Stem Cell Research & Therapy

RESEARCH

Open Access

OCT4 translationally promotes AKT signaling as an RNA-binding protein in stressed pluripotent stem cells

Wenjie Chen[1,2†], Xinyu Chen[1†], Cheng Chen[1,3†], Shiqi She[1,4], Xia Li[2], Lina Shan[5], Xiaobing Zhang[1], Songsong Dan[1], Yisha Wang[1], Yan-Wen Zhou[6], Qingyi Cao[1], Wenxin Wang[7], Jianwen Hu[8], Yaxun Wei[9], Yaqiang Xue[9], Yi Zhang[9], Songying Zhang[2*], Ying-Jie Wang[1,10*] and Bo Kang[1*]

Abstract

Background Despite numerous studies addressing the molecular mechanisms by which pluripotent stem cells (PSCs) maintain self-renewal and pluripotency under normal culture conditions, the fundamental question of how PSCs manage to survive stressful conditions remains largely unresolved. Post-transcriptional/translational regulation emerges to be vital for PSCs, but how PSCs coordinate and balance their survival and differentiation at translational level under extrinsic and intrinsic stress conditions is unclear.

Methods The high-throughput sequencing of cross-linking immunoprecipitation cDNA library (HITS-CLIP) was employed to decipher the genome-wide OCT4-RNA interactome in human PSCs, a combined RNC-seq/RNA-seq analysis to assess the role of OCT4 in translational regulation of hypoxic PSCs, and an OCT4-protein interactome to search for OCT4 binding partners that regulate cap-independent translation initiation. By taking the Heterozygous Knocking In N-terminal Tags (HKINT) approach that specifically disrupts the 5′-UTR secondary structure and tagging its protein product of the mRNA from one allele while leaving that from the other allele intact, we examined the effect of disrupting the OCT4/5′-UTR interaction on translation of AKT1 mRNA.

Results We revealed OCT4 as a *bona fide* RNA-binding protein (RBP) in human PSCs that bound to the 5′-UTR, 3′-UTR and CDS regions of mRNAs. Multiple known proteins participating in IRES-mediated translation initiation were detected in the OCT4-protein interactome, and a combined RNC-seq/RNA-seq analysis further confirmed a crucial role of OCT4 in translational regulation of PSCs in response to hypoxic stress. Remarkably, OCT4 bound to the GC-rich elements in the 5′-UTR of *AKT1* and multiple PI3K/AKT-pathway-gene mRNAs, and promoted their translation initiation via IRES-mediated pathways under stress conditions. Specifically disrupting the AKT1 mRNA 5′-UTR structure and the OCT4/5′-UTR interaction by the HKINT approach significantly reduced the translation level of AKT1 that led to a higher susceptibility of PSCs to oxidative stress-induced apoptotic death and prioritized differentiation toward ectoderm and endoderm.

[†]Wenjie Chen, Xinyu Chen and Cheng Chen contributed equally to this work.

*Correspondence:
Songying Zhang
zhangsongying@zju.edu.cn
Ying-Jie Wang
yingjiewang@zju.edu.cn
Bo Kang
kangbo@zju.edu.cn
Full list of author information is available at the end of the article

Conclusions Our results reveal OCT4 as an anti-stress RBP for translational regulation that critically coordinates the survival and differentiation of PSCs in response to various stressors.

Keywords OCT4, AKT, RNA-binding protein (RBP), Translation initiation, Oxidative stress, Pluripotent stem cells (PSCs)

Introduction

Human embryonic stem cells (hESCs) hold great promise for regenerative medicine. Deciphering the detailed regulatory networks that dictate the self-renewal and directed differentiation of hESCs under physiological and pathological conditions is key to their safe and efficacious clinical applications [1]. The acquisition and maintenance of pluripotency is tightly controlled by the regulatory circuit comprising the core stemness transcription factors (mainly OCT4, SOX2 and NANOG), which are finely regulated by reversible post-translational modifications such as phosphorylation. Among the major phosphorylation regulators, the PI3K/AKT-mediated phosphorylation signaling cascade plays an indispensable and unique role in maintaining the self-renewal and survival of hESCs [2].

PI3K/AKT signaling pathway elicited by IGF and Heregulin can affect Activin A/SMAD and FGF/ERK/WNT signaling pathway, respectively, to activate stemness factors and inhibit differentiation signals, thereby maintaining hESC self-renewal [3, 4]. When the PI3K/AKT pathway is kept active in hESCs, the role of AKT in self-renewal maintenance is primarily mediated by inhibiting ERK to maintain the activity of GSK3β, and phosphorylating SMAD2/3 to promote the expression of NANOG, in contrast to mESCs where AKT mediates the phosphorylation and inhibition of GSK3 that leads to β-catenin-mediated TCF3 activation and NANOG maintenance [2]. When the PI3K/AKT pathway is down-regulated, the relieved ERK and WNT pathways can work together to inhibit the activity of GSK3β, facilitating the entry of β-catenin into the nucleus, enhancing the expression of MixL1 and other mesoendoderm differentiation genes, and promoting the differentiation of hESCs toward mesoendoderm [5, 6]. On the other hand, over-activated PI3K/AKT pathway will not only promote the rapid proliferation of pluripotent stem cells (PSCs), but also inhibit the intrinsic cellular quality control mechanisms such as apoptosis and autophagy pathway, resulting in the accumulation of DNA damages, hence reducing the self-renewal ability and differentiation potential of PSCs [7, 8]. Therefore, in order to safeguard the self-renewal ability and differentiation potential of hPSCs, the PI3K/AKT activity must be maintained at a relatively stable and balanced level [3]. However, so far, little is known about the molecular mechanisms governing the homeostasis of PI3K/AKT signaling components in hPSCs.

Our previous work shed some light on potential regulatory mechanism for AKT1, the best known and characterized isoform of the AKT family gene. We discovered that in pluripotent human embryonic carcinoma cells (hECCs), the master stemness transcription factor OCT4 (Octamer-binding protein 4, encoded by *POU5F1* gene) can bind to the promoter of *AKT1* and suppress its transcription. Thus, when endogenous OCT4 was knocked down by RNA interference, the mRNA level of AKT1 was dramatically elevated [9]. In contrast, total AKT1 protein levels remained unchanged or even slightly decreased in OCT4-silenced hECCs. These results indicated a disconnection between the transcriptional versus post-transcriptional/translational regulation of AKT1 expression by OCT4, and called for a closer look into potential post-transcriptional/translational regulatory mechanisms for AKT-pathway-genes in hPSCs.

In mammals, post-transcriptional regulation mechanisms universally govern gene expression across cell types, developmental stages and metabolic conditions by controlling pre-mRNA splicing and maturation, mRNA transport, storage, editing, turnover and translation, and influence the type, abundance and location of expressed proteins [10, 11]. Post-transcriptional gene regulations are usually mediated by numerous RNA-binding proteins (RBPs) that impact on a variety of physiologic processes [12, 13]. Emerging evidence indicated that some transcription factors (TFs) may also act as RBPs and thereby playing dual roles in transcriptional and post-transcriptional regulations. For instance, hnRNPM and p54nrb serving as TFs to activate the transcription of FGF1 in the nucleus, can also act as the 5′ internal ribosome entry site (IRES) sequence binding protein to take the FGF1 mRNAs out of the nucleus for translation [14]. SOX2 was shown to bind to RNAs in a high-affinity/low-specificity manner in murine ESCs (mESCs) [15] and can simultaneously associate with RNA and DNA to form ternary RNA/SOX2/DNA complexes during somatic cell reprogramming [16]. Previously, systematic RBP screening studies implicated OCT4 as a potential RBP in HeLa cells [17] and hESCs [18]. Furthermore, another work identified 803 potential RBPs in the nucleus of mESCs that encompassed OCT4. It was further shown that the conjectured sequence of the RNA binding domain (RBD) of OCT4 may reside at the α1 helix of OCT4 (ELEQFAK),

while the DNA binding domain (DBD) of OCT4 is on the α3 helix, indicating that the binding of OCT4 to RNA can be separated from its binding to DNA [19]. However, that study only defined the general RBP roles of OCT4 in the nucleus, and there was no demonstration of its specific roles in post-transcriptional regulations.

Here, by multiomics approaches, we revealed that OCT4 bound to the GC-rich elements in the 5′-UTR of multiple PI3K/AKT-pathway-gene mRNAs, and promoted their translation initiation via IRES-mediated EIF4F-independent pathway in stressed PSCs. Conditions and events in the microenvironment of PSCs that impair their ability to self-renew and differentiate is defined as stress. There are two types of stressed conditions: extrinsic stress refers to any detrimental change in the biological system caused by adverse environmental factors such as unphysiological hypoxia, non-optimal temperature, nutrients, chemical or biological signals, mechanical forces, etc., while intrinsic stress arises from internal metabolic challenges such as the accumulation of waste products, generation of reactive metabolites such as reactive oxygen species (ROS), or DNA alterations resulting from repeated cell division [20, 21]. We show here that under hypoxic condition, OCT4 knockdown significantly dampened IRES-dependent AKT1 translation when the cap-dependent translation pathway in PSCs was blocked. Furthermore, specific disruption of OCT4-AKT1 mRNA interaction led to a higher susceptibility of PSCs to oxidative stress-induced apoptotic death and prioritized differentiation toward ectoderm and endoderm. Thus, our results reveal OCT4 as a novel player for translational regulation that critically coordinates and balances the survival and differentiation of PSCs under extrinsic and intrinsic stress conditions.

Materials and methods

Cell lines

The NCCIT cells were purchased from American Type Culture Collection (ATCC, Rockville, USA). NCCIT cells were cultured at 37 °C in DMEM (21063-029, Invitrogen, California, USA) supplemented with 10% fetal bovine serum (GIBCO 10099 or Pufei 1101-500) in a humidified 5% CO_2 incubator (3111, Thermo Fisher Scientific, Massachusetts, USA). The sources and culture conditions for the hESC lines H1 and H9 were as previously described [22–24]. The TAP-OCT4 H1 cell line was constructed as described in Supplementary file 6: Fig. S2, and the TAP-AKT1 H9 cell lines were constructed as described in Fig. 8A.

Antibodies

The rabbit monoclonal anti-OCT4 antibody (2890S), anti-pan Akt antibody (4691), anti-Lamin B1 antibody (17416), anti-TBP antibody (44059), anti-Histone H3 antibody (4499), anti-HSP90 (4877) antibody, anti-HNRNPA1 antibody (8443) and anti-Tubulin antibody (2148) used in this study were purchased from Cell Signaling Technology. Anti-OCT4 antibody (SC-5279) was purchased from Santa Cruz Biotechnology. Anti-EIF3G antibody (ab191422) was purchased from Abcam. Anti-GAPDH [HRP] antibody (A00192) was purchased from GenScript. Anti-FLAG M2 antibody (F1804) and anti-FLAG M2 magnetic beads (M8823) were purchased from Sigma-Aldrich.

siRNA knockdown of OCT4

Transfection of small interfering RNA (siRNA) was performed using OligofectamineTM (Invitrogen 12252-011) according to the manufacturer's instructions with slight modifications as previously described [9]. Briefly, NCCIT, H1 or H9 cells were dissociated using 0.25% trypsin–EDTA or 0.25% EDTA one day before transfection and plated into a 12-well plate in 900 µl culture medium. The next day, when the cultures were reaching 30% confluence, cells were transfected with scramble siRNA or a mixture of three OCT4 siRNAs (synthesized by Invitrogen) with equal molar ratio. The sequences of those siRNAs are described in Supplementary file 6: Fig. S1A.

HITS-CLIP

The HITS-CLIP experiment was performed as described before [25]. Briefly, cells were washed with ice-cold PBS for 3 times and UV cross-linking was performed with UV irradiation type C (254 nm) at 400 mJ/cm^2. Cross-linked cells were scraped off the plates and collected by centrifugation at 1000×*g* for 5 min. Cell lysis was performed in ice-cold lysis buffer (1×PBS, 0.1% SDS, 0.5% NP-40 and 0.5% sodium deoxycholate) supplemented with a 1% RNase inhibitor (Takara) and 2% protease inhibitor cocktail (Roche) for 30 min. Cell lysates were cleared by centrifugation at 10,000 rpm for 10 min at 4 °C and the supernatants were used for immunoprecipitation.

For DNA digestion, RQ1 (Promega) was added to the lysate that was incubated at 37 °C for 3 min. For immunoprecipitation, 600 µl lysate was incubated with 15 µl anti-FLAG M2 beads overnight at 4 °C. After being applied to magnet to remove the supernatant, the beads were sequentially washed with wash buffer (1×PBS, 0.1% SDS, 0.5% NP-40 and 0.5% sodium deoxycholate), high-salt wash buffer (5×PBS, 0.1% SDS, 0.5% NP-40 and 0.5% sodium deoxycholate), and PNK buffer (50 mM Tris pH = 7.4, 10 mM $MgCl_2$ and 0.5% NP-40) for two times, respectively.

The on-bead digestion was performed by adding MNase (Thermo), followed by incubation at 37 °C for 10 min. After being washed with PNK buffer as described

above, dephosphorylation and phosphorylation of the sample was performed with calf intestinal alkaline phosphatase (CIP, NEB) and polynucleotide kinase (PNK, NEB), respectively.

The immunoprecipitated protein-RNA complexes were eluted from the beads by heat denaturing and resolved on a Novex 4–12% Bis–Tris precast polyacrylamide gel (Invitrogen). The protein-RNA complexes were cut from the gel and RNA was extracted with Trizol after digesting the proteins.

The cDNA libraries were prepared using the KAPA Stranded RNA-Seq Library Preparation Kit (Kapa biosystems). The cDNAs were purified and amplified, and PCR products corresponding to 200–500 bps were purified, quantified and stored at −80 °C before sequencing. For high-throughput sequencing, the libraries were prepared following the manufacturer's instructions and applied to Illumina HiSeq X Ten system for 150 nt paired-end sequencing by ABLife, Inc. (Wuhan, China).

Analysis of HITS-CLIP/CLIP-Seq data

HITS-CLIP/CLIP-Seq libraries were sequenced on an Illumina platform. 3′-end adaptors and 5′-end 4 nt barcodes were removed using a FASTX toolkit (http://hannonlab.cshl.edu/ fastx_toolkit). Tags ≥ 14 nt in length were kept for further analysis, while the shorter tags were discarded, which were usually aligned to multiple loci in the genome. Tags ranging from 14–19 nt (no mismatch allowed) and tags ≥ 20 nt (two mismatches allowed) were aligned to the Human genome ((human_gencode_v23)) using Bowtie [26]. Only tags with a unique genomic location were kept for further analysis. Identical sequences with the same 4 nt random barcode were regarded as one tag to exclude PCR duplication. Regions with at least two overlapping tags were grouped and subjected to in silico random CLIP simulations to discard transient and non-specific OCT4-RNA interactions [25]. In brief, observed tags were randomly aligned to a given gene region and the maximum numbers of overlapping tags were calculated. This simulation was repeated 500 times. P-value was calculated by counting the frequency of maximum cluster height from the 500 repeats. Peak heights with $P < 0.01$ were used as thresholds to identify significant binding sites/peaks. OCT4 binding sites/peaks subtracting the Input binding sites/peaks were considered as OCT4 specific binding sites/peaks.

RIP-qPCR

Cells were cross-linked on ice with UV irradiation type C (254 nm) at 400 mJ per cm^2 in the presence of ice-cold PBS (4 ml per 15-cm dish). Cells were scraped off and pelleted at 1000 × *g* at 4 °C and stored at −80 °C until further use. Cells lysis was performed in ice-cold Wash buffer (1 × PBS, 0.1% SDS, 0.5% NP-40 and 0.5% sodium deoxycholate) supplemented with a 200 U/mL RNase inhibitor (Takara) and protease inhibitor cocktail (Roche) and incubated on ice for 30 min. Cell lysates were cleared by centrifugation at 10,000 rpm for 10 min at 4 °C. RQ I (promega, 1U/μl) was added to a final concentration of 1 U/μl and incubated in a water bath for 3 min at 37 °C. The reaction was subsequently cooled for 5 min on ice before being stopped by the addition of EDTA.

For immunoprecipitation, the supernatant was incubated overnight at 4 °C with 10 μg FLAG-antibody and control IgG-antibody. The immunoprecipitates were further incubated with protein A/G Dynabeads for 2 h at 4 °C. After being applied to magnet and removal of the supernatants, the beads were sequentially washed with lysis buffer, high-salt buffer (250 mM Tris 7.4, 750 mM NaCl, 10 mM EDTA, 0.1% SDS, 0.5% NP-40 and 0.5 deoxycholate), and PNK buffer (50 mM Tris, 20 mM EGTA and 0.5% NP-40) for two times, respectively. The beads were resuspended in the Elution buffer (50 nM Tris 8.0, 10 mM EDTA and 1% SDS), and the suspension was incubated for 20 min in a heat block at 70 °C to release the IPed RBP with crosslinked RNA, and vortexed. The magnetic beads were removed on the separator and the supernatant was transferred to a clean 1.5 ml microfuge tube. Proteinase K (Roche) was added into the 1% input and IPed RBP with crosslinked RNA, with a final concentration of 1.2 mg/ml. Samples were incubated for 120 min at 55 °C. The RNA was purified with Trizol reagent (Life technologies). cDNAs were reverse transcribed from equal amount of adapted RNA acquired in the above steps. Normalization of RIP results was carried out by diluting all samples to the same cDNA concentration.

