

Journal Pre-proof



Translational and Post-translational Control of Human Naïve versus Primed Pluripotency

Cheng Chen, Xiaobing Zhang, Yisha Wang, Xinyu Chen, Wenjie Chen, Songsong Dan, Shiqi She, Weiwei Hu, Jie Dai, Jianwen Hu, Qingyi Cao, Qianyu Liu, Yinghua Huang, Baoming Qin, Bo Kang, Ying-Jie Wang

PII: S2589-0042(21)01615-1

DOI: <https://doi.org/10.1016/j.isci.2021.103645>

Reference: ISCI 103645

To appear in: *ISCIENCE*

Received Date: 28 October 2020

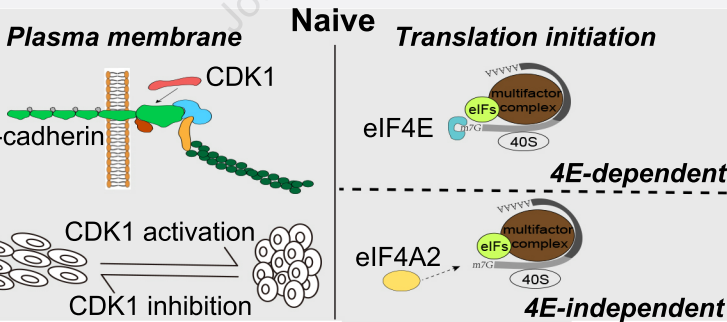
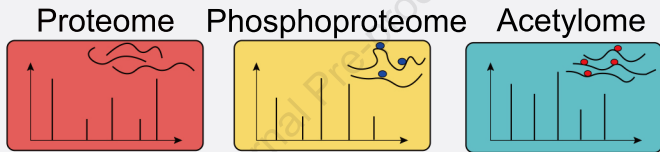
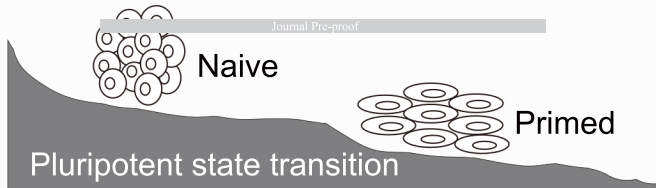
Revised Date: 22 October 2021

Accepted Date: 10 December 2021

Please cite this article as: Chen, C., Zhang, X., Wang, Y., Chen, X., Chen, W., Dan, S., She, S., Hu, W., Dai, J., Hu, J., Cao, Q., Liu, Q., Huang, Y., Qin, B., Kang, B., Wang, Y.-J., Translational and Post-translational Control of Human Naïve versus Primed Pluripotency, *ISCIENCE* (2022), doi: <https://doi.org/10.1016/j.isci.2021.103645>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 The Author(s).



Translational and Post-translational Control of Human Naïve versus Primed Pluripotency

Cheng Chen,^{1,2,8} Xiaobing Zhang,^{1,8} Yisha Wang,¹ Xinyu Chen,¹ Wenjie Chen,¹ Songsong Dan,¹ Shiqi She,^{1,3} Weiwei Hu,⁴ Jie Dai,⁴ Jianwen Hu,⁴ Qingyi Cao,¹ Qianyu Liu,⁵ Yinghua Huang,⁶ Baoming Qin,⁶ Bo Kang,^{1,*} and Ying-Jie Wang^{1,7,9*}

¹State 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, Zhejiang 310003, China

²Shaoxing People's Hospital; Shaoxing Hospital, Zhejiang University School of Medicine, Shaoxing, Zhejiang 312000, China

³Zhejiang Museum of Natural History, Hangzhou, Zhejiang 310014, China

⁴Shanghai Bioprofile Technology Co., Ltd., Shanghai 200241, China

⁵College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China

⁶Laboratory of Metabolism and Cell Fate, Center for Cell Lineage and Development, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

⁷Cancer Center, Zhejiang University, Hangzhou, Zhejiang 310058, China

⁸These authors contributed equally

⁹Lead Contact

* Correspondence: yingjiewang@zju.edu.cn (YJW), kangbo@zju.edu.cn (BK)

