9

10 **DECLARATION OF INTEREST**

11 M.I. is a scientific adviser (without salary) of xFOREST Therapeutics. H.S. is an
12 outside director of aceRNA Technologies, Co., Ltd. K.T. is on the scientific advisory
13 board (without salary) of I Peace, Inc. S.Y. is a scientific advisor (without salary) of
14 iPS Academia Japan.

15

16

17

1 **REFERENCES**

- 2 Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. *Nat Rev Genet*
3 *15*, 163-175. 10.1038/nrg3662.
- 4 Buccitelli, C., and Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene
5 expression control. *Nat Rev Genet* *21*, 630-644. 10.1038/s41576-020-0258-4.
- 6 Chandrasekaran, V., Juskiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and
7 Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. *Nat Struct*
8 *Mol Biol* *26*, 1132-1140. 10.1038/s41594-019-0331-x.
- 9 Chen, Q., and Hu, G. (2017). Post-transcriptional regulation of the pluripotent state. *Curr Opin*
10 *Genet Dev* *46*, 15-23. 10.1016/j.gde.2017.06.010.
- 11 D'Orazio, K.N., and Green, R. (2021). Ribosome states signal RNA quality control. *Mol Cell* *81*, 1372-
12 1383. 10.1016/j.molcel.2021.02.022.
- 13 Darnell, A.M., Subramaniam, A.R., and O'Shea, E.K. (2018). Translational Control through
14 Differential Ribosome Pausing during Amino Acid Limitation in Mammalian Cells. *Mol Cell* *71*, 229-
15 243.e211. 10.1016/j.molcel.2018.06.041.
- 16 Dendooven, T., Luisi, B.F., and Bandyra, K.J. (2020). RNA lifetime control, from stereochemistry to
17 gene expression. *Curr Opin Struct Biol* *61*, 59-70. 10.1016/j.sbi.2019.10.002.
- 18 Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.K., Alvarez,
19 M., Talukder, S., Pan, Q., Mazzone, E.O., et al. (2011). An alternative splicing switch regulates
20 embryonic stem cell pluripotency and reprogramming. *Cell* *147*, 132-146. 10.1016/j.cell.2011.08.023.
- 21 Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael,
22 I.P., Nachman, E.N., et al. (2013). MBNL proteins repress ES-cell-specific alternative splicing and
23 reprogramming. *Nature* *498*, 241-245. 10.1038/nature12270.
- 24 Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding
25 proteins. *Nat Rev Mol Cell Biol* *19*, 327-341. 10.1038/nrm.2017.130.
- 26 Kainov, Y.A., and Makeyev, E.V. (2020). A transcriptome-wide antitermination mechanism sustaining
27 identity of embryonic stem cells. *Nat Commun* *11*, 361. 10.1038/s41467-019-14204-z.
- 28 Khan, Z., Ford, M.J., Cusanovich, D.A., Mitrano, A., Pritchard, J.K., and Gilad, Y. (2013). Primate
29 transcript and protein expression levels evolve under compensatory selection pressures. *Science* *342*,
30 1100-1104. 10.1126/science.1242379.
- 31 Lee, A.S., Kranzusch, P.J., Doudna, J.A., and Cate, J.H. (2016). eIF3d is an mRNA cap-binding protein
32 that is required for specialized translation initiation. *Nature* *536*, 96-99. 10.1038/nature18954.
- 33 Leppek, K., Fujii, K., Quade, N., Susanto, T.T., Boehringer, D., Lenarčič, T., Xue, S., Genuth, N.R.,