Motif analysis

OCT4 binding motifs were identified using MEME as previously described [27]. Range of motif widths was set to 4 and 20 as the minimum motif width and maximum motif width, respectively.

Protein purification and mass spectrometry

TAP-OCT4 H1 cells expressing FLAG-tagged endogenous OCT4 and parental H1 cells were expanded to five 14 cm diameter dishes, washed with PBS, scraped off, and the whole cell lysates were prepared as previously described [28]. 60 μl of anti-FLAG M2 beads (Sigma) were added to 1.5 ml of whole cell lysates in No-Stick microcentrifuge tubes (Alpha Laboratories) without addition of any nucleases. Beads were washed five times for 5 min each with HEPES buffer containing 0.1% NP-40 and the bound proteins were eluted four times for 15 min at 4 °C with 0.2 mg/ml FLAG-tripeptide (Sigma). The elutes were in-solution digested by trypsin and analyzed

by mass spectrometry (Shanghai Bioprofile Inc.). For immunoprecipitation of endogenous OCT4 complexes, cell lysates were prepared as previously described [29]. In brief, 60 μg of OCT4 antibody (sc-5279X, Santa Cruz) were added to cell lysates and incubated for 3 h, then 100 μl Protein G magnetic beads (Pierce, Thermo Scientific) were added followed by further incubation for 2 h at 4 °C. Beads were washed five times for 5 min each with HEPES buffer containing 0.1% NP-40 at 4 °C, the IP complexes were on-beads digested by trypsin and analyzed by mass spectrometry (Shanghai Bioprofile Inc.) as described previously [23, 24].

qRT-PCR analysis

Total mRNAs were isolated using TRIzol reagent. Reverse transcription reaction was performed using the PrimeScript RT reagent kit (TaKaRa RR037A) according to the manufacturer's instructions. All the primers used were listed in Supplementary file 4: Dataset S4. qRT-PCR was performed as previously described [9].

Immunoblotting

Immunoblotting was performed as previously described [24, 30]. In brief, cultured cells with or without treatment were lysed on ice in Lysis Buffer (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, 1 mM Na3VO4, 1 mg/ml leupeptin, Cell Signaling Technology 9803) supplemented with Complete Protease Inhibitor Cocktail (Santa Cruz Biotechnology sc-29130) and phosphatase inhibitors (PhosSTOP, Roche 04906845001). The whole-cell lysates were boiled with SDS-PAGE sample loading buffer, separated by SDS-PAGE, blotted on PVDF or nitrocellulose membranes and probed with the indicated antibodies. The signals were visualized using the Immobilon Western Chemiluminescent HRP Substrate (Millipore WBKLS0100).

Luciferase reporter assays

For luciferase assays, the psiCheck-2 vector (Promega) was first cut by XhoI and ApaI to drop out the Firefly Luciferase promoter HSV-TK and ligated by a multiple cloning site containing fragment. The recombinant plasmid was designated as pRF-bicistronic vector. 5′-UTRs of AKT1 variants were cloned into the pRF-bicistronic vector downstream from the Renilla luciferase stop codon into the XhoI and NotI sites, and transfected with Lipofectamin Stem reagent into 15,000 H1 cells plated in 96-well plates. After 48 h, the expression of Firefly luciferase indicative of 5′-UTR translational activity was detected by double luciferase detection kit (Beyotime Biotechnology) according to manufacturer's instructions. The mean value of the technical replicates was used as the ratio of Renilla to Firefly luciferase activity for each culture.

Determination of global protein synthesis rate

Nascent protein synthesis in a given period as a reflection for global protein synthesis rate was determined using the Click-iT Plus OPP Alexa Fluor 488 Protein Synthesis Assay Kit (ThermoFisher Scientific, Cat. #C10456) as described previously [23]. Briefly, hESCs were incubated for 30 min in culture medium supplemented with O-propargyl-puromycin (OPP) (20 mg/ml final concentration). Cells were then washed with DPBS, fixed with 1% PFA for 15 min on ice, and permeabilized using PBS supplemented with 3% FBS and 0.1% saponin for 5 min at room temperature. The copper-catalyzed azide-alkyne cycloaddition was performed following the manufacturer's protocol. The Alexa Fluor 488 fluorescence intensities of the cells indicative of newly synthesized proteins were quantified and analyzed by the Operetta CLS High-Content Analysis System.

Transcriptome and full-length translating mRNA sequencing

The total RNA and RNC-RNA extraction from hESC cells was performed as previously described [31]. A pooled sample from two independent experiments was used for subsequent RNA-seq regarding mRNA and RNC-RNA, respectively. For mRNA and RNC-mRNA, the polyA + mRNA was selected by RNA Purification Beads (Illumina). The library was constructed by using the Illumina TruSeq RNA sample Prep Kit v2 and sequenced by the Illumina HiSeq-2000 for 50 cycles. High quality reads that passed the Illumina quality filters were kept for the sequence analysis. Briefly, the reads were mapped to RefSeq-RNA reference sequence with FANSe2 algorithm. Alternative splice variants were merged. The expression levels were estimated by using the RPKM unit. The mRNA length information was also acquired from RefSeq.

Statistical analysis

All quantitative data were presented as mean ± S.D. of three independent experiments. The statistical significance of compared measurements was evaluated with one-way ANOVA and LSD test using SPSS 19.0, and the difference was considered significant and increasing when $P < 0.05$ (*), < 0.01(**), < 0.001(***), or < 0.0001(****).

Results

OCT4 positively modulates AKT1 protein level via post-transcriptional mechanisms

Previously, we have found some clues about the dual-regulatory mechanism of OCT4 on AKT1 transcription and translation. Our data showed that in hECC line NCCIT, OCT4 proteins bind to the promoter of *AKT1*

and suppress its transcription. Genetic silencing of OCT4 by RNA interference greatly increased the AKT1 mRNA level but slightly decreased its protein level [9], indicating a differential regulation of AKT1 expression by OCT4 at the transcriptional versus post-transcriptional level.

To confirm and generalize the above result, we knocked down OCT4 in hECC cell line NCCIT and hESC cell lines H1 and H9 using a combination of three pairs of siRNAs against the *POU5F1* (OCT4) gene for 48 h (Supplementary file 6: Fig. S1A, B, C). As expected, both OCT4 mRNA (Fig. 1A) and protein (Fig. 1B) levels were drastically attenuated. Meanwhile, the mRNA levels of AKT1 were dramatically up-regulated in all three cell lines (Fig. 1A), while its protein levels were slightly up-regulated (NCCIT, Fig. 1B, Supplementary file 6: Fig. S1D, E) or slightly down-regulated (H1 and H9, Fig. 1B, Supplementary file 6: Fig. S1D, E), depending on the duration and efficiency of OCT4 silencing. These results confirmed that genetic silencing of total OCT4 or OCT4A in hESCs can lead to up-regulated AKT1 mRNA levels but relatively unchanged AKT1 protein levels. In addition, when H9 and H1 cells were treated with varying concentrations of Actinomycin D (ACTD), a general inhibitor for DNA transcription, for up to 48 h, AKT1 mRNA level was reduced by 50% to 90% (Fig. 1C, Supplementary file 6: Fig. S1F) while its protein level still kept constant (Figs. 1C, 1D, Supplementary file 6: Fig. S1F, G). Given the half-life of AKT1 protein is between 24 h (Supplementary file 6: Fig. S25) and 36 h [32], the maintenance of AKT1 protein level for up to 48 h in the presence of ACTD may indicate its enhanced translation to compensate for degradation. Thus, there is a notable disconnection between transcriptional and post-transcriptional/translational regulation of AKT1 expression by OCT4, calling for a closer look into the potential post-transcriptional/translational regulation of AKT-pathway-genes in hPSCs. As OCT4 is a transcription factor best known for its nuclear localization and functionality while most post-transcriptional/translational processes take place in the cytoplasm, it is important to determine if a fraction of OCT4 is present in the cytoplasm in hESCs. To this end, we conducted a careful subcellular fractionation experiment to obtain cytosolic fraction, nuclear extraction and nuclear pellet extraction (enriched in nuclear matrix proteins). We loaded samples adjusted by their volume ratios so that the band intensity of OCT4 protein in a certain fraction is proportional to its total quantity in that fraction. As shown in the immunoblots, the cytosolic fraction largely free of nuclear protein histone H3, nucleoplasm protein TBP and nuclear matrix protein Lamin B1 did contain OCT4, which approximately accounted for 10% of total OCT4 protein in H1 cells (Fig. 1E).

Genome-wide deciphering the OCT4-RNA interactome in vivo by HITS-CLIP

Given the indication that OCT4 may be an RBP [17, 19] and therefore presumably involved in gene post-transcriptional regulations, we set out to systematically decipher the OCT4-RNA interactome by high-throughput sequencing of cross-linking immunoprecipitation (CLIP) cDNA library (HITS-CLIP) [25], utilizing a CRISPR/Cas9-edited H1 cell line (designated as 'TAP-OCT4 H1 cells') in which a TAP (Tandem Affinity Purification) tag was knocked in to a position right after the 5′-UTR of the endogenous *POU5F1* locus (Supplementary file 6: Fig. S2A) to facilitate the isolation and enrichment of OCT4-RNA complexes. This is mainly because so far none of the commercially-available anti-OCT4 antibody can specifically immunoprecipitate (IP) OCT4 protein. Such TAP-OCT4 H1 cell line had endogenous level of OCT4 (Supplementary file 6: Fig. S2B) and maintained all attributes of native H1 cells, and therefore offered unique advantages and high specificity in purifying OCT4-associated complexes and in deciphering the OCT4-protein/RNA/DNA interactomes. The TAP tag used in this study contains 3×FLAG and a Strep-Tag, but Strep-Tag turned out to be highly nonspecific, so we chose 3×FLAG as the target tag. The initial cellular fraction from which the OCT4-RNA complexes were enriched was designated as 'input' and used as a control for OCT4-HITS-CLIP because it can eliminate potential variations caused by genotype diversity of samples (e.g., using H1 parental cells as control), uneven background (e.g., some proteins and RNAs in input may be captured by FLAG M2 beads at background levels but can't be captured by IgG beads) and culture conditions (in our experiments, the identical cell sample was divided into two aliquots for input and FLAG-IP HITS-CLIP).

First, TAP-OCT4 H1 cells were UV cross-linked to fix the protein-nucleic acid interaction. OCT4-RNA complexes were then immunoprecipitated from the TAP-OCT4 H1 whole cell lysates using the FLAG M2 antibody followed by MNase digestion. A narrow smear above the OCT4 over-digestion control was separated in NuPAGE gel and used to generate the HITS-CLIP library for high-throughput sequencing (Supplementary file 6: Fig. S2C). HITS-CLIP reads obtained from both immunoprecipitated (IPed) TAP-OCT4 and input libraries were analyzed in parallel, two independent biological replicates were performed (Fig. 2A). Overall, ~28 million and ~29 million raw reads were collected from two OCT4-CLIP libraries, respectively (Supplementary file 6: Fig. S2D). After subtraction of adaptors, barcodes, multiple mapped reads and PCR duplicates, a total of 1,092,524 and 1,864,424 final tags from the OCT4-CLIP were uniquely mapped to the human genome (human_gencode_v23)

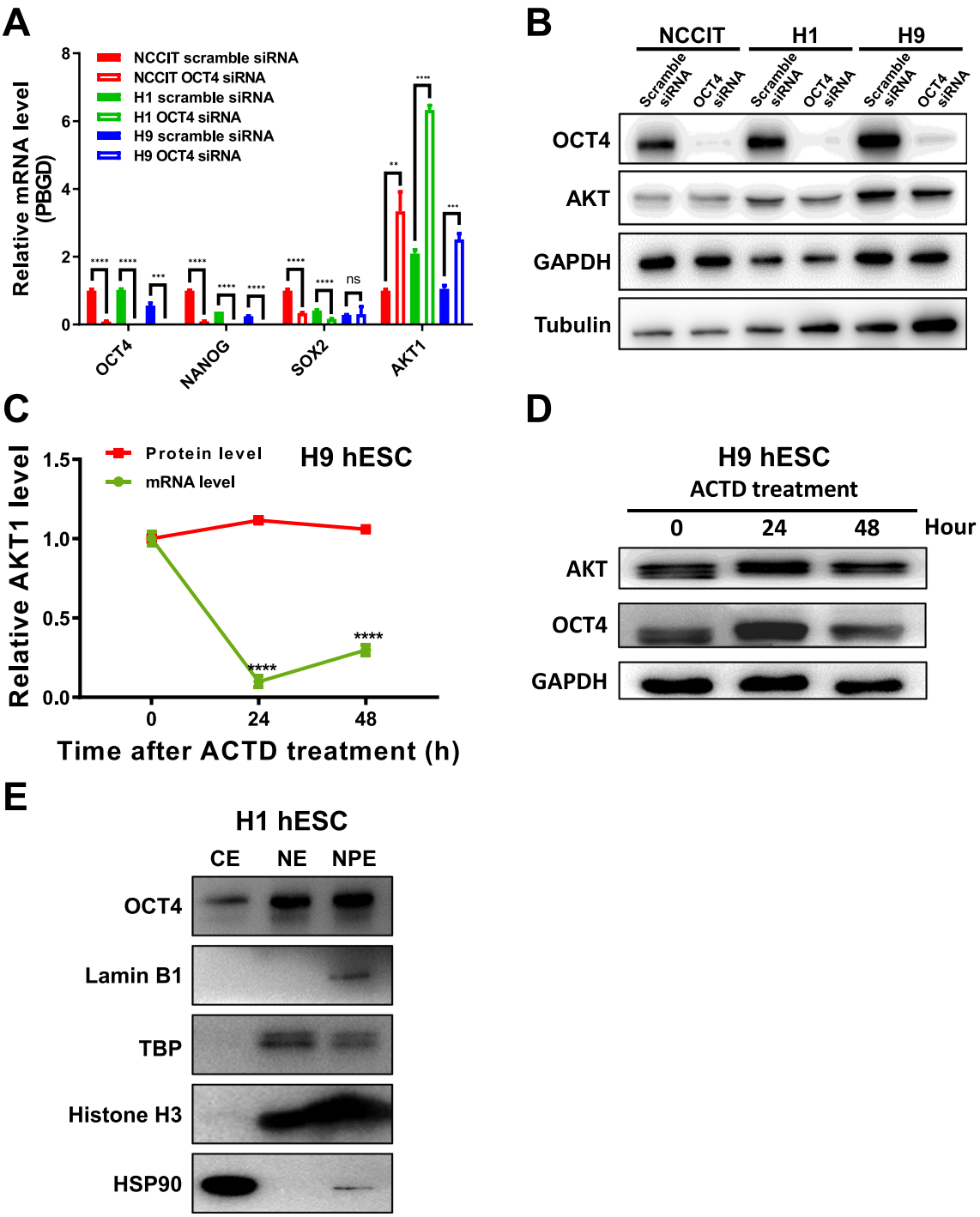


Fig. 1 Post-transcriptional regulation of AKT1 by OCT4. **A** NCCIT, H1 and H9 cells were transfected with scramble siRNA or OCT4 siRNAs for 48 h, and the mRNA levels of OCT4, NANOG, SOX2 and AKT1 were determined by qRT-PCR. **B** NCCIT, H1 and H9 cells were transfected with scramble siRNA or OCT4 siRNAs. After 48 h, cells were lysed and the whole cell lysates were subjected to immunoblotting with indicated antibodies. Data shown were from one of two independent experiments that gave similar results, and the densitometric analyses were presented in Supplementary file 6: Fig. S1D and E. **C**, **D** H9 cells were treated with 5 ng/ml Actinomycin D (ACTD) for varying period, samples were collected for qRT-PCR (C) and immunoblotting (**D**). The band intensities of AKT and OCT4 in **D** were normalized by those of GAPDH, and plotted in (**C**). Data in (**A**) and **C** were presented as mean ± SD of three independent experiments. In **C**, values at 24 h and 48 h were statistically compared with those at 0 h. Unless otherwise specified, the statistical significance of compared measurements was evaluated with one-way ANOVA and LSD test using SPSS 19.0. ns, not significant, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. **E** Relative quantity of OCT4 proteins in cytoplasmic extract (CE), nuclear extraction (NE) and nuclear pellet extraction (NP, enriched in nuclear matrix proteins) of H1 cells. Full-length blots/gels are presented in Supplementary file 7: Fig. S28

(Supplementary file 6: Fig. S2E). Pearson correlation for reads counts of OCT4-CLIP and Input-CLIP was calculated from two replicates (Supplementary file 6: Fig. S3A). Cluster analysis for 4 samples according to sample correlation coefficient showed reproducibility between two OCT4-CLIP groups and difference between OCT4-CLIP groups and Input-CLIP groups (Supplementary file 6: Fig. S3B, C).

