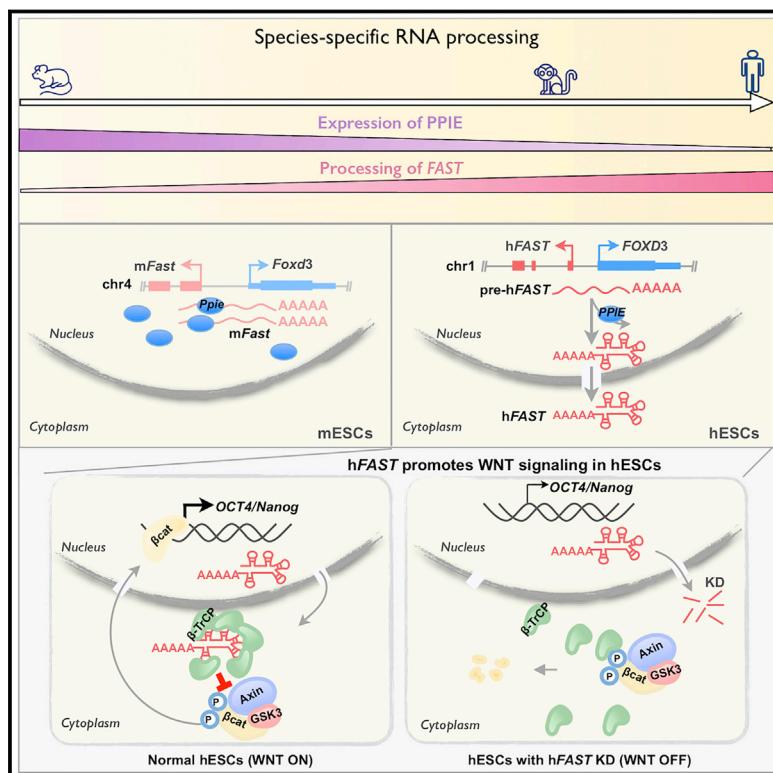


Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells

Graphical Abstract



Authors

Chun-Jie Guo, Xu-Kai Ma,
Yu-Hang Xing, ..., Gordon G. Carmichael,
Li Yang, Ling-Ling Chen

Correspondence

linglingchen@sibcb.ac.cn

In Brief

A pair of lncRNA orthologs exhibits different subcellular localization in human and murine ESCs because of differential RNA processing, which, in turn, leads to their functional divergence in the context of pluripotency regulation. The findings highlight how conserved lncRNAs may achieve functional evolution through non-conserved RNA processing.

Highlights

- Subcellular localization of conserved lncRNAs is different in hESCs and mESCs
- Cytoplasmic *hFAST* but not nuclear *mFast* promotes WNT signaling in hESC pluripotency
- PPIE regulates distinct *FAST* processing in hESCs and mESCs
- RNA processing and localization contribute to lncRNA functional evolution

Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells

Chun-Jie Guo,^{1,7} Xu-Kai Ma,^{2,7} Yu-Hang Xing,^{1,6} Chuan-Chuan Zheng,¹ Yi-Feng Xu,¹ Lin Shan,¹ Jun Zhang,¹ Shaohua Wang,³ Yangming Wang,³ Gordon G. Carmichael,⁴ Li Yang,^{2,5} and Ling-Ling Chen^{1,5,8,*}

¹State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of the Chinese Academy of Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

²CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, University of the Chinese Academy of Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

³Beijing Key Laboratory of Cardiometabolic Molecular Medicine, Institute of Molecular Medicine, Peking University, 100871 Beijing, China

⁴Department of Genetics and Genome Sciences, UCONN Health, Farmington, CT 06030, USA

⁵School of Life Science and Technology, ShanghaiTech University, 100 Haik Road, Shanghai 201210, China

⁶Present address: Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

⁷These authors contributed equally

⁸Lead Contact

*Correspondence: linglingchen@sibcb.ac.cn

<https://doi.org/10.1016/j.cell.2020.03.006>

SUMMARY

Long noncoding RNAs (lncRNAs) evolve more rapidly than mRNAs. Whether conserved lncRNAs undergo conserved processing, localization, and function remains unexplored. We report differing subcellular localization of lncRNAs in human and mouse embryonic stem cells (ESCs). A significantly higher fraction of lncRNAs is localized in the cytoplasm of hESCs than in mESCs. This turns out to be important for hESC pluripotency. *FAST* is a positionally conserved lncRNA but is not conserved in its processing and localization. In hESCs, cytoplasm-localized *hFAST* binds to the WD40 domain of the E3 ubiquitin ligase β -TrCP and blocks its interaction with phosphorylated β -catenin to prevent degradation, leading to activated WNT signaling, required for pluripotency. In contrast, *mFast* is nuclear retained in mESCs, and its processing is suppressed by the splicing factor PPIE, which is highly expressed in mESCs but not hESCs. These findings reveal that lncRNA processing and localization are previously under-appreciated contributors to the rapid evolution of function.

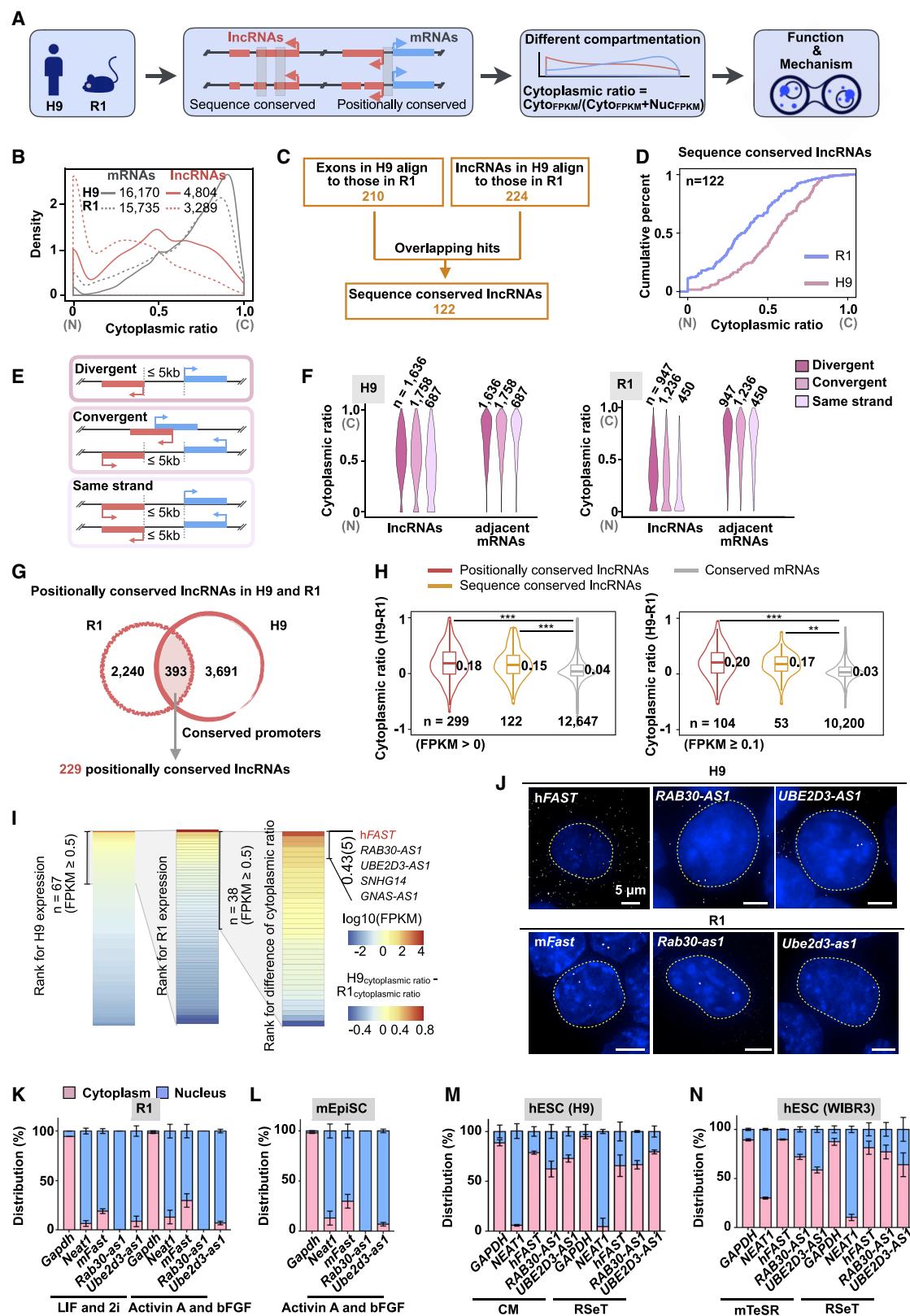
INTRODUCTION

Pervasive transcription of the eukaryotic genome leads to expression of a broad collection of protein-coding and noncoding RNAs. The conservation of sequences of mRNAs as well as those of translated proteins among species is high and functionally significant during evolution (Makalowski et al., 1996). In contrast, long noncoding RNAs (lncRNAs) in general lack high sequence (Hezroni et al., 2015; Kutter et al., 2012; Necsulea et al., 2014; Ulitsky et al., 2011) or secondary structure conserva-

tion (Kutter et al., 2012; Ulitsky, 2016). lncRNA conservation can also occur at the position and mechanism-of-action levels (Diederichs, 2014; Johnsson et al., 2014; Ulitsky, 2016). Transcription of positionally conserved lncRNAs with nearby conserved coding genes among different species has been recognized as an indicator of potential functional significance (Amaral et al., 2018; Hezroni et al., 2015; Necsulea et al., 2014; Ulitsky et al., 2011). Although mechanisms of action are thought to be associated with specific RNA conformations in cells, analysis of RNA structure still remains a challenge because of the lack of experimental data for essential structural modules (Ulitsky, 2016). Thus, although increasing numbers of lncRNAs have been recognized to play important roles in diverse cellular processes (Yao et al., 2019), concerns still remain regarding to what degree their functions are conserved.

The subcellular localization of lncRNAs is related to their function (Carlevaro-Fita and Johnson, 2019; Chen, 2016). Both cis-elements and trans-factors can affect subcellular localization (Chin and Lévy, 2017; Lubelsky and Ulitsky, 2018; Miyagawa et al., 2012; Zhang et al., 2014) and processing (Licitra and Darnell, 2010; Valencia et al., 2008), which can result in different RNA isoforms with different subcellular fates (Melé et al., 2017; Schlackow et al., 2017). So far, however, whether lncRNA processing is conserved and how processing contributes to its compartmentalization and function in different species remain unexplored.

Embryonic stem cells (ESCs) are derived from the inner cell mass of mammalian blastocyst embryos and possess the ability of self-renewal and differentiation to specific cell types. Although human ESCs (hESCs) and mouse ESCs (mESCs) share some common properties in pluripotency, including high alkaline phosphatase and telomerase activity (Koestenbauer et al., 2006), they have distinct morphologies, stemness markers, growth conditions, and extrinsic signals to maintain their pluripotent state (Koestenbauer et al., 2006; Pera and Tam, 2010). Whether lncRNAs contribute to these differences between hESCs and mESCs is unknown.



(legend on next page)

Here we profiled conserved lncRNAs from hESCs and mESCs to examine whether their processing and subcellular localization are conserved. Surprisingly, we found that the localization pattern of lncRNAs between hESCs and mESCs is different. Conserved lncRNAs in hESCs are more frequently spliced with increased cytoplasmic localization, and this is required for pluripotency regulation and modulated by differential expression of *trans*-factors. Such distinct RNA processing, localization, and subsequent function have not been sufficiently appreciated previously.

RESULTS

Conserved lncRNAs Have Distinct Subcellular Localization Patterns in hESCs and mESCs

To explore whether sequence or positionally conserved lncRNAs undergo conserved RNA processing and subcellular localization, we retrieved sequence-conserved and positionally conserved lncRNAs in hESC H9 and mESC R1/E (R1) lines and compared their subcellular localization patterns by calculating the cytoplasmic ratio ($\text{cyto}_{\text{FPKM}} / [\text{cyto}_{\text{FPKM}} + \text{nuc}_{\text{FPKM}}]$) of each lncRNA, using mRNAs as controls (Figures 1A and S1A).

We first profiled the subcellular localization of all RNAs in H9 and R1 ESCs. Cytosolic and nuclear RNAs were examined by fractionation efficiency (Figure S1B), followed by total RNA sequencing after depletion of ribosomal RNAs (Ribo- RNA-seq). Repeated fractionated samples were highly correlated ($R > 0.95$) (Figure S1C). Analysis of cytoplasmic ratios of 4,804 expressed H9 lncRNAs and 3,289 expressed R1 lncRNAs revealed significant subcellular localization differences (Figure 1B; Table S1), whereas the subcellular localization of expressed mRNAs is similar between these two cell types (Figure 1B; Table S1).

We next compared the subcellular localization patterns of sequence-conserved lncRNAs in H9 and R1 ESCs. By aligning human lncRNAs to the R1 genome or vice versa, 122 lncRNAs

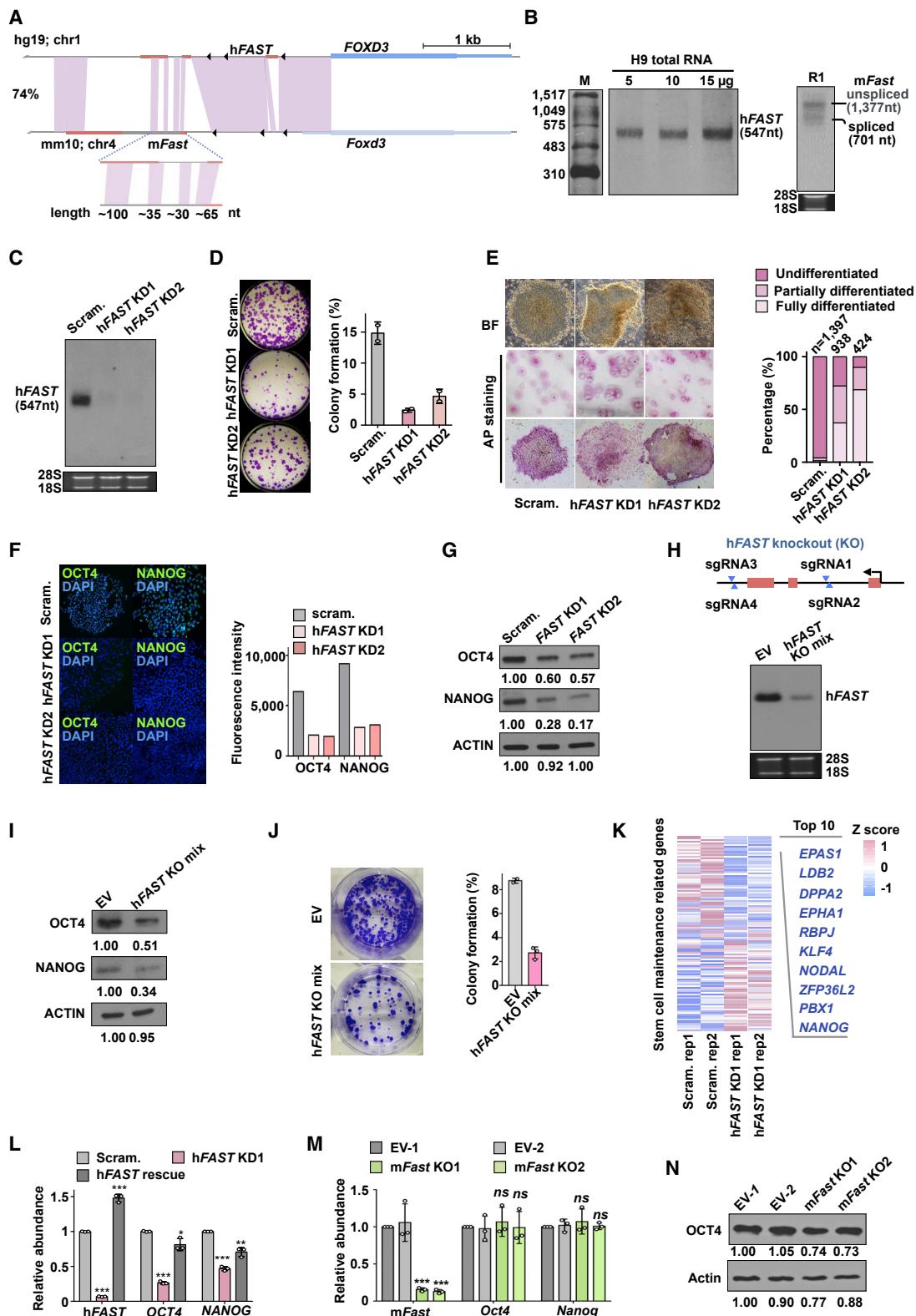
with 20% or more sequence similarity (Tang et al., 2017) were defined as sequence-conserved lncRNAs (Figures 1C and S1A; Table S2). Calculation of the cytoplasmic ratio of these 122 lncRNAs showed that their subcellular localization patterns differ between H9 and R1 cells (Figure 1D; Table S2), although several well-known lncRNAs (*MALAT1*, *NEAT1*, *FIRRE*, and *NORAD*) exhibited similar patterns (Figures S1B and S1D). To exclude any concern regarding this 20% threshold, we further analyzed 15%, 30%, 50%, and 80% sequence similarities and found that conserved lncRNAs under each condition ($n = 130$, 108, 60, and 24, respectively) exhibited stronger cytoplasmic localization patterns in H9 cells than in R1 cells (Figure S1E).

Further, we compared the subcellular localization patterns of positionally conserved lncRNAs. We classified GENCODE lncRNA annotations into three sub-groups according to relative transcription directionality and position to nearby protein coding genes: divergent, convergent, and same-strand lncRNAs (Figures 1E and S1A). In general, these positionally transcribed lncRNAs in H9 cells preferred to locate in the cytoplasm compared with the corresponding group of lncRNAs in R1 cells, whereas mRNAs in both cells were preferentially cytoplasmic (Figure 1F). We further obtained 229 lncRNAs that are positionally conserved relative to their adjacent conserved mRNAs and promoters (Figures 1G and S1A; Table S3). Of note, 151 of 229 lncRNAs (Figure 1G) were confirmed with conserved transcription start sites (TSSs) by analyzing the CAGE data (Fort et al., 2014; Figure S1F).

Remarkably, both sequence-conserved and positionally conserved lncRNAs have distinct subcellular localization patterns in H9 and R1 cells, whereas their conserved mRNAs have similar subcellular distribution patterns (Figure 1H). One example, the *FOXD3* antisense transcript1 (FAST), is shown in Figure S1G. The distinct subcellular localization patterns of selected lncRNAs (with $\text{FPKM} \geq 0.5$), including *FAST*, *RAB30-AS1*, *UBE2D3-AS1*, and others, in H9 and R1 cells (Figure 1I)

Figure 1. Conserved lncRNAs Have Distinct Subcellular Localization Patterns in hESCs and mESCs

- (A) Schematic of the compartmentalization analysis of conserved lncRNAs in H9 and R1 cells.
 - (B) Subcellular distribution of mRNAs and lncRNAs in H9 and R1 cells. The lncRNA distribution is shown by the cytoplasmic ratio, calculated by $\text{cyto}_{\text{FPKM}} / (\text{cyto}_{\text{FPKM}} + \text{nuc}_{\text{FPKM}})$.
 - (C) Identification of 122 sequence-conserved lncRNAs in H9 and R1 cells by genome alignment and sequence comparison.
 - (D) Sequence-conserved lncRNAs are more preferentially localized to the cytoplasm in H9 cells than in R1 cells.
 - (E) Three types of lncRNA—divergent, convergent, and the same strand—are classified according to their transcription directionality and position relative to nearby protein-coding genes.
 - (F) The subcellular distribution of the three types of lncRNAs in (E) is more cytoplasmic in H9 cells and more nuclear in R1 cells; as controls, their neighboring mRNAs are mostly cytoplasmic in both cells.
 - (G) Identification of 229 positionally conserved lncRNAs with conserved promoters in H9 and R1 cells.
 - (H) The localization difference of both sequence-conserved and positionally conserved lncRNAs between H9 and R1 cells is significantly higher than that of conserved mRNAs (from the NCBI HomoloGene database). The p values were calculated by two-tailed unpaired Student's t test. ** $p < 0.01$, *** $p < 0.001$.
 - (I) Identification of conserved lncRNAs with $\text{FPKM} \geq 0.5$ in H9 and R1 cells that have a large compartmentalization difference. hFAST is a positionally conserved lncRNA with the largest compartmentalization difference.
 - (J) Validation of subcellular localization of conserved lncRNAs with distinct subcellular localization patterns in H9 and R1 cells by RNA FISH. hFAST, mFast, *Rab30-as1*, and *Ube2d3-as1* were performed by smFISH, *RAB30-AS1* and *UBE2D3-AS1* were performed by Dig-labeled RNA FISH. Nuclei are indicated by yellow dotted lines, and the scale bar represents 5 μm here and in all panels in this study.
 - (K) Conserved lncRNAs in R1 mESCs (naive) or differentiated into EpiLCs (primed) are mainly localized in the nucleus, as revealed by qRT-PCR. *Gapdh* and *Neat1* are cytoplasmic and nuclear markers, respectively, here and throughout this study. Cytoplasmic and nuclear RNAs from equal cell numbers were assayed here and throughout this study.
 - (L) Conserved lncRNAs in mEpiSCs are mainly localized in the nucleus, as revealed by qRT-PCR.
 - (M and N) Conserved lncRNAs in H9 (M) and WIBR3 (N) cells cultured with conditioned medium (CM, primed) or mTeSR (primed) or RSeT (naive-like) are mainly cytoplasm localized, as revealed by qRT-PCR.
- Data in (K)–(N) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. See also Figure S1 and Tables S1, S2, and S3.



(legend on next page)

were validated by RNA fluorescence *in situ* hybridization (FISH) (Figure 1J).

These distinct localization patterns of lncRNAs did not seem to be cell-line-specific but, rather, species-specific, as shown by the cytoplasmic localization of the examined *FAST*, *RAB30-AS1*, and *UBE2D3-AS1* in additional hESC lines (H1 and CT1) and their nuclear retention in other mESC lines (E14) (Figure S1H). Further, different pluripotent (primed versus naive-like) states of mESCs and hESCs under different culture conditions had little effect on the distinct localization patterns of lncRNAs (Figures 1K–1N and S1I). Together, these data reveal that some conserved lncRNAs have distinct subcellular localizations, suggesting altered processing and perhaps different functional potential in human and mouse ESCs.