SUMMARY

Deciphering the regulatory network for human naïve and primed pluripotency is of fundamental theoretical and applicable significance. Here, by combining quantitative proteomics, phosphoproteomics and acetylproteomics analyses, we revealed RNA processing and translation as the most differentially-regulated processes between naïve and primed human embryonic stem cells (hESCs). While glycolytic primed hESCs rely predominantly on eIF4E-mediated cap-dependent pathway for protein translation, naïve hESCs with reduced mTORC1 activity are more tolerant to eIF4E inhibition, and their bivalent metabolism allows for translating selective mRNAs via both eIF4E-dependent and eIF4E-independent/eIF4A2-dependent pathways to form a more compact naïve proteome. Globally up-regulated proteostasis and down-regulated post-translational modifications help to further refine the naïve proteome that is compatible with the more rapid cycling of naïve hESCs, where CDK1 plays an indispensable coordinative role. These findings may assist in better understanding the unrestricted lineage potential of naïve hESCs and in further optimizing conditions for future clinical applications.

INTRODUCTION

The successful isolation of inner cell mass (ICM) cells from blastocysts and propagating them in culture as embryonic stem cells (ESCs) offers exciting opportunities for studying early developmental mechanisms, drug screening, disease modeling, and cell-replacement therapies (Boroviak and Nichols, 2014; Ilic and Ogilvie, 2017). Unlike mouse ESCs (mESCs) that can be maintained in a naive ICM-like state in the presence of LIF (leukemia inhibitory factor) and 2i (MEK and GSK3 inhibitors) in vitro (Martello and Smith, 2014), human ESCs (hESCs) cultured in conventional conditions are in a so-called “primed” pluripotency that resembles the post-implantation epiblasts (Niakan et al., 2012). Naïve ESCs are characterized by dome-shaped colonies, dependence on JAK/STAT signaling, and increased homogeneity and single-cell survival while primed ESCs display flat colony morphology, dependence on TGF β /activin signaling and low single-cell clonogenicity (Nichols and Smith, 2009; Wu and Izpisua Belmonte, 2016).

To convert primed hESCs back to a naïve state represents an appealing research endeavor that is of both theoretical and applicable importance. The naïve hESCs have multiple advantageous features: they offer a unique model to study human preimplantation biology including the regulation of lineage decisions, responsiveness to signaling changes, genetic and epigenetic alterations induced by environmental cues; compared to primed hESCs, naïve hESCs have higher homogeneity, pluripotency and self-renewal capacity that are expected to lead to more efficient directed differentiation towards germ layer derivatives (Collier and Rugg-Gunn, 2018; Wu and Izpisua Belmonte, 2016); in addition, naïve hESCs also provide an opportunity for interspecies blastocyst complementation to produce transplantable patient-specific organs in large animal hosts (De Los Angeles et al., 2018). More excitingly, recent work from several laboratories revealed the capabilities of naïve hESCs in generating trophectoderm and descendant trophoblast cell types and in forming three-dimensional blastocyst-like structures (Guo et al., 2021; Io et al., 2021; Yanagida et al., 2021), which will have a transformative impact on future human reproductive research.

Over the past years, by combining 2i/LIF with different components, several groups have formulated protocols to induce naive pluripotency in hESC, both from existing primed hESCs and by direct derivation from the blastocysts (Collier and Rugg-Gunn, 2018; Theunissen et al., 2014; Warriar et al., 2017). Although naïve-like hESCs derived from different induction conditions differ in some properties (Liu et al., 2017; Warriar *et al.*, 2017), they generally share several key features: they display global DNA hypomethylation compared to primed hESCs; they have reduced levels of the repressive histone mark H3K27me3 at >3000 Polycomb-associated gene promoter; they exhibit preferential activity of the OCT4 distal enhancer, whereas primed hESCs use the proximal enhancer; female naïve-like hESCs have two active X-chromosomes with localized depletion of the H3K27me3 mark (Collier and Rugg-Gunn, 2018; Theunissen et al., 2016). Besides, several genes or cell surface proteins are considered as relatively specific biomarkers for naïve-like hESCs (Taei et al., 2020; Trusler et al., 2018; Weinberger et al., 2016). So far, most naïve-like hESCs are induced in the presence of feeder cells that usually are of murine origin (Collier and Rugg-Gunn, 2018) and the partially feeder-dependent PXGL hESC induction system derived from the t2iLGö system (Bredenkamp et al., 2019) is considered as one of the best systems that can induce the so-called ‘bona fide naïve hESCs’. However, the presence of non-human factors in the culture