The uniquely mapped OCT4-CLIP tags in two independent CLIP experiments showed a broad range of binding tags to sense targets (2,570,801), suggesting the transcriptome-wide activity of OCT4 in regulating RNA metabolism (Supplementary file 6: Fig. S3D). OCT4-CLIP tags encompassed 5′-UTR (4.41%), coding sequences (CDS, 24.66%), introns (31.60%), 3′-UTR (6.56%), intergenic regions (17.01%) and noncoding gene exons (2.63%) (Fig. 2B, Supplementary file 6: Fig. S3D). In addition, OCT4 binding tags were significantly enriched in 5′-UTR ($P=0.002260$, t-test), 3′-UTR ($P=0.014984$, t-test) and CDS ($P=0.016299$, t-test), compared with the input control (Fig. 2C), indicating the preferential binding of OCT4 to 5′-UTRs. The fact that OCT4 binding tags enriched after transcriptional initiation site supported this assumption (Fig. 2D; Supplementary file 6: Fig. S3E), and specific binding of OCT4 at start codon indicated that OCT4 may participate in the translation initiation process (Fig. 2E; Supplementary file 6: Fig. S3F).

OCT4 tags were further normalized and clustered as binding sites as described in the Materials and methods and Supplementary file 6: Fig. S4A. Overall, we identified 25,120 and 29,211 binding sites of 9654 and 10,740 genes from two replicates, respectively (Supplementary file 6: Fig. S4B). The average width of binding peaks was ~ 200 bp (Supplementary file 6: Fig. S4C, D). Cases of OCT4 binding to each gene region were shown as snapshots or experimentally validated by RIP-qPCR from two biological replicates (Supplementary file 6: Figs. S5A, S6). Totally 1,486 overlapped genes (obtained from the overlapped peaks) between two OCT4-CLIP biological replicates were detected (Fig. 2F). Overlapped peaks from two OCT4-CLIP replicates exhibited significant enrichment at 5′-UTR, CDS and 3′-UTR (Supplementary file 6: Fig. S5C–E), in agreement with the reads distribution results above.

Using all the HITS-CLIP reads obtained from OCT4 binding sites, unbiased search for motifs enriched in OCT4 binding sites revealed over-represented GC-rich (5′-GCCG-3′) sequences predominantly present at the 5′-UTRs (Supplementary file 6: Fig. S7). Then we used overlapped peaks to determine over-represented motifs, the top 8 motifs and their distribution on genome were presented (Fig. 2G; Supplementary file 6: Fig. S7D). Among them, 2 GC-rich hexamers (5′-CTGCCG-3′) (Z score, 274.6) and (5′-GAGCCG-3′) (Z score, 300.2) were highly enriched in the 5′-UTRs (Figs. 2H-J). Notably, these two specific OCT4 binding motifs were more enriched at the translation initiation site, further supporting the role of OCT4 in 5′-UTR regulation and translation initiation (Figs. 2K and 2L).

OCT4 binds to RNAs of multiple PI3K/AKT-pathway genes

Gene Ontology analysis of OCT4 sense binding targets revealed significantly enriched terms in RNA metabolism, translation, various stimuli and stress responses, suggesting that OCT4 may have profound regulatory roles in physiological processes such as protein phosphorylation, intrinsic apoptotic signaling pathway and stem cell maintenance (Fig. 3A). GO cellular component and molecular function analysis were performed with OCT4-bound RNAs, indicating that many of them have DNA binding, RNA binding and kinase binding functions, and they spread over most of the cell compartments (Supplementary file 6: Fig. S8A, B).

OCT4-bound RNAs were clustered into several pathways by KEGG pathway analysis. These included Focal adhesion, PI3K/Akt signaling pathway and

(See figure on next page.)

Fig. 2 OCT4 preferentially binds to the 5′-UTR of targeted transcripts. **A** The PCR products of the duplicate OCT4-IP and input samples subjected to HITS-CLIP were analyzed. The main bands were between 200 and 500 bp, and were purified and used to prepare the library for high-throughput sequencing. **B** Percentages of 5′-UTR, CDS, 3′-UTR, nc_exon, introns, intergenic region, and antisense CLIP-seq tags in OCT4-IP versus Input samples. **C** OCT4-IP CLIP tags are significantly enriched in 5′-UTR, CDS and 3′-UTR. **D** The distribution of OCT4-IP versus Input reads in the vicinity of transcriptional initiation site, transcriptional termination site, taking the transcriptional initiation site (TSS) and transcriptional termination site (TTS) as the origin, the distribution of reads in the upstream and downstream 1 kb range was calculated. **E** The distribution of OCT4-IP versus Input reads in the vicinity of start codon and stop codon, taking the start codon and stop codon as the origin, respectively, the distribution of reads in the upstream and downstream 1 kb range was calculated. **F** The Venn diagram showing the number of genes overlapping between the two repeated OCT4 CLIP-seq analysis. **G** The top 8 motifs of OCT4 binding tags analyzed using MEME software. **H** Overrepresented OCT4-binding motifs identified by CLIP-seq. Histogram of Z scores indicates the enrichment of hexamers in CLIP-seq clusters compared to randomly chosen regions of similar sizes in the same genes. Z score of the top hexamer is indicated. **I**, **J** Distribution of the two overrepresented OCT4 binding motifs in binding genes. **K**, **L** Distribution of the two overrepresented binding motifs (insert) relative to the start codon (indicated as 0) are indicated by green and red curves for two repeats, respectively

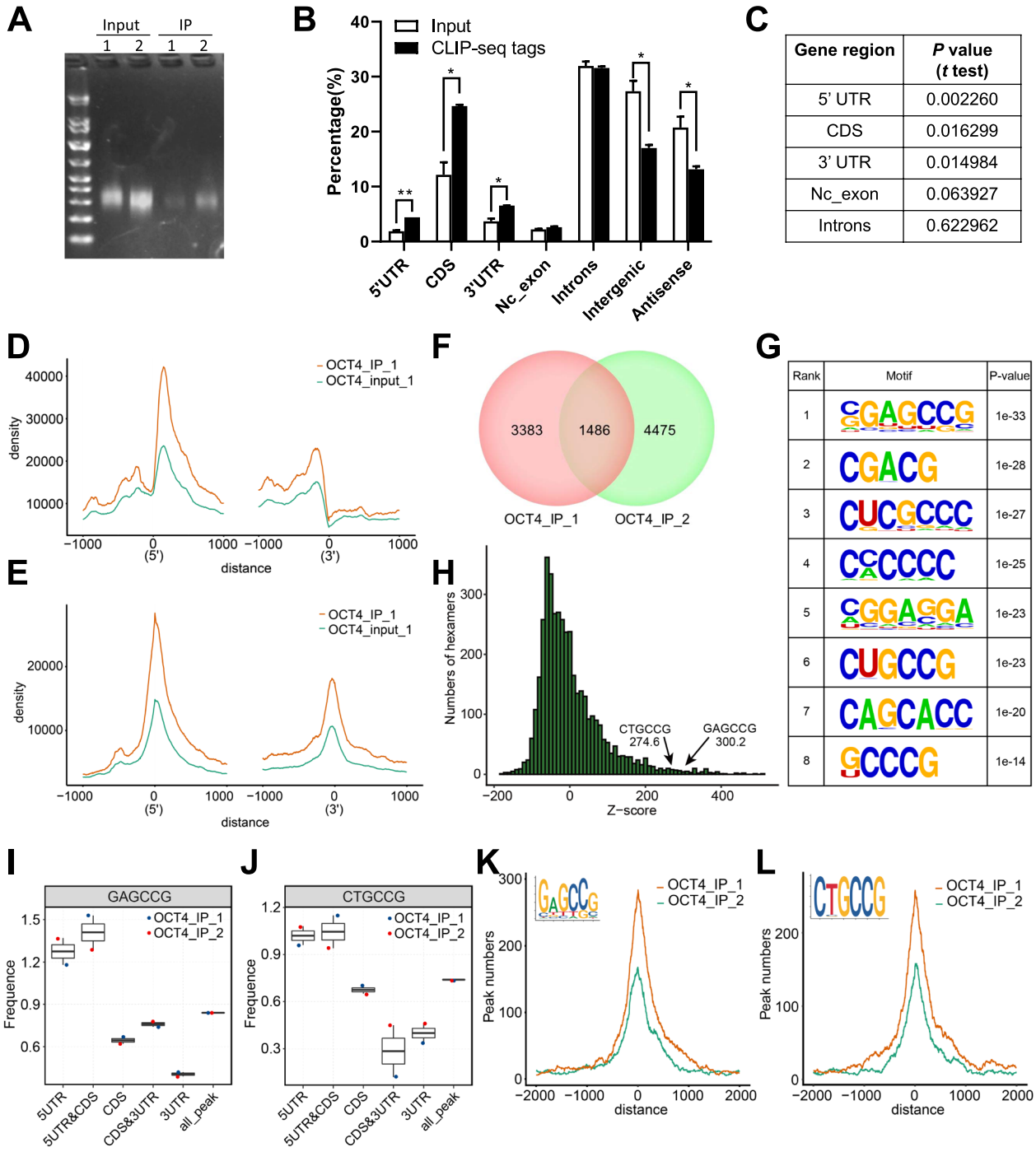


Fig. 2 (See legend on previous page.)

PPAR signaling pathway (Fig. 3B). Consistent with its role in AKT regulation, OCT4 bound to the transcripts of multiple PI3K/AKT-pathway genes, including PIK3CD, PIK3R2, PDPK1, AKT1, KRAS, BAD and EIF4E (Figs. 3C, 3D; Supplementary file 6: Fig. S8C). Furthermore, OCT4 bound to the sense transcripts of several PI3K/AKT-pathway genes in close proximity to the 5′-UTRs, implicating its role in translation initiation processes (Fig. 3E; Supplementary file 6: Figs. S5, S9). Next, we conducted RIP-qPCR using IgG and FLAG M2 beads as negative controls to further verify the binding of OCT4 protein with ten of its mRNA targets identified

by HITS-CLIP (AKT1, PIK3R2, EPHA2, YWHAQ, FOS, GADD45A, PML, EIF4E, FGFR4, RAB13) (Supplementary file 6: Fig. S6). These results raised an interesting possibility that OCT4 may regulate the translation of a group of PI3K/AKT-pathway genes by directly binding to their 5′-UTRs.

OCT4 interacts with proteins regulating cap-independent and/or cap-dependent but EIF4F-independent translation initiation

It has been shown that besides TFs, multiple transport-related proteins such as nuclear pore transporter (Nucleoporin, Nup) were also in the OCT4-interacting protein list, indicating that OCT4 may participate in nuclear and cytoplasmic transport of mRNAs [33, 34]. However, most OCT4-protein interactome work carried out thus far utilized exogenously expressed OCT4 proteins from the nuclear fraction, primarily focusing on its roles as a TF, and few studies have focused on the cytosolic OCT4. In two widely-cited OCT4-protein interactome papers, Ding et al. [33] used mESC nuclear extract to identify OCT4 interactome in nucleus and Pardo et al. [28] used whole cell lysate to identify OCT4 interactome in mESCs. In both papers, Benzonase, a genetically engineered endonuclease was employed to fragment genome DNA which significantly improved the enrichment efficiency of nuclear TFs. Since our focus was on cytosolic pool of OCT4, we used mild conditions for cell lysis and modified the Pardo's method by omitting the addition of any nucleases to the lysis buffer. Such conditions not only retained the protein/RNA interactions, but also enabled precipitation of the nuclear TFs with genome DNA, and hence a relative enrichment of the cytosolic proteins and other soluble nuclear proteins (such as splicing factors) in the supernatant, and therefore allowed us to minimize the interference of the TF roles of OCT4 and to focus more on its novel RBP roles. To obtain some clues for its potential involvement in translation regulation, endogenous TAP-OCT4 proteins in the whole cell lysates of the TAP-OCT4 H1 cells obtained by the above-described manner were IPed by an anti-FLAG-M2 antibody attached to magnetic beads, followed by thorough washes and a final release of all bound proteins with 3 × FLAG peptides, and mass spectrometric identification of the OCT4-binding proteins. Reaction products from the intermediate steps were subjected to immunoblotting to determine the efficiency, yield, and specificity of the affinity purification process (Fig. 4A; Supplementary file 6: Fig. S10A). To independently verify candidate TAP-OCT4-interacting proteins, we subjected the whole cell lysates of the TAP-OCT4 H1 cells to an OCT4 antibody (SC-5279) that is expected to have lower binding affinity and specificity than anti-FLAG-M2 (Fig. 4B). Both samples were followed by identification of OCT4-binding proteins via high-performance liquid chromatography coupled to tandem mass spectrometry (Supplementary file 6: Fig. S10B, C). A total of 187 peptides were identified in anti-FLAG IP complexes, including 142 effective peptides and 75 proteins (Fig. 4C; Supplementary file 1: Dataset S1), and a total of 1117 peptides were identified in anti-OCT4 IP complexes, including 905 effective peptides and 412 proteins (Fig. 4D; Supplementary file 2: Dataset S2). 28 putative OCT4-binding proteins were identified in both IPs, of which only one protein (ANXA2, annexin A2) was detected in the negative control IP/MS where TAP-OCT4-protein-free parental H1 cell lysates were IPed with the FLAG M2 beads. Hence, the remaining 27 proteins were considered as authentic OCT4-interacting proteins (Figs. 4E, 4F; Supplementary file 1: Dataset S1).

The subcellular locations, molecular functions, and biological processes of the above identified OCT4-interacting proteins were further analyzed by GO annotation. Nucleus, cytoplasm and membrane were the most representative cell component classifications among the identified proteins (Fig. 5A), highlighting the presence

(See figure on next page.)

Fig. 3 Enriched OCT4 binding to PI3K/AKT-pathway-gene transcripts. **A** Gene Ontology (GO) analysis of OCT4 binding targets. Significantly enriched GO terms of genes with OCT4 binding were identified using the Bingo software (hypergeometric test with Benjamini and Hochberg false discovery rate correction). The x axis indicates the enrichment *P* value on a −log10 scale; the y axis indicates number of genes with OCT4 binding on a log2 scale. The size of each point is proportional to the ratio of OCT4-bound genes associated with one GO term to all genes associated with this GO term. **B** Enrichment analysis of the KEGG pathway. The size and color of the dots represents the number of enriched genes and the adjusted *P* values, respectively. **C** OCT4 binding targets enriched in PI3K/AKT pathway. **D** Circos diagram shows enriched OCT4 binding sites at PI3K/AKT pathway transcripts: from outer to inner, there are circle 1, human chromosomes (indicated as chr01-23 with different colors); circle 2, heat map displaying all of human genes; circle 3, binding density of Input counts is normalized to tag per million (TPM); circle 4, binding density of OCT4 indicating all the OCT4 binding sites across the transcriptome counts is normalized to tag per million (TPM); circle 5, heat map view of genes with OCT4 binding; circle 6, red link lines indicate PI3K/AKT pathway transcripts with OCT4 binding. **E** Distribution and percentage of OCT4 binding sites in genes. Binding sites are shown as wiggle plots (dark blue for the OCT4 library). CDS regions are boxed in black. The 5′-UTR and 3′-UTR are boxed in green and grey, respectively. Introns are indicated as lines. The cyan boxes above gene structures indicate predicted OCT4 binding peaks. The x axis indicates genomic positions in chromosomes. The y axis indicates normalized HITS-CLIP/CLIP-seq abundance. HITS-CLIP/CLIP-seq tag counts were normalized to tag per million (TPM) to adjust for differences of two HITS-CLIP/CLIP-seq libraries in sequencing depth