FAST Is a Positionally Conserved lncRNA that Is Specifically Expressed in hESCs and mESCs

Subcellular localization of lncRNAs is highly associated with function (Chen, 2016). To test whether distinctly localized lncRNAs between hESCs and mESCs lead to distinct biological effects, we designed short hairpin RNAs (shRNAs) to knock down five lncRNAs with the highest specific cytoplasmic localization in H9 cells and single guide RNAs (sgRNAs) to knock out the corresponding nuclear localized ones in R1 cells, followed by examining ESC self-renewal upon their loss (Figures S2A–S2D). Although depletion of all five nuclear lncRNAs had little effect on R1 cell pluripotency, as revealed by expression of the key pluripotency transcription factors *OCT4* and *NANOG* (Figures S2A and S2C), impaired expression of five tested cytoplasmic lncRNAs—*hFAST*, *RAB30-AS1*, *UBE2D3-AS1*, *SNHG14*, and *GNAS-AS1*—in H9 cells resulted in altered expression of *OCT4* and *NANOG* (Figures S2B and S2D). These results suggest that distinctly localized lncRNAs play different roles in cells and that some cytoplasmic lncRNAs might regulate hESC pluripotency.

Among these examined lncRNAs, *FAST* is transcribed in the opposite direction from its nearby coding gene *FOXD3* (Figure 2A). GENCODE annotation reveals that *hFAST* has five iso-

forms (Figure S2E), but we detected only one major isoform, which contains three exons in H9 cells (Figure 2B). *hFAST* is 547 nt in length and expressed at ~140 copies per H9 cell (Figures 2B and S2F). GENCODE annotation reveals that *mFast* has one isoform (Figure S2E), but we detected two isoforms, one of which is the major but unspliced 1,377-nt isoform, and another is a 701-nt spliced isoform (Figures 2B and S2E) in R1 cells. Analysis of these expressed hESC and mESC *FAST* RNAs shows conservation at three levels: exonic structure, sequence, and position (Ulitsky, 2016). First, *hFAST* contains an ~65-bp conserved region in the second exon and the second intron junction that matches the first exon and the first intron in *mFast*. Second, *hFAST* has an ~230-nt conserved sequence to the unspliced, major isoform of *mFast* and 74% genomic sequence conservation between human and mouse. Finally, *FAST* is positionally conserved, as shown by having the same transcriptional direction and conserved upstream and downstream sequences (Figure 2A).

Although *hFAST* is quite abundant in H9 cells, other ESC lines and induced pluripotent stem cells (iPSCs) (Choi et al., 2015; Figures S2G and S2H), it is not expressed or only expressed at low levels in many other non-pluripotent cells and tissues (Figures S2I and S2J). This suggests that it may have a specific role in pluripotency or differentiation. We found that *FAST* expression was rapidly decreased during H9 ectoderm, trophoblast, and mesoderm differentiation (Figure S3A) and upon R1 spontaneous differentiation (Figure S3B). Although *hFAST* is mainly localized in the cytoplasm, we excluded the possibility that it could encode functional peptides by inserting a FLAG tag into its predicted open reading frame (ORF) (Figure S3C). Together, these findings suggest that *hFAST* plays a role in maintenance of hESC pluripotency, consistent with impaired pluripotent gene expression upon *hFAST* depletion (Figure S2D).

hFAST, but Not *mFast*, Is Required for Maintenance of Pluripotency

To examine the effect of *hFAST* on hESC pluripotency maintenance, we depleted *hFAST* by two shRNAs (Figures 2C and

Figure 2. *hFAST*, but not *mFast*, Is Required for Stem Cell Pluripotency

- (A) Conservation of the *hFAST* and *mFast*, *FOXD3* and *Foxd3* region between human and mouse. A purple shadow marks the conserved regions between human and mouse, analyzed by the EMBOSS Matcher (Madeira et al., 2019). The expanded region shows the conserved sequences between *hFAST* and *mFast*. The triangles show TSSs revealed by CAGE datasets (Fort et al., 2014).
- (B) Detection of *FAST* in H9 and R1 cells by northern blotting (NB). Left: *hFAST* has a major isoform (547 nt) in H9 cells, shown by 10% denaturing PAGE NB. Right: *mFast* has two isoforms (1,377 nt and 701 nt) in R1 cells, shown by 1% agarose NB.
- (C) *hFAST* knockdown (KD) in H9 cells was validated by 1% agarose NB.
- (D) *hFAST* KD impaired the colony formation ability of H9 cells. Left: cell colonies were stained by 0.1% crystal violet. Right: statistics of colony formation efficiency.
- (E) *hFAST* KD promoted H9 cell differentiation. Left: alkaline phosphatase (AP) staining of *hFAST* KD cells. Right: statistics of differentiation efficiency.
- (F and G) *hFAST* KD led to reduced *OCT4* and *NANOG* expression, as shown by IF (F, left; right, quantification of IF) and WB (G).
- (H) Knockout (KO) *hFAST* by CRISPR/Cas9. Top: sgRNAs used to deplete *hFAST*. Bottom: *hFAST* KO efficiency in H9 cells was validated by 1% agarose NB. *hFAST* KO mixture cells were assayed because of the severe damage to pluripotency upon *hFAST* KO.
- (I and J) CRISPR/Cas9-mediated loss of *hFAST* in H9 led to reduced *OCT4* and *NANOG*, as shown by WB (I), and impaired colony formation ability (J, left; right, statistics of colony formation).
- (K) *hFAST* KD led to altered expression of stem cell maintenance-related genes in H9 cells, shown by a heatmap of these genes from RNA-seq.
- (L) Overexpression (OE) of *hFAST* in *hFAST* KD H9 cells rescued *OCT4* and *NANOG* expression, revealed by qRT-PCR.
- (M and N) Loss of *mFast* by CRISPR/Cas9 had no detectable effect on mESC pluripotency, revealed by qRT-PCR (M) and WB (N).
- Data in (J), (L), and (M) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All p values were calculated using two-tailed unpaired Student's t test. *p < 0.05; **p < 0.01; ***p < 0.001; ns, no significant difference. Quantification of WB in (G), (I), and (N) was calculated by the software Quantity-One. See also Figures S2 and S3 and Table S4.

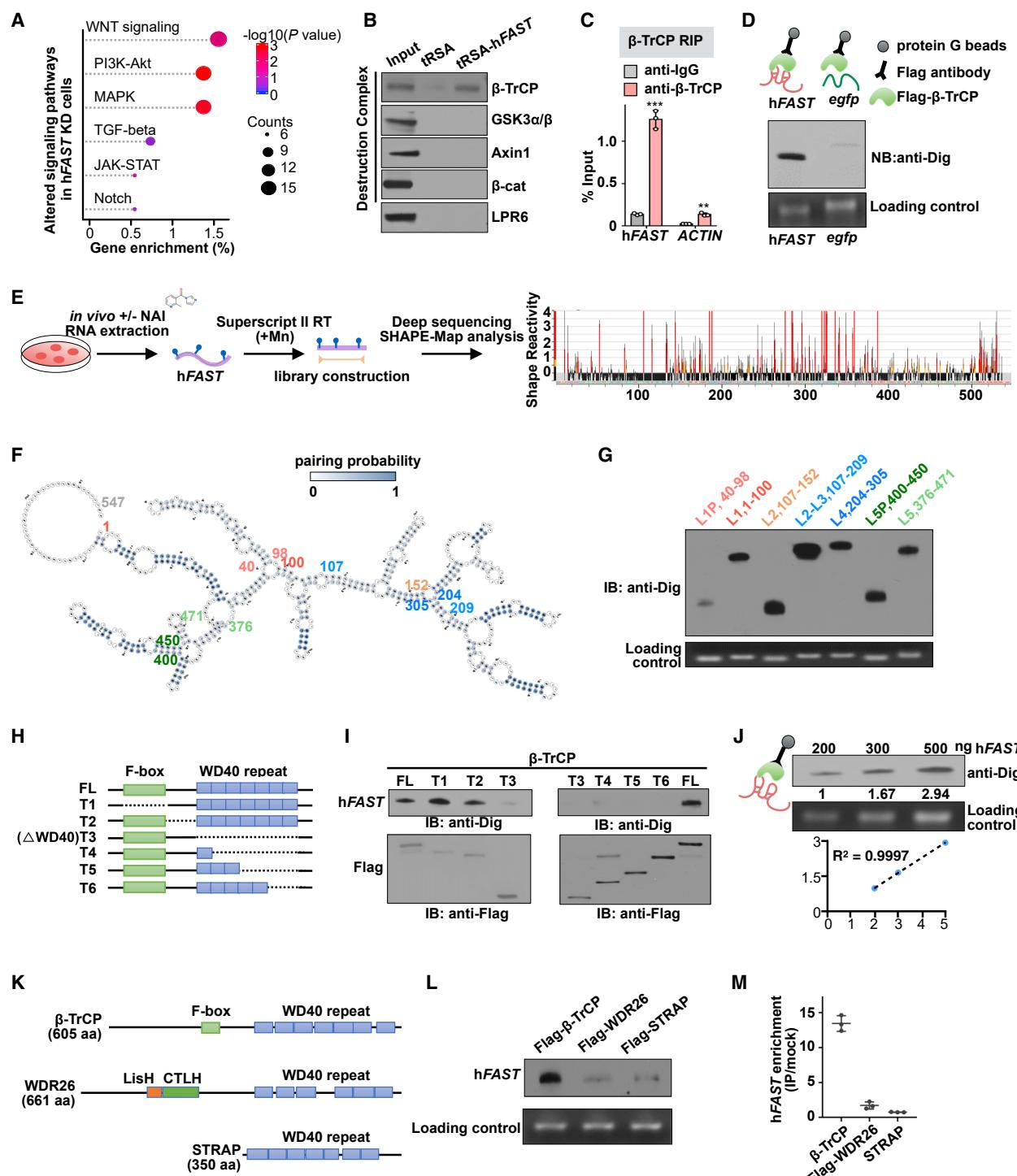


Figure 3. hFAST Maintains hESC Pluripotency by Binding β-TrCP via Multiple Loops

(A) Bubble plot showing dysregulated genes related to stem-cell-related pathways in hFAST KD by RNA-seq. The x axis shows the ratio of the number of enriched mRNAs in each pathway divided by the number of total dysregulated RNAs (one-sided Fisher's exact test). The y axis shows the category of pathways.

(B) hFAST binds to β-TrCP. Proteins in the WNT pathway were analyzed for their interaction with hFAST by tRSA-hFAST pulldown assay followed by WB.

(C) β-TrCP interacts with hFAST. A native RIP assay was performed in H9 cells using anti-β-TrCP or anti-IgG antibodies followed by qRT-PCR. ACTIN mRNA was used as a control.

(legend continued on next page)

S3D). hFAST KD significantly impaired H9 colony formation ability (**Figure 2D**), and hFAST knockdown (KD) colonies showed strong spontaneous differentiation under mouse embryonic fibroblast (MEF) and MEF-free conditions after 3–4 passages (**Figure 2E**). Immunofluorescence (IF), western blot (WB), and qRT-PCR analyses showed that hFAST KD impaired expression of the core transcription factors *OCT4* and *NANOG* (**Figures 2F, 2G, and S2D**). hFAST KD in CT1 and H1 hESCs also reduced *OCT4* and *NANOG* expression (**Figure S3E**) and impaired the colony formation ability of CT1 cells (**Figure S3F**). Similar results were also observed using CRISPR/Cas9 to knock out hFAST in H9 cells (**Figures 2H–2J**).

To ask whether hFAST could affect expression of other genes besides *OCT4* and *NANOG*, we collected RNAs from duplicate experiments using scrambled or hFAST KD cells for RNA-seq analyses and found that many genes related to maintenance of stem cell pluripotency were downregulated (**Figure 2K; Table S4**). Comparison of our hFAST KD RNA-seq datasets with other published datasets in H1 and H9 cells and human induced pluripotent stem cells (hiPSCs) ([Choi et al., 2015](#)) also revealed that hFAST KD impaired expression of many pluripotency-related genes, with 67% overlap of the downregulated genes in these additional datasets (**Figures S3G and S3H**). Importantly, expression of *OCT4* and *NANOG* could be rescued to normal levels by ectopic expression of hFAST in H9 cells (**Figure 2L**).

Because of the different subcellular localization patterns, we hypothesized that FAST in hESCs and mESCs has different functions. Generation of two mFast KO clones by CRISPR/Cas9 showed no effect on *Oct4* and *Nanog* expression at the mRNA or protein levels (**Figures 2M and 2N**). Collectively, these data show that hFAST, but not mFast, plays a critical role in pluripotency maintenance.

hFAST Binds the β-TrCP WD40 Domain via Multiple Stem Loops

Maintenance of pluripotency is affected by a number of signaling pathways ([Pera and Tam, 2010](#)). We analyzed six pathways known to be required for stemness and found that WNT signaling showed significant enrichment with altered expression of WNT target genes in hFAST KD cells (**Figures 3A and S4A**). We performed TRSA-RNA pulldown assay ([Xing et al., 2017](#)) to examine

whether some key factors in WNT signaling, such as β-TrCP, GSK-3α/β, Axin1, β-catenin, and LPR6, can interact with hFAST. The results showed that the E3 ubiquitin ligase β-TrCP was specifically associated with hFAST (**Figure 3B**). RNA immunoprecipitation (RIP) confirmed its interaction with hFAST (**Figure 3C**). Also, purified FLAG-β-TrCP expressed from *E. coli* (**Figure S4B**) could directly bind to hFAST *in vitro* (**Figure 3D**).

To find out how hFAST interacts with β-TrCP, we used *in vitro* assays (**Figure S4C**) to first arbitrarily generate a series of hFAST fragments ~100 nt in length, with 30- to 40-nt overlapping sequences between the adjacent fragments (**Figure S4D**). Each *in vitro* transcribed digoxin (Dig)-labeled hFAST fragment was incubated with partially purified FLAG-β-TrCP from 293FT cells, and eluted RNAs were resolved by denaturing PAGE, followed by immunoblotting with anti-Dig antibodies (**Figure S4C**). We found that two 100-nt fragments (F1 and F6) from the 5' and 3' ends of hFAST specifically interacted with β-TrCP (**Figure S4D**). Longer hFAST fragments that span different regions (F1–F3, F3–F5, F5–F6, and F6–F7) resulted in similar observations but also revealed an additional binding domain spanning F3–F5 (**Figure S4E**).

These results inspired us to further explore hFAST *in vivo* structural conformation by carrying out *in-cell* selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) ([Smola et al., 2015](#); **Figure 3E**). SHAPE-MaP revealed that hFAST tended to form several independent stem loops (**Figures 3F and S4F**), including fragments F1 and F6 (**Figures S4D and S4E**). This conformation was determined in cells but could be different in the absence of cellular proteins. Nevertheless, generation of hFAST fragments according to its secondary loop regions revealed by *in-cell* SHAPE-MaP (**Figures 3F and S4G**), followed by performing binding assay with β-TrCP, revealed that hFAST is a multivalent β-TrCP binding platform via five individual loop regions (**Figures 3G and S4G**).

Next we asked which domain of β-TrCP interacts with hFAST. β-TrCP belongs to the Fbw (F-box/WD40 repeat-containing) protein family, with an F-box motif at the N terminus and seven WD40 repeats at the C terminus ([Fuchs et al., 2004](#)). We generated a series of β-TrCP truncations lacking the F-box motif, the WD40 repeats, the linker between these two domains, or truncations of different repeats from the WD40 domain (**Figure 3H**). We

(D) hFAST binds to β-TrCP *in vitro*. Top: diagram of the *in vitro* binding assay. FLAG-β-TrCP was purified from *E. coli*. hFAST and egfp RNAs were *in-vitro*-transcribed with Dig-labeled dNTPs. Bottom: hFAST binds to FLAG-β-TrCP *in vitro*, as revealed by NB.

(E) In-cell SHAPE-MaP assay of hFAST in H9 cells. Left: diagram of the SHAPE-MaP assay for hFAST in cells. Right: SHAPE-MaP profile of hFAST.

(F) Schematic of hFAST secondary structure calculated from SHAPE-MaP. hFAST tends to form multiple loops, L1–L5, shown in (G), with their start and end sites labeled. The color in circles represents the pairing probability calculated by Superfold. A darker color indicates higher pairing probability.

(G) Partially purified FLAG-β-TrCP interacts with multiple loops in hFAST, shown in (F), as shown by *in vitro* binding assay shown in **Figure S4C**.

(H) Schematic of β-TrCP truncations used in (I). T1, deletion of the F-box domain of β-TrCP; T2, deletion of the internal linker of β-TrCP; T3, deletion of the WD40 repeat domain of β-TrCP; T4–T6, truncations contain different repeats of the WD40 domain. All mutants were fused with a FLAG tag at the N terminus.

(I) The complete WD40 repeat domain of β-TrCP is required for hFAST binding, as shown by *in vitro* binding assay shown in **Figure S4C**.

(J) hFAST binds FLAG-β-TrCP with a linear correlation. Different amounts of hFAST were incubated with equal amounts of FLAG-β-TrCP, followed by quantification of hFAST associated with FLAG-β-TrCP by NB.

(K) Schematic comparison of domains of three WD40 repeat-containing proteins β-TrCP, WDR26, and STRAP.

(L) hFAST specifically binds to FLAG-β-TrCP, but not FLAG-WDR26 or FLAG-STRAP, as shown by *in vitro* binding assay shown in **Figure S4C**.

(M) hFAST specifically binds to β-TrCP, but not FLAG-WDR26 or STRAP, as shown by RIP assay using anti-β-TrCP, anti-STRAP, or anti-FLAG antibodies followed by qRT-PCR. hFAST enriched by each antibody was normalized to anti-immunoglobulin G (IgG; mock).

Error bars in (C) and (M) represent SD in triplicate experiments. The p values were calculated using two-tailed unpaired Student's t test; **p < 0.01, ***p < 0.001. See also **Figure S4**.

performed assay, as illustrated in **Figure S4C**, and found that β-TrCP interacted with hFAST mainly via the WD40 repeat domain (**Figure 3I**) and with a linear correlation (**Figure 3J**). The direct interaction between hFAST and β-TrCP was confirmed by electrophoretic mobility shift assays (EMSA) (**Figure S4H**). The spliced and unspliced isoforms of mFast failed to interact with mouse β-TrCP (**Figure S4I**), suggesting that mFast is not involved in WNT signaling, consistent with the loss-of-function studies shown in **Figures 2M** and **2N**.

Is hFAST sufficiently abundant to affect levels of β-TrCP involved in WNT signaling in hESCs? It has been reported that each HeLa cell has 9,602 copies of β-TrCP, as assayed by quantitative proteomics ([Hein et al., 2015](#)). WB of β-TrCP in HeLa and H9 cells suggested ~8,173 copies of β-TrCP in H9 cells (**Figure S4J**). IF imaging further showed that 40% of β-TrCP was localized to the cytoplasm (**Figure S4K**), indicating ~3,269 copies of β-TrCP in the cytoplasm per H9 cell. Given the observations that each H9 cell contains ~140 copies of hFAST (**Figure S2F**), mostly in the cytoplasm (**Figure 1K**), and that each binds five β-TrCP molecules (**Figures 3F** and **3G**), then, theoretically, more than 20% of the cytosolic β-TrCP can be bound by hFAST. Importantly, hFAST barely interacted with other WD40 domain-containing proteins, such as WDR26 or STRAP (**Figures 3K–3M** and **S4L**), that act in the mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K)-Akt, and transforming growth factor β (TGF-β) pathways ([Seong et al., 2005](#); [Ye et al., 2016](#); [Zhu et al., 2004](#)), confirming the specific interaction between hFAST and β-TrCP.

hFAST Maintains hESC Pluripotency by Protecting β-Catenin from Ubiquitination

WNT signaling in hESCs pluripotency is stage specific and cell context dependent ([Nusse and Clevers, 2017](#); [Sato et al., 2004](#); [Sokol, 2011](#); [Wray and Hartmann, 2012](#)). To investigate WNT signaling in our hESCs, we treated H9, H1, and CT1 cells with BIO (6-bromoindirubin-3'-oxime) and CHIR99021, inhibitors of GSK3α/β, to activate WNT signaling ([Lian et al., 2013](#); [Sato et al., 2004](#)). We found that the levels of β-catenin (**Figure S5A**), the WNT target genes AXIN2 and CD44, and the pluripotency-related genes OCT4 and NANOG (**Figure S5B**) increased after BIO or CHIR99021 treatment. Also, depleting components of the destruction complex is known to activate the WNT pathway ([Major et al., 2007](#)). Axin1 KD showed the same results as BIO or CHIR99021 treatment (**Figure S5C**). Depletion of β-catenin also led to decreased expression of OCT4 and NANOG (**Figure S5D**). Together, these results confirm that WNT signaling plays a positive role in regulating hESC pluripotency.

Next we investigated how the interaction between hFAST and β-TrCP regulates hESC pluripotency. The WD40 domain of β-TrCP not only has RNA binding activity ([Castello et al., 2012](#); [Jin et al., 2016](#)) but is also required for recognition of phosphorylated β-catenin ([Wu et al., 2003](#)). Because hFAST and phosphorylated β-catenin can both interact with the WD40 repeat domain of β-TrCP, we speculated that hFAST might inhibit β-TrCP recognition of phosphorylated β-catenin, leading to β-catenin accumulation in the nucleus to keep WNT signaling active in normal cells (**Figure 4A**). In this model, depletion of hFAST would facilitate β-TrCP recognition of phosphorylated

β-catenin and result in β-catenin degradation by proteasomes (**Figure 4A**). Experimentally, hFAST KD by shRNAs or CRISPR/Cas9 led to decreased β-catenin protein levels (**Figure S5E**) in the cytoplasm and nucleus (**Figure 4B**). Further, loss of hFAST enhanced the ubiquitination of β-catenin (**Figure 4C**), suggesting that hFAST inhibits β-catenin ubiquitination and degradation by binding β-TrCP.

Axin1 is a skeleton component of the destruction complex that controls β-catenin stability in WNT signaling ([Nusse and Clevers, 2017](#)). We speculated that the interaction between β-TrCP and the destruction complex should be enhanced in the absence of hFAST (**Figure 4A**). *In vitro* immunoprecipitation (IP) of Axin1 in hFAST KD cells showed enhanced interaction between β-TrCP and the destruction complex (**Figure 4D**). Further, *in vitro* IP of FLAG-β-TrCP in hFAST overexpression (OE) 293FT cells revealed that hFAST prevented β-TrCP from recognizing phosphorylated β-catenin (**Figure 4E**). As a result, hFAST KD impaired expression of the WNT target genes AXIN2 and CD44 in hESCs (**Figure 4F**); OE of hFAST (**Figure 4G**), but not mFast (**Figure S5F**), rescued AXIN2 and CD44 expression. Activation of WNT signaling by BIO or CHIR99021 in multiple hESC lines rescued hFAST KD phenotypes (**Figures 4H**, **4I**, **S5G**, and **S5H**), further establishing WNT signaling as the mechanism of hFAST pluripotency maintenance in hESCs. Collectively, these results led us to conclude that loss of hFAST results in reduced WNT signaling, driving hESCs out of the pluripotent state.