may limit the use of hESCs for therapeutic purposes. Recent improvements in hESC culture have enabled the commercial development of completely defined, feeder-free culture conditions such as mTeSR1 for primed hESCs, and RSeT™ Feeder-Free Medium for naïve hESCs that are highly tractable and have less abnormal karyotypes (Ward *et al.*, 2017), addressing the challenge of delivering the clinical promise of hESCs and ultimately leading to Good Manufacturing Practice (GMP)-compatible services and products. More recently, another feeder-free naïve hESC induction system (FINE) was developed that offered more robust naïve marker expression including 8-cell stage-specific transcripts (Szczerbinska *et al.*, 2019).

To further improve the induction and cultivation conditions for human naïve pluripotency that can faithfully recapitulate the key attributes of ICM cells, it is important to gain deeper insights into the roles of the major signaling and metabolic pathways in the acquisition and maintenance of the different states of pluripotency. Corresponding to 2i, MEK/ERK and GSK3 signaling pathways are well established to be down-regulated in naïve hESCs (Collier and Rugg-Gunn, 2018). Conversely, LIF/JAK/STAT3 (Chen *et al.*, 2015), LPA/YAP (Qin *et al.*, 2016), WNT, TGF- β pathways (Taei *et al.*, 2020) are the main signaling pathways known to promote human naïve pluripotency. These signaling pathways are expected to affect a battery of downstream proteins via multiple post-translational modifications (PTMs) such as phosphorylation, acetylation, ubiquitination, etc. (Wang *et al.*, 2014a), but systematical comparison of PTM profiles between naïve and primed hESCs is lacking. In recent years, there has been a growing appreciation that energy balance and metabolic status are closely associated with stem cell functions, and metabolic reprogramming and regulation has become one of the research focuses in stem cell biology (Harvey *et al.*, 2019; Tsogtbaatar *et al.*, 2020). Although consensus has been reached that the primed hESCs exclusively rely on glycolysis for energy supply and cell proliferation while naïve hESCs utilize a bivalent metabolic system (i.e., both glycolysis and oxidative phosphorylation (OXPHOS)) (Tsogtbaatar *et al.*, 2020), the potential functional roles and specific contributions of distinct metabolic modes in regulating naïve versus primed pluripotency remain largely unknown. In this study, by multi-omics analyses, we identified CDK1 as a crucial kinase for maintaining human naïve pluripotency, and revealed RNA processing and translation as the most differentially-regulated processes between naïve and primed hESCs. We further discovered that under stressed conditions, glycolytic primed hESCs rely predominantly on eIF4E-mediated cap-dependent pathway for protein translation, while bivalent metabolism in naïve hESCs promotes both cap-dependent and eIF4A2-dependent/cap-independent translation that confer them higher capability for rapid cell cycling and propagation.

RESULTS

Induction of Primed hESCs into Naïve-like hESCs

Given its advantages mentioned above, in this study, we mainly employed the RSeT™ Feeder-Free Medium (designated as ‘RSet-ff’) (Ward *et al.*, 2017) to induce naïve-like hESCs, where dome-shaped colonies of both H1 (Figure S1A) and H9 hESCs (Figure 1A) were formed within 5-7 days. Besides morphological features, examination of several cellular and

molecular hallmarks for naïve pluripotency (Collier and Rugg-Gunn, 2018; Tai *et al.*, 2020) further demonstrated our success in acquiring the naïve-like hESCs (Figures 1B and S1B). RSet-ff naïve-like hESCs had a much shorter doubling time (Figures 1C and S1C) and enhanced single-cell clonogenicity (Figures 1D and S1D) than their primed counterparts. Furthermore, immunofluorescence microscopy confirmed a much higher expression level of the naïve pluripotency marker KLF4 in RSet-ff naïve-like cells (Figures 1E and S1E). Consistent with the notion that female naïve hESCs have two active X-chromosomes (XaXa) while female primed hESCs undergo one X chromosome inactivation (XaXi) caused by H3K27me₃, the H3K27me₃ foci-positive nuclei were detected predominantly in primed but not naïve-like H9 hESCs (Figures 1F). In addition, the CpG methylation level at the DNMT3L promotor was reduced by nearly 35% (Figure S1F and S2C), indicating a global DNA hypomethylation state present in RSet-ff naïve-like hESCs.