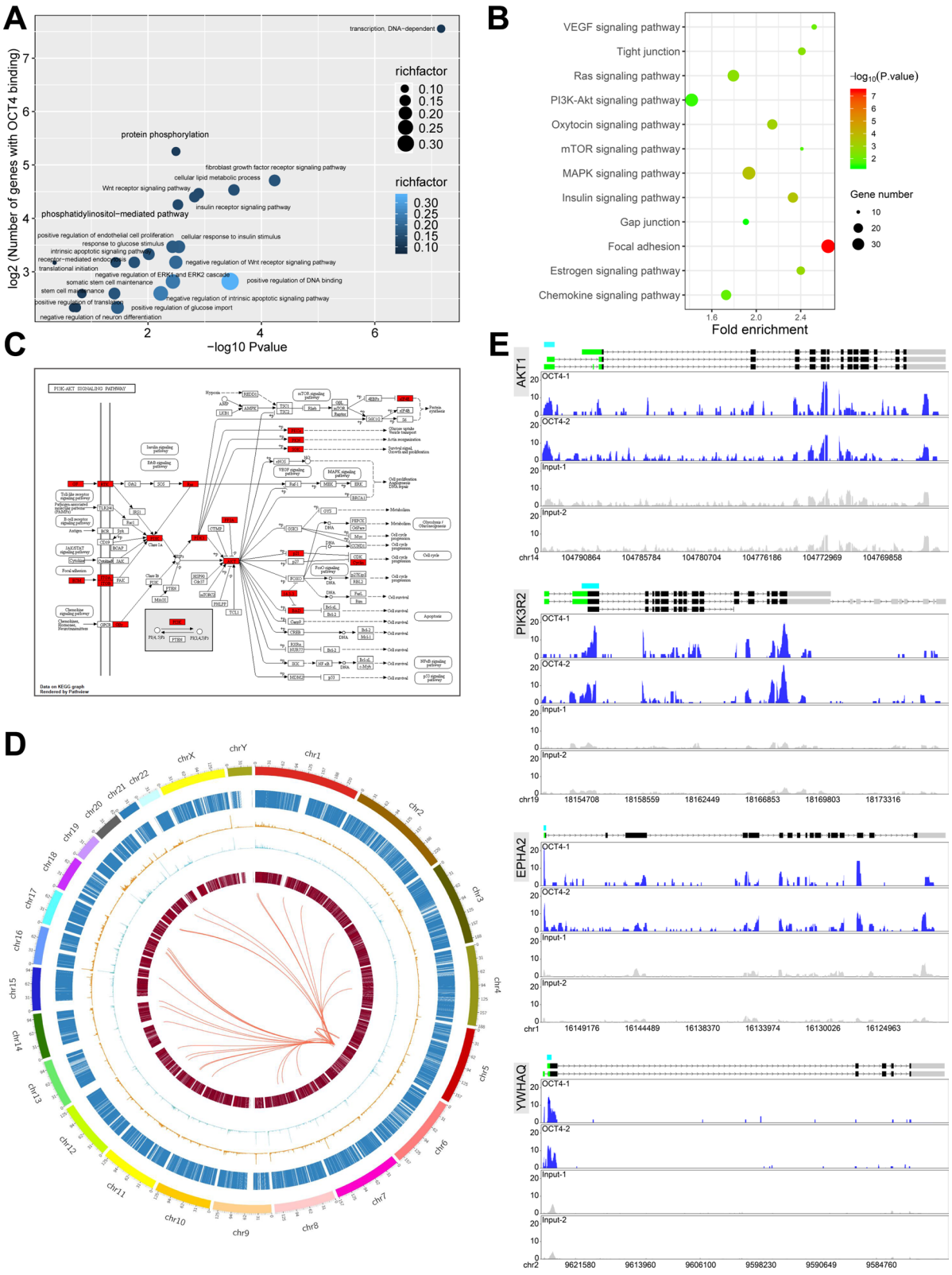


Fig. 3 (See legend on previous page.)

of a significant proportion of the OCT4-interacting proteins at extra-nuclear locations. In terms of molecular functions, the most representative groups of the identified proteins were translation factor binding, RNA binding and protein binding (Supplementary file 6: Fig. S12A). With respect to biological processes, cell–cell adhesion, protein folding, translational elongation and response to stress were the most representative processes (Fig. 5B). Genes encoding those OCT4-interacting proteins were clustered into several pathways by KEGG pathway analysis. These included protein processing in endoplasmic reticulum, RNA transport, RNA degradation and phagosome pathways (Supplementary file 6: Fig. S12B).

Proteomic network analysis by STRING (http://string-db.org/) further revealed that multiple OCT4-interacting proteins were associated with 4 groups, including translation, transcription, response to stress and transport (Fig. 5C). This was in agreement with a GO analysis, which identified proteins highly enriched in fundamental processes involved in RNA metabolism and ribosomal RNA processing. Particularly, several well-defined translation initiation proteins (e.g., 40S ribosomal components and eIF subunits) were identified as OCT4-interacting proteins, supporting a role for OCT4 in regulating mRNA translation. In addition, among the 27 OCT4-interacting proteins, there were a classical IRES trans-acting factor (ITAF) HNRNPA1, and, EIF3G, an important component of EIF3 (Fig. 4F). There is evidence that the binding of EIF3 with PABP potentiates ribosome recruitment to the IRES of XIAP mRNA [35], and initiates the translation of *nanos* mRNA [36], via cap-independent pathway. In addition, EIF3D can mediate cap-dependent but EIF4F-independent translation under certain stress conditions [37–39]. Therefore, the detection of EIF3 subunits as OCT4-binding proteins may indicate their participation in cap-independent and/or cap-dependent but EIF4F-independent translation initiation in hESCs.

Taken together, these data placed OCT4 into an extended RBP network that is associated with cap-independent and/or cap-dependent but EIF4F-independent translation initiation processes. Thus, it is of particular interest to find out if OCT4 regulates the translation initiation of the PI3K/AKT-pathway genes.

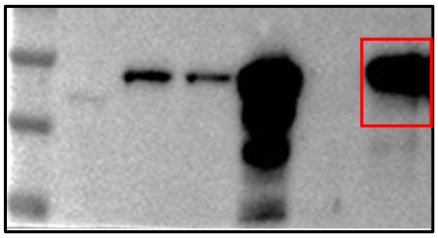
PI3K/AKT-pathway gene mRNAs are commonly bound by OCT4 and OCT4-interacting RBPs

Among the RBPs identified as OCT4-interacting proteins, HNRNPA1 is one of the best characterized RBPs related to translation initiation process. It was reported that HNRNPA1 is an ITAF that regulates the translation of cyclin D1 and c-Myc via IRES-mediated initiation under conditions of reduced eIF4F complex formation [40]. Therefore, OCT4 may function as part of the cap-independent and/or cap-dependent but EIF4F-independent translation initiation complex promoting the translation of PI3K/AKT-pathway genes. To confirm the proteomic data, we examined two of the key OCT4-interacting proteins, EIF3G and HNRNPA1, in TAP-OCT4 immune complexes by Western blotting. FLAG M2 antibody beads were employed to IP non-tagged OCT4 in H1 cells (as a negative control) and TAP (3×FLAG)-OCT4 in TAP-OCT4 H1 cells. As shown in the elute fractions, TAP-OCT4 was successfully IPed from the TAP-OCT4 H1 cells but not H1 cells (red arrow), and both EIF3G (green arrow) and HNRNPA1 (blue arrow) proteins were co-IPed from the TAP-OCT4 H1 cells much more than those from H1 cells (Supplementary file 6: Fig. S11). Furthermore, we compared previously reported HNRNPA1-RNA interactome [41] and EIF3G-RNA interactome [42] with OCT4-RNA interactome in this study and were able to identify 34 mRNAs shared by all three RBPs, and a total of 634 mRNAs that are bound by OCT4 together with HNRNPA1 and/or EIF3G (Fig. 5D; Supplementary file 3: Dataset S3). Among the 634 genes, at least 13 were PI3K/AKT-pathway genes (Supplementary file 6: Fig. S12C). We then performed wiggle plot for all OCT4, HNRNPA1 and EIF3G binding peaks on PI3K/AKT-pathway genes' mRNAs, and identified several partially-overlapping or adjacent peaks co-bound by OCT4, HNRNPA1 and EIF3G (data used for wiggle plot were downloaded from ENCODE database (www.encodeproject.org/)) (Figs. 5E, 5F; Supplementary file 6: Fig. S13),

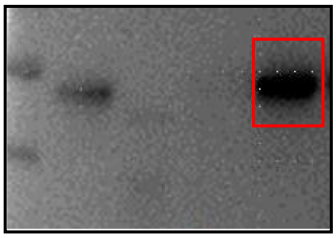
(See figure on next page.)

Fig. 4 Identification of cap-independent translation initiation components in OCT4-protein interactome in hESCs. **A** TAP-OCT4 H1 whole cell lysates were immunoprecipitated with the FLAG M2 antibody, the immune complexes eluted with FLAG-tripeptide (3×F Elute), together with the post-IP supernatant (SP), the whole cell lysate (input), cell pellet, and recombinant His-OCT4 protein, were subjected to SDS-PAGE and immunoblotting with anti-OCT4. The TAP-OCT4 band is indicated. **B** TAP-OCT4 H1 whole cell lysates were immunoprecipitated with the anti-OCT4 antibody (SC-5279), the immune complexes were washed, eluted, and subjected to SDS-PAGE and immunoblotting. The OCT4 band is indicated. **C**, **D** Summary of mass spectrometry analysis for OCT4-interacting proteins immunoprecipitated by FLAG M2 antibody **C** or by anti-OCT4 antibody (**D**). **E** The Venn diagram showing the number of proteins overlapping between the two OCT4-protein interactome analyses in (**C**) and (**D**). **F** List of 27 OCT4-interacting proteins commonly identified in two mass spectrometry analyses shown in (**C**) and (**D**). OCT4 (*POU5F1*) was consistently detected in all analyses and therefore also included in the table. Full-length blots/gels are presented in Supplementary file 7: Fig. S28

A



B



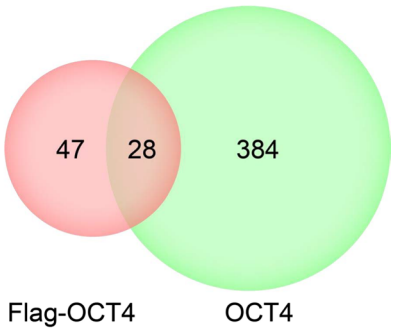
C

Sample IPed by	Flag M2 antibody
Protein group	75
Unique peptide	142
Peptide-Spectrum matches	187

D

Sample IPed by	OCT4 antibody
Protein group	412
Unique peptide	905
Peptide-Spectrum matches	1117

E



F

Protein IDs	Gene names	Protein names	Number of proteins	Unique peptides	Unique sequence coverage [%]	Mol. weight [kDa]	Sequence length	Score
P50990	**CCT8**	**T-complex protein 1 subunit theta**	**1**	**1**	**1.8**	**59.6**	**548**	**6.3**
P68104	**EEF1A1**	**Elongation factor 1-alpha 1**	**3**	**2**	**7.8**	**50.2**	**462**	**12.6**
O75821	**EIF3G**	**Eukaryotic translation initiation factor 3 subunit G**	**1**	**1**	**3.1**	**35.6**	**320**	**6.3**
Q04637	**EIF4G1**	**Eukaryotic translation initiation factor 4 gamma 1**	**1**	**1**	**1.1**	**175.5**	**1599**	**5.9**
P21333	**FLNA**	**Filamin-A**	**2**	**10**	**6**	**280.7**	**2647**	**86.1**
P09651	**HNRNPA1**	**Heterogeneous nuclear ribonucleoprotein A1**	**1**	**1**	**2.7**	**38.7**	**372**	**6.2**
P07900	HSP90AA1	Heat shock protein HSP 90-alpha	1	1	2	84.7	732	7.1
P08238	HSP90AB1	Heat shock protein HSP 90-beta	6	11	23.5	83.3	724	234.1
Q58FF3	HSP90B2P	Putative endoplasmin-like protein	1	1	4.3	45.9	399	13
P11021	HSPA5	78 kDa glucose-regulated protein	1	3	8.3	72.3	654	26
P11142	HSPA8	Heat shock cognate 71 kDa protein	3	9	14.4	70.9	646	89.6
P38646	HSPA9	Stress-70 protein, mitochondrial	1	1	1.3	73.7	679	6.4
P10809	HSPD1	60 kDa heat shock protein, mitochondrial	1	3	8.2	61.1	573	19
P20700	LMNB1	Lamin-B1	1	1	2	66.4	586	6.6
P35579	MYH9	Myosin-9	1	1	0.6	226.5	1960	6.5
P60660	MYL6	Myosin light polypeptide 6	1	1	9.9	16.9	151	24.8
P49321	NASP	Nuclear autoantigenic sperm protein	1	1	1.8	85.2	788	6
P22234	PAICS	Multifunctional protein ADE2	1	1	2.4	47.1	425	5.9
O00264	PGRMC1	Membrane-associated progesterone receptor component 1	1	1	8.2	21.7	195	5.9
Q01860	POU5F1	POU domain, class 5, transcription factor 1	2	16	60.3	38.6	360	323.3
Q06830	PRDX1	Peroxiredoxin-1	1	2	16.1	22.1	199	13.3
P04843	RPN1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase	1	1	1.8	68.6	607	5.9
P42677	RPS27	40S ribosomal protein S27	1	1	16.7	9.5	84	34.2
P23396	RPS3	40S ribosomal protein S3	1	1	3.7	26.7	243	6.64
Q96EB6	SIRT1	NAD-dependent protein deacetylase sirtuin-1	1	5	10.6	81.7	747	47.8
P53985	SLC16A1	Monocarboxylate transporter 1	1	1	2.2	53.9	500	6.17
Q71U36	TUBA1A	Tubulin alpha-1A chain	9	7	29.5	50.1	451	78.3
Q13885	TUBB2A	Tubulin beta-2A chain	5	5	12.1	49.9	445	70.4

Fig. 4 (See legend on previous page.)

indicating that OCT4/HNRNPA1/EIF3G may function as part of the cap-independent and/or cap-dependent but EIF4F-independent translation initiation complex that

defines the translation level of the PI3K/AKT-pathway genes.

OCT4 regulates translation ratio of PI3K/AKT-pathway genes in hypoxic hESCs

Hypoxia represents a major extrinsic stress for hESCs [21], and hypoxic conditions tend to promote mesodermal differentiation toward mesoderm-derived lineages such as cardiomyocytes [43]. To gain insights into the dynamic roles of OCT4 in regulating the mRNA translation of PI3K/AKT pathway genes as well as the global genes in response to hypoxic stress, we conducted a combined ribosome nascent-chain complex-bound RNA sequencing (RNC-seq) and mRNA sequencing (RNA-seq) [31] with hESCs transfected with scramble siRNA (siNC) or OCT4 siRNA (siOCT4) that were maintained either under 20% O_2 normoxia (N20, S20, respectively) or under 1% O_2 hypoxia (N1, S1, respectively).

First, samples from two independent experiments for RNA-seq and RNC-seq were used to generate libraries for high-throughput sequencing (Supplementary file 6: Fig. S14A). Reads obtained from libraries were analyzed, and two independent replicates (−1, −2) were performed (Supplementary file 6: Fig. S14B, C). Principal Component Analysis (PCA) for all samples according to reads per kilo base per million (RPKM) of randomly selected genes showed reproducibility of two replicates and differences among the four groups (Supplementary file 6: Fig. S15A). And Pearson correlation calculated from two replicates for reads counts of all groups demonstrated good correlation (Supplementary file 6: Fig. S15B).

To globally compare the impact of OCT4 in regulating gene transcription versus mRNA translation under hypoxia, the RPKM values from either RNA-seq or RNC-seq of hypoxic siNC (N1) or hypoxic siOCT4 (S1)-treated cells were normalized by RPKM of their normoxic counterparts (N20 and S20, respectively) to yield normalized RPKM of N1/N20 and S1/S20, respectively, and the ratios of (S1/S20)/(N1/N20) were calculated. Genes with the ratio values > 2 and $P < 0.05$ were considered as up-regulated (red), genes with ratio values < 0.5 and $P < 0.05$ were considered as down-regulated (green). It was shown that at the transcriptional level, 6283 genes were down-regulated and 2727 genes up-regulated by siOCT4 under hypoxia, while at the translational level, 843 genes were down-regulated and 567 genes up-regulated (Fig. 6A).

Next, the RPKM values from RNC-seq were divided by RPKM values from RNA-seq to yield the translation ratio (TR), i.e., the ratio of translating mRNA versus total transcribed mRNA [31] (Supplementary file 5: Dataset S5). The ratios of (S1 TR/S20 TR) / (N1 TR/N20 TR) were calculated. The percentages of up-regulated genes (red), down-regulated genes (green) and not significantly-changed genes (grey) in all detected genes (left) or in ten RIP-qPCR validated PI3K/AKT-pathway genes (right) were shown in Fig. 6B. In contrast to 22% of the global genes that had decreased TRs in siOCT4-treated hypoxic hESCs, 60% of the PI3K/AKT-pathway genes had declined TRs, indicating that OCT4 is preferentially required for maintaining or promoting the translation of PI3K/AKT-pathway genes in response to hypoxic stress (Fig. 6B). The duplicate TR values of ten PI3K/AKT-pathway genes in differently-treated cells, the ratios of (S1/S20) / (N1/N20) and the associated p values were presented in Fig. 6C.

Gene Ontology (GO) biological process analysis of genes with down-regulated TR in siOCT4-treated hypoxic hESCs revealed significantly enriched terms in translational initiation, response to stresses, apoptosis and autophagy, suggesting that OCT4 may play profound regulatory roles in these processes under stressed conditions (Fig. 6D), while up-regulated TR was mainly associated with genes regulating cell division and DNA repair processes (Fig. 6E). GO molecular function analysis revealed that genes with down-regulated TR were associated with RNA-binding, kinase activity, cell adhesion and heat shock protein binding (Supplementary file 6: Fig. S16A), while genes with up-regulated TR were involved in DNA binding, helicase activity and transcription

(See figure on next page.)