We also asked whether exogenously expressed hFAST could regulate WNT signaling in other cell types. 293FT cells have no detectable hFAST (**Figures S5I** and **S2J**). Exogenous expression of hFAST in 293FT cells increased β-catenin levels, and in these cells, hFAST could interact with β-TrCP (**Figures S5I–S5K**). Consistently, WNT signaling target genes were also upregulated (**Figure S5L**). On the other hand, exogenous expression of mFast had no effect on expression of β-catenin and WNT target genes (**Figures S5I**, **S5J**, and **S5L**). These findings are in accordance with the observation that mFast did not bind to β-TrCP *in vitro* (**Figure S4I**).

Distinct Processing and Localization of FAST Is Regulated by trans-Factors

RNA splicing affects subcellular localization ([Valencia et al., 2008](#)). Consistently, splicing is more suppressed in nuclear retained lncRNAs in mESCs compared with hESCs (**Figure S6A**). Further, mFast is retained in the nucleus in mESCs, whereas hFAST is fully spliced and exported to the cytoplasm (**Figures 1J–1N**, **S1H**, and **2B**). These distinct localizations of lncRNAs suggested that RNA splicing or additional *trans*-factors might contribute to their distinct localization between H9 and R1 cells. To test this possibility, we introduced exogenous hFAST or mFast into H9 or R1 cells, respectively, by lentivirus infection. hFAST and mFast were stably expressed at comparable levels (**Figure S6B**); intriguingly, they both localized to the cytoplasm in H9 cells, whereas both were more nuclear in R1 cells (**Figures S6C–S6F**). These results were validated by single-molecule (sm) FISH (**Figures 5A** and **5B**), where hFAST was strikingly retained in the nucleus in R1 cells, and mFast became cytosolic in H9 cells. Transient expression of hFAST and mFast also exhibited the same results (**Figure S6G**). Such distinct subcellular

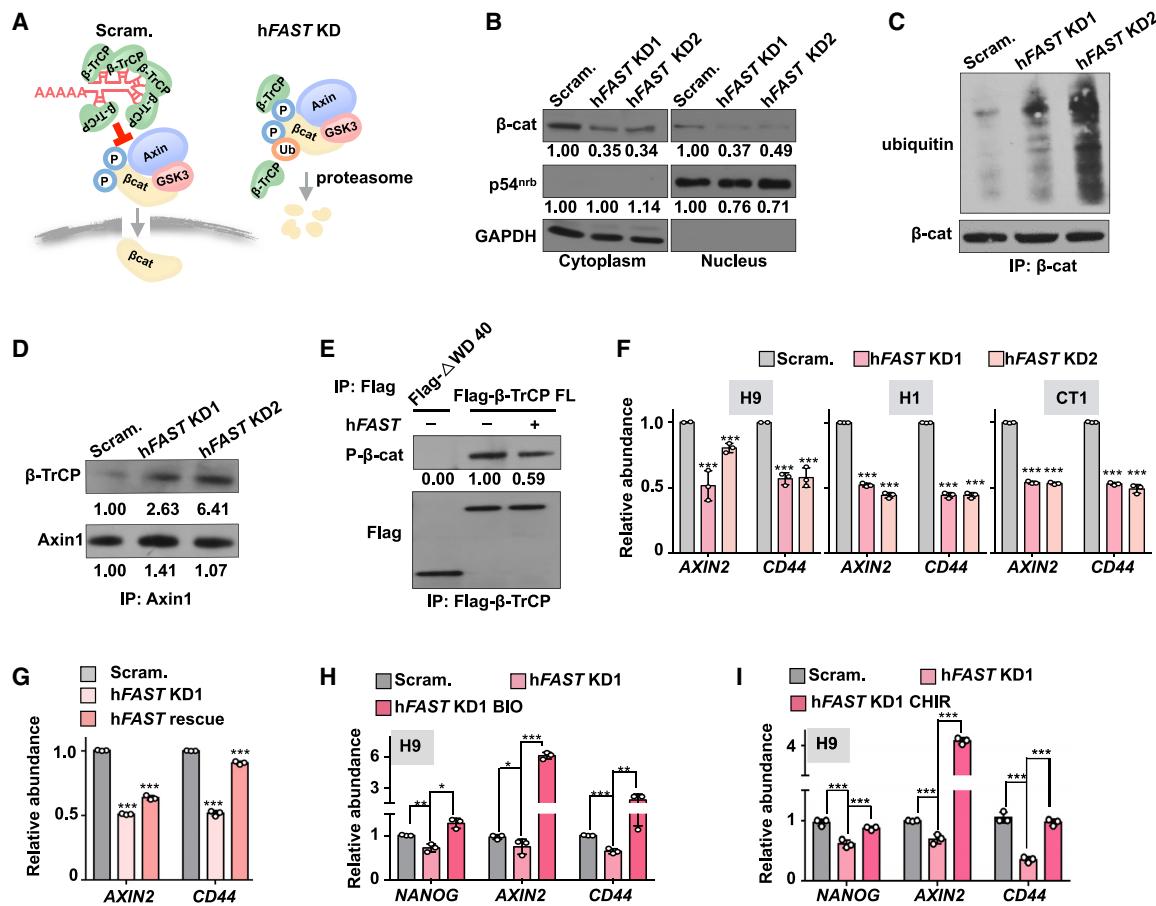


Figure 4. hFAST Protects β-Catenin from Ubiquitination to Keep WNT Activation in hESCs

(A) A proposed model of hFAST in WNT signaling. hFAST was proposed to block β-TrCP binding to the destruction complex in H9 cells (left), and depletion of hFAST would result in increased binding between β-TrCP and the destruction complex (right).

(B) hFAST KD led to reduced β-catenin (β-cat) expression in the cytoplasm and nucleus, as shown by WB in H9 hFAST KD cells. p54^{nrb} (NONO) and GAPDH are markers for the nucleus and cytoplasm.

(C) hFAST KD led to increased ubiquitination of β-cat. Assays were performed by anti-β-cat IP in hFAST KD cells, followed by WB with anti-ubiquitin (top) and anti-β-cat (bottom) in hFAST KD cells.

(D) hFAST blocked the interaction between β-TrCP and the destruction complex. Assays were performed by anti-Axin1 IP in hFAST KD cells, followed by WB with anti-β-TrCP (top) and anti-Axin1 (bottom) in hFAST KD cells.

(E) hFAST OE led to reduced interaction between β-TrCP and phosphorylated β-cat. Assays were performed by anti-FLAG-β-TrCP IP in 293FT cells, followed by WB with anti-phosphorylated β-cat in hFAST OE cells (top) and anti-FLAG (bottom). ΔWD40 is a negative control.

(F) hFAST KD impaired WNT target gene AXIN2 and CD44 expression, as revealed by qRT-PCR in hFAST KD hESC lines (H9, H1, and CT1).

(G) hFAST OE in hFAST KD H9 cells rescued WNT target gene AXIN2 and CD44 expression, revealed by qRT-PCR.

(H and I) Activation of the WNT pathway by 2 μM BIO (H) or 4 μM CHIR99021 (I) treatment in hFAST KD H9 cells rescued NANOG, AXIN2, and CD44 expression, revealed by qRT-PCR.

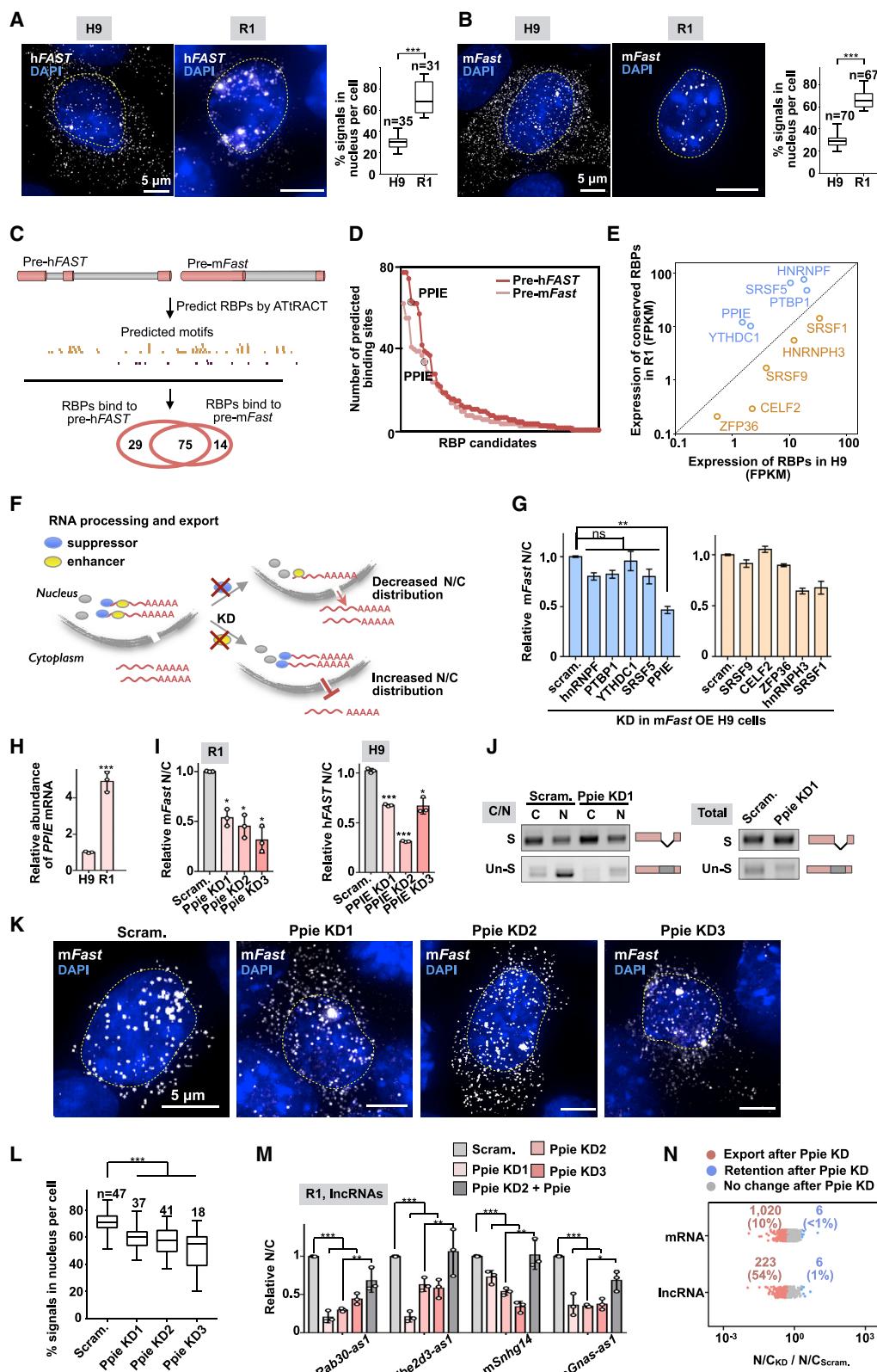
Data in (F–I) are presented as mean ± SD. Error bars represent SD in triplicate experiments. All p values were calculated using two-tailed unpaired Student's t test; *p < 0.05, **p < 0.01, ***p < 0.001. Quantifications were analyzed by Quantity-One. See also Figure S5.

localization of exogenously expressed hFAST and mFast in hESCs and mESCs suggests that the differentially expressed *trans*-factor(s), but not *cis*-elements, contribute primarily to differential localization.

Differential Expression of Peptidylprolyl Isomerase E (PPIE) Regulates FAST Processing and Export

To explore which *trans*-factor(s) might regulate distinct FAST processing and export in H9 and R1 cells, we first computationally analyzed potential hFAST and mFast binding proteins. We fetched hFAST and mFast pre-RNA sequences to predict their potential binding proteins using the ATtRACT database and found that 75 common proteins might interact with both hFAST and mFast (Figure 5C; Table S5). According to the highly enriched predicted binding sites and known functions in RNA processing, we selected 10 RNA binding protein (RBPs) that have the potential to bind both hFAST and mFast but are differentially expressed in H9 and R1 cells, as shown by RNA-seq data

ally analyzed potential hFAST and mFast binding proteins. We fetched hFAST and mFast pre-RNA sequences to predict their potential binding proteins using the ATtRACT database and found that 75 common proteins might interact with both hFAST and mFast (Figure 5C; Table S5). According to the highly enriched predicted binding sites and known functions in RNA processing, we selected 10 RNA binding protein (RBPs) that have the potential to bind both hFAST and mFast but are differentially expressed in H9 and R1 cells, as shown by RNA-seq data



(legend on next page)

(Figures 5D and 5E; Table S5). We speculated that those more highly expressed in R1 cells may suppress RNA processing and export, whereas those more highly expressed in H9 cells may enhance RNA splicing and export (Figure 5F). Accordingly, if RBPs suppress lncRNA export, then their KD should decrease the nucleus/cytoplasm ratio (N/C); if these RBPs enhance lncRNA export, then KD should increase N/C distribution (Figure 5F). After KD (Figure S7A), N/C distribution was monitored by qRT-PCR (Figure 5G). Among all examined factors, KD of the differently expressed PPIE led to the most dramatically decreased mFast N/C distribution (Figure 5G) in H9 mFast OE cells, indicating that PPIE is at least one factor that suppresses mFast export to the cytoplasm.

How does PPIE regulate RNA export in these cells? PPIE belongs to the peptidyl-prolyl *cis-trans* isomerase (PPIase) family, and several members of this family have been identified as part of the spliceosome and are essential for splicing (Bertram et al., 2017; Bessonov et al., 2008; Schiene-Fischer, 2015). PPIE is highly conserved between human and mouse and has an RNA recognition motif (RRM) and PPIase cyclophilin-type domain (Figure S7B). The expression level of mouse Ppie in R1 cells is much higher than that of human PPIE in H9 cells (Figures 5H, S7C, S7D, and S7J), which accords with our hypothesis that highly expressed Ppie in R1 cells suppresses lncRNA processing and export. PPIE is an abundant protein (Figures S7C and S7D) that is localized to the nucleus in H9 and R1 cells (Figure S7E).

Consistent with predicted PPIE binding to pre-hFAST or pre-mFast (Figures 5D and S7F), PPIE interacted with hFAST or mFast mature or precursor transcripts (Figures S7G–S7I). PPIE KD (Figure S7J) resulted in a decreased endogenous mFast or hFAST N/C distribution (Figure 5I). Consistently, loss of Ppie increased the level of the mFast spliced isoform and promoted

its export to the cytoplasm (Figures 5J and S7K). Pre-mFast was also reduced upon Ppie KD (Figure S7M), further indicating that loss of Ppie increased mFast processing. Re-introducing Ppie into Ppie KD R1 cells rescued mFast nuclear processing and localization (Figures S7L and S7M). smRNA FISH of exogenously expressed mFast confirmed reduced nuclear localization after Ppie KD in R1 cells (Figures 5K and 5L). Importantly, Ppie KD also reduced the N/C distribution of a number of other positionally conserved lncRNAs in R1 cells, which can be largely rescued by re-introducing Ppie into Ppie KD cells (Figure 5M). RNA-seq analysis of fractionated Ppie KD R1 cells further revealed that Ppie had a broader effect on nuclear retention of lncRNAs than mRNAs (Figure 5N; Table S6). Taken together, these results suggest that the differentially expressed splicing suppressor PPIE contributes to the distinct localization of a subgroup of lncRNAs, including FAST, between H9 and R1 cells.

Localization and Processing of lncRNAs Is Associated with PPIE during Evolution

FAST is highly conserved between human and monkey, having 94% sequence similarity (Figure 6A). Northern blotting revealed that cemFAST (534 nt) is highly expressed and processed in the crab-eating macaque ESC M21 line (Figure 6B). Similar to that in hESCs (Figures 1J–1N), the processed cemFAST is largely localized to the cytoplasm of M21 cells (Figure 6C). Loss of cemFAST impaired the expression of NANOG and OCT4 in M21 cells (Figures 6D and 6E), indicating impaired pluripotency. These results suggest that processing, localization, and function of FAST are conserved among primates but distinct from that of rodents. Consistently, PPIE expression in M21 as well as rhesus macaque ESC LYON-1 cells is slightly higher than in hESCs but to a much lower degree than in mESCs, as

Figure 5. Differential Expression PPIE Regulates FAST Processing and Export in H9 and R1 Cells

- (A) Distinct localization of hFAST in H9 and R1 cells. Left: smFISH validation of hFAST OE and localization in H9 and R1 cells. Right: statistics of hFAST signals in each hFAST OE cell nucleus. 35 and 31 cells were analyzed in hFAST OE H9 and R1 cells.
- (B) Distinct localization of mFast in H9 and R1 cells. Left: smFISH validation of mFast OE and localization in H9 and R1 cells. Right: statistics of mFast signals in the nucleus. 70 and 67 cells were analyzed in hFAST OE H9 and R1 cells.
- (C) A pipeline for predicting RBP candidates that bind both hFAST and mFast. 75 RBPs were predicted to bind both pre-hFAST and pre-mFast using the ATTRACT database.
- (D) Number of predicted RBP binding sites on pre-hFAST and pre-mFast.
- (E) Expression of RBPs in H9 and R1 cells, revealed by FPKM of Ribo–RNA-seq samples in this study. RBPs that are highly expressed in R1 cells are labeled in blue, and those in H9 are labeled in yellow.
- (F) Schematic of proposed RBP function in regulating RNA processing and export.
- (G) Screening RBPs that are critical for RNA processing and export. Shown is KD of each RBP in mFast OE H9 cells, followed by nuclear and cytoplasmic fractionation and qRT-PCR analysis of mFast N/C distribution. Also shown is KD of suppressor (left) and enhancer (right) candidates of RNA processing and export.
- (H) Relative abundance of human PPIE and mouse Ppie mRNAs in H9 and R1 cells, shown by qRT-PCR.
- (I) KD of Ppie (left) or PPIE (right) promoted mFast or hFAST export to the cytoplasm in R1 and H9 cells, respectively, shown by reduced N/C distribution by qRT-PCR.
- (J) Ppie KD promoted mFast processing and export in R1 cells. Left: the spliced isoform was increased and exported to the cytoplasm, and the unspliced isoform was reduced in the nucleus, shown by nuclear and cytoplasmic fractionation followed by RT-PCR. Right: spliced and unspliced isoforms in total RNAs by RT-PCR.
- (K and L) Ppie KD promoted exogenously expressed mFast export to the cytoplasm revealed by smFISH in R1 cells (K). Also shown are statistics of mFast signals in the nucleus (L).
- (M) Ppie KD promoted the export of conserved lncRNAs to the cytoplasm, shown by reduced N/C distribution by qRT-PCR in R1 cells. The altered localization can be largely or fully rescued by re-introducing Ppie to R1 Ppie KD cells.
- (N) Ppie KD promoted the export of lncRNAs (54%) and some mRNAs (10%) genome wide revealed by RNA-seq in R1 cells. lncRNAs and mRNAs with comparable expression levels ($0.5 < \text{fold change} < 2$) after Ppie KD were used in the analysis.
- Data in (G)–(I) and (M) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All p values in (A), (B), (G)–(I), (L), and (M) were calculated using two-tailed unpaired Student's t test; *p < 0.05, **p < 0.01, ***p < 0.001. See also Figures S6 and S7 and Tables S5 and S6.

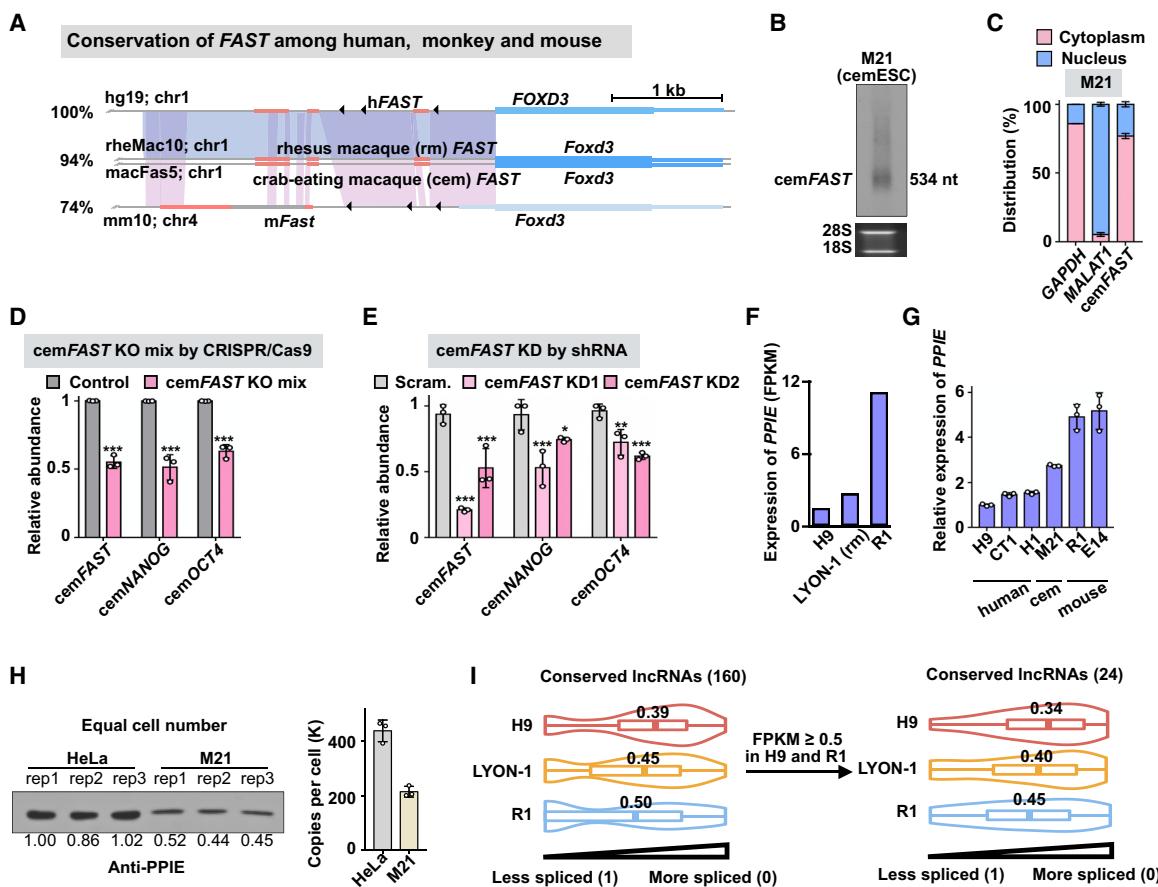


Figure 6. Differential Expression of PPIE Affects RNA Processing during Evolution from Mouse to Monkey and Human

(A) Conservation of *FAST* among human, monkey, and mouse. A purple shadow marks the conservation regions in these species, analyzed by the EMBOSS Matcher (Madeira et al., 2019). A blue shadow marks the conservation regions between human and monkey. Similarities of a 3.2-kb-long sequence upstream of *FOXD3* containing *FAST* between monkey and human (94%) as well as mouse and human (74%) are shown on the left. The triangles show TSSs revealed by CAGE datasets (Fort et al., 2014).