We also induced naïve-like hESCs via another two approaches. The PXGL hESC induction system (Bredenkamp *et al.*, 2019) (Figure S2A), one of the best systems that can induce the ‘bona fide naïve hESCs’ and further generate blastocyst trophectoderm (Guo *et al.*, 2021), gave rise to characteristic SUSD2+/CD24 low or negative naïve-like hESCs that were unambiguously distinguished from their SUSD2-/CD24+ primed counterparts (Figure 1G). Remarkably, the extent of upregulation of the key naïve genes and pluripotency genes over primed hESCs were rather similar between the PXGL- and RSet-ff-induced naïve-like hESCs (Figure S2D). Likewise, the stabilized (i.e., passage 5 onward) naïve-like hESCs derived from the FINE feeder-free naïve induction system (Szczerbinska *et al.*, 2019) (Figure S2B) exhibited significantly-upregulated expression of naïve genes over primed hESCs (Figure 1H).

Collectively, we established three tractable, reliable and efficient systems to induce naïve-like hESCs (hereafter referred to as naïve hESCs) that can be used for further investigations.

Global Profiles of Proteome, Acetylome and Phosphoproteome in Primed and Naïve hESCs

To gain novel insights into the differential protein translational and post-translational regulatory mechanisms in the two hESC pluripotency states, we used label-free LC-MS/MS approaches to quantify and compare the proteome, acetylome and phosphoproteome of primed and RSet-ff naïve H9 hESCs in biological triplicates. As depicted in Figure 2A, H9 cells were extracted, digested and enriched by antibody beads or TiO₂ for acetylome or phosphoproteome, respectively, then analyzed by LC-MS/MS. In total, we quantified 4210 proteins, 4982 acetylation sites and 11058 phosphosites, with equal amounts of the total protein from the two hESC states being analyzed (Tables S1-S3). Principal-component analysis (PCA) of three omics showed considerable differences between primed and naïve cells that were generally kept consistent in three replicates (Figure S3A). Relative standard deviations (RSDs) of quantitative values between replicates were less than 0.2, suggestive of high reproducibility (Figure S3B).

With respect to more detailed phosphorylation and acetylation patterns, we totally identified 1004 proteins with 1 phosphosite, 519 proteins with 2 phosphosites, 316 proteins with 3 phosphosites, and 228 proteins with 4 phosphosites (Figure S3C). Corresponding acetylation

data were also represented. The dominant phosphorylated amino acids were serine (85%), followed by threonine (14%), and only 0.9% of the phosphor-sites were identified as tyrosine (Figure S3D). The differential proteins and PTM sites were selected according to the following criteria: the modification site localization probability > 0.75 ; the quantification ratio > 1.5 or < 0.66 (naïve versus primed); the p-value < 0.05 .

Multi-omics Analyses Reveal Differences in Proteome, Histone Acetylation and Kinase Activity between Primed and Naïve hESCs

Firstly, we analyzed the proteomic data and sought to identify the most differentially-regulated biological processes between the two pluripotency states. We subjected all the differentially-expressed proteins between the two states to an integrated Molecular Complex Detection (MCODE) analysis and the protein-protein interaction (PPI) network analysis via the STRING database. 6 modules (comprising 84 proteins) within which the protein members are highly interconnected, were identified from global PPI network (Figure 2B). One or two key members in a module mediated connections with other modules. Then, we tried to associate the majority of these modules with known gene annotation terms and physiological processes. For example, the largest module was mainly comprised of proteins participating in ribosome RNA (rRNA) processing. Additionally, a series of function annotations including spliceosome, RNA transport, immune effector process, proteasome and protein folding chaperone were linked to other 5 modules. Of note, 3 out of these 6 modules were associated with RNA-related post-transcriptional processes, namely splicing, transport and translation initiation, and 2 out of them were related to proteostasis, namely proteasome and protein folding chaperone. Further, heatmap and cluster analyses were performed for each single module described above, showing consistent down-regulation of RNA processing-related 3 modules and up-regulation of the other 3 modules in naïve pluripotency (Figure S4A). As a parallel analysis, the global connections among various KEGG pathways for all differential proteins were established (Figure S4B). Among them, three significantly enriched pathways governing translation and energy metabolism, namely RNA transport, ribosome biogenesis in eukaryotes, and glycolysis/gluconeogenesis, were chosen for further analysis, and the PPI network for each chosen pathway was shown in Figure S4C. Thus, various analyses of proteomic data consistently indicate a global repression of RNA processing and translation pathways in human naïve pluripotency. Together with the up-regulated proteostasis machinery, the down-regulated translational machinery is likely to give rise to a reduced total protein level and proteome size in naïve hESCs.