Fig. 5 PI3K/AKT-pathway gene mRNAs are commonly bound by OCT4 and OCT4-interacting RBPs. **A** Gene ontology (GO) cellular localization analysis of the identified OCT4-interacting proteins in hESCs. **B** GO biological process analysis of the identified OCT4-interacting proteins in hESCs. **C** Computational construction of OCT4-interacting protein network using STRING database (https://string-db.org). Interactions include direct (physical) and indirect (functional) associations. Light-blue edge: known interaction from curated databases; purple edge: experimentally determined interaction; green edge: text mining; black edge: co-expression; dark-blue edge: protein homology. (**D**) The Venn diagram showing the number of peaks overlapping among the OCT4-, HNRNPA1- and EIF3G-bound targets. (**E**, **F**) The genomic structures of AKT1 and EIF4E are shown in the top panel. Dark blue for the OCT4 library, CDS regions are boxed in black. The 5′-UTR and 3′-UTR are boxed in green and grey, respectively. Introns are indicated as lines. Wiggle plots of four HITS-CLIP/CLIP seq (OCT4 CLIP-1,2, HNRNPA1 CLIP-1,2 (from ENCODE database) and EIF3G CLIP-1,2 (from ENCODE database)) are shown below the gene structures. OCT4 binding peaks in the 5′-UTR of AKT1 and EIF4E are indicated by cyan boxes. The x axis indicates genomic positions in chromosomes. The y axis indicates normalized HITS-CLIP/CLIP-seq abundance. HITS-CLIP/CLIP seq counts were normalized to tag per million (TPM) to adjust for differences of two libraries in sequencing depth

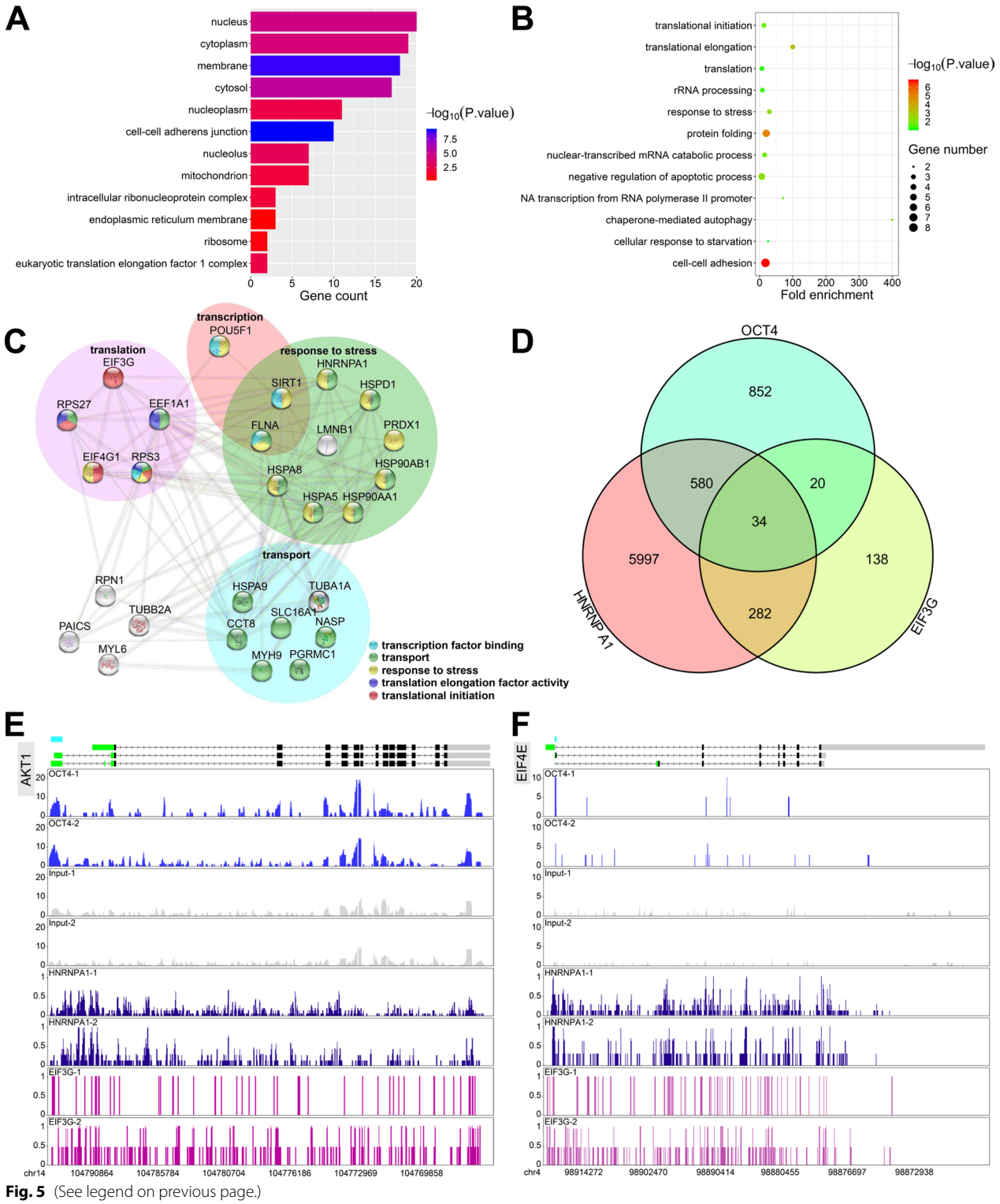


Fig. 5 (See legend on previous page.)

factor activity (Supplementary file 6: Fig. S16B). GO cellular component analysis showed that genes with down-regulated TR were mainly located in the cytoplasm and those with up-regulated TR were mainly in nucleus and nucleoplasm (Supplementary file 6: Fig. S16C, D).

The above genes with down-regulated TR were clustered into multiple pathways by KEGG pathway analysis

that included PI3K/AKT signaling, mTOR signaling, insulin signaling pathways, apoptosis and phagosome pathways (Fig. 6F). In comparison, the genes with up-regulated TR were mainly involved in basal transcription factors, ubiquitin mediated proteolysis, cell cycle and PI3K/AKT signaling pathways (Fig. 6G). Since PI3K/AKT signaling pathway was identified from gene clusters with both down-regulated and up-regulated TR, we took a closer look at this pathway (KEGG PATHWAY hsa04151). Among 354 genes belonging to the PI3K/AKT signaling pathway, 62 genes had down-regulated TR and 47 genes had up-regulated TR in siOCT4-treated hypoxic hESCs (Supplementary file 6: Fig. S17B, C). The key node genes AKT1, AKT2, PIK3R2 and PIK3CD had down-regulated TR while AKT3, PIK3R1, PIK3CA and PIK3CB had up-regulated TR (Supplementary file 6: Fig. S17). EIF4E, a key factor for cap-dependent translation initiation, also had up-regulated TR (Fig. 6C, Supplementary file 6: Fig. S17), indicating a potential negative role of OCT4 in regulating cap-dependent translation. As shown above, the mRNAs of AKT1 and PIK3R2 were identified (Fig. 3E) and verified (Figure S6) to be bound by OCT4 proteins, and OCT4 knockdown significantly suppressed their TR (Fig. 6C, Supplementary file 6: Fig. S17B), we therefore decided to take AKT1 as an example to find out the molecular details on how OCT4 regulates the translation of the PI3K/AKT-pathway genes.

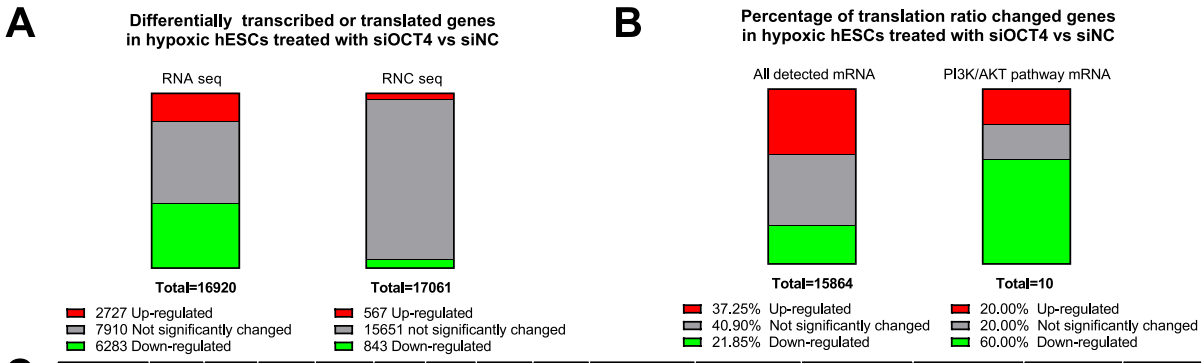
OCT4 promotes IRES-mediated translation initiation of AKT1

As OCT4 bound to 5′-UTR of AKT1 mRNA, we examined how it may affect the translation initiation of AKT1. Previous broad-spectrum genomic screening of IRES loci indicated that AKT1 may contain potential IRES sequences [44], suggesting that its translation initiation may be mediated by IRES structures in a cap-independent manner. To test this hypothesis, we utilized the dual-luciferase reporter system to determine whether the AKT1 5′-UTR can mediate cap-independent translation initiation. Since there are three known AKT1 transcripts resulting from alternative splicing (Supplementary file 6: Fig. S18A), we conducted RT-PCR and clone sequencing, using specific primers for AKT1v1 and common primers for AKT1v2 and v3 to detect the expression of AKT1 transcripts in hESC cell lines H1 and H9. The results showed that AKT1v1 and AKT1v3 were the main splicing forms in hESCs (Supplementary file 6: Fig. S18B, C). To gain insights into how the AKT1 5′-UTR may mediate translation initiation, we simulated the secondary structures of the 5′-UTR for AKT1v1 and v3 using the Mfold webserver, and both AKT1v1 and v3 transcripts exhibited stem-loop secondary structures at their 5′-UTRs and adjacent regions (Supplementary file 6: Fig. S18D, E), supporting the possibility for their IRES-mediate translation initiation.

As several OCT4 binding peaks on AKT1 5′-UTR were detected in two biological replicates of the OCT4 HITS-CLIP analysis (Fig. 7A), we chose the unique peak identified in both assays for further analysis. As this unique peak sequence was encompassed in AKT1v2 and AKT1v3 but not AKT1v1 (Fig. 7B), and AKT1v3 was the main splicing variant present in hESCs (Supplementary file 6: Fig. S18B, C), we modeled the secondary structure of this OCT4-bound sequence in the context of AKT1v3 mRNA (Fig. 7C). Indeed, the OCT4-bound segment of the AKT1v3 mRNA was completely situated at the 5′-UTR and formed typical stem-loops indicative of IRES-like structure. The dual-luciferase reporter system

(See figure on next page.)

Fig. 6 OCT4 critically regulates the translation ratio of PI3K/AKT-pathway gene mRNAs in hESCs exposed to hypoxia. **A** Parental WT H9 cells were transfected with scramble siRNA (siNC) or OCT4 siRNA (siOCT4) under 20% O_2 normoxia. After 48 h, a half of the cells were maintained under normoxia, and the other half of the cells were switched to 1% O_2 hypoxia and cultured for another 24 h, the last 8 h of which were treated with 60 μM 4E1RCat. Samples were harvested and subjected to genome-wide mRNA-seq (left) and RNC-seq (right), respectively. The RPKM values of hypoxic siNC (N1) or hypoxic siOCT4 (S1)-treated cells were normalized by RPKM values of their normoxic counterparts (N20 and S20, respectively) to yield normalized RPKM N1/N20 and S1/S20, respectively, and the ratios of (S1/S20)/(N1/N20) were calculated. Genes with ratio values > 2 and P < 0.05 were considered as up-regulated (red), genes with ratio values < 0.5 and *P* < 0.05 were considered as down-regulated (green), and both are differentially transcribed or translated genes between siOCT4- and siNC-treated H9 cells. **B** The RPKM values from RNC-seq were divided by RPKM values from mRNA-seq to yield translation ratio (TR), i.e., the ratio of translating mRNA versus total transcribed mRNA. The ratios of (S1 TR/S20 TR)/(N1 TR/N20 TR) were calculated. Genes with ratio values > 2 and *P* < 0.05 were considered as up-regulated, genes with ratio values < 0.5 and *P* < 0.05 were considered as down-regulated. The percentages of up-regulated genes (red), down-regulated genes (green) and not significantly-changed genes (grey) in all detected genes (left) or in ten RIP-qPCR validated PI3K/AKT-pathway genes (right) were shown. **C** The duplicate (− 1, − 2) TR values of ten RIP-qPCR validated PI3K/AKT-pathway genes in hypoxic siNC (N1), normoxic siNC (N20), hypoxic siOCT4 (S1), normoxic siOCT4 (S20)-treated cells, the ratios of S1/S20, N1/N20, (S1/S20)/(N1/N20) and associated *P* values were calculated and presented. **D**, **E** Gene ontology (GO) biological process analysis for identified genes with down-regulated **D** or up-regulated **E** TR by siOCT4 under 1% O_2 hypoxia. The size and color of the dots represented the number of enriched genes and the adjusted P values, respectively. **F**, **G** KEGG pathway analysis for identified genes with down-regulated **F** or up-regulated **G** TR by siOCT4 under 1% O_2 hypoxia. The size and color of the dots represented the number of enriched genes and the adjusted P values, respectively



C

Gene	N1-1	N1-2	N20-1	N20-2	S1-1	S1-2	S20-1	S20-2	S1-1/S20-1	S1-2/S20-2	N1-1/N20-1	N1-2/N20-2	(S1/S20)/(N1/N20)	p value
AKT1	0.70	0.62	0.77	0.66	0.39	0.40	0.97	1.02	0.40	0.40	0.91	0.94	0.43	**0.0006**
EIF4E	0.97	1.15	7.11	7.25	3.18	3.25	1.64	1.52	1.94	2.13	0.14	0.16	13.79	**0.0013**
EPHA2	1.41	1.35	0.32	0.27	0.43	0.43	0.84	0.81	0.51	0.53	4.45	5.03	0.11	**0.0023**
FGFR4	1.27	1.07	0.55	0.45	0.36	0.35	0.99	1.10	0.36	0.32	2.32	2.35	0.15	**0.0001**
FOS	1.62	1.66	0.78	0.64	1.51	1.35	1.02	1.34	1.48	1.00	2.08	2.59	**0.53**	**0.0432**
GADD45A	4.11	2.79	0.63	0.46	1.53	1.40	1.36	1.26	1.13	1.11	6.49	6.12	0.18	**0.0007**
PIK3R2	0.72	0.55	0.76	0.45	0.52	0.46	1.02	1.05	0.52	0.44	0.95	1.22	0.44	**0.0247**
PML	0.68	0.68	0.41	0.34	0.72	0.55	0.78	0.80	0.92	0.68	1.67	2.03	0.43	**0.0195**
RAB13	1.85	0.96	1.88	1.07	0.82	0.92	1.19	1.18	0.69	0.78	0.98	0.89	**0.78**	**0.0448**
YWHAQ	0.91	0.90	2.90	3.38	2.03	2.04	1.26	1.31	1.61	1.56	0.31	0.27	5.46	**0.0004**

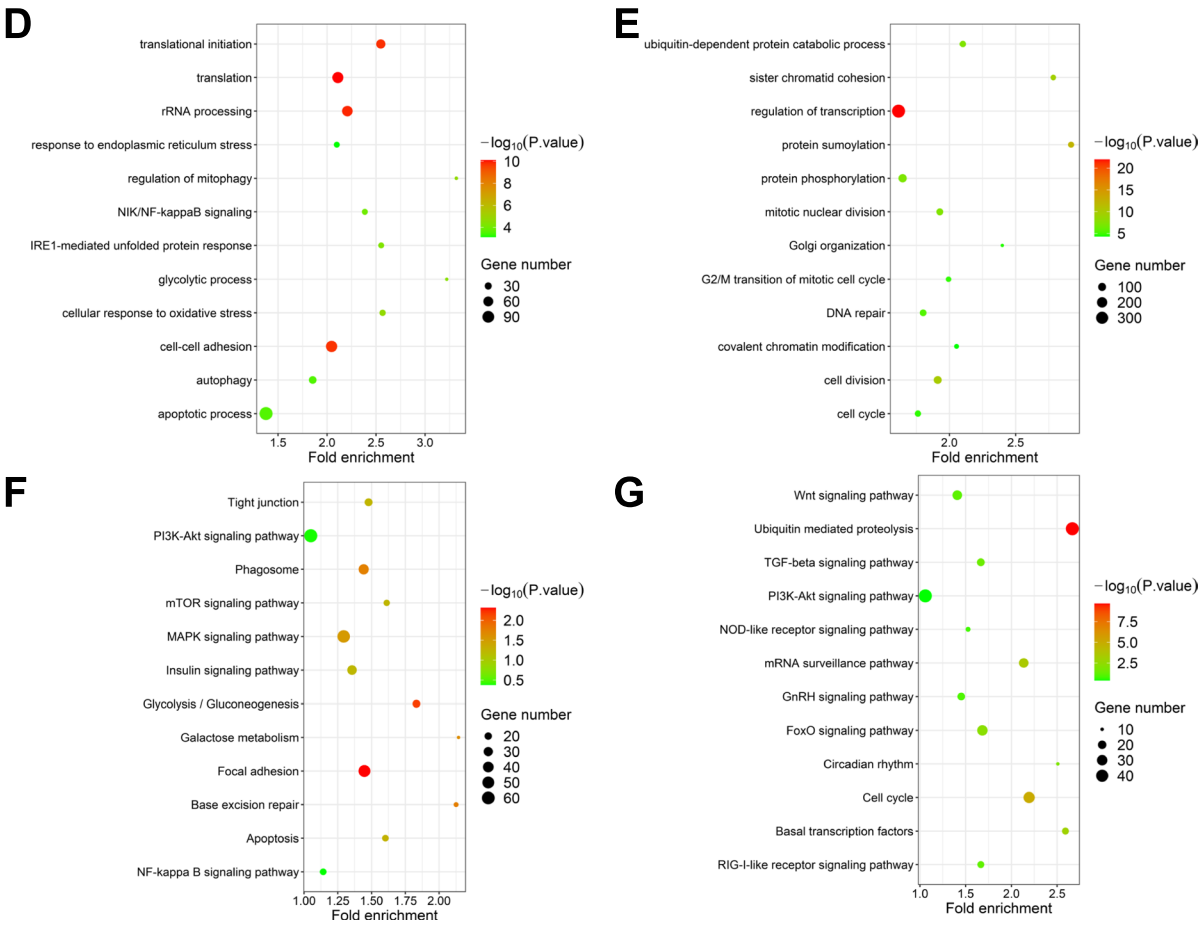


Fig. 6 (See legend on previous page.)

was employed to assess the capabilities of AKT1v1 and v3 5′-UTR in mediating translation initiation (Fig. 7D). The result showed that the 5′-UTR of AKT1v3 but not that of AKT1v1 promoted translation of the FLUC reporter (Fig. 7E), indicating that the OCT4-bound IRES-like structures at the AKT1v3 5′-UTR are likely to mediate translation initiation in hESCs.