(B) cem*FAST* is highly processed in the crab-eating macaque ES cell M21 line, shown as a major isoform (534 nt) by NB.

(C) cem*FAST* mainly localized at cytoplasm in M21 cells, as revealed by qRT-PCR. GAPDH and MALAT1 are markers for the cytoplasm and nucleus, respectively.

(D and E) cem*FAST* KO by CRISPR/Cas9 or KD by shRNAs impaired expression of the pluripotency genes *OCT4* and *NANOG*, shown by qRT-PCR.

(F and G) Reduced PPIE expression during evolution, shown by gradually decreased mRNA levels in examined ESCs derived from mouse, monkey, and human, revealed by RNA-seq data (F) (Fiddes et al., 2018) and RT-PCR (G). LYON-1 is a rhesus macaque ESC line.

(H) Quantification of PPIE in M21 cells by WB (left). Shown are statistics of PPIE copies per cell compared with HeLa cells (right). See also Figures S7C and S7D for PPIE in H9 and R1 cells. PPIE copy number per HeLa cell was extracted from quantitative proteomics (Hein et al., 2015).

(I) Conserved lncRNAs tend to have increased splicing capability from mouse to monkey and human ESC lines, revealed by splicing score (see also Figure S6A) by analyzing RNA-seq datasets from the H9, LYON-1, and R1 cell lines (GEO: GSE106245).

Data in (C)–(E) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All p values were calculated using two-tailed unpaired Student's t test; *p < 0.05, **p < 0.01, ***p < 0.001. See also Figure S7.

shown by mRNA and protein levels (Figures 6F–6H, S7C, S7D, and S7J). In addition to *FAST*, a number of other conserved lncRNAs tended to have increased splicing capability from mouse to monkey and human ES cell lines (Figure 6I).

DISCUSSION

Analysis of lncRNA conservation has revealed that lncRNAs are rapidly evolving (Nitsche and Stadler, 2017; Ulitsky, 2016). The less constrained conservation and rapid evolutionary turnover of lncRNAs make it difficult to dissect their

functions and mechanisms of action. Subcellular localization is clearly related to function (Chen, 2016; Ulitsky, 2016). Here we found that, in human and mouse stem cells, most conserved lncRNAs exhibit distinct patterns of subcellular localization compared to conserved mRNAs, unlikely dependent on culture conditions (Figures 1 and S1). Importantly, the different processing and localization of conserved lncRNAs lead to distinct functions in ESCs derived from monkey and human compared to mESCs (Figures 2, 3, 4, 5, 6, and S2–S7), suggesting a new layer of understanding regarding the rapid evolution of lncRNAs and that the relatively high

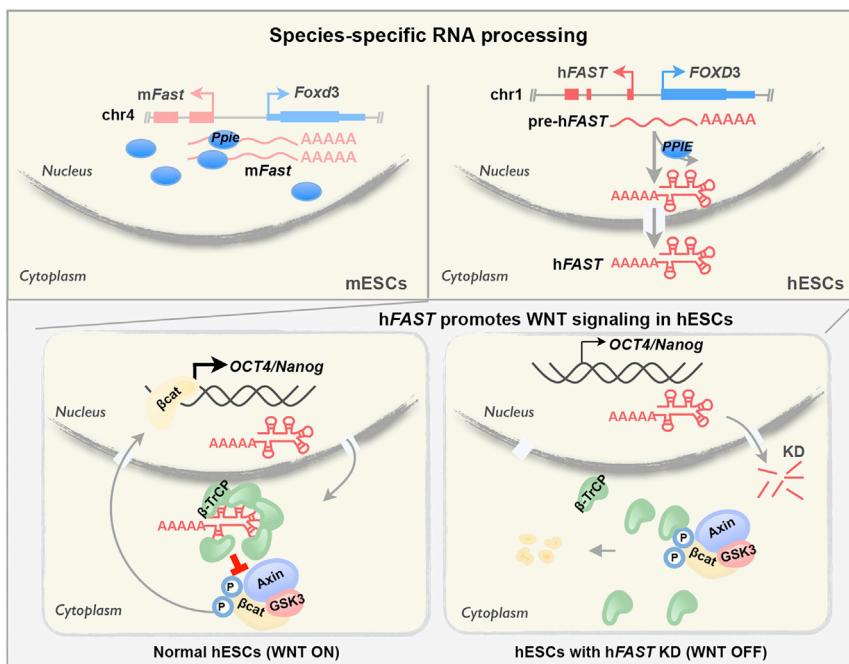


Figure 7. A Model of Distinct RNA Processing in Modulating the Non-conserved *FAST* Function in Pluripotency

Top: processing of the positionally conserved *FAST* lncRNA is not conserved in hESCs and mESCs. *mFast* is nuclear retained and partially processed in mESCs because of the high expression of *PPIE*. In hESCs, *PPIE* is expressed at a lower level, and *hFAST* is fully processed and localized to the cytoplasm. Bottom: in the cytoplasm of hESCs, *hFAST* binds to the WD40 domain of β -TrCP and blocks its interaction with phosphorylated β -catenin to prevent degradation, leading to activated WNT signaling, required for pluripotency.

evolutionary plasticity of lncRNAs can support species-specific gene expression programs (Figure 7).

Intriguingly, we found that conserved lncRNAs in mESCs in general exhibited enhanced nuclear retention compared with those in hESCs (Figures 1C–1N), and this correlated with less efficient splicing (Figures S6A and 6I). Seeking mechanisms that contribute to this specific lncRNA localization pattern revealed that differential expression of *trans*-regulators, such as *PPIE*, acts to regulate these differences (Figures 5, 6, S6, and S7). *PPIE* expression is decreased from mouse to primates (Figures 6F–6H and S7C–S7D), and this is accompanied by increased processing of conserved lncRNAs (Figures 1, 6C, and 6I). These observations suggest the possibility that gaining processing of lncRNA, such as *FAST*, is achieved by escaping the inhibitory effect of *PPIE* in the processing machinery during evolution. It remains to be determined how *PPIE* suppresses RNA processing in different types of cells. Given the abundance of *PPIE* (i.e., ~80,000 per H9 cell [Figure S7C] versus ~280,000 per R1 cell [Figure S7D]), it is possible that *PPIE* may form different, yet-to-be-defined complexes in H9 and R1 cells that are required for distinct RNA processing, even though *PPIE* alone could interact with pre-*hFAST* in H9 cell lysates (Figure S7H). Although the splicing suppressor *PPIE* likely accounts for the nuclear localization of many lncRNAs in mESCs (Figures 5, S6, and S7), other RBPs may also affect differential localization and remain to be identified.

Fully processed *hFAST* is located in the cytoplasm of hESCs (Figures 1J–1N and 6A) and is a previously uncharacterized lncRNA required for WNT signaling and pluripotency by blocking the interaction between β -TrCP and phosphorylated β -catenin, resulting in suppressed β -catenin degradation and WNT activation (Figures 3 and 4). However, *mFast* does not interact with β -TrCP or affect mESC pluripotency (Figures 2M, 2N, S4I,

and S2C). These findings suggest another new mechanism that differs in hESCs and mESCs. However, although expressed at low levels and primarily located in the nucleus, we cannot yet conclude that *mFast* has no function, only that it does not appear to play a role in WNT signaling or pluripotency maintenance. Because recent studies

have shown that lncRNAs can recruit or block transcriptional factors from WNT target genes (Di Cecilia et al., 2016; Ma et al., 2016), additional uncharacterized lncRNAs may act to regulate WNT signaling in mESCs.

The difference in subcellular localization of conserved lncRNAs between hESCs, monkey ESCs, and mESCs is remarkable (Figures 1, 5, and 6). In addition to the observation that differential nuclear retention of lncRNAs in mESCs and hESCs is functionally associated with pluripotency maintenance, as shown for *hFAST* in WNT signaling regulation (Figures 2, 3, and 4), it is possible that some cells may display stronger nuclear retention of lncRNAs than others and that such differentially processed lncRNAs may contribute to different functional outputs. Future studies are needed to investigate more general ramifications of this phenomenon in pluripotency and in other contexts.

STAR★METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Human cell lines
 - Monkey cell lines
 - Mouse cell lines
 - Bacterial strains
- METHOD DETAILS
 - Cell Transfection and Lentivirus Infection
 - Differentiation of hESCs
 - Reversion of Primed hESCs to a Naive-Like State
 - Differentiation of Epiblast-Like Cells (EpiLCs)

- Activation of WNT Signaling Pathway
- Plasmid Constructions
- Lentivirus Production and Cell Infection
- Protein Expression and Purification
- RNA Isolation, RT-qPCR and Northern Blotting (NB)
- Isolation of Cytoplasmic and Nuclear RNA
- *In vitro* RNA Transcription and Purification
- tRSA RNA Pull-down Assay and Western Blotting
- RNA Immunoprecipitation (RIP)
- *In vitro* RNA Protein Binding Assay
- Dig-labeled RNA Pull-down Assay
- Electrophoretic Mobility Shift Assay
- RNA *In Situ* Hybridization and smFISH
- Knockout mFast and hFAST by Multiplex CRISPR/Cas9 Assembly System
- Measurement of hFAST Copy Number
- Colony Formation Assay
- Alkaline Phosphatase Staining
- In-Cell SHAPE Probing
- SHAPE-MaP Reverse Transcription
- SHAPE-MaP Library Preparation and Sequencing
- Polyadenylated RNA Separation, rRNA Depletion for RNA-seq
- Library Preparation and Deep Sequencing

● QUANTIFICATION AND STATISTICAL ANALYSIS

- RNA-seq Data Processing
- Imaging Analysis of lncRNA Nuclear Localization
- Identification of Conserved lncRNAs
- Conservation Analysis of FAST
- Analysis of RNA Localization
- Stem Cell Maintenance Related Gene Analysis
- Stem-related Signaling Pathway Analysis
- Calculation of the SHAPE Reactivity
- hFAST Secondary Structure Modeling
- Prediction of hFAST ORFs
- Analysis of FAST Associated *trans*-factors
- Conservation Analysis of PPIE
- RNA Localization Analysis upon PPIE Knockdown
- Statistical Analysis

● DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2020.03.006>.

ACKNOWLEDGMENTS

We thank Fang Nan for SHAPE-MaP analysis, Jinsong Li for providing M21 ESCs, Naihe Jing for EpiSCs, and all lab members for discussion. This work was supported by the Ministry of Science and Technology of the People's Republic of China, China (2016YFA0100701), the Chinese Academy of Sciences, China (XDB19020104), the National Natural Science Foundation of China, China (31861143025, 31830108, 31821004, 31925011, 31725009, and 31730111), and the HHMI International Research Scholar Program, USA (55008728).

AUTHOR CONTRIBUTIONS

Conceptualization, L.-L.C.; Methodology, C.-J.G., X.-K.M., L.Y., and L.-L.C.; Investigation, C.-J.G., Y.-H.X., C.-C.Z., L.S., Y.-F.X., J.Z., S.W., L.Y., and L.-L.C.; Formal Analysis, C.-J.G., X.-K.M., L.Y., and L.-L.C.; Writing – Original Draft, C.-J.G., X.-K.M., Y.-H.X., and L.-L.C.; Writing – Review & Editing, L.-L.C., L.Y., G.G.C., and Y.W.; Funding Acquisition, L.-L.C. and L.Y.; Resources, C.-J.G. and X.-K.M.; Supervision, L.-L.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 23, 2019

Revised: January 5, 2020

Accepted: March 5, 2020

Published: April 6, 2020

REFERENCES

- Amaral, P.P., Leonardi, T., Han, N., Viré, E., Gascoigne, D.K., Arias-Carrasco, R., Büscher, M., Pandolfini, L., Zhang, A., Pluchino, S., et al. (2018). Genomic positional conservation identifies topological anchor point RNAs linked to developmental loci. *Genome Biol.* 19, 32.
- Bertram, K., Agafonov, D.E., Liu, W.T., Dybkov, O., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H., and Lührmann, R. (2017). Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* 542, 318–323.
- Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and Lührmann, R. (2008). Isolation of an active step I spliceosome and composition of its RNP core. *Nature* 452, 846–850.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* 224, 213–232.
- Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.
- Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., and Lewis, S.; AmiGO Hub; Web Presence Working Group (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* 25, 288–289.
- Carlevaro-Fita, J., and Johnson, R. (2019). Global Positioning System: Understanding Long Noncoding RNAs through Subcellular Localization. *Mol. Cell* 73, 869–883.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steimetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
- Chen, L.L. (2016). Linking Long Noncoding RNA Localization and Function. *Trends Biochem. Sci.* 41, 761–772.
- Chen, L.L., and Carmichael, G.G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* 35, 467–478.
- Chen, L.L., DeCerbo, J.N., and Carmichael, G.G. (2008). Alu element-mediated gene silencing. *EMBO J.* 27, 1694–1705.
- Chen, T., Xiang, J.F., Zhu, S., Chen, S., Yin, Q.F., Zhang, X.O., Zhang, J., Feng, H., Dong, R., Li, X.J., et al. (2015). ADAR1 is required for differentiation and neural induction by regulating microRNA processing in a catalytically independent manner. *Cell Res.* 25, 459–476.
- Chin, A., and Lécuyer, E. (2017). RNA localization: Making its way to the center stage. *Biochim. Biophys. Acta, Gen. Subj.* 1861 (11 Pt B), 2956–2970.

- Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., et al. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat. Biotechnol.* 33, 1173–1181.
- Deigan, K.E., Li, T.W., Mathews, D.H., and Weeks, K.M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. USA* 106, 97–102.
- Di Cecilia, S., Zhang, F., Sancho, A., Li, S., Aguiló, F., Sun, Y., Rengasamy, M., Zhang, W., Del Vecchio, L., Salvatore, F., and Walsh, M.J. (2016). RBM5-AS1 Is Critical for Self-Renewal of Colon Cancer Stem-like Cells. *Cancer Res.* 76, 5615–5627.
- Diederichs, S. (2014). The four dimensions of noncoding RNA conservation. *Trends Genet.* 30, 121–123.
- Fiddes, I.T., Lodewijk, G.A., Mooring, M., Bosworth, C.M., Ewing, A.D., Mantalas, G.L., Novak, A.M., van den Bout, A., Bishara, A., Rosenkrantz, J.L., et al. (2018). Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis. *Cell* 173, 1356–1369.e22.
- Fort, A., Hashimoto, K., Yamada, D., Salimullah, M., Keya, C.A., Saxena, A., Bonetti, A., Voineagu, I., Bertin, N., Kratz, A., et al.; FANTOM Consortium (2014). Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat. Genet.* 46, 558–566.
- Fuchs, S.Y., Spiegelman, V.S., and Kumar, K.G. (2004). The many faces of beta-TrCP E3 ubiquitin ligases: reflections in the magic mirror of cancer. *Oncogene* 23, 2028–2036.
- Giudice, G., Sánchez-Cabo, F., Torroja, C., and Lara-Pezzi, E. (2016). ATtRACT—a database of RNA-binding proteins and associated motifs. Database. Published online April 7, 2016. <https://doi.org/10.1093/database/baw035>.
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstruction of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146, 519–532.
- Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* 163, 712–723.
- Hezroni, H., Koppstein, D., Schwartz, M.G., Avrutin, A., Bartel, D.P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep.* 11, 1110–1122.
- Hu, S.B., Xiang, J.F., Li, X., Xu, Y., Xue, W., Huang, M., Wong, C.C., Sagum, C.A., Bedford, M.T., Yang, L., et al. (2015). Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus. *Genes Dev.* 29, 630–645.
- Jin, W., Wang, Y., Liu, C.P., Yang, N., Jin, M., Cong, Y., Wang, M., and Xu, R.M. (2016). Structural basis for snRNA recognition by the double-WD40 repeat domain of Gemin5. *Genes Dev.* 30, 2391–2403.
- Johnsson, P., Lipovich, L., Grandér, D., and Morris, K.V. (2014). Evolutionary conservation of long non-coding RNAs: sequence, structure, function. *Biochim. Biophys. Acta* 1840, 1063–1071.
- Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30.
- Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360.
- Koestenbauer, S., Zech, N.H., Juch, H., Vanderzwalm, P., Schoonjans, L., and Dohr, G. (2006). Embryonic stem cells: similarities and differences between human and murine embryonic stem cells. *Am. J. Reprod. Immunol.* 55, 169–180.
- Kutter, C., Watt, S., Stefflova, K., Wilson, M.D., Goncalves, A., Ponting, C.P., Odom, D.T., and Marques, A.C. (2012). Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet.* 8, e1002841.
- Lam, A.Q., Freedman, B.S., Morizane, R., Lerou, P.H., Valerius, M.T., and Bonventre, J.V. (2014). Rapid and efficient differentiation of human pluripotent stem cells into intermediate mesoderm that forms tubules expressing kidney proximal tubular markers. *J. Am. Soc. Nephrol.* 25, 1211–1225.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Moffat, J., et al. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* 141, 872–883.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Lian, X., Zhang, J., Azarin, S.M., Zhu, K., Hazeltine, L.B., Bao, X., Hsiao, C., Kamp, T.J., and Palecek, S.P. (2013). Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat. Protoc.* 8, 162–175.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
- Licalatosi, D.D., and Darnell, R.B. (2010). RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* 11, 75–87.
- Liu, C.X., Li, X., Nan, F., Jiang, S., Gao, X., Guo, S.K., Xue, W., Cui, Y., Dong, K., Ding, H., et al. (2019). Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity. *Cell* 177, 865–880.e21.
- Lubelsky, Y., and Ulitsky, I. (2018). Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 555, 107–111.
- Ma, Y., Yang, Y., Wang, F., Moyer, M.P., Wei, Q., Zhang, P., Yang, Z., Liu, W., Zhang, H., Chen, N., et al. (2016). Long non-coding RNA CCAL regulates colorectal cancer progression by activating Wnt/β-catenin signalling pathway via suppression of activator protein 2α. *Gut* 65, 1494–1504.
- Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basukar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., and Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47 (W1), W636–W641.
- Major, M.B., Camp, N.D., Berndt, J.D., Yi, X., Goldenberg, S.J., Hubbert, C., Biechele, T.L., Gingras, A.C., Zheng, N., MacCoss, M.J., et al. (2007). Wilms tumor suppressor WTX negatively regulates WNT/beta-catenin signaling. *Science* 316, 1043–1046.
- Makowski, W., Zhang, J., and Boguski, M.S. (1996). Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res.* 6, 846–857.
- Mélé, M., Mattioli, K., Mallard, W., Shechner, D.M., Gerhardinger, C., and Rinn, J.L. (2017). Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. *Genome Res.* 27, 27–37.
- Miyagawa, R., Tano, K., Mizuno, R., Nakamura, Y., Ijiri, K., Rakwal, R., Shibato, J., Masuo, Y., Mayeda, A., Hirose, T., and Akimitsu, N. (2012). Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA* 18, 738–751.
- Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepfer, A.M., Hinkle, G., Piqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K., et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 124, 1283–1298.
- Mustoe, A.M., Busan, S., Rice, G.M., Hajdin, C.E., Peterson, B.K., Ruda, V.M., Kubica, N., Nutiu, R., Baryza, J.L., and Weeks, K.M. (2018). Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* 173, 181–195.e18.
- Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Baker, J.C., Grützner, F., and Kaessmann, H. (2014). The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* 505, 635–640.
- Nitsche, A., and Stadler, P.F. (2017). Evolutionary clues in lncRNAs. *Wiley Interdiscip. Rev. RNA* 8, e1376.
- Nusse, R., and Clevers, H. (2017). Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* 169, 985–999.

- Pera, M.F., and Tam, P.P. (2010). Extrinsic regulation of pluripotent stem cells. *Nature* **465**, 713–720.
- Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., and Yamamoto, T. (2014). Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* **4**, 5400.
- Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat. Med.* **10**, 55–63.
- Schiene-Fischer, C. (2015). Multidomain Peptidyl Prolyl cis/trans Isomerases. *Biochim. Biophys. Acta* **1850**, 2005–2016.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., and Proudfoot, N.J. (2017). Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Mol. Cell* **65**, 25–38.
- Seong, H.A., Jung, H., Choi, H.S., Kim, K.T., and Ha, H. (2005). Regulation of transforming growth factor-beta signaling and PDK1 kinase activity by physical interaction between PDK1 and serine-threonine kinase receptor-associated protein. *J. Biol. Chem.* **280**, 42897–42908.
- Smola, M.J., Rice, G.M., Busan, S., Siegfried, N.A., and Weeks, K.M. (2015). Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669.
- Sokol, S.Y. (2011). Maintaining embryonic stem cell pluripotency with Wnt signaling. *Development* **138**, 4341–4350.
- Tang, Z., Wu, Y., Yang, Y., Yang, Y.T., Wang, Z., Yuan, J., Yang, Y., Hua, C., Fan, X., Niu, G., et al. (2017). Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutionary conservation in *Sus scrofa*. *Sci. Rep.* **7**, 43166.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.
- Ulitsky, I. (2016). Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat. Rev. Genet.* **17**, 601–614.
- Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537–1550.
- Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci. USA* **105**, 3386–3391.
- Wang, Y., Hu, S.B., Wang, M.R., Yao, R.W., Wu, D., Yang, L., and Chen, L.L. (2018). Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat. Cell Biol.* **20**, 1145–1158.
- Wray, J., and Hartmann, C. (2012). WNTing embryonic stem cells. *Trends Cell Biol.* **22**, 159–168.
- Wu, G., Xu, G., Schulman, B.A., Jeffrey, P.D., Harper, J.W., and Pavletich, N.P. (2003). Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. *Mol. Cell* **11**, 1445–1456.
- Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., Yan, Z., Li, D., and Li, J. (2013). Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659–662.
- Xing, Y.H., Yao, R.W., Zhang, Y., Guo, C.J., Jiang, S., Xu, G., Dong, R., Yang, L., and Chen, L.L. (2017). SLERT Regulates DDX21 Rings Associated with Pol I Transcription. *Cell* **169**, 664–678.e16.
- Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G., and Chen, L.L. (2011). Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16.
- Yang, H., Liu, Z., Ma, Y., Zhong, C., Yin, Q., Zhou, C., Shi, L., Cai, Y., Zhao, H., Wang, H., et al. (2013). Generation of haploid embryonic stem cells from *Macaca fascicularis* monkey parthenotes. *Cell Res.* **23**, 1187–1200.
- Yao, R.W., Wang, Y., and Chen, L.L. (2019). Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* **21**, 542–551.
- Ye, Y., Tang, X., Sun, Z., and Chen, S. (2016). Upregulated WDR26 serves as a scaffold to coordinate PI3K/AKT pathway-driven breast cancer cell growth, migration, and invasion. *Oncotarget* **7**, 17854–17869.
- Yin, Q.F., Yang, L., Zhang, Y., Xiang, J.F., Wu, Y.W., Carmichael, G.G., and Chen, L.L. (2012). Long noncoding RNAs with snoRNA ends. *Mol. Cell* **48**, 219–230.
- Yin, Q.F., Chen, L.L., and Yang, L. (2015). Fractionation of non-polyadenylated and ribosomal-free RNAs from mammalian cells. *Methods Mol. Biol.* **1206**, 69–80.
- Zhang, B., Gunawardane, L., Niazi, F., Jahanbani, F., Chen, X., and Valadkhan, S. (2014). A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. *Mol. Cell. Biol.* **34**, 2318–2329.
- Zhu, Y., Wang, Y., Xia, C., Li, D., Li, Y., Zeng, W., Yuan, W., Liu, H., Zhu, C., Wu, X., and Liu, M. (2004). WDR26: a novel Gbeta-like protein, suppresses MAPK signaling pathway. *J. Cell. Biochem.* **93**, 579–587.