- 1 Ban, N., and Barna, M. (2020). Gene- and Species-Specific Hox mRNA Translation by Ribosome
- 2 Expansion Segments. *Mol Cell* *80*, 980-995.e913. 10.1016/j.molcel.2020.10.023.
- 3 Lu, R., Markowitz, F., Unwin, R.D., Leek, J.T., Airoidi, E.M., MacArthur, B.D., Lachmann, A., Rozov,
- 4 R., Ma'ayan, A., Boyer, L.A., et al. (2009). Systems-level dynamic analyses of fate change in murine
- 5 embryonic stem cells. *Nature* *462*, 358-362. 10.1038/nature08575.
- 6 Matsumoto, M., Matsuzaki, F., Oshikawa, K., Goshima, N., Mori, M., Kawamura, Y., Ogawa, K.,
- 7 Fukuda, E., Nakatsumi, H., Natsume, T., et al. (2017). A large-scale targeted proteomics assay
- 8 resource based on an in vitro human proteome. *Nat Methods* *14*, 251-258. 10.1038/nmeth.4116.
- 9 Mauger, D.M., Cabral, B.J., Presnyak, V., Su, S.V., Reid, D.W., Goodman, B., Link, K., Khatwani, N.,
- 10 Reynders, J., Moore, M.J., and McFadyen, I.J. (2019). mRNA structure regulates protein expression
- 11 through changes in functional half-life. *Proc Natl Acad Sci U S A* *116*, 24075-24083.
- 12 10.1073/pnas.1908052116.
- 13 Ohta, S., Nishida, E., Yamanaka, S., and Yamamoto, T. (2013). Global splicing pattern reversion
- 14 during somatic cell reprogramming. *Cell Rep* *5*, 357-366. 10.1016/j.celrep.2013.09.016.
- 15 Palazzo, A.F., and Lee, E.S. (2018). Sequence Determinants for Nuclear Retention and Cytoplasmic
- 16 Export of mRNAs and lncRNAs. *Front Genet* *9*, 440. 10.3389/fgene.2018.00440.
- 17 Pelletier, J., Thomas, G., and Volarević, S. (2018). Ribosome biogenesis in cancer: new players and
- 18 therapeutic avenues. *Nat Rev Cancer* *18*, 51-63. 10.1038/nrc.2017.104.
- 19 Pohl, C., and Dikic, I. (2019). Cellular quality control by the ubiquitin-proteasome system and
- 20 autophagy. *Science* *366*, 818-822. 10.1126/science.aax3769.
- 21 Roumeliotis, T.I., Williams, S.P., Gonçalves, E., Alsinet, C., Del Castillo Velasco-Herrera, M., Aben,
- 22 N., Ghavidel, F.Z., Michaut, M., Schubert, M., Price, S., et al. (2017). Genomic Determinants of
- 23 Protein Abundance Variation in Colorectal Cancer Cells. *Cell Rep* *20*, 2201-2214.
- 24 10.1016/j.celrep.2017.08.010.
- 25 Salomonis, N., Schlieve, C.R., Pereira, L., Wahlquist, C., Colas, A., Zambon, A.C., Vranizan, K.,
- 26 Spindler, M.J., Pico, A.R., Cline, M.S., et al. (2010). Alternative splicing regulates mouse embryonic
- 27 stem cell pluripotency and differentiation. *Proc Natl Acad Sci U S A* *107*, 10514-10519.
- 28 10.1073/pnas.0912260107.
- 29 Teixeira, F.K., and Lehmann, R. (2019). Translational Control during Developmental Transitions.
- 30 *Cold Spring Harb Perspect Biol* *11*. 10.1101/cshperspect.a032987.
- 31 Tresenrider, A., Morse, K., Jorgensen, V., Chia, M., Liao, H., van Werven, F.J., and Ünal, E. (2021).
- 32 Integrated genomic analysis reveals key features of long undecoded transcript isoform-based gene
- 33 repression. *Mol Cell* *81*, 2231-2245.e2211. 10.1016/j.molcel.2021.03.013.

- 1 Umemura, Y., Koike, N., Ohashi, M., Tsuchiya, Y., Meng, Q.J., Minami, Y., Hara, M., Hisatomi, M.,
2 and Yagita, K. (2017). Involvement of posttranscriptional regulation of Clock in the emergence of
3 circadian clock oscillation during mouse development. *Proc Natl Acad Sci U S A* *114*, E7479-E7488.
4 [10.1073/pnas.1703170114](https://doi.org/10.1073/pnas.1703170114).
- 5 Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li,
6 L.H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy
7 human tissues. *Mol Syst Biol* *15*, e8503. [10.15252/msb.20188503](https://doi.org/10.15252/msb.20188503).

8

9