Stress conditions such as hypoxia are known to shift the translation mode from cap-dependent to IRES-mediated pathway [45–48]. In a similar manner to that of Fig. 6, we exposed OCT4 siRNA- or scramble siRNA-transfected hESCs into 1% O_2 hypoxia, treating them with or without 4E1RCat, an inhibitor that can specifically block cap-mediated EIF4F-dependent translation by interfering the formation of the EIF4F complex [49]. Remarkably, when cells were exposed to 1% O_2 hypoxia, 4E1RCat increased rather than decreased AKT protein level in scramble siRNA-transfected cells indicative of greatly-stimulated and rebounded AKT protein synthesis via IRES pathway when EIF4F-dependent pathway is blocked (Fig. 7F). Such rebounded increase in AKT translation was completely abolished in OCT4 siRNA-transfected cells (Fig. 7F), supporting a key role for OCT4 in IRES-mediated AKT translation under hypoxia. Furthermore, under hypoxia, OCT4 siRNA slightly decreased the global mRNA translation rate. 4E1RCat alone significantly reduced global mRNA translation that was further aggravated by OCT4 siRNA (Fig. 7G). This result indicated that the global translation events controlled by OCT4 and EIF4F are largely non-overlapping, further supporting a role of OCT4 in EIF4F-independent translation pathway. Taken together, under hypoxic condition, OCT4 siRNA treatment reduced the translation ratio of over 20% of all mRNAs that are enriched in PI3K/AKT-pathway genes including AKT1 (Fig. 6), completely blocked the rebounded increase in AKT1 protein level stimulated by 4E1RCat/1% O_2, and slightly inhibited the global protein synthesis (Fig. 7G). Based on all the results above, a working model was proposed to explain how OCT4 may work as an RBP and a probable ITAF in regulating IRES-mediated translation initiation (Fig. 7H).

OCT4-promoted AKT1 translation counteracts hypoxic and oxidative stresses and regulates lineage specification

Since OCT4 knockdown has many biological effects besides regulating AKT1 translation, approaches that can specifically block or disrupt the binding of OCT4 protein to AKT1 mRNA are the key to further mechanistic investigation. To this end, by using CRISPR/Cas9 system, we knocked in a TAP tag to the exon 2 of *AKT1* gene in wild-type parental H9 cells to alter the secondary structure of the TAP-AKT1 mRNA 5′-UTR that can presumably reduce the binding of OCT4 proteins. Remarkably, despite much effort, we failed to obtain any homozygotic TAP-AKT1 knock-in clones. A heterozygotic H9 clone (A11) with a wild-type AKT1 allele and a $Puro^R$ cassette-containing TAP-$Puro^R$-AKT1 knock-in allele, two heterozygotic H9 clones (A1101, A1103) with a wild-type (WT) AKT1 allele and a TAP-AKT1 knock-in allele (Fig. 8A, Supplementary file 6: Fig. S19) were selected for further analyses. We termed this technical approach as Heterozygous Knocking In N-terminal Tags (HKINT) and view it as a powerful tool to study the functional roles and regulatory machinery of the 5′-UTR of a specific gene of interest. As expected, the removal of the $Puro^R$ cassette, the genomic integrity and the successful expression of the knocked-in TAP tag of H9 clones A1101 and A1103 were confirmed (Fig. 8B, Supplementary file 6: Fig. S19). Consistent with our hypothesis, in immunoblots of all examined heterozygotic TAP-AKT1 clones, the TAP-AKT1 band was significantly weaker than the unedited native WT AKT1 band (Fig. 8C, Supplementary

(See figure on next page.)

Fig. 7 OCT4 promotes IRES-like structure-mediated translation initiation of AKT1. **A** OCT4-binding sites on AKT1 mRNA identified from two independently repeated HITS-CLIP experiments. **B** Genome view of relative positions of identified OCT4-binding sites on three AKT1 splicing variants. **C** Predicted secondary structure of the 5′-UTR of AKT1v3 by using the Mfold webserver. The identified OCT4-binding region is marked in light blue. **D** Schematic representation of the plasmids used in (**E**). **E** Relative luciferase activity indicative of IRES activity in H1 cells transfected with a series of plasmids in (**D**). Firefly luciferase (FLUC) activity was measured and normalized by Renilla luciferase (RLUC) activity. Data were presented as mean ± SD of three independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$ compared with the control groups. **F** H9 and H1 cells transfected with OCT4 siRNA- or scramble siRNA in normoxia for 48 h were placed under hypoxic (1% O_2) or normoxic (20% O_2) conditions for another 24 h, the last 8 h of which were treated with or without 30 μM 4E1RCat. Samples were lysed and the whole cell lysates were immunoblotted with indicated antibodies (upper panels). The intensities of the AKT bands were quantified by densitometry and normalized by those of the GAPDH bands. The normalized AKT protein levels were plotted in the bottom panel. Data were presented as mean ± SD of three independent experiments. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. **G** H9 cells transfected in similar manners as in **F** were placed under hypoxic (1% O_2) condition and treated with 30 μM 4E1RCat for 4 h. The nascent protein synthesis (translation) rates of treated cells were determined as described in the Methods, and presented as a scatter dot plot. Each dot represented the measurement for a single cell, and randomly selected 300 cells from three independent experiments (with 100 cells from each experiment) were analyzed for each treatment group. The lines represented mean ± SD of the measurements for 300 cells. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. **H** Working model for OCT4 as an RBP and a probable ITAF regulating IRES-like structure-mediated translation initiation. Full-length blots/gels are presented in Supplementary file 7: Fig. S28

A

Sample	Symbol	#Chr	Start	End	PeakID	Tags	Strand	Length	maxHeight	Summit	regionInfo	GeneID
OCT4-CLIP-1	AKT1	chr14	104795483	104795725	chr14_72099	14	-	243	10	104795558	five_prime_UTR	ENSG00000142208.15
OCT4-CLIP-1	AKT1	chr14	104792597	104792722	chr14_72090	11	-	126	9	104792675	five_prime_UTR	ENSG00000142208.15
OCT4-CLIP-2	AKT1	chr14	104795444	104795762	chr14_16099	27	-	319	19	104795616	five_prime_UTR	ENSG00000142208.15

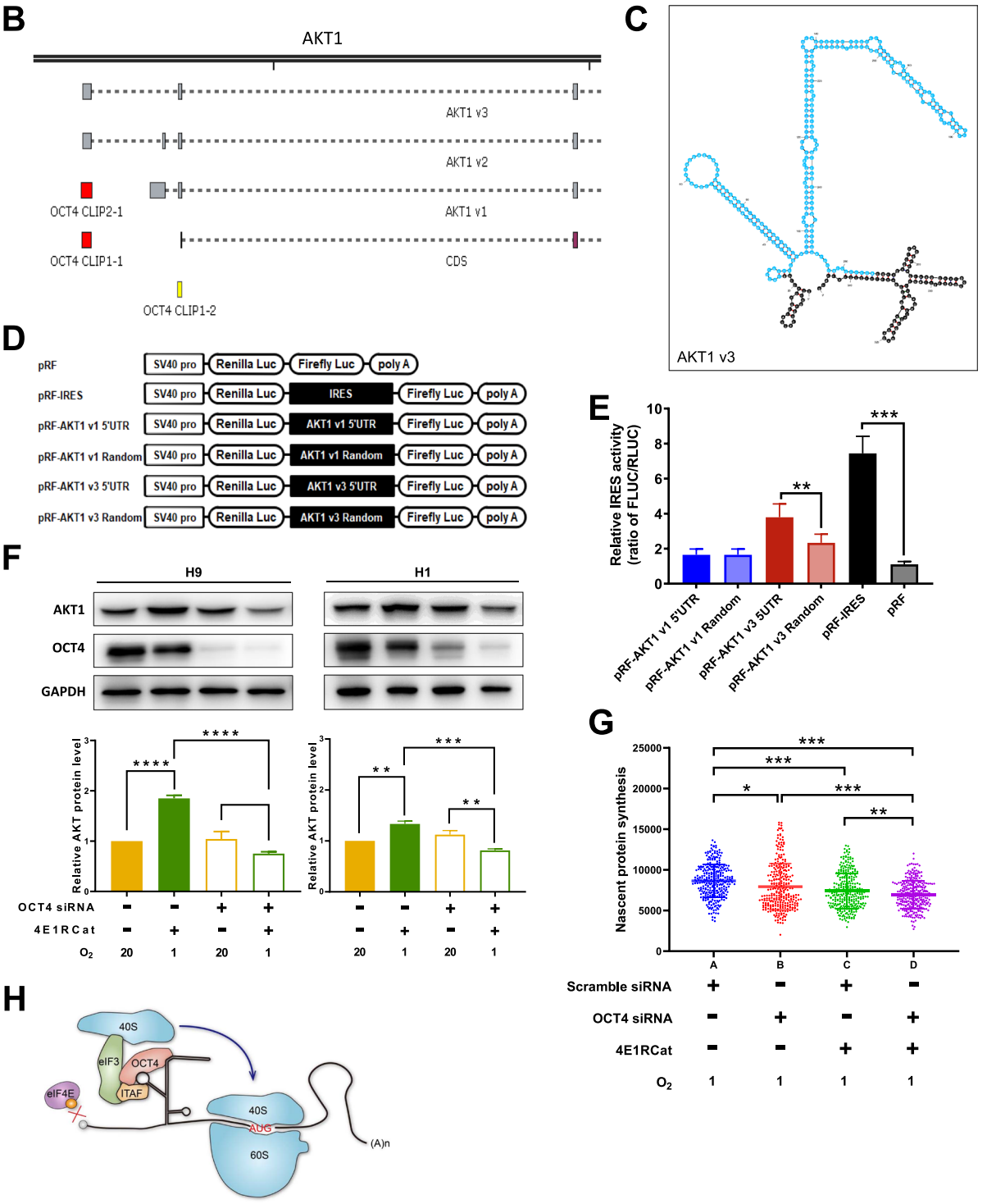


Fig. 7 (See legend on previous page.)

file 6: Fig. S20). Densitometric quantification showed that under routine normoxic culture conditions, the TAP-AKT1 protein levels were approximately 32% and 31% of the WT AKT1 protein levels in clones A1101 and A1103, respectively (Fig. 8D, Supplementary file 6: Fig. S20). In contrast, the corresponding TAP-AKT1 mRNA levels were approximately 109% and 347% of the WT AKT1 mRNA levels, respectively (Fig. 8E, Supplementary file 6: Fig. S21), indicating that the dramatically suppressed protein expression of the TAP-AKT1 is not due to suppressed transcription of the TAP-AKT1 mRNA, but most likely caused by its reduced translation. Modeling of the secondary structures of a portion or full-length WT versus TAP knock-in AKT1 v3 mRNA (Fig. 8F, Supplementary file 6: Figs. S22, S23) indicated that alteration of the IRES-like structure of WT 5′-UTR by the neighboring knocked-in TAP tag may reduce its binding with OCT4 proteins, leading to reduced translation of the TAP-AKT1 mRNA.

To examine dynamic changes of the TAP-AKT1 versus WT AKT1 translation under hypoxic stress, we exposed parental H9 and two heterozygotic TAP-AKT1 H9 clone A1101 and A1103 cells to 1% O_2 hypoxia for varying period. In parental H9 cells, there was a dramatic increase in WT AKT1 protein level at 8 h post hypoxic treatment that was returned back to original level after 24 h of exposure (Fig. 8G, Supplementary file 6: Fig. S24). Similar patterns of dynamic changes were seen with both WT AKT1 and TAP-AKT1 proteins in both A1101 and A1103 cells (Fig. 8G, Supplementary file 6: Fig. S24). Remarkably however, over the whole period of hypoxic treatment, the ratios of TAP-AKT1/WT AKT1 kept decreasing in both clones (Fig. 8H, Supplementary file 6: Fig. S24). Cycloheximide, a widely used inhibitor that is known to block translational elongation, reduced the TAP-AKT1 and WT AKT1 protein levels at similar rates (Supplementary file 6: Fig. S25), suggesting that the faster decline of the TAP-AKT1 over WT AKT1 protein level under hypoxic stress is most likely a result of preferentially-reduced translation initiation of the TAP-AKT1 mRNA due to its altered 5′-UTR structure.

(See figure on next page.)