STAR★METHODS**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-Oct-3/4	Santa Cruz	Cat# sc-5279; RRID: AB_628051
Anti-β-Actin	Sigma	Cat# A3854; RRID: AB_262011
Mouse monoclonal anti-FLAG	Sigma	Cat# F1804; RRID: AB_262044
Rabbit monoclonal anti-β-TrCP	CST	Cat# 4394S; RRID: AB_10545763
Mouse monoclonal anti-β-catenin	BD PharMingen	Cat# 610154; RRID: AB_397555
Rabbit monoclonal anti-GSK3α/β	CST	Cat# 5676; RRID: AB_10547140
Rabbit polyclonal anti-Phospho-β-catenin	CST	Cat# 9561; RRID: AB_331729
Rabbit monoclonal anti-Axin1	CST	Cat# 2087; RRID: AB_2274550
Rabbit monoclonal anti-LPR6	CST	Cat# 3395; RRID: AB_1950408
Mouse monoclonal anti-p54nrb	BD Biosciences	Cat# 611278; RRID: AB_398806
Mouse monoclonal anti-GAPDH	Abways	Cat# AB0038; RRID: AB_2828027
Rabbit polyclonal anti-Nanog	Abcam	Cat# ab21624; RRID: AB_446437
Rabbit polyclonal anti-Ubiquitin	CST	Cat# 3933; RRID: AB_2180538
Mouse monoclonal anti-PPIE	Santa Cruz	Cat# sc-100700; RRID: AB_2169267
Rabbit polyclonal anti-STRAP	ABclonal	Cat# A5964; RRID: AB_2766689
Bacterial and Virus Strains		
BL21	Transgen Biotech	Cat# CD801
T1	Transgen Biotech	Cat# CD501-01
Chemicals, Peptides, and Recombinant Proteins		
BIO	Sigma	Cat# B1686
CHIR99021	Selleck	Cat# S1236
PD0325901	Selleck	Cat# S1036
LIF	Millipore,	Cat# ESG1107
Activin A	R&D	Cat# 338-AC-010
Recombinant Human FGF basic (146 aa) Protein, CF	R&D	Cat# 233-FB 001MG/CF
BMP4	R&D	Cat# 314-BP-010
Y-27632	Tocris	Cat# 1254
Mercaptoethanol	Sigma	Cat# M-7522
monothioglycerol	Sigma	Cat# M1753
Bovine Serum Albumin	Sigma	Cat# B2064
Pen-strep	Millipore	Cat# TMS-AB2-C
Nucleosides	Millipore	Cat# ES-008-D
insulin	Tocris	Cat# 560922
Transferrin	Roche	Cat# 10652202001
DMEM	GIBCO	Cat# 11965
Neurobasal	GIBCO	Cat# 21103049
ES-DMEM	GIBCO	Cat# 12430
DMEM/F-12	GIBCO	Cat# 11330
Ham's F-12 Nutrient Mixture	GIBCO	Cat# 11765047
IMDM	GIBCO	Cat# 12440061
FBS	GIBCO	Cat# 10438-026
ES-FBS	GIBCO	Cat# 16000-044
KnockOut Serum Replacement	GIBCO,	Cat# 10828

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RSeT Feeder-Free Medium	STEMCELL	Cat# 05975
mTeSR1	STEMCELL	Cat# 85850
L-Glutamin	GIBCO	Cat# 25030081
NEAA	GIBCO	Cat# 11140
B27 supplement minus vitamin A	GIBCO	Cat# 12587010
N-2 Supplement	GIBCO	Cat# 17502048
DPBS	Invitrogen	Cat# 14190-135
TrypLE Express Enzyme (1X)	GIBCO	Cat# 12604021
Accutase Cell Dissociation Reagent	GIBCO	Cat# A1110501
Dispase	Invitrogen	Cat# 17105041
Matrigel	Corning-BioCoat	Cat# 356231
Fibronectin Human, Plasma	GIBCO	Cat# 33016015
Gelatin	Sigma	Cat# G1890-
Polybrene	Sigma	Cat# TR-1003
FuGENE HD	Promega	Cat# PRE2311
Lipofectamine 2000 Reagent	Thermo	Cat# 11668019
Lenti-Concentrin	ExCell Bio	Cat# EMB810A-1
Crystal violet	Beyotime	Cat# C0121
TRIzol Reagent	Ambion	Cat# 15596018
Protease Inhibitor Cocktail, mini-Tablet	MedChem Express	Cat#: HY-K0011
Ribonucleoside Vanadyl Complex	NEB	Cat# S1402S
Dynabeads Protein G	Invitrogen	Cat# 1003D
Dynabeads MyOne Streptavidin C1	GIBCO	Cat# 65001
Glycerol	ABCONE	Cat# G46055
HEPES	ABCONE	Cat# H33755
TWEEN 20	ABCONE	Cat# P87875
Triton X-100	ABCONE	Cat# X10010
Agarose	ABCONE	Cat# A47902
Bovine Serum Albumin	ABCONE	Cat# A23088
2-methylnicotinic acid imidazolide (NAI)	EMD Millipore	Cat# 03-310
Critical Commercial Assays		
DNA-free kit	Ambion	Cat# AM1907
DIG Northern Starter Kit	Roche	Cat# 12039672910
Alkaline Phosphatase Detection Kit	Millipore	Cat# SCR004
PrimeScript™ RT Master Mix (Perfect Real Time)	TaKaRa	Cat# RR036A
SuperScript III Reverse Transcriptase	Invitrogen	Cat# 18080044
SuperScript II Reverse Transcriptase	Invitrogen	Cat# 18064071
RiboMAX Large Scale RNA Production System	Promega	Cat# P1300
One-tube General Sample DNAup for PCR	Sangon Biotech	Cat# B518401-0500
2×T5 Super PCR Mix (Colony)	Tsingke Biological Technology	Cat# TSE005
StarPrep Gel Extraction Kit StarPrep	GenStar	Cat# D205-04
2×TransTaq® High Fidelity (HiFi) PCR SuperMix II	TransGen Biotech	Cat# AS131-21
RiboMinus Transcriptome Isolation Kit	Invitrogen	Cat# K155002
Deposited Data		
All H9 and R1 RNA-seq datasets	This paper	GEO: GSE143496 NODE: OEP000734
SHAPE-Map data of hFAST RNA	This paper	GEO: GSE143496 NODE: OEP000734
All original unprocessed data related to this paper	This paper	https://dx.doi.org/10.17632/f4nkpt5h82.1

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
H9	WiCell	Cat# WA09
H1	WiCell	Cat# WA01
CT1	UCConn Stem Cell Core	NIHhESC-10-0068
WIBR3	Whitehead Institute	NIHhESC-10-0079
M21	Yang et al., 2013	N/A
R1/E	ATCC	Cat# SCRC-1036
E14	ATCC	Cat# CRL-1821
NIH/3T3	Stem Cell Bank, Chinese Academy of Sciences	Cat#: GNM6
EpiSC	Modified from Brons et al. (2007)	N/A
293FT	Thermo Fisher	Cat#: R70007
Oligonucleotides		
shRNA target sequences: hFAST shRNA1: AGAAGCCATAACTGGCTACTT	This paper	N/A
shRNA target sequences: hFAST shRNA2: GCTGGGATGTGGATTAAATTCT	This paper	N/A
sgRNA target sequences: mFast sgRNA1: AATAGGTTAACATCGATAGCTA	This paper	N/A
sgRNA target sequences: mFast sgRNA2: CGGGAAACCCCAGCTGATGG	This paper	N/A
Other primers and shRNA target sequences, see Table S7	This paper	N/A
Recombinant DNA		
pLKO.1-TRC vector	Moffat et al., 2006	Addgene Plasmid # 10878
px330 vector	Wu et al., 2013	Addgene Plasmid # 98750
pX330A 1x4	Sakuma et al., 2014	Addgene Plasmid #58768
pcDNA3	Invitrogen	Cat#: A150228
p23-hFAST	This paper	N/A
p23-mFast	This paper	N/A
P23-Flag-PPIE	This paper	N/A
P23-Flag-Ppie	This paper	N/A
P23-Flag-β-TrCP	This paper	N/A
P23-Flag-m-β-TrCP	This paper	N/A
P23-Flag-WDR26	This paper	N/A
P23-Flag-STRAP	This paper	N/A
pcDNA3-hFAST-Flag-KI-1	This paper	N/A
pcDNA3-hFAST-Flag-KI-2	This paper	N/A
pcDNA3-hFAST-Flag-KI-3	This paper	N/A
pC013-His-Flag-β-TrCP	This paper	N/A
Software and Algorithms		
GraphPad Prism	GraphPad Software	https://www.graphpad.com/scientificsoftware/prism/
softWoRx 7.0	GE Healthcare	http://incelldownload.gehealthcare.com/bin/download_data/SoftWoRx/7.0.0/SoftWoRx.htm
Fiji/ImageJ	Fiji/ImageJ	https://imagej.net/Fiji
Trimmomatic Version: 0.35	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
HISAT2 Version: 2.0.9	Kim et al., 2015	https://ccb.jhu.edu/software/hisat2/index.shtml
Samtools Version: 0.1.19	(Li et al., 2009)	http://samtools.sourceforge.net/
UniProt	The UniProt Consortium	https://www.uniprot.org:443/
R Version: 3.4.1	https://www.r-project.org	https://www.r-project.org
FPKM	Trapnell et al., 2010	N/A
LiftOver	UCSC Genome Browser Utilities	https://genome-store.ucsc.edu/
blastn Version: 2.2.30+	Camacho et al., 2009	https://blast.ncbi.nlm.nih.gov/Blast.cgi
featureCounts Version: 1.5.1	Liao et al., 2014	http://bioinf.wehi.edu.au/featureCounts/
ShapeMaper Version: 1.2	Smola et al., 2015	http://chem.unc.edu/rna/software.html
ATtRACT database	Giudice et al., 2016	https://attract.cnic.es/
EMBOSS Matcher	Madeira et al., 2019	https://www.ebi.ac.uk/Tools/psa/emboss_matcher/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ling-Ling Chen (linglingchen@sibcb.ac.cn). All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human cell lines

Human cell lines including H9 and H1 were purchased from WiCell, CT1 was purchased from UConn Stem Cell Core, WIBR3 was obtained from Whitehead Institute, 293FT cells were purchased from ThermoFisher, and were originally authenticated using STR profiling.

H9 (human female embryo origin), H1 (human male embryo origin) and CT1 (human female embryo origin) were maintained in DMEM/F-12 supplemented with 20% KnockOut Serum Replacement, 1mM Glutamin, 0.1mM NEAA and 0.1mM mercaptoethanol and 4 ng/mL β-FGF and cultured with irradiated-MEF feeder cells with daily changed hESC cultured medium and passaged weekly (Chen et al., 2015). The primed H9, H1 and CT1 cells were cultured in conditioned medium (CM). The naive-like state H9 cells were culture in RSeT medium. CM was made by our lab and performed as described below. 2.7×10^7 Irradiated-MEF feeder cells were cultured in T525 dish with 125mL hESC without FGF medium for 12 days and changed and collected medium every day. CM was finally prepared when these collected media were mixed and then added β-FGF to a final concentration of 4ng/mL.

WIBR3 (human female embryo origin) cells were originally derived by Rudolf Jaenisch's Lab (Lengner et al., 2010). The primed state WIBR3 cells were cultured in mTesR medium. The naive-like state WIBR3 cells were culture in RSeT medium. Primed or naive-like state cells were passaged every 4 days with Accutase, and medium was changed every day.

293FT cells (human fetus origin) were maintained in DMEM supplemented with 10% Fetal Bovine Serum (FBS). We maintained cell lines at 37°C in a 5% CO₂ cell culture incubator and tested all cell lines routinely for mycoplasma contamination.

Monkey cell lines

Monkey ES cell line M21 was isolated from ICSI-derived blastocysts in Macaca fascicularis monkeys in Jinsong Li lab (Yang et al., 2013). M21 cells maintenance was performed as previously described (Yang et al., 2013) with DMEM/F-12 supplemented with 20% KnockOut Serum Replacement, 2mM L-Glutamin, 0.1mM NEAA and 0.1mM mercaptoethanol, 2 mM Thiazovivin and 10 mM Y-27632 and 10ng/mL β-FGF. The culture medium was changed daily, and ESC colonies were split every 3-5 days manually or disaggregated by Accutase. We maintained cell lines at 37°C in a 5% CO₂ cell culture incubator and tested all cell lines routinely for mycoplasma contamination.

Mouse cell lines

Mouse cell lines including R1/E, E14 were purchased from the American Type Culture Collection (ATCC; <https://www.atcc.org:443/>). NIH/3T3 cells were kindly provided by Stem Cell Bank, Chinese Academy of Sciences. EpiSCs were derived from the late epiblast layer, which is dissected from E5.75 pre-gastrula stage mouse embryo as previously described (Brons et al., 2007) and kindly pro-

vided by Naihe Jing lab. R1 (mouse male embryo origin) and E14 (mouse male embryo origin) cells were cultured in DMEM (GIBCO, 12430) supplemented with 15% Fetal Bovine Serum (GIBCO, 16000-044), 1mM Glutamin, 1% NEAA, 1% Nucleosides and 1% penicillin/streptomycin, 0.1mM mercaptoethanol and 1 μ M PD0325901, 3 μ M CHIR99021 and 1000 units/mL mouse LIF. EpSCs were maintained in chemically defined medium (CDM) supplemented with 20ng/mL Activin A and 12ng/mL bFGF as previously described (Brons et al., 2007). The composition of CDM was 50% IMDM plus 50% F12 NUT-MIX, supplemented with 7 μ g/mL insulin, 15 μ g/mL transferrin, 450 μ M of monothioglycerol and 5 mg/mL bovine serum albumin fraction. NIH/3T3 (mouse embryonic fibroblast origin) were maintained in DMEM supplemented with 10% Fetal Bovine Serum (FBS). We maintained cell lines at 37°C in a 5% CO₂ cell culture incubator and tested all cell lines routinely for mycoplasma contamination.

Bacterial strains

E.coli expression strain BL21 [Tranetta (DE3) chemically competent cell] were procured from Transgen Biotech (Cat# CD801) and were grown in LB culture at 37°C.

METHOD DETAILS

Cell Transfection and Lentivirus Infection

Human, mouse and monkey cell lines were maintained using standard protocols from ATCC. To transfect hESCs with plasmids, H9 cells were split into small colonies with dispase and passaged on Matrigel coated dish with CM one day in advance. Plasmids were carried out with FuGENE HD (Roche) according to the manufacturer's protocol. To transfect R1 cells with plasmids, R1 cells were digested into single cell with Type E. Plasmids were carried out with Lipofectamine 2000 Reagent (Thermo) according to the manufacturer's protocols. To infect hESCs with lentivirus, H9, H1 or CT1 cells were digested with Accutase to single cells and grown on Matrigel with CM containing 5 μ L concentrated lentivirus, 5 μ g/mL polybrene and 10 μ M Y-27632 at 37°C. To infect 293FT, NIH/3T3 or R1/E cells with lentivirus, cells were incubated with culture medium containing 10 μ L concentrated lentivirus and 5 μ g/mL polybrene at 37°C for one day. Cells were infected the second time with pre-warmed medium containing 5 μ g/mL polybrene on the plates at the second day. Four days after the first infection, protein or total RNAs were collected for further analyses.

Differentiation of hESCs

Differentiation of hESCs into ectoderm, mesoderm cells and trophoblasts was performed as previously described (Chen et al., 2015; Lam et al., 2014; Chen and Carmichael, 2009).

Reversion of Primed hESCs to a Naive-Like State

The reversion of primed hESCs to a naive-like state was performed by culturing H9 cells under RSeT Feeder-Free Medium (STEMCELL) according to the manufacturer's protocol. H9 cells were split into small colonies with dispase and passaged on Matrigel coated dish with CM for 24 hours. Then CM was removed by adding 2mL RSeT Feeder-Free Medium and the cells were incubated at 37°C under hypoxic conditions. After 6 days reversion in RSeT Feeder-Free Medium, H9 cells were passaged by TrypLE and maintained in RSeT Feeder-Free Medium with 5 μ M Y27632 for 24 hours, followed by replacing RSeT Feeder-Free Medium every day. The reverting efficiency was monitored by increased expression of naive-like associated genes (an example was shown in Figure S1I).

Differentiation of Epiblast-Like Cells (EpiLCs)

Differentiation of EpiLCs from mouse ESCs was performed as previously described (Hayashi et al., 2011). EpiLCs were induced by plating 2.5×10^5 R1 ESCs on a 6-well plate coated with 16.7 mg/mL human plasma fibronectin in N2B27 medium containing 20 ng/mL activin A, 12 ng/mL bFGF and KSR (1%). The medium was changed every day maintained for two days.

Activation of WNT Signaling Pathway

Activation of WNT signaling pathway was performed as previously described (Lian et al., 2013; Sato et al., 2004). H9 cells were treated with 2 μ M BIO (sigma) or 4 μ M CHIR99021 (Selleck) under CM for 3 days.

Plasmid Constructions

To knock down hFAST, other lncRNAs or mRNAs, DNA sequences for shRNAs were designed by the GPP Web Portal online tool (<https://portals.broadinstitute.org/gpp/public/>). DNA oligos of shRNA sequences for targeted RNAs or scramble shRNAs were individually cloned into pLKO.1-TRC vector (Moffat et al., 2006). To construct plasmids to detect hFAST protein-coding ability, predicted ORFs were individually cloned into pcDNA3 vector. Flag sequence (GACTACAAAGACGATGACGATAAG) was knocked-in the three continuous nucleotides downstream of the predicted small ORFs and plasmids were constructed by mutant PCR. To construct plasmids for overexpressing hFAST or mFast, the full length sequence flanked by its natural intron, splicing sites and exons was PCR amplified from the genomic DNA of H9 or R1 cells and was cloned into the p23-phage vector. To construct plasmids to overexpress flag-tagged β -TrCP, its CDS sequence was amplified from H9 cDNA library. Truncated β -TrCP plasmids were constructed using Mut Express MultiS Fast Mutagenesis Kit (Vazyme C213-01). To construct plasmids for protein expression and purification, the full length of β -TrCP CDS sequence with Flag-tagged at its N-terminal was amplified from the H9 cDNA library and cloned into the pC013 vec-

tor. The full length of human or mouse PPIE CDS sequence with the Flag-tagged at its N-terminal was amplified from H9 or R1 cDNA libraries and cloned into the pET-28a vector. All plasmids were confirmed by Sanger sequencing. Primers for plasmid constructions were listed in [Table S7](#).

Lentivirus Production and Cell Infection

To produce lentiviral particles, 5×10^6 293FT cells in a 10 cm dish were co-transfected with 10 μg pLKO.1 shRNA, 7.5 μg of psPAX2 and 3 μg pMD2.G plasmids. The supernatant containing viral particles was harvested twice at 48h and 72h after transfection, then filtered through Millex-GP Filter Unit (0.22 μm pore size, Millipore). Viral particles were concentrated about 100-fold by Lenti-Concentration Virus Precipitation Solution (ExCell Bio), resuspended in PBS containing 0.1% BSA, and stored at -80°C until use.

Protein Expression and Purification

For protein purification, Flag-tagged β -TrCP was cloned into the bacterial expression pC013 vector (6x His/Twin Strep SUMO, a pET-based expression vector) and was transformed into *E. coli* expression strain BL21 [Transetta (DE3) chemically competent cell (TRANSGEN BIOTECH, CD801)]. After transformation, a single colony was inoculated in 10 mL LB media supplemented with 100 $\mu\text{g}/\text{L}$ ampicillin at 250 rpm, 37°C . After overnight growth, the culture was diluted 100-fold into 1L LB medium supplemented with 100 $\mu\text{g}/\text{L}$ ampicillin. Absorbance was monitored at a wavelength of 600 nm, and upon reaching an optical density (OD600) of 0.6 - 0.8, IPTG was added to the LB medium at the final concentration of 0.2 mM for the induction of protein expression. After overnight incubation at 250 rpm, 16°C , cell pellets were harvested by centrifugation (5,000 rpm, 10 min, 4°C), resuspended in the lysis buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM DTT, pH 8.0) supplemented with 1 mg/mL lysozyme on ice for 30 min, and sonicated for 10 min (5 s on/off) on ice. After centrifugation at 10,000 rpm for 30 min at 4°C , the supernatant cell lysate was applied to StrepTactin Sepharose (GE) and incubated with rotation for 2 hours at 4°C . The protein-bound Sepharose was washed three times with lysis buffer. The Sepharose was resuspended in SUMO digest buffer (30 mM Tris-HCl, 500 mM NaCl 1 mM DTT, 0.15% Igepal (NP-40), pH 8.0) along with 100 Units of SUMO protease (ThermoFisher) and incubated overnight at 4°C with rotation. The concentration of purified protein was determined using Modified Bradford Protein Assay Kit (Sangon Biotech, C503041) and checked by Coomassie blue staining.

RNA Isolation, RT-qPCR and Northern Blotting (NB)

Total RNA from each cultured cell line or cultured cells with different treatments was extracted with Trizol Reagent (Invitrogen) according to the manufacturer's protocol. For RT-qPCR, after treatment with DNase I (Invitrogen, DNA-free kit), the cDNA synthesis was carried out using PrimeScript RT Master Mix (TaKaRa) according to the manufacturer's protocol. RT-qPCR was carried out using SYBR Green Realtime PCR Master Mix (TOYOBO) and a StepOnePlus real-time PCR system (Applied Biosystems). The relative expression of different sets of genes was quantified to *Actin* mRNA. The relative N/C ratio was calculated by normalizing cytoplasmic RNA to *gapdh* mRNA and nuclear RNA to *U6* RNA.

NB was carried out according to the manufacturer's protocol (DIG Northern Starter Kit, Roche). RNA was loaded on native agarose gels or denatured PAGE gels. Digoxigenin (Dig) labeled antisense probes were generated using T7 RNA polymerase or SP6 RNA polymerase by *in vitro* transcription with the RiboMAX Large Scale RNA Production System (Promega). All primers used for RT-qPCR or probes were listed in [Table S7](#).

Isolation of Cytoplasmic and Nuclear RNA

Fractionation of the cytoplasmic and nuclear RNA in H9 and R1 cells was performed as described ([Chen et al., 2008](#)) with modifications. 1×10^6 cells were rinsed twice with ice-cold PBS, and centrifuged at 1,000 rpm for 3 min. Cell pellet was suspended by gentle pipetting in 200 mL lysis buffer (10 mM Tris pH 8.0, 140 mM NaCl, 1.5 mM MgCl₂, 0.5% Igepal, 2 mM Ribonucleoside Vanadyl Complex), and incubated on ice for 5 min. During the incubation, one fifth of the lysate was added to 1 mL Trizol for total RNA extraction. The rest of the lysate was centrifuged at 1,000 g for 3 min at 4°C to pellet the nuclei and the supernatant was the cytoplasmic fraction. To obtain pure cytoplasmic RNA, the supernatant fraction was centrifuged at 14,000 rpm for 10 min at 4°C and then collected the supernatant carefully to a new tube and extracted the RNA with Trizol. To obtain pure nuclear RNA, nuclear pellets underwent additional wash with 200 μL lysis buffer, one additional wash by adding 0.5% deoxycholic acid into the lysis buffer. Finally, the purified nuclei were resuspended in 200 μL lysis buffer followed by extraction with Trizol. Fractionated RNA from the same amounts of cells was used for cDNA production, RT-qPCR, PCR and RNA-seq analysis.

In vitro RNA Transcription and Purification

Linear RNA was *in vitro* transcribed from T7 expression vector prepared by RiboMax large RNA production system (Promega) according to the manufacturer's protocol with slight modifications. Briefly, 1 μg PCR-amplified T7-DNA fragments were incubated with 2 μL T7 RNA polymerase enzyme and 0.5 mM Dig labeled dNTPs (Roache). *In vitro* transcription was carried out for 3 hr at 37°C , followed by DNase I treatment for 30 min at 37°C to remove DNA templates. Transcribed RNA was precipitated with ethanol and washed with 75% ethanol and resuspended in RNase-free water. For *in vitro* transcribed RNA purification, 10 μg transcribed RNA was resolved on denaturing urea polyacrylamide gel and visualized by Ethidium bromide staining. Corresponding bands on denaturing urea polyacrylamide gel were excised for further purification.

tRSA RNA Pull-down Assay and Western Blotting

tRSA RNA pull-down assays were carried out as described (Xing et al., 2017) with modifications. Full-length hFAST was cloned into pcDNA3 plasmid with the tRSA tag at its 5' end. RNA product was *in vitro* transcribed using the T7 RiboMAX Large-Scale RNA Production System (Promega). 10 mg per reaction of synthetic RNA was denatured for 5 min at 65°C in RNA Structure buffer (10 mM HEPES pH 7.0, 10 mM MgCl₂) and slowly cooled down to room temperature. Then, folded RNA was incubated with 50 mL of streptavidin Dynabeads (Invitrogen) for 20 min at 4°C in the presence of 2 U/mL RNasin (Promega). H9 cells (2x10⁷) were harvested and resuspended in 1 mL of lysis buffer [10 mM HEPES pH 7.0, 200 mM NaCl, 1 mM DTT, 1% Triton X-100, protease inhibitor cocktail (MedChem Express)] followed by sonication for 4x10s with an interval of 1 min on ice and then centrifuged at 14,000 rpm for 10 min at 4°C. The supernatant was pre-cleared with 50 mL of streptavidin Dynabeads for 20 min at 4°C followed by the addition of 20 mg/mL yeast tRNA for 20 min at 4°C. Then the pre-cleared lysate was added to folded RNAs and incubated for 3.5 hr at 4°C followed by washing 4x5 min with wash buffer [10 mM HEPES pH 7.0, 400 mM NaCl, 1 mM DTT, 1% Triton X-100, protease inhibitor cocktail (MedChem Express), 2 mM RVC]. To harvest the protein complex, 50 mL of 1x SDS loading buffer was added and boiled for 10 min at 100°C. Retrieved proteins were analyzed by WB with primary antibodies.

RNA Immunoprecipitation (RIP)

Cells growing in 10 cm dishes were rinsed twice with ice-cold PBS, harvested in 10 mL ice-cold PBS and then centrifuged at 1,000 rpm for 5 min at 4°C. Cells were resuspended in 1 mL RIP buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5% Igepal, 1 mM PMSF, 1X protease inhibitor cocktail (MedChem Express) and 2 mM RVC) and subjected to three rounds of gentle sonication. Cell lysates were centrifuged at 12,000 rpm for 15 min at 4°C and the supernatants were precleared with 15 mL Dynabeads Protein G (Invitrogen) to get rid of non-specific binding. Then, the precleared lysates were used for IP with anti-Flag antibodies (Sigma) or anti-β-TrCP antibodies (CST). IP was carried out for 2 hours at 4°C. The beads were washed three times with high salt buffer and two times with the RIP buffer, followed by extraction with elution buffer (100 mM Tris, pH 6.8, 4% SDS, and 10mM EDTA) at room temperature for 10 min. One-third of the eluted sample was used for WB and the remaining was used for RNA extraction. The RNA enrichment was assessed by RT-qPCR.

In vitro RNA Protein Binding Assay

In vitro synthesized and purified linear hFAST and egfp RNAs were heated for 5 min at 65°C in RNA folding buffer (10 mM HEPES and 10 mM MgCl₂) and slowly cooled down to room temperature. Purified flag-tagged-β-TrCP protein were incubated with Dynabeads Protein G (Invitrogen) for 1h at 4°C in 0.2 mL binding buffer (50 mM HEPES at pH 7.0, 150 mM NaCl, 10 mM MgCl₂, 0.1 mM DTT, 0.5 mM PMSF, 2mM RVC). After washing with binding buffer, equal molecular amounts of folded RNAs were added to incubate with binding beads for 2h at 4°C in 0.2 mL binding buffer. After incubation and washing, RNAs were extracted with Trizol (Life technologies) and the abundance of RNAs bound to flag-tagged-β-TrCP protein was revealed by denatured PAGE gel.

Dig-labeled RNA Pull-down Assay

Dig-RNA pull-down assays were carried out as described (Hu et al., 2015) with modifications. 1 × 10⁷ 293FT cells or 3T3 cells expressing flag-β-TrcP or flag-β-TrcP truncations were used to immunoprecipitate with anti-Flag (Sigma) antibodies. After immunoprecipitation and wash, one-fourth beads were saved for WB. The rest was equilibrated in binding buffer [50 mM Tris-HCl at pH 7.4, 150 mM NaCl, 0.05% Igepal, 0.5% NP-40, 0.5 mM PMSF, 2 mM RVC, protease inhibitor cocktail (MedChem Express)] and incubated with 300 ng of Dig-labeled RNA for 4h at 4°C. For all hFAST truncated fragments, the same mole transcripts were used for protein binding assays. After washing 4 × 5 min with binding buffer, the bound RNAs were extracted and analyzed by denatured PAGE gels.

Electrophoretic Mobility Shift Assay

The Dig-labeled RNAs were annealed by heating at 65°C for 5 min, then slowly cooled down to room temperature. 10ng of Dig-labeled RNAs were used for each EMSA reaction. Protein-RNA binding was carried out with the indicated amount of purified protein and annealed Dig-labeled RNA in binding buffer [100 mM HEPES pH7.5, 200 mM KCl, 10 mM MgCl₂, 10 mM DTT (DL-Dithiothreitol)]. Binding reactions were incubated at room temperature for 25 min, then immediately loaded onto a 1.5% agarose gel. The gel was transferred to nylon membrane, and imaged by incubation with Anti-Digoxigenin-AP Solution.

RNA In Situ Hybridization and smFISH

RNA *In Situ* Hybridization was carried out as previously described with *in vitro* transcribed digoxin labeled antisense probes (Yin et al., 2012). Briefly, H9 and R1 cells were fixed with 3.6% PFA and 10% acetic acid for 10 min, followed by permeabilization with 0.5% Triton X-100 for 5 min. Then cells were subjected to incubation with denatured probes in hybridization buffer (50% formamide in 2xSSC) at 50°C overnight. After hybridization, anti-Dig primary antibodies and fluorescent secondary antibodies were sequentially added to visualize signals with DeltaVision Elite imaging system (GE Healthcare). The nuclei were counterstained with DAPI.

To detect hFAST and mFast, smFISH was carried out as previously described (Wang et al., 2018). Probes were designed by Stellaris Probe Designer and labeled with Cyanine 3-dUTP (Enzo Life) at the 3' ends of hFAST and mFast. H9 and R1 cells were fixed with

4% PFA for 10 min, followed by permeabilization with 0.5% Triton X-100 for 5 min. Then cells were subjected to incubation with smFISH probes at 37°C overnight. The nuclei were counterstained with DAPI. Images were taken with DeltaVision Elite imaging system (GE Healthcare).

Knockout mFast and hFAST by Multiplex CRISPR/Cas9 Assembly System

All-in-one CRISPR/Cas9 vector (Addgene) expressing 4 gRNAs and a Cas9 nuclease was used to knock out mFast or hFAST (Wu et al., 2013). Two pairs of sgRNAs that specifically target the genomic loci of mFast or hFAST were constructed followed by Multiplex CRISPR/Cas9 Assembly System Kit protocol. Transfection plasmids in R1 cells were carried out with Lipofectamine 2000 Reagent (Thermo) according to the manufacturer's instructions. Genomic DNA and total RNA of selected single clones were extracted either for the genotyping validation with appropriate sets of primer listed in Table S7 or for knockout efficiency validation by qRT-PCR.

Measurement of hFAST Copy Number

A serial dilution of the linearized plasmid pcDNA3-hFAST was used for RT-qPCR to generate a standard curve to calculate hFAST absolute abundance. Copy number of the diluted plasmid pcDNA3-hFAST was calculated by the DNA/RNA Copy Number Calculator from website (<http://endmemo.com/bio/dnacopynum.php>). To measure the hFAST copy number in H9 cells, total RNA extracted from 1×10^6 cells was reverse transcribed into cDNAs for qPCR analysis, and the copy number of hFAST was quantitated from the standard curve.

Colony Formation Assay

2.5×10^3 scramble or EV and hFAST KD or KO H9 or CT1 cells were calculated by Countess II FL Automated Cell Counter (Therm Fisher) and seeded on irradiated-MEF feeder cells with 10 μM Y-27632 (Tocris) at the first day, then cultured for 7 days by changing the culture medium daily. Survival colonies were stained by 0.1% Crystal Violet at room temperature for 2 minutes followed washing by DPBS twice. Pictures were taken and the average number of colonies per well were measured. The colony formation ability was presented by the ratio of survival colonies to the total seeded cells at the first day.

Alkaline Phosphatase Staining

Alkaline phosphatase staining was carried out using Alkaline Phosphatase Detection Kit (Millipore) according to the manufacturer's protocol. H9 colonies were seeded on the irradiated-MEF feeder cells or on Matrigel coated dishes, and then cultured for 5 days at 37°C. H9 cells were fixed with 4% PFA for 1 minute after wash twice with DPBS, and were then incubated with 1 x TBST [20mM Tris-HCl (pH7.4), 0.15M NaCl, 0.05% Tween-20, 0.5% Triten X 100] for 5 minutes after wash off 4% PFA. The cells were incubated with Naphthol/Fast Red solution mix in dark for 15 minutes, followed by wash off the solution and add 1mL DPBS for pictures taken and further analysis.

In-Cell SHAPE Probing

In-cell SHAPE probing was performed in H9 cells as previously described (Liu et al., 2019) with modifications. Briefly, H9 cells were cultured on 10 cm dishes for two days to reach 1×10^6 cells. After washed with DPBS, H9 cells were incubated with 900 μL of cell culture medium and 100 μL of 10x SHAPE Chemical in DMSO with the final concentration of NAI (EMD Millipore) at 100 mM for 10 minutes at 37°C. After removed probing medium, RNAs were isolated by 1 mL Trizol reagent (Invitrogen) according to the manufacturer's protocol. The same procedure was also performed in parallel for the untreated control samples, but with the addition of only DMSO. In the denaturing control (DC) reaction, RNAs were suspended in a denaturing buffer containing 50% formamide and were incubated at 95°C before modification with SHAPE reagents.

SHAPE-MaP Reverse Transcription

Isolated RNAs were treated with DNase I (Ambion, DNA-freeTM kit) to remove possible DNA contamination. About 50-100 ng of RNAs were obtained under each treatment, and were then used for SHAPE-MaP reverse transcription by adding 1 μL (200 U/μL) of SuperScript II (Invitrogen), 6 mM Mn²⁺ and gene-specific primers for hFAST. Mn²⁺ was removed using G-25 micro-spin columns (GE Healthcare) after SHAPE-MaP reverse transcription. Second-strand synthesis was performed with Q5 hot start high-fidelity DNA polymerase and nested PCR was performed to further improve DNA yield. The resulting PCR products were further isolated with PureLink micr spin columns (Life Technologies). Primers for SHAPE-MaP reverse transcription and second-strand synthesis (1st round PCR and nested PCR reactions) were listed in Table S7.

SHAPE-MaP Library Preparation and Sequencing

SHAPE-MaP libraries were prepared from 1 ng of DNAs, and size-selected with AmpureXP beads (Agencourt) with a 1:1 (bead to sample) ratio to obtain library DNA products spanning 100-400 bp in length. Final libraries were quantified using Agilent Bioanalyzer 2100 and QuBit high-sensitivity dsDNA assay. Deep sequencing was performed by Illumina NextSeq 500 at CAS-MPG Partner Institute for Computational Biology Omics Core, Shanghai, China. About 15-25 million mapped sequencing reads were obtained for each sample, with 88% of bases at or above Q30.

Polyadenylated RNA Separation, rRNA Depletion for RNA-seq

For RNA-seq samples from H9 and R1/E cells, total/cytoplasmic/nuclear RNAs were processed with the RiboMinus kit (Human/Mouse Module, Invitrogen) to deplete ribosomal RNAs for sequencing (Ribo- RNA-seq). For hFAST KD total RNA-seq samples, polyadenylated (poly(A)⁺) RNA preparation was carried out as previously described (Yang et al., 2011; Yin et al., 2015). Total RNAs (8 µg) were incubated with oligo(dT) magnetic beads to isolate poly(A)⁺ RNAs, which were bound to beads, and selection was performed three times to ensure the purity of poly(A)⁺ RNA populations.

Library Preparation and Deep Sequencing

Ribo- and poly(A)⁺ RNA-seq libraries were prepared using Illumina TruSeq Stranded Total RNA LT Sample Prep Kit. All libraries were size-selected with AmpureXP beads (Agencourt) and quantified using Agilent Bioanalyzer 2100 and QuBit high-sensitivity dsDNA assay. Size-selected libraries were subjected to deep sequencing with Illumina HiSeq X ten at CAS-MPG Partner Institute for Computational Biology Omics Core, Shanghai, China. Raw read qualities were evaluated by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq Data Processing

RNA-seq reads were first processed by using trimomatic (Bolger et al., 2014) (version 0.35, parameters: ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:80) to remove low quality bases and adaptor sequences, and then mapped by using HISAT2 (Kim et al., 2015) (version 2.0.5, parameters: -no-softclip-score-min L,-16,0-mp 7,7-rfg 0,7-rdg 0,7-dta -k 1-max-seeds 20) to GRCh37/hg19 human reference genome with the GENCODE gene annotation (version 28lift27) or GRCm38/mm10 mouse reference genome with the GENCODE gene annotation (version M17) or Mmu10/rheMac10 with Ensembl gene annotation (version 98). The HISAT2-mapped sam files were transferred to bam format by samtools (Li et al., 2009) (version 0.1.19). Gene expression levels were calculated with FPKM (Trapnell et al., 2010) (Fragments Per Kilobase of transcript per Million mapped reads) by normalizing gene counts from featureCounts (Liao et al., 2014) (version 1.5.1, parameters: -fraction O -T 10 -exon -g gene_id).

Imaging Analysis of lncRNA Nuclear Localization

Raw images were processed with softWoRx 7.0 and ImageJ/Fiji (Schindelin et al., 2012). Deconvolution images were generated by softWoRx 7.0. Nuclear foci statistics were performed using the “JACoP” plugin for ImageJ/Fiji (Bolte and Cordelières, 2006). Briefly, split RGB images into two separate channels containing DAPI (blue) and RNA signals (red) using “Split Channels.” Then, calculate the colocalization of DAPI and RNA signals using “JACoP” plugin with “M1 & M2 coefficients.” M2 is defined as the ratio of the ‘summed intensities of pixels from the red image for which the intensity in the blue channel is above zero’ to the ‘total intensity in the red channel’, refer to the percentage of nuclear foci of RNA signals.

Identification of Conserved lncRNAs

Sequence-conserved lncRNAs were detected by two methods (Figure S1A). On the one hand, H9 expressed lncRNAs were mapped to mouse genome by the UCSC LiftOver tool (parameters: -bedPlus = 12 -tab -minMatch = 0.8 -minBlocks = 0.5 -multiple). Sequence-conserved lncRNA candidates were defined with ≥ 20% exon sequence conservation after the LiftOver alignment. On the other hand, sequences of paired lncRNAs expressed in H9 and R1 were compared by Blast (Camacho et al., 2009) (parameters: -task blastn -word_size 6 -strand plus -outfmt 6). The sequence-conserved lncRNA candidates were defined with ≥ 20% sequence similarity between paired lncRNAs. Only overlapped lncRNAs identified in both methods were defined as sequence-conserved lncRNAs.

Positionally conserved lncRNAs between human and mouse were identified in the same subgroups (divergent, convergent or the same strand lncRNAs) with the distance between the lncRNA and its neighbor mRNA no longer than 5 kb, together with the same conserved adjacent protein coding genes and the conserved promoters. Promoters were defined as upstream and downstream 500 bps from transcription start site (TSS). TSSs were fetched from either gene annotations or CAGE datasets (Fort et al., 2014). Conserved protein coding genes were fetched from NCBI HomoloGene database (version build68; <https://www.ncbi.nlm.nih.gov/homologene>).

Conservation Analysis of FAST

3,200 bps sequences upstream of human, rhesus monkey and mouse FOXD3 genes were extracted. Then EMBOSS Matcher (Madeira et al., 2019) was used to pairwisely compare these sequences. According to the comparison results from EMBOSS, conserved regions were extracted manually. Finally, conservation of FAST were evaluated in three aspects (Ulitsky, 2016) including: 1) conserved exonic structure, 2) conserved sequence, and 3) positionally conserved transcription.

Analysis of RNA Localization

We applied a “Cytoplasmic Ratio” to indicate the subcellular localization pattern of each expressed RNA in a given cell type from related RNA-seq dataset. The RNA concentrations in the cytoplasmic and the nuclear fractions of a given cell type were experimentally measured after cell fractionation. A constant cytoplasmic factor was calculated by $\text{concentration}_{\text{cyto}} / (\text{concentration}_{\text{cyto}} + \text{concentration}_{\text{nuc}})$, and a constant nuclear factor was calculated by $\text{concentration}_{\text{nuc}} / (\text{concentration}_{\text{cyto}} + \text{concentration}_{\text{nuc}})$. Expression of cytoplasmic RNAs was normalized by multiplying $\text{FPKM}_{\text{cyto}}$ by the constant cytoplasmic factor in a given cell type, and expression of nuclear RNAs was normalized by multiplying FPKM_{nuc} by the constant nuclear factor. Finally, the cytoplasmic ratio of an expressed RNA in a given cell type was defined as dividing normalized $\text{FPKM}_{\text{cyto}}$ by the sum of normalized $\text{FPKM}_{\text{cyto}}$ and normalized FPKM_{nuc} .

Stem Cell Maintenance Related Gene Analysis

All stem cell maintenance-related genes were fetched from AmiGO ([Carbon et al., 2009](#)). Expression levels of these genes in the scramble or hFAST knockdown H9 cells and additional hiPS and ES cells ([Choi et al., 2015](#)) were shown as heatmap by pheatmap package in R (version 3.4.1) from related RNA-seq samples.

Stem-related Signaling Pathway Analysis

Genes in six stem-related signaling pathways (WNT, PI3K-Akt, MAPK, TGF-beta, JAK-STAT and Notch signaling pathway) were collected from KEGG database ([Kanehisa and Goto, 2000](#)). Fold changes (FC) between the scramble and the hFAST knockdown RNA-seq samples were used to define upregulated ($\text{FC} \geq 1.5$), unchanged ($0.667 < \text{FC} < 1.5$) or downregulated ($\text{FC} \leq 0.667$) genes. Dys-regulated, either upregulated or downregulated, genes were identified after hFAST knockdown. Bubble plot, drawn by R (version 3.4.1), indicates dys-regulated genes in each individual stem-related signaling pathway after hFAST knockdown from related RNA-seq datasets.

Calculation of the SHAPE Reactivity

SHAPE reactivity profiles were created from deep sequencing datasets for hFAST by using ShapeMapper (version 1.2) with default parameters ([Smola et al., 2015](#)). Briefly, low quality score sequence reads were trimmed. Paired-end reads without overlapping sequences were analyzed as two single reads, while paired-end reads with overlapped sequences were merged into single reads. Pre-processed reads were then mapped to target sequences by Bowtie2 ([Langmead and Salzberg, 2012](#)) (version 2.1.0, parameters: -p 24 -L 15 -D 20 -R 3-local-sensitive-local-ma 2-mp 6,2-rdg 5,1-rfg 5,1-dpad 100-maxins 800-ignore-quals-no-unal). Mutation rate (MutR) of each position was calculated by dividing the total number of mismatches and indels (deletions and insertions) by the number of all mapped reads to the position. SHAPE reactivities of all positions were calculated by $[(\text{ModifiedMutR} - \text{UntreatedMutR})/\text{DenaturedMutR}]$ and normalized by a model-free box-plot approach ([Deigan et al., 2009](#)). More than 80% positions with at least 1000-read depth in three (modified, untreated and denatured) samples ([Mustoe et al., 2018](#)) were used for the hFAST secondary structure modeling.

hFAST Secondary Structure Modeling

With SHAPE reactivity values determined by ShapeMapper, hFAST secondary structure was modeled by Superfold (version 1.0).