Fig. 8 OCT4-promoted AKT1 translation counteracts hypoxic and oxidative stresses and regulates lineage specification. **A** Heterozygous Knocking In N-terminal Tags (HKINT) by CRISPR/Cas9 system. In this strategy, a 171 bp TAP tag, followed by a 647 bp intron, and a Puro^R expression cassette with LoxP sequence on both sides was knocked into the exon 2 of AKT1 gene in wild-type parental H9 cells. A long ssDNA donor fragment flanked at both ends by the 1 kb homology arms was used as a template to induce the homology-directed repair mechanism after the CRISPR/Cas9-mediated double-strand break. After Puro screening, clones with the full-length cassette knocked-in were treated with Cre recombinase to remove the Puro^R expression cassette. A heterozygotic H9 clone (A11) with a wild-type AKT1 allele (blue) and a Puro^R cassette-containing TAP-Puro^R-AKT1 knock-in allele (black), two heterozygotic H9 clones (A1101, A1103) with a wild-type AKT1 allele (blue) and a TAP-AKT1 knock-in allele (red) were selected for further analyses. **B** Agarose gel image showing the PCR amplicons using primers spanning the Puro^R cassette and the TAP tag, respectively to confirm the removal of the Puro^R cassette and genomic integrity of H9 clones A1101 and A1103. **C** The whole cell lysates of wild-type parental H9, Cre-treated H9, H9 clones A11, A1101, A1102 and A1103 were subjected to immunoblotting with anti-AKT1. The TAP-AKT1 and unedited native AKT1 bands were marked in red and blue, respectively. **D** Parental H9 and heterozygotic TAP-AKT1 H9 clones A1101 and A1103 cells cultured under normoxia were harvested and the whole cell lysates were subjected to immunoblotting. Band intensities of TAP-AKT1 and WT-AKT1 in H9, A1101 and A1103 cells were quantified by ImageJ software and plotted as relative protein levels. More details were included in Supplementary file 6: Fig. S20. **E** The relative mRNA levels of WT-AKT1 in parental H9 and in H9 clones A1101 and A1103, and those of TAP-AKT1 in A1101 and A1103 were calculated and presented. More details were included in Supplementary file 6: Fig. S21. **F** Schematic representation for altered secondary structure of AKT1 v3 mRNA 5′-UTR induced by TAP knock-in. OCT4 protein is presumed to bind to the identified OCT4-binding region (green) of the 5′-UTR transcribed by the AKT1 WT allele, but not the altered 5′-UTR transcribed by the TAP-AKT1 KI allele. The remaining part of the 5′-UTR is marked in grey, the TAP sequence is marked in cyan, AKT1 CDS is marked in yellow, and the ATG start codon is marked in red. **G** Parental H9 and heterozygotic TAP-AKT1 H9 clone A1101 and A1103 cells exposed to 1% O_2 hypoxia for varying period were harvested and the whole cell lysates were subjected to immunoblotting with the indicated primary antibodies. **H** Band intensities of TAP-AKT1 and WT-AKT1 from three independent experiments were quantified by ImageJ software, and the ratios of TAP-AKT1/WT AKT1 band intensities were plotted as mean ± SD (n = 3). More details were included in Supplementary file 6: Fig. S24 **I** Parental H9, TAP-AKT1 H9 clone A1101 and A1103 cells cultured under normoxia (20% O_2) were analyzed by flow cytometry. Dead cells, living cells and cells underwent early or late apoptotic cell death were plotted as percentages of total cells. **J** Parental H9, A1101 and A1103 cells cultured under normoxia (20% O_2) were treated with 1000 μM H_2O_2 for 2 h, and analyzed by flow cytometry. Dead cells, living cells and cells underwent early or late apoptotic cell death were plotted as percentages of total cells. The primary data for (**I**) and **J** were presented in Supplementary file 6: Fig. S26. (K) Cells of parental H9 and heterozygotic TAP-AKT1 H9 clones A1101 and A1103 cultured under 20% O_2 normoxia in mTeSR™1 media were switched to STEMdiff™ Trilineage Ectoderm, Mesoderm, or Endoderm Medium (lineage inducer +), respectively. The Trilineage Medium were replaced with fresh medium daily and the induction proceeded for five to seven days. Cells grown in mTeSR™1 media throughout (lineage inducer –) served as a control. Cells were harvested and their RNA samples were analyzed by qRT-PCR using primers that amplify marker genes for three germ layers. The expression levels of germ layer marker mRNAs were normalized by that of GAPDH, and the derived values were further normalized by those of "H9 lineage inducer –" samples, with the latter being set as 1 (Supplementary file 6: Fig. S27). The ratios of "lineage inducer + "/ "lineage inducer –" for three cell lines were expressed as mean ± SD of triplicate measurements from one of three independent experiments, which gave similar results. *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$. Full-length blots/gels are presented in Supplementary file 7: Fig. S28

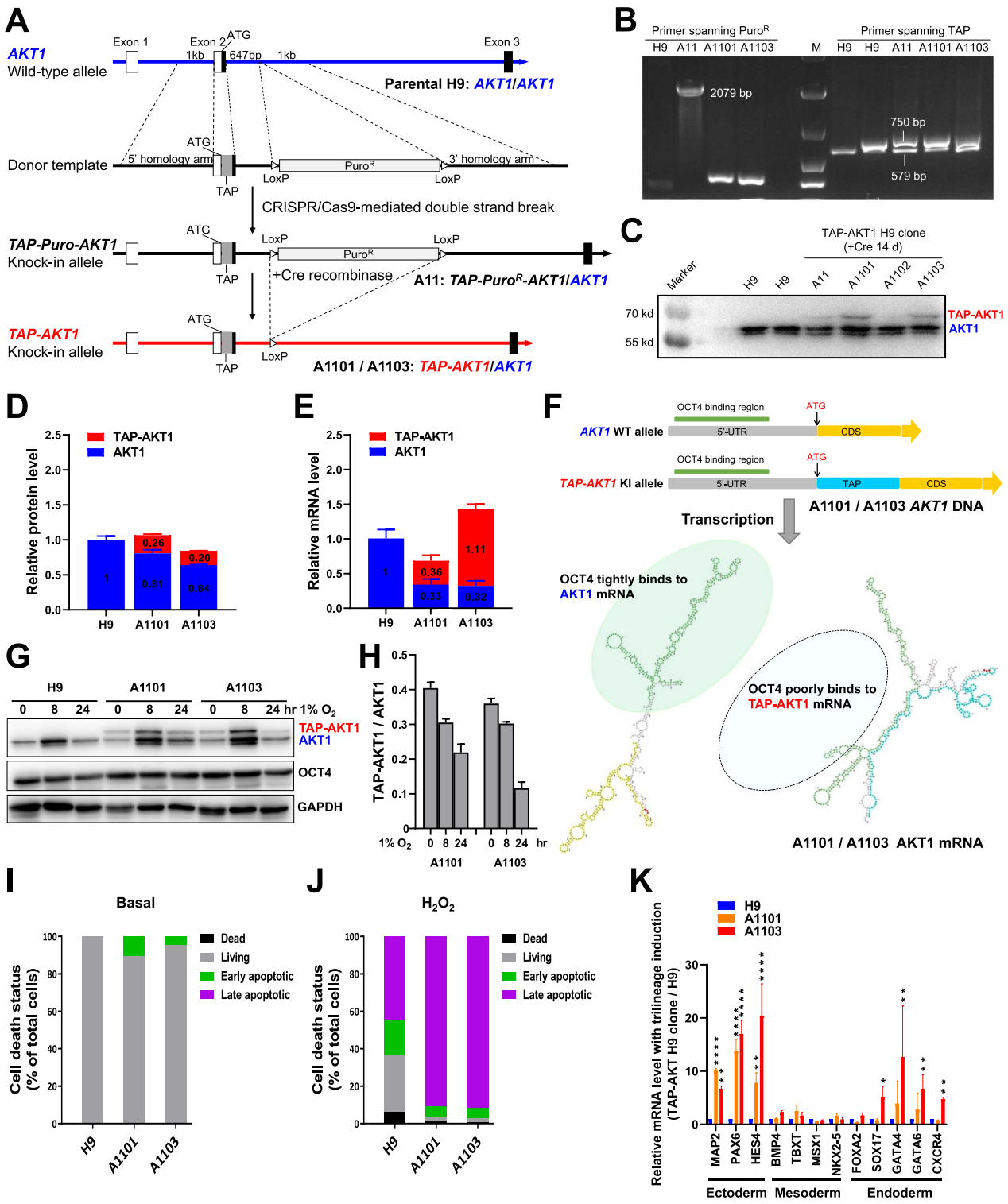


Fig. 8 (See legend on previous page.)

To evaluate the impact of reduced binding of OCT4 to TAP-AKT1 mRNA 5′-UTR and reduced AKT1 translation on cell survival in response to oxidative stress, a common intrinsic stress for PSCs [50, 51], we determined the cell death status of H9, A1101 and A1103 cells grown at routine normoxic (basal) conditions (Fig. 8I,

Supplementary file 6: Fig. S26) or treated with 1000 μM H_2O_2 for 2 h (Fig. 8J, Supplementary file 6: Fig. S26). Under basal conditions, there were significantly more early apoptotic cells in the two heterozygotic TAP-AKT1 clones than in H9 cells (Fig. 8I, Supplementary file 6: Fig. S26), and H_2O_2 treatment led to significantly elevated proportions of late apoptotic cells in the TAP-AKT1 clones (Fig. 8J, Supplementary file 6: Fig. S26). Thus, it was evident that the heterozygotic TAP-AKT1 H9 clones were more susceptible to oxidative stress than parental H9 cells.

We went on further to examine if the two heterozygotic TAP-AKT1 H9 clones may have altered potential of differentiation toward three germ layers. When normoxic parental H9 cells cultured in mTeSR™1 media (lineage inducer -) were switched to STEMdiff™ Trilineage Ectoderm, Mesoderm, or Endoderm Medium (lineage inducer +) and cultured for five to seven days, the mRNA levels of all the trilineage markers were significantly elevated (Supplementary file 6: Fig. S27), indicating successful lineage-specific differentiation. Interestingly, after normalization by the trilineage marker mRNA levels in parental H9 cells, the ectoderm markers in both TAP-AKT1 clones and the endoderm markers in clone A1103 were found to increase dramatically (Fig. 8K). In contrast, the mesoderm markers in both TAP-AKT1 clones remained unchanged (Fig. 8K). Therefore, the heterozygotic TAP-AKT1 H9 clones were more inclined to differentiate into ectoderm, and to a lesser extent, endoderm, but not mesoderm.

Discussion

Despite numerous studies addressing the molecular mechanisms by which PSCs maintain self-renewal and pluripotency under normal culture conditions [52–55], the fundamental question of how PSCs manage to survive stressful conditions remains largely unresolved [21]. We revealed in this study that OCT4 is a key anti-stress RBP that regulates mRNA translation under stress conditions to promote PSC survival and self-renewal. Using all the HITS-CLIP/CLIP-Seq tags obtained from OCT4 binding sites, unbiased search for motifs enriched in OCT4 binding sites revealed over-represented two GC-rich hexamers (5′-CTGCCG-3′) and (5′-GAGCCG-3′) sequences that are predominantly present at the 5′-UTR, supporting the potential roles of OCT4 in 5′-UTR-mediated translation initiation. Ontology analysis of the OCT4 sense binding targets revealed significantly enriched terms in RNA metabolism, translation, various stimuli and stress responses, suggesting that OCT4 may have profound regulatory roles in stress-related mRNA metabolism and translation. Such notion was well corroborated by subsequent combined RNC-seq and RNA-seq analyses using OCT4 siRNA-treated hESCs under hypoxia versus normoxia.

OCT4-bound RNAs were enriched in several pathways by KEGG pathway analysis. These included Focal adhesion, PI3K/AKT signaling pathway and PPAR signaling pathway. Consistent with its role in AKT regulation, OCT4 binds to many transcripts of the PI3K/AKT-pathway genes, including PIK3R2, PIK3CD, PDPK1 and YWHAQ. We further showed that OCT4 binds to several PI3K/AKT-pathway genes' sense transcripts around the 5′-UTR, suggesting that it may coordinately affect the protein levels of a group of the PI3K/AKT-pathway genes by commonly controlling their cap-independent translation initiation. Again, the combined RNC-seq and RNA-seq analyses further supported a crucial role of OCT4 in promoting mRNA translation of PI3K/AKT1-pathway genes in stressed hESCs. Although it remains to be rigorously determined if OCT4 is an ITAF, we propose that OCT4, by binding to the GC-rich motifs, can facilitate the IRES-mediated cap-independent translation of many cell survival genes represented by the PI3K/AKT-pathway genes. A previous study showed that p53, a downstream target gene of AKT1, can inhibit the expression of the rRNA methyltransferase fibrillarin, altering the rRNA methylation pattern, producing ribosomes that can preferentially recognize IRES structures, and thereby promoting the IRES-mediated translation initiation [56]. This finding, together with our current finding, indicates that AKT1 may be a key signal node to coordinate the IRES translation initiation for hESC survival and self-renewal.

It is well established that m7G cap-dependent ribosome scanning is a classic form of translation initiation in eukaryotes [57]. Activation of PI3K/AKT pathway, via the mTOR/4EBP phosphorylation cascade, increases the availability of EIF4E that is the rate-limiting factor for cap-dependent translation [58, 59]. Therefore, AKT1 may promote both cap-dependent and cap-independent translation initiation pathways via phosphorylating over a hundred substrate proteins [60]. When the classical translation initiation mechanism is blocked under stressed or certain conditions, some mRNAs will translate through the cap-independent mechanism to maintain the efficiency of protein synthesis and ensure the survival and recovery of the cells [11, 23, 45, 61]. Among multiple known cap-independent translation initiation pathways, the IRES pathway is the most related to stress response and embryonic development [62, 63]. The IRES structures are mainly located at the 5′-UTR of an mRNA. During the initiation of cap-independent translation, when IRESs interact with ribosomal subunits, specific ITAFs are usually required for stabilizing or remodeling the IRES structures [45, 64] and OCT4 could well serve

as an ITAF in stressed PSCs. It was found that a substantial fraction of the IRES-containing mRNAs encode transcription factors, transporters, receptors, growth factors, apoptosis-related proteins, heat shock proteins, cyclins and translation factors [62], and most ITAF-modulated mRNAs encode key regulatory factors for cell adaptation and survival [62], consistent with the primary roles of AKT1 and other PI3K/AKT-pathway genes [65]. Such notion was also supported by the OCT4-protein interactome in hESC whole cell lysates that identified the 40S ribosomal components and cap-independent and/or cap-dependent but EIF4F-independent translation initiation factors such as EIF3 and EEF1A2, but not typical cap-dependent EIF4F-driven translation initiation factors such as EIF4E [66]. Furthermore, identification of the classical ITAF, HNRNPA1, as an OCT4-interacting protein raised the possibility that several ITAFs comprising OCT4/HNRNPA1 may work in concert to initiate the IRES-mediated cap-independent translation process. Three main factors are known to affect the functions of ITAFs: nucleocytoplasmic shuttling, post-translational modifications and the composition of ITAF complex [62]. Based on its known subcellular localization and transcriptional roles, OCT4 is most likely a nuclear-cytoplasmic shuttling ITAF. There have been a few studies reporting that a TF could also serve as an ITAF [67], and our present work indicates OCT4 may also serve such dual roles- both as a TF controlling gene transcription and as an ITAF regulating mRNA translation. In the case of AKT1 regulation in PSCs, OCT4 suppresses its transcription while promoting its translation. It is conceivable that under normal conditions, OCT4 binds to the promoter of *AKT1* as a TF, restricting its transcription to keep the total AKT1 protein at a relatively low level, thereby slowing down cell proliferation and enabling the cell quality control system to ensure the genome stability of PSCs. Under certain stress conditions (such as hypoxia, ROS, starvation), OCT4 pre-bound at the AKT1 5′-UTR IRES-like structures may work together with other ITAFs and EIF3, mediating the recruitment of the 40S ribosomal subunit and translation initiation (Fig. 7H), and thereby rapidly raising the total AKT1 protein level that may promote PSC survival and self-renewal.

By using the Heterozygous Knocking In N-terminal Tags (HKINT) approach, we simultaneously achieved disrupting the 5′-UTR secondary structure and tagging its protein product of the mRNA from one allele while leaving that from the other allele unedited. One of the main advantages of the HKINT approach is that the two sets of mRNA molecules transcribed from the two alleles of a specific gene and their translational products/processes can be compared within the same cells that have otherwise identical genetic background and cellular contexts. The introduction of the TAP tag into one allele allows densitometric comparison of the TAP-tagged protein of interest with its unedited native protein on immunoblots, and thereby assessing the effect of disrupting the mRNA 5′-UTR secondary structure on its translation level. By multiomics comparison of the protein/RNA binding partners of the TAP knock-in mRNA versus its native mRNA, the HKINT system can also serve as a useful tool to unveil the regulatory machinery of the 5′-UTR of a specific gene of interest. Furthermore, as we have done for AKT1 in this work, the HKINT system allows for assessing the biological consequences of disrupting a specific mRNA 5′-UTR structure and hence its binding with specific RBPs. With this system, we discovered that disrupting the AKT1 mRNA 5′-UTR structure and hence the OCT4/5′-UTR interaction significantly reduced its translation level that led to a higher susceptibility of hESCs to oxidative stress-induced apoptotic death and prioritized differentiation toward ectoderm and endoderm. A previous study showed that excessive intracellular ROS can trigger MAPK-ERK1/2 signaling-mediated neural ectoderm differentiation of hPSCs [51], raising an interesting possibility that OCT4-mediated AKT1 translation may suppress ectoderm/endoderm differentiation under oxidative stress conditions.

Based on our previous estimation [68], there are approximately 7×10^5 OCT4 protein molecules in each pluripotent stem cell. Thus, 7×10^4 OCT4 protein molecules in the cytoplasm can sufficiently contribute to post-transcriptional regulation in hESCs. In fact, there is clear evidence that OCT4 protein can shuttle between cytoplasm and nucleus in ESCs [69] and ECCs [9]. Compared to its canonical TF role that usually takes longer time to function, OCT4 residing in the cytoplasm may rapidly bind to certain mRNAs in response to specific stress signals and can thereby serve as a fast-acting anti-stress RBP. Interestingly, Fukuda et al. found that OCT4 is mainly localized at the cytoplasm in early preimplantation phases, with no major nuclear localization until the 8 to 16-cell stage [70]. It raises a possibility that OCT4 could mainly function as an RBP in very early developmental stage, and its TF role takes the center stage at a later point.

Although this work is centered on mRNA 5′-UTR regulation, the binding of OCT4 to 3′-UTR, CDS, and noncoding RNAs revealed here warrants further investigation given the evidence that there is reciprocal regulation between OCT4 and long noncoding RNAs [71]. Future work is required to dissect out the relative contribution of OCT4 as an RBP/ITAF versus as a TF to establishing and maintaining pluripotency for normal PSCs, adult stem cells, as well as malignant cancer stem cells in vitro and in vivo. Also, it is important to

elucidate the mechanisms by which the RBP/ITAF and TF roles of OCT4 are connected and coordinated during embryogenesis, lineage-specific differentiation, and more broadly, cell fate conversion. In addition, it remains to be understood how genes transcriptionally and post-transcriptionally controlled by OCT4 (such as AKT1) remodel their expression mechanisms when OCT4 expression is virtually lost during differentiation.