Prediction of hFAST ORFs

The ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) webtool was used to predict possible ORFs of hFAST with default parameters.

Analysis of FAST Associated trans-factors

The sequences of hFAST and mFast were uploaded onto ATtRACT (A daTabase of RNA binding proteins and AssoCiated moTifs) database ([Giudice et al., 2016](#)). Only RBPs with bound motifs longer than 5 nt were considered as candidate trans-factors that may impact FAST processing. Predicted RBPs that can bind both hFAST and mFast were sorted according to their predicted binding sites and were used in the subsequent analyses.

Conservation Analysis of PPIE

Conservation between human PPIE protein and mouse Ppie protein were analyzed by UniProt Align web tool (<https://www.uniprot.org/align/>) with default parameters, which is based on Clustal Omega.

RNA Localization Analysis upon Ppie Knockdown

After RNA-seq data processing, we get gene expression level (FPKM) from six ribo-sample: scramble cytoplasmic ($\text{FPKM}_{\text{scr_cyto}}$), scramble nuclear ($\text{FPKM}_{\text{scr_nuc}}$), scramble total cell ($\text{FPKM}_{\text{scr_total}}$), Ppie KD cytoplasmic ($\text{FPKM}_{\text{kd_cyto}}$), Ppie KD nuclear ($\text{FPKM}_{\text{kd_nuc}}$) and Ppie KD total cell ($\text{FPKM}_{\text{kd_total}}$) RNA-seq datasets. mRNAs and lncRNAs with $\text{FPKM}_{\text{scr_total}} \geq 0.5$ or $\text{FPKM}_{\text{kd_total}} \geq 0.5$ are used in the downstream analysis. Fold changes (FC; $\text{FPKM}_{\text{kd_total}}/\text{FPKM}_{\text{scr_total}}$) between the scramble and the Ppie knockdown RNA-seq samples were used to define upregulated ($\text{FC} \geq 2$), unchanged ($0.5 < \text{FC} < 2$) or downregulated ($\text{FC} \leq 0.5$) genes. Unchanged genes

were selected for further different subcellular localization analysis. Next, expression levels of *Gapdh* in each sample were used to normalize expression levels in each sample, for example: before normalization, expression level of a gene x in scramble cytoplasmic sample is $\text{FPKM}_{\text{scr_cyto_x}}$, after normalization, expression level of the gene x is $\text{FPKM}'_{\text{scr_cyto_x}} = \text{FPKM}_{\text{scr_cyto_x}} / \text{FPKM}_{\text{scr_cyto_Gapdh}}$. After normalization to *Gapdh*, N/C is calculated as $\text{FPKM}'_{\text{scr_nuc}} / \text{FPKM}'_{\text{scr_cyto}}$ to represent subcellular localization in the cells. Finally, localization changes (LC; $(N/C_{\text{kd}}) / (N/C_{\text{scr}})$) between scramble and Ppie KD samples were defined export after KD ($LC \leq 0.5$), unchanged ($0.5 < LC < 2$) or nuclear retention ($LC \geq 2$) genes.

Statistical Analysis

Statistical significance for comparisons was generally assessed by Student's t test with exceptions described below. Statistical significance for stem-related signaling pathway analysis (Figure 3A) was assessed using Fisher's Exact test with R (version 3.4.1). *P* values below 0.05 were marked by one asterisk, while two asterisks indicate *P* value < 0.01 and three asterisks indicate *P* value < 0.001.

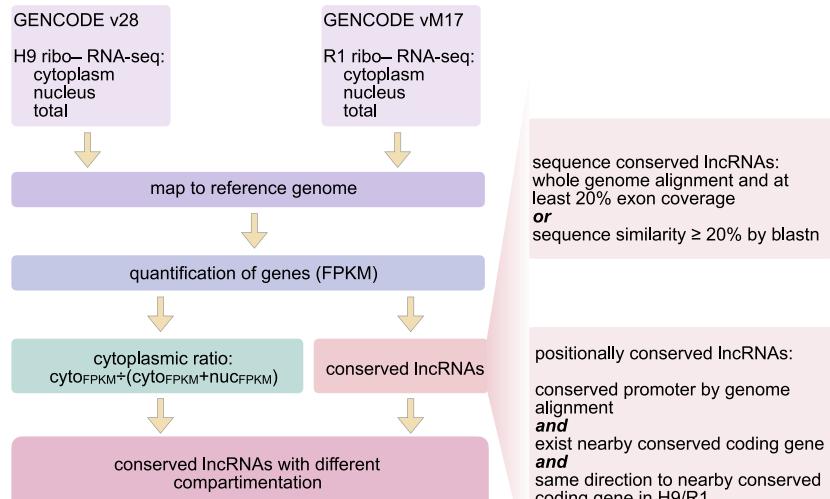
DATA AND CODE AVAILABILITY

The accession number for ribo- RNA-seq samples from H9 and R1 cells, total/cytoplasmic/nuclear RNA and polyadenylated [poly(A)⁺] RNA-seq samples from hFAST KD/scramble H9 cells is GEO: GSE143496 and NODE: OEP000734. All original unprocessed data related to this paper were uploaded to Mendeley Data in the website: <https://dx.doi.org/10.17632/f4nkpt5h82.1>.

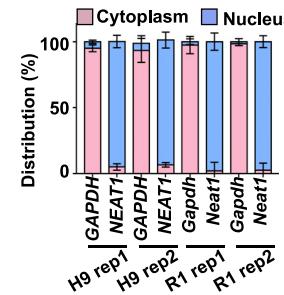
Supplemental Figures

Cell

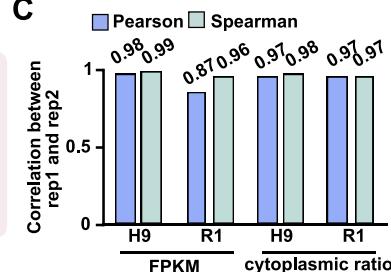
A



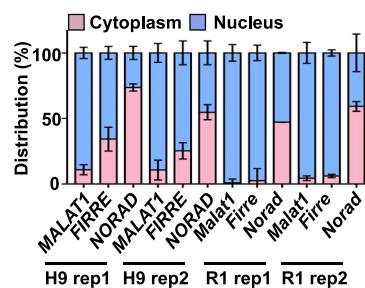
B



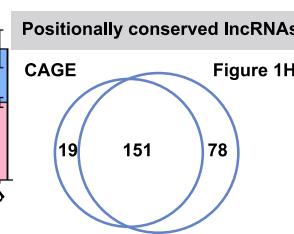
C



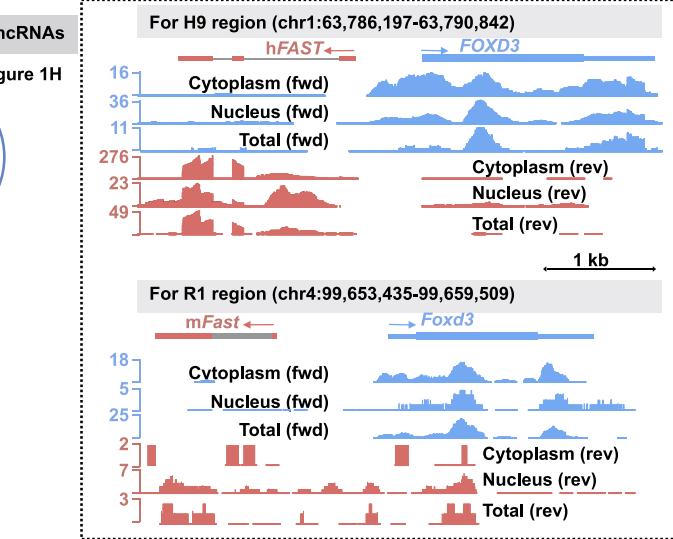
D



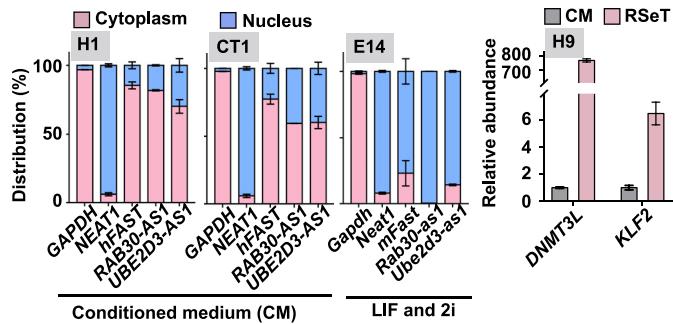
F



G



H



(legend on next page)

Figure S1. Analysis of Subcellular Localization Patterns of Conserved lncRNAs, Related to Figure 1

- (A) A workflow to select conserved lncRNAs with different compartmentalization.
- (B) Validation of nuclear and cytoplasmic fractionation efficiency in H9 and R1 cells. Successful nuclear and cytoplasmic fractionation in H9 and R1 cells was confirmed by examining RNAs with well-established subcellular localization.
- (C) Pearson and Spearman correlation of RNA expression and RNA cytoplasmic ratio, defined as $\text{Cyto}_{\text{FPKM}} / (\text{Cyto}_{\text{FPKM}} + \text{Nuc}_{\text{FPKM}})$, between two replicates of RNA-seq datasets from H9 and R1 cells were calculated.
- (D) Validation of the subcellular localization of known sequence conserved lncRNAs from H9 and R1 RNA-seq samples.
- (E) The subcellular distribution of sequence conserved lncRNAs in H9 and R1 cells with different sequence similarity thresholds (15%, 30%, 50% and 80%).
- (F) Overlap of positionally conserved lncRNAs identified in Figure 1H and those with conserved promoters revealed by CAGE datasets (Fort et al., 2014)
- (G) Wiggle tracks of hFAST and FOXD3 as well as mFast and Foxd3 in H9 and R1 cells. Blue wiggle tracks indicate the protein coding gene FOXD3 and Foxd3 whereas pink wiggle tracks indicate the lncRNA gene hFAST and mFast.
- (H) Conserved lncRNAs in hESC H1 and CT1 cells are mainly cytoplasmic, whereas they are mainly nuclear in mESC E14, revealed by RT-qPCR.
- (I) Primed H9 cells were reverted to a naive-like state by cultured in RSeT medium, shown by the enhanced expression of naive-associate genes, DNMT3L and KLF2. Data in (B), (D), (H), and (I) are presented as mean \pm SD. Error bars represent SD in triplicate experiments.

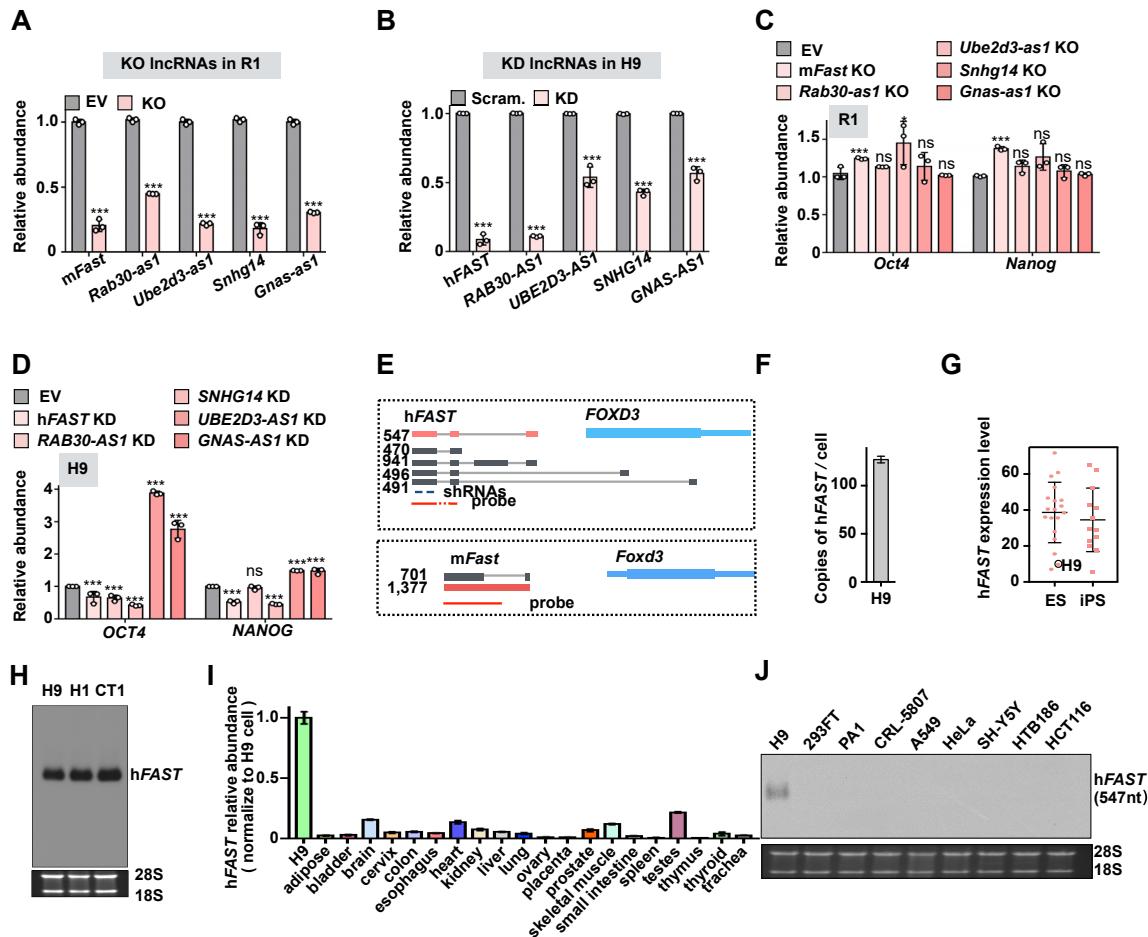


Figure S2. Functional Analysis of Conserved lncRNAs in H9 and R1 Cells, Related to Figure 2

- (A) KO of distinctly localized lncRNAs by CRISPR/Cas9 in R1 cells, revealed by RT-qPCR.
 - (B) KD of distinctly localized lncRNAs by shRNAs in H9 cells, revealed by RT-qPCR.
 - (C) KO of distinctly localized lncRNAs has little effects on pluripotency gene expression in R1 cells, revealed by RT-qPCR.
 - (D) KD of distinctly localized lncRNAs in H9 cells alters pluripotency gene expression, revealed by RT-qPCR.
 - (E) Schematic of hFAST and mFast and their corresponding divergently transcribed protein coding genes. Red lines indicate positions of NB probes, blue lines indicate positions of two shRNAs used in the study.
 - (F) The copy number of hFAST per H9 cell measured by RT-qPCR.
 - (G) hFAST is highly expressed in all examined human ES and iPS cells by RNA-seq data (Choi et al., 2015).
 - (H) Validation of hFAST expression in different hESC lines (H9, H1 and CT1), revealed by 1% agarose NB.
 - (I) Expression of hFAST in different human tissues is low as revealed by RT-qPCR after normalization to that in H9 cells.
 - (J) Validation of hESC-specific hFAST expression in other cell lines by 1% agarose NB.
- Data in (A)-(D), (F), (G), and (I) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All P values were calculated using two-tailed unpaired Student's t test; *p < 0.05; **p < 0.001; ns means no significant difference.

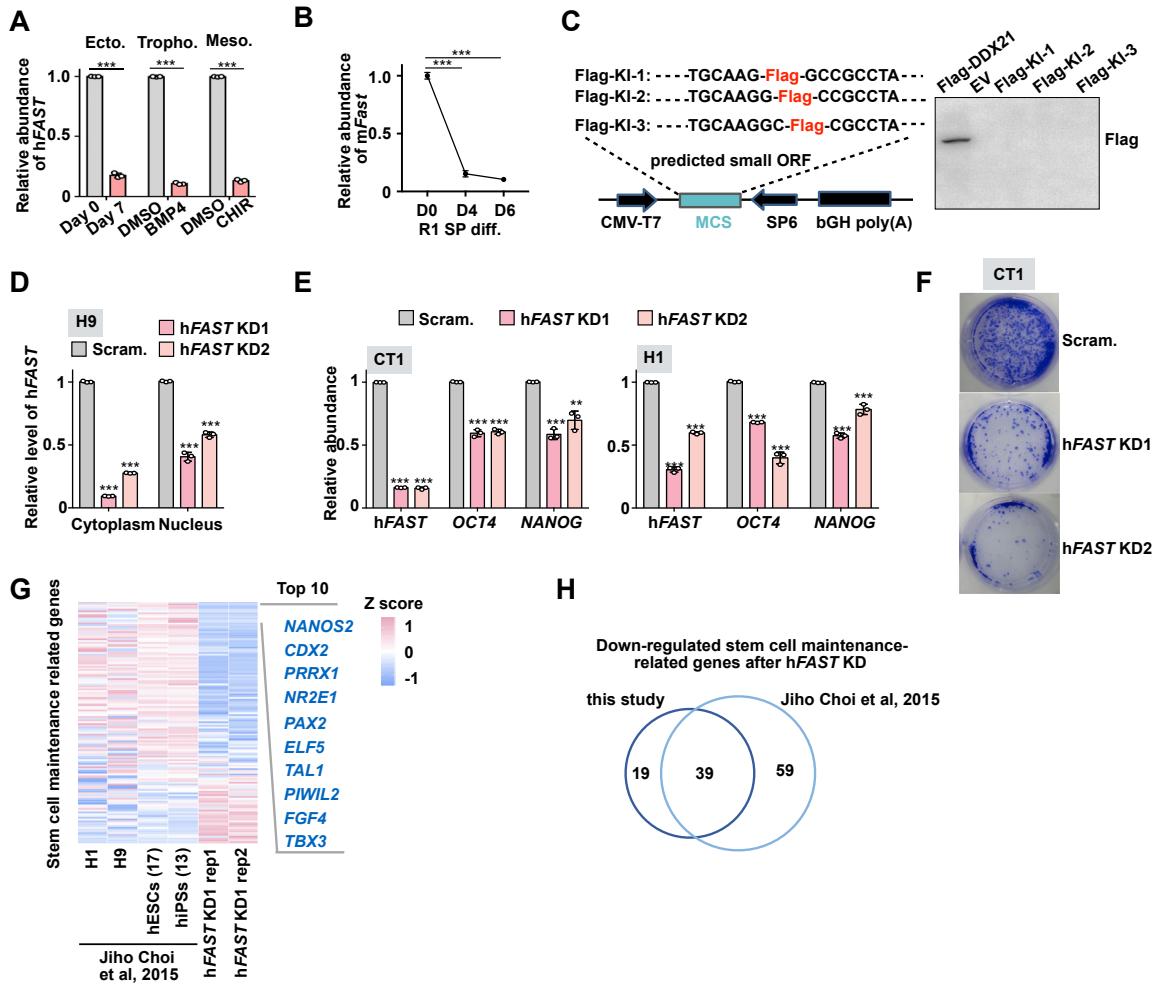
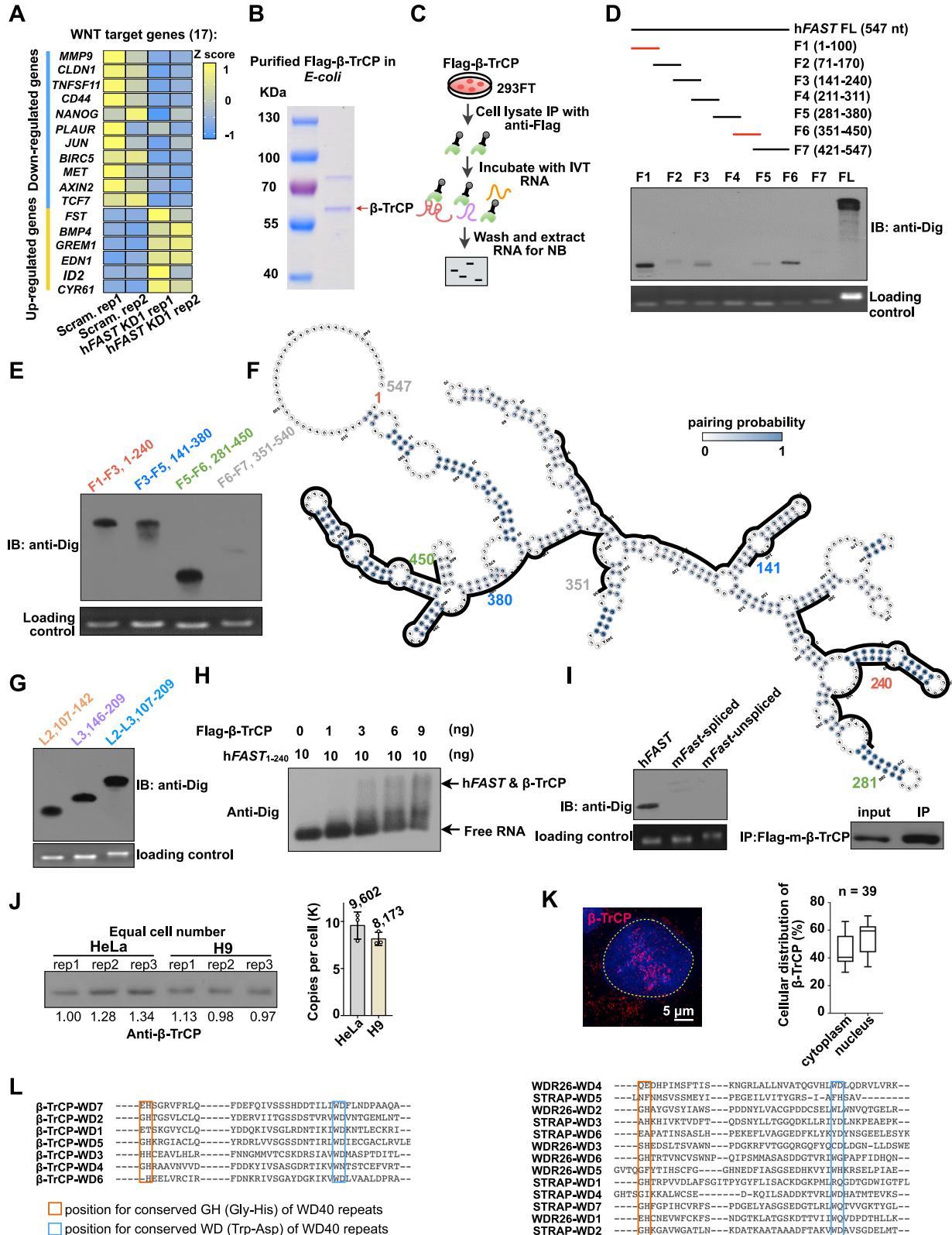