Conclusions

In conclusion, we revealed that OCT4 is a novel anti-stress RBP and a probable ITAF promoting IRES-mediated translation initiation of the PI3K/AKT-pathway genes vital for PSC self-renewal and survival under stressed conditions.

Data and code availability

CLIP and RNC/RNA sequencing data reported in this study have been submitted to the NCBI Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE154655 and GSE186550. Mass spectrometry proteomics data have been submitted to ProteomeXchange via the PRIDE database under accession number PXD033840.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13287-025-04229-1.

Additional file1
Additional file2
Additional file3
Additional file4
Additional file5
Additional file6
Additional file7

Acknowledgements
We are grateful to Drs. Baojian Liao and Guangjin Pan (Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences) for technical assistance in generating the TAP-OCT4 H1 cell line, and Wei Tang (ABLife, Inc.) for coordinating the HITS-CLIP/CLIP-Seq assays. This work was supported by grants from the National Natural Science Foundation of China (grant numbers 32170739, 31601103, 32100597, 31401104), the National Key Research and Development Program of China (grant numbers 2016YFA0100303, 2016YFA0101201, 2018YFC1004800), the Natural Science Foundation of Hunan Province, China (2021JJ40899), a grant from Independent Task of the State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, School of Medicine, Zhejiang University, and intramural grants from the First Affiliated Hospital, School of Medicine, Zhejiang University (grant numbers b2019, b3444).

Author contributions
WC, BK, SZ and YJW conceived and designed the study; WC, XC, CC, SS, XL, LS, XZ, SD, YW and WW carried out the experiments; WC, BK and YJW wrote the manuscript; WC, BK, YWZ, QC, JH, YW, YX and YZ analyzed the data and performed the statistical analysis; WC, BK, YZ and YJW contributed to review and editing of the manuscript. YJW, SZ, YWZ and BK contributed to financial support. All the authors reviewed and approved the final version of the manuscript.

Declarations

Ethics approval and consent to participate
Human (h) ESC line H1 (SCSP-301) cells (https://www.cellbank.org.cn/search-detail.php?id=483) and H9 (SCSP-302) cells (https://www.cellbank.org.cn/search-detail.php?id=484) were acquired from the Cell Bank of the Chinese Academic of Sciences, Shanghai, China, under Material Transfer Agreements. The Cell Bank obtained prior ethical approval from WiCell Research Institute for the collection of hESC lines. The handling, storage and preparation of those cells was conducted following the protocols provided by the Cell Bank. All experiments using the obtained hESC lines in this study were approved by the Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University, and were performed according to institutional guidelines. Title of the approved project: Endogenous OCT4 promoter-driven reporter system to study self-renewal of pluripotent and cancer stem cells. Name of the institutional approval committee or unit: Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University. Approval number: 2019–377. Date of approval: February 8, 2019.

Consent for publication
Not applicable.

Competing interest
The authors declare that they have no conflict of interest.

Author details
[1]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, Zhejiang, China. [2]Department of Obstetrics and Gynecology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou 310016, Zhejiang, China. [3]Shaoxing People's Hospital, Shaoxing Hospital, Zhejiang University School of Medicine, Shaoxing 312000, Zhejiang, China. [4]Zhejiang Museum of Natural History, Hangzhou 310014, Zhejiang, China. [5]Department of Colorectal SurgerySir Run Run Shaw Hospital,, School of Medicine, Zhejiang University, Hangzhou 310016, Zhejiang, China. [6]Department of Infectious Diseases, the Second Xiangya Hospital, Central South University, Changsha 410011, Hunan, China. [7]School of Medicine, Zhejiang University, Hangzhou 310058, Zhejiang, China. [8]Shanghai Bioprofile Technology Co., Ltd., Shanghai 200241, China. [9]Center for Genome Analysis, ABLife Inc., Wuhan 430075, Hubei, China. [10]Cancer Center, Zhejiang University, Hangzhou 310058, Zhejiang, China.

Received: 6 August 2024 Accepted: 11 February 2025
Published online: 23 February 2025

References

1. Trounson A, DeWitt ND. Pluripotent stem cells progressing to the clinic. Nat Rev Mol Cell Biol. 2016;17(3):194–200.
2. Yu JSL, Cui W. Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. Development. 2016;143(17):3050–60.
3. Zhao H, Jin Y. Signaling networks in the control of pluripotency. Curr Opin Genet Dev. 2017;46:141–8.
4. Godoy-Parejo C, Deng C, Liu W, Chen G. Insulin stimulates PI3K/AKT and cell adhesion to promote the survival of individualized human embryonic stem cells. Stem Cells. 2019;37(8):1030–41.
5. Singh AM, Reynolds D, Cliff T, Ohtsuka S, Mattheyses AL, Sun Y, et al. Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. Cell Stem Cell. 2012;10(3):312–26.
6. Yu JS, Ramasamy TS, Murphy N, Holt MK, Czapiewski R, Wei SK, et al. PI3K/mTORC2 regulates TGF-beta/Activin signalling by modulating Smad2/3 activity via linker phosphorylation. Nat Commun. 2015;6:7212.

7. Magalhaes-Novais S, Bermejo-Millo JC, Loureiro R, Mesquita KA, Domingues MR, Maciel E, et al. Cell quality control mechanisms maintain stemness and differentiation potential of P19 embryonic carcinoma cells. Autophagy. 2020;16(2):313–33.
8. Wu X, Fleming A, Ricketts T, Pavel M, Virgin H, Menzies FM, et al. Autophagy regulates Notch degradation and modulates stem cell development and neurogenesis. Nat Commun. 2016;7:10533.
9. Lin Y, Yang Y, Li W, Chen Q, Li J, Pan X, et al. Reciprocal regulation of Akt and Oct4 promotes the self-renewal and survival of embryonal carcinoma cells. Mol Cell. 2012;48(4):627–40.
10. Corbett AH. Post-transcriptional regulation of gene expression and human disease. Curr Opin Cell Biol. 2018;52:96–104.
11. Chua BA, Van Der Werf I, Jamieson C, Signer RAJ. Post-transcriptional regulation of homeostatic, stressed, and malignant stem cells. Cell Stem Cell. 2020;26(2):138–59.
12. Sarshad AA, Juan AH, Muler AIC, Anastasakis DG, Wang X, Genzor P, et al. Argonaute-miRNA complexes silence target mRNAs in the nucleus of mammalian stem cells. Mol Cell. 2018;71(6):1040-50.e8.
13. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol. 2018;19(5):327–41.
14. Ainaoui N, Hantelys F, Renaud-Gabardos E, Bunel M, Lopez F, Pujol F, et al. Promoter-dependent translation controlled by p54nrb and hnRNPM during myoblast differentiation. PLoS ONE. 2015;10(9):e0136466.
15. Holmes ZE, Hamilton DJ, Hwang T, Parsonnet NV, Rinn JL, Wuttke DS, et al. The Sox2 transcription factor binds RNA. Nat Commun. 2020;11(1):1805.
16. Hou L, Wei Y, Lin Y, Wang X, Lai Y, Yin M, et al. Concurrent binding to DNA and RNA facilitates the pluripotency reprogramming activity of Sox2. Nucleic Acids Res. 2020;48(7):3869–87.
17. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell. 2012;149(6):1393–406.
18. Dvir S, Argoetti A, Lesnik C, Roytblat M, Shriki K, Amit M, et al. Uncovering the RNA-binding protein landscape in the pluripotency network of human embryonic stem cells. Cell Rep. 2021;35(9):109198.
19. He C, Sidoli S, Warneford-Thomson R, Tatomer DC, Wilusz JE, Garcia BA, et al. High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. Mol Cell. 2016;64(2):416–30.
20. Tower J. Stress and stem cells. Wiley Interdiscip Rev Dev Biol. 2012;1(6):789–802.
21. Darwish T, Swaidan NT, Emara MM. Stress factors as possible regulators of pluripotent stem cell survival and differentiation. Biology (Basel). 2023;12(8):1119.
22. Cheng J, Li W, Kang B, Zhou Y, Song J, Dan S, et al. Tryptophan derivatives regulate the transcription of Oct4 in stem-like cancer cells. Nat Commun. 2015;6:7209.
23. Chen C, Zhang X, Wang Y, Chen X, Chen W, Dan S, et al. Translational and post-translational control of human naive versus primed pluripotency. iScience. 2022;25(1):103645.
24. Chen W, Chen X, Zhang X, Chen C, Dan S, Hu J, et al. DNA repair proteins cooperate with SOX2 in regulating the transition of human embryonic stem cells to neural progenitor cells. Biochem Biophys Res Commun. 2022;586:163–70.
25. Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, et al. Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol Cell. 2009;36(6):996–1006.
26. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.
27. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202–8.
28. Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell. 2010;6(4):382–95.
29. Brumbaugh J, Hou Z, Russell JD, Howden SE, Yu P, Ledvina AR, et al. Phosphorylation regulates human OCT4. Proc Natl Acad Sci USA. 2012;109(19):7162–8.
30. Chen C, Liu Q, Chen W, Gong Z, Kang B, Sui M, et al. PRODH safeguards human naive pluripotency by limiting mitochondrial oxidative phosphorylation and reactive oxygen species production. EMBO Rep. 2024;25:2015–44.
31. Lian X, Guo J, Gu W, Cui Y, Zhong J, Jin J, et al. Genome-wide and experimental resolution of relative translation elongation speed at individual gene level in human cells. PLoS Genet. 2016;12(2):e1005901.
32. Basso AD, Solit DB, Chiosis G, Giri B, Tsichlis P, Rosen N. Akt forms an intracellular complex with heat shock protein 90 (Hsp90) and Cdc37 and is destabilized by inhibitors of Hsp90 function. J Biol Chem. 2002;277(42):39858–66.
33. Ding J, Xu H, Faiola F, Ma'ayan A, Wang J. Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res. 2012;22(1):155–67.
34. van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, et al. An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell. 2010;6(4):369–81.
35. Thakor N, Smith MD, Roberts L, Faye MD, Patel H, Wieden HJ, et al. Cellular mRNA recruits the ribosome via eIF3-PABP bridge to initiate internal translation. RNA Biol. 2017;14(5):553–67.
36. Ramat A, Garcia-Silva MR, Jahan C, Nait-Saidi R, Dufourt J, Garret C, et al. The PIWI protein Aubergine recruits eIF3 to activate translation in the germ plasm. Cell Res. 2020;30(5):421–35.
37. Lee AS, Kranzusch PJ, Doudna JA, Cate JH. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. Nature. 2016;536(7614):96–9.
38. Guan BJ, van Hoef V, Jobava R, Elroy-Stein O, Valasek LS, Cargnello M, et al. A unique ISR program determines cellular responses to chronic stress. Mol Cell. 2017;68(5):885-900 e6.
39. de la Parra C, Ernlund A, Alard A, Ruggles K, Ueberheide B, Schneider RJ. A widespread alternate form of cap-dependent mRNA translation initiation. Nat Commun. 2018;9(1):3068.
40. Jo OD, Martin J, Bernath A, Masri J, Lichtenstein A, Gera J. Heterogeneous nuclear ribonucleoprotein A1 regulates cyclin D1 and c-myc internal ribosome entry site function through Akt signaling. J Biol Chem. 2008;283(34):23274–87.
41. Bruun GH, Doktor TK, Borch-Jensen J, Masuda A, Krainer AR, Ohno K, et al. Global identification of hnRNP A1 binding sites for SSO-based splicing modulation. BMC Biol. 2016;14:54.
42. Lee AS, Kranzusch PJ, Cate JH. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. Nature. 2015;522(7554):111–4.
43. Kaitsuka T, Hakim F. Response of pluripotent stem cells to environmental stress and its application for directed differentiation. Biology (Basel). 2021;10(2):84.
44. Weingarten-Gabbay S, Elias-Kirma S, Nir R, Gritsenko AA, Stern-Ginossar N, Yakhini Z, et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science. 2016;351(6270):aad4939.
45. Leppek K, Das R, Barna M. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat Rev Mol Cell Biol. 2018;19(3):158–74.
46. Lang KJ, Kappel A, Goodall GJ. Hypoxia-inducible factor-1alpha mRNA contains an internal ribosome entry site that allows efficient translation during normoxia and hypoxia. Mol Biol Cell. 2002;13(5):1792–801.
47. Conte C, Riant E, Toutain C, Pujol F, Arnal JF, Lenfant F, et al. FGF2 translationally induced by hypoxia is involved in negative and positive feedback loops with HIF-1alpha. PLoS ONE. 2008;3(8):e3078.
48. Braunstein S, Karpisheva K, Pola C, Goldberg J, Hochman T, Yee H, et al. A hypoxia-controlled cap-dependent to cap-independent translation switch in breast cancer. Mol Cell. 2007;28(3):501–12.
49. Cencic R, Hall DR, Robert F, Du Y, Min J, Li L, et al. Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. Proc Natl Acad Sci USA. 2011;108(3):1046–51.
50. Cao H, Zhi Y, Xu H, Fang H, Jia X. Zearalenone causes embryotoxicity and induces oxidative stress and apoptosis in differentiated human embryonic stem cells. Toxicol In Vitro. 2019;54:243–50.
51. Hu Q, Khanna P, Ee Wong BS, Lin Heng ZS, Subhramanyam CS, Thanga LZ, et al. Oxidative stress promotes exit from the stem cell state and spontaneous neuronal differentiation. Oncotarget. 2018;9(3):4223–38.
52. Liu G, David BT, Trawczynski M, Fessler RG. Advances in pluripotent stem cells: history, mechanisms, technologies, and applications. Stem Cell Rev Rep. 2020;16(1):3–32.

53. Yilmaz A, Benvenisty N. Defining human pluripotency. Cell Stem Cell. 2019;25(1):9–22.
54. De Los AA, Ferrari F, Xi R, Fujiwara Y, Benvenisty N, Deng H, et al. Hallmarks of pluripotency. Nature. 2015;525(7570):469–78.
55. Smith A. Propagating pluripotency—the conundrum of self-renewal. BioEssays. 2024;46:e2400108.
56. Marcel V, Ghayad SE, Belin S, Therizols G, Morel AP, Solano-Gonzalez E, et al. p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. Cancer Cell. 2013;24(3):318–30.
57. Merrick WC, Pavitt GD. Protein synthesis initiation in eukaryotic cells. Cold Spring Harbor Perspectives Biol. 2018;10(12):a033092.
58. Roux PP, Topisirovic I. Signaling pathways involved in the regulation of mRNA Translation. Mol Cell Biol. 2018;38(12).
59. Hsieh AC, Truitt ML, Ruggero D. Oncogenic AKTivation of translation as a therapeutic target. Br J Cancer. 2011;105(3):329–36.
60. Manning BD, Toker A. AKT/PKB signaling: navigating the network. Cell. 2017;169(3):381–405.
61. Shatsky IN, Terenin IM, Smirnova VV, Andreev DE. Cap-independent translation: What's in a name? Trends Biochem Sci. 2018;43(11):882–95.
62. Godet AC, David F, Hantelys F, Tatin F, Lacazette E, Garmy-Susini B, et al. IRES trans-acting factors, key actors of the stress response. Int J Mol Sci. 2019;20(4):924.
63. Xue S, Tian S, Fujii K, Kladwang W, Das R, Barna M. RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. Nature. 2015;517(7532):33–8.
64. Komar AA, Hatzoglou M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle. 2011;10(2):229–40.
65. Chu N, Salguero AL, Liu AZ, Chen Z, Dempsey DR, Ficarro SB, et al. Akt kinase activation mechanisms revealed using protein semisynthesis. Cell. 2018;174(4):897-907.e14.
66. Jia X, He X, Huang C, Li J, Dong Z, Liu K. Protein translation: biological processes and therapeutic strategies for human diseases. Signal Transduct Target Ther. 2024;9(1):44.
67. Conte C, Ainaoui N, Delluc-Clavieres A, Khoury MP, Azar R, Pujol F, et al. Fibroblast growth factor 1 induced during myogenesis by a transcription-translation coupling mechanism. Nucleic Acids Res. 2009;37(16):5267–78.
68. Zhou Y, Chen X, Kang B, She S, Zhang X, Chen C, et al. Endogenous authentic OCT4A proteins directly regulate FOS/AP-1 transcription in somatic cancer cells. Cell Death Dis. 2018;9(6):585.
69. Oka M, Moriyama T, Asally M, Kawakami K, Yoneda Y. Differential role for transcription factor Oct4 nucleocytoplasmic dynamics in somatic cell reprogramming and self-renewal of embryonic stem cells. J Biol Chem. 2013;288(21):15085–97.
70. Fukuda A, Mitani A, Miyashita T, Kobayashi H, Umezawa A, Akutsu H. Spatiotemporal dynamics of OCT4 protein localization during preimplantation development in mice. Reproduction. 2016;152(5):417–30.
71. Zhou RT, Ni YR, Zeng FJ. The roles of long noncoding RNAs in the regulation of OCT4 expression. Stem Cell Res Ther. 2022;13(1):383.

Publisher's Note