Figure S3. hFAST Is Required for Pluripotency of hESCs, Related to Figure 2

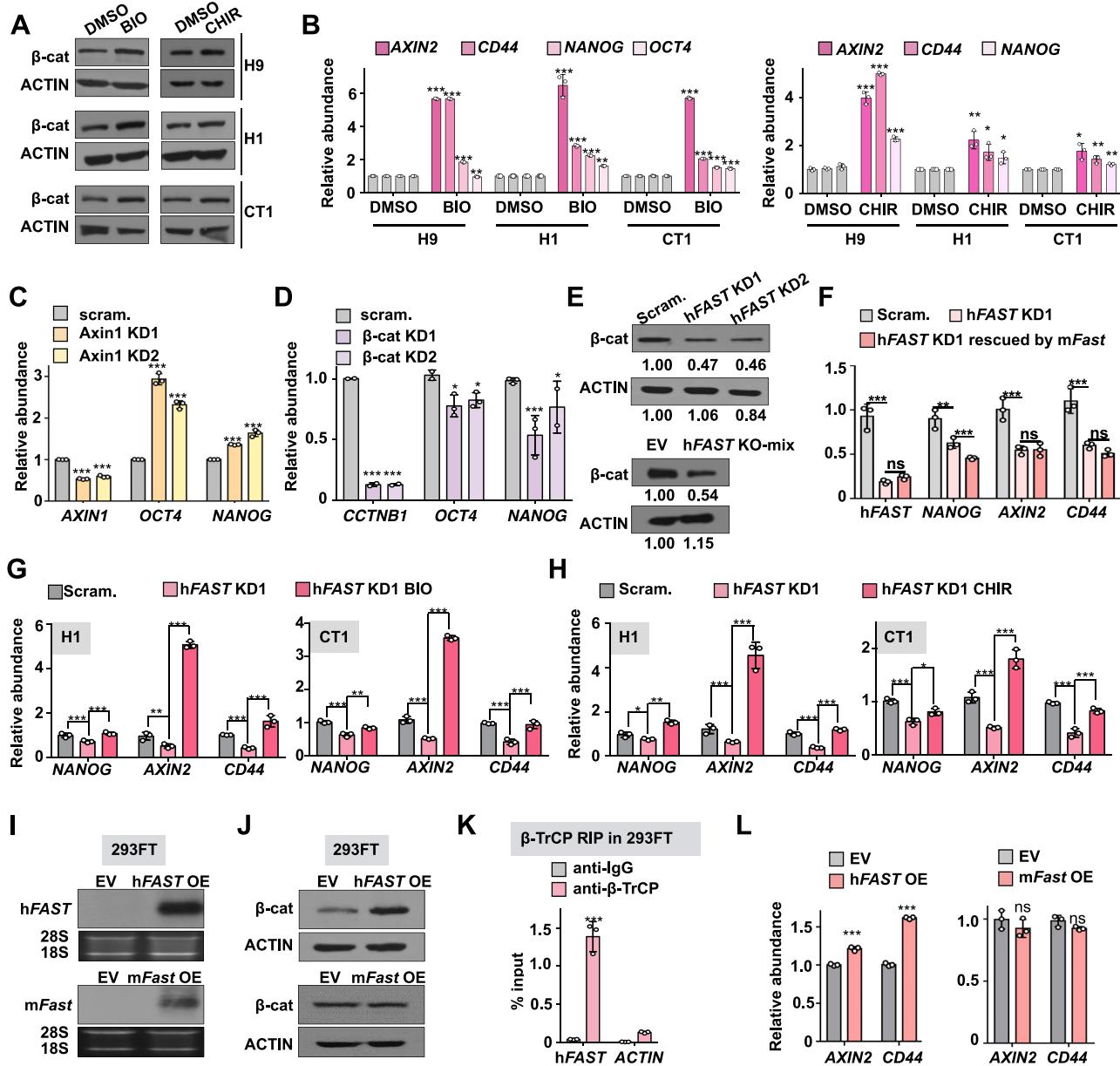
- (A) Expression of hFAST is rapidly decreased during ectoderm, trophoblast and mesoderm differentiation, revealed by RT-qPCR.
- (B) Expression of mFast is dramatically reduced upon R1 spontaneous differentiation, revealed by RT-qPCR.
- (C) hFAST does not encode a peptide. Left: an experimental strategy to examine hFAST coding capability by Flag-tag knocking-in the three continuous nucleotides downstream of the predicted small ORFs, named Flag-KI-1, Flag-KI-2, Flag-KI-3. Right: WB with anti-Flag antibodies showed that hFAST does not encode a peptide in H9 cells transfected with the individual plasmids.
- (D) KD of hFAST by shRNAs led to reduced hFAST expression, with that localized to the cytoplasm more strikingly depleted in H9 cells. The relative level of hFAST was measured by RT-qPCR.
- (E) KD of hFAST led to reduced expression of OCT4 and NANOG in hESC lines CT1 (left), and H1 (right), revealed by RT-qPCR.
- (F) hFAST KD impaired the colony formation ability of CT1 cells. Cell clones were stained by 0.1% crystal violet.
- (G) hFAST KD resulted in altered expression of genes required for stem cell maintenance between hFAST KD H9 cells and other published pluripotent stem cells (Choi et al., 2015), shown by heatmap of expression levels of such genes from RNA-seq datasets.
- (H) Comparison of the altered genes between H9 and hFAST KD cells in this study, as well as between published hESCs and hiPSCs (Choi et al., 2015) and hFAST KD H9 cells, showed 67% overlapped downregulated genes in these two sets of comparison.
- In (A), (B), (D), and (E), error bars represent SD in triplicate experiments. All *P* values were calculated using two-tailed unpaired Student's t test; ***P* < 0.01; ****p* < 0.001.



(legend on next page)

Figure S4. WNT Signaling Is Required for Pluripotency of Cultured hESCs, Related to Figure 3

- (A) Altered expression of WNT target genes in hFAST KD H9 cells, shown by RNA-seq.
- (B) Purification of His-Flag- β -TrCP (red arrow) in *E.coli*, revealed by SDS-PAGE and Coomassie Blue staining.
- (C) Diagram of the experimental pipeline for binding assay between Flag- β -TrCP and Dig-labeled hFAST fragments shown in Figures S4D, S4E, S4G, S4I, Figures 3G, 3I, 3J, and 3L.
- (D) Identification of arbitrarily fragmented hFAST that interacts with Flag- β -TrCP by assay shown in Figure S4C.
- (E) Identification of hFAST fragments that interact with Flag- β -TrCP by assay in Figure S4C.
- (F) Enlarged view of hFAST secondary structure by in-cell SHAPE-MaP. Fragments in Figure S4E are shown the start and end sites. Black lines indicate conserved regions between human and mouse. The color in circles represents the pairing probability calculated by Superfold.
- (G) Identification of hFAST L3 fragments that interact with Flag- β -TrCP by assay in Figure S4C.
- (H) Dig-labeled hFAST₁₋₂₄₀ binds to purified His-Flag- β -TrCP at different concentrations, revealed by EMSA.
- (I) mFast does not interact with β -TrCP, revealed by assay shown in Figure S4C (Left). Right: IP efficiency of Flag- β -TrCP.
- (J) The relative abundance of β -TrCP in H9 cells, compared to that in HeLa cells by WB (Left). Right: statistics of PPIE copies per cell. β -TrCP copy number per HeLa cell was extracted from quantitative proteomics (Hein et al., 2015). Error bars represent SD in triplicate experiments.
- (K) Subcellular localization of β -TrCP in H9 cells, revealed by IF (Left). Right: statistics of β -TrCP signals distribution in each cell, n = 39 H9 cells.
- (L) Sequences of WD40 repeats of β -TrCP, WDR26 and STRAP. Red box shows conserved GH (Gly-His) of WD40 repeats; blue box shows conserved WD (Trp-Asp) of WD40 repeats.

**Figure S5. hFAST Regulates the WNT Signaling Pathway, Related to Figure 4**

- (A) Activation of the WNT pathway in hESCs by different chemicals. Activation of WNT pathway by 2 μ M BIO or 4 μ M CHIR99021 led to increased β -cat in H9, H1 and CT1 cells, revealed by WB.
- (B) Activation of the WNT pathway by BIO or CHIR99021 led to enhanced expression of pluripotent genes (OCT4 and NANOG) and WNT target genes (AXIN2 and CD44) in examined hESCs (CT1, H1 and H9), revealed by RT-qPCR. Each sample was treated with 2 μ M BIO or 4 μ M CHIR99021 for 3 days in MEF-free condition; control cells were treated with equal amount of DMSO.
- (C) Activation of the WNT pathway by Axin1 KD led to increased expression of pluripotent genes (OCT4 and NANOG) in H9 cells, revealed by RT-qPCR.
- (D) β -cat KD led to reduced NANOG and OCT4 expression in H9 cells, revealed by RT-qPCR.
- (E) hFAST KD or hFAST KO in H9 cells led to reduced β -cat expression, as shown by WB.
- (F) mFast OE in hFAST KD H9 cells did not rescue expression of WNT target genes, revealed by RT-qPCR.
- (G)(H) Activation of the WNT pathway by 2 μ M BIO (G) or 4 μ M CHIR99021 (H) in hFAST KD H1 and CT1 cells rescued expression of WNT target genes, revealed by RT-qPCR.
- (I) hFAST (upper panels) and mFast (bottom panels) OE in 293FT cells, shown by NB.
- (J) hFAST OE (upper panels), but not mFast OE (bottom panels), resulted in increased β -cat expression, revealed by WB.

(legend continued on next page)

(K) β -TrCP interacted with hFAST in 293FT hFAST OE cells. Native RIP of β -TrCP in 293FT cells using anti- β -TrCP or anti-IgG antibodies followed by RT-qPCR. ACTIN mRNA was used as a control.

(L) OE of hFAST but not mFast promoted WNT target genes expression in 293FT cells, revealed by RT-qPCR.

Data in (B)-(D), (F)-(H) and (K)-(L) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All P values were calculated using two-tailed unpaired Student's t test; *p < 0.05; **p < 0.01; ***p < 0.001; ns means no significant difference.

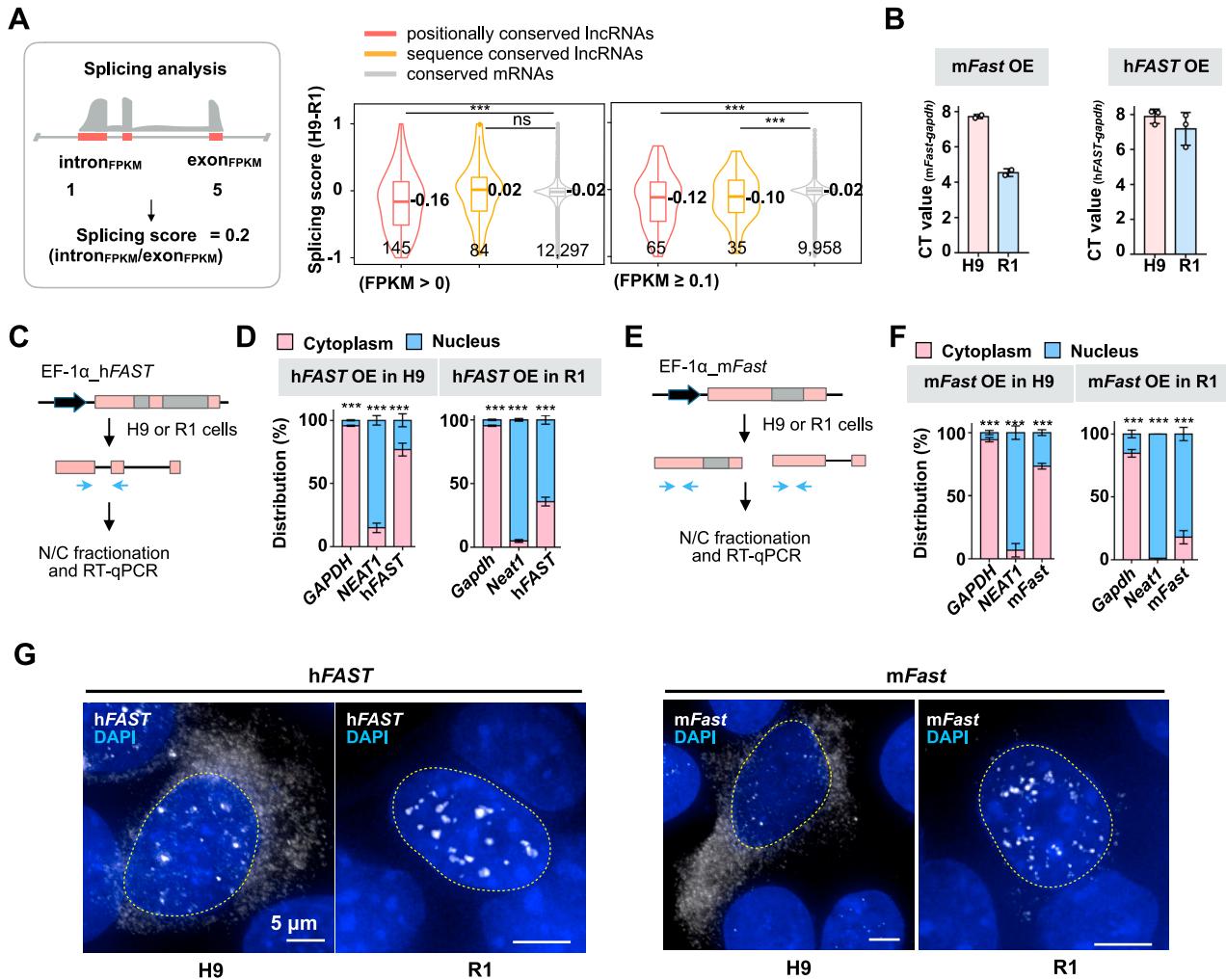


Figure S6. Distinct Processing and Localization of hFAST and mFast in H9 and R1 Cells Is Regulated by trans-Factors, Related to Figure 5

(A) Splicing status of positionally conserved lncRNAs, sequence-conserved lncRNAs and conserved mRNAs. Left: splicing score is defined as $\text{intron}_{\text{FPKM}} / \text{exon}_{\text{FPKM}}$ to indicate the splicing status of a given RNA. Note that a higher value indicates a higher degree of unspliced RNA. Middle: splicing status of positionally conserved lncRNAs is less conserved than that of sequence conserved lncRNAs and mRNAs when all expressed lncRNAs are considered. Right: splicing status of both sequence and positionally conserved lncRNAs with $\text{FPKM} \geq 0.1$ are statistically less conserved than that of conserved mRNAs. All P values were calculated using two-tailed unpaired Student's t test. *** $P < 0.001$. ns means no significant difference.

(B) OE of hFAST and mFast in both H9 and R1 cells achieved comparable levels, normalized by GAPDH mRNA, revealed by RT-qPCR.

(C) Schematic of hFAST OE in H9 and R1 cells following nuclear and cytoplasmic fractionation. Blue arrows indicate primers to detect processed hFAST in Figure S6F.

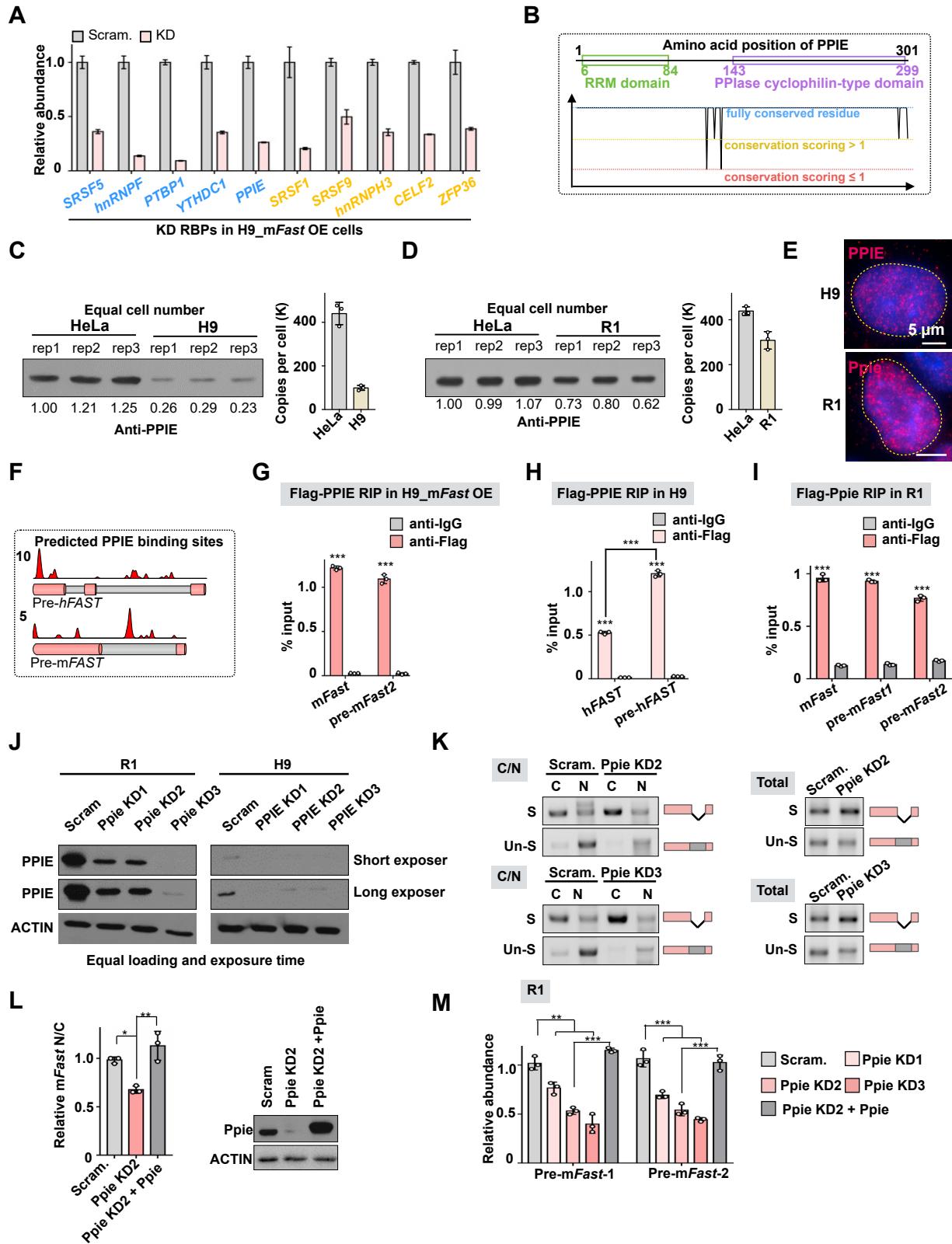
(D) hFAST OE by lentivirus resulted in distinct subcellular localization patterns in H9 and R1 cells. hFAST OE in H9 cells was localized to cytoplasm that recapitulated endogenous hFAST. In contrast, hFAST OE in R1 was localized to nucleus, revealed by RT-qPCR.

(E) Schematic of mFast OE in H9 and R1 cells following nuclear and cytoplasmic fractionation. Blue arrows indicate primers to detect processed and unprocessed mFast in Figure S6F.

(F) mFast OE by lentivirus resulted in distinct subcellular localization patterns in H9 and R1 cells. mFast OE in R1 cells led to nuclear localization that recapitulated endogenous mFast. In contrast, mFast OE in H9 cells resulted in cytoplasmic localization.

(G) Transient OE of hFAST or mFast in H9 and R1 cells, respectively also displayed similar patterns as those observed by stable expression shown in Figures 5A, 5B, and S6C-S6F.

Data in (B), (D), and (F) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All P values were calculated using two-tailed unpaired Student's t test; *** $p < 0.001$.



(legend on next page)

Figure S7. Ppie Is a Suppressor of mFast Processing, Related to Figures 5 and 6

- (A) KD efficiency of individual RBPs by shRNAs in H9 mFast OE cells, revealed by RT-qPCR. Each of RBP was knocked down by two shRNAs virus mixture.
- (B) PPIE is highly conserved between human and mouse. PPIE contains RRM and PPIase cyclophilin-type domains.
- (C)(D) Quantification of PPIE in H9 (C) and R1 (D) cells by WB (left panels). Statistics of PPIE copies per H9 (C) and R1 (D) cell, compared to that in HeLa cell by WB (right panels). PPIE copy number per HeLa cell was extracted from quantitative proteomics (Hein et al., 2015).
- (E) Nuclear localization of PPIE in H9 and R1 revealed by IF.
- (F) Predicted PPIE binding sites on pre-hFAST and pre-mFast.
- (G) and (H) PPIE interacts with mFast (G) and hFAST (H) in H9 mFast OE cells. Native RIP of Flag-PPIE in H9 mFast OE cells using anti-Flag or anti-IgG antibodies followed by RT-qPCR. In (H), PPIE bound more pre-hFAST than hFAST.
- (I) Ppie interacts with mFast in R1 cells. Native RIP of Flag-Ppie in R1 cells using anti-Flag or anti-IgG antibodies followed by RT-qPCR.
- (J) KO efficiency of PPIE in H9 and R1 cells with three shRNAs revealed by WB. Of note, Ppie expression in R1 cells is much higher than that in H9 cells; KD Ppie resulted in similar level of Ppie in R1 cells to the natural status of PPIE in H9 cells.
- (K) Ppie KD by different shRNAs led to the same promotion of mFast processing and export. See Figure 5J for details.
- (L) The enhanced mFast export was rescued by re-introducing Ppie to Ppie KD R1 cells, shown by RT-qPCR (left). WB confirmed Ppie KD and OE in R1 cells (right).
- (M) Ppie KD by different shRNAs all led reduced pre-mFast abundance, which could be rescued by re-introducing Ppie to R1 Ppie KD cells, revealed by RT-qPCR. Data in (A), (C), (D), (G)–(I), (L), and (M) are presented as mean \pm SD. Error bars represent SD in triplicate experiments. All *P* values were calculated using two-tailed unpaired Student's t test; **p* < 0.05; ***p* < 0.01; ****p* < 0.001.