Next, we separately analyzed the acetylome and phosphoproteome data. By using the MoMo software (<http://meme-suite.org/tools/momo>), the position-specific frequencies of ten flanking amino acid residues up- and down-stream of the detected acetylation sites in naïve hESCs were compared. Five acetyl-K motifs were identified and KY was the most conserved one (Figure S4D). Besides histone proteins, acetylation modifications have been identified in numerous non-histone proteins participating in various cellular processes (Narita et al., 2019). In this study, highly acetylated non-histone proteins in naïve state were significantly enriched in pathways concerning cell proliferation. In contrast, less acetylated non-histone proteins were mainly involved in metabolism of amino acid and fatty acid (Figures 2C and S4E). Given the crucial role of histone acetylation as a major epigenetic marker in regulating early

development (Moussaieff et al., 2015; Ware et al., 2009), we summarized the quantitative acetylation data for all histones in Figure 2D. In general, the overall degree of histone acetylation in naïve hESCs was similar to that in primed hESCs. The naïve hESCs had an increase in H3K9Ac, an epigenetic marker representing transcriptionally active chromatin and is involved in pluripotency maintenance (Qiao et al., 2015). The unique H2BK24Ac, as well as the prominently-increased H2BK15Ac and H4K31Ac in naïve hESCs may warrant special attention and further in-depth investigation.

For phosphosites, we identified 6 significantly enriched phosphor-S (pS) motifs and three phosphor-T (pT) motifs in naïve hESCs (Figure S4F). Due to limited number of phosphor-Y (pY, phosphorylated tyrosine) sites, no common motifs can be identified. The frequency analysis for amino acids around the phosphosites revealed the preferred residues following pS11 were P12 and E14. TP is known to be by far the most common pT motif, which was also confirmed in this study. Furthermore, a machine learning algorithm, IKAP (Mischnik et al., 2016), was used to infer the upstream kinases that phosphorylate multiple downstream substrates whose phosphorylation sites and levels were revealed by the quantitative phosphoproteomics after normalization of the phosphopeptide abundance by the protein abundance. Out of the analyzed 170 kinases, 46 differentially-regulated kinases were clustered (Figure 2E). Only 9 kinases were relatively more active in naïve hESCs, including CDK1, CDK2, ATR, PKCG, RSK2, HIPK2, TAK1, MNK1 and SRC. We then constructed a kinase-to-kinase interaction network for all the differentially-regulated kinases, rooted on centrally positioned kinases MAPK1 and GSK3B that are known to be suppressed by 2i in almost all naïve culture system (Takashima et al., 2014; Theunissen *et al.*, 2016) (Figure 2F). Strikingly, a quarter of the clustered kinases had direct or indirect connection to CDK1, a kinase with potentially up-regulated activities in naïve hESCs (Figures 2E and 2F).

In summary, we identified a series of differential acetylation sites on the histones and constructed the kinase-to-kinase interaction network for multiple upstream kinases that may critically shape the phosphoproteome. Remarkably, four members of the CDK family (CDK1, CDK2, CDK4 and CDK6) stood out as potential key regulators for naïve pluripotency.

CDK1 Activity Is Crucial for Maintaining Naïve Pluripotency

Given the above analyses, we next focused on the CDK family in the two pluripotency states. Our pilot experiments using specific small-molecule inhibitors showed that, CDK1 rather than other CDK family members is critical for human naïve pluripotency. CDK1 is an unique mitosis cyclin-dependent kinase (M-CDK) indispensable for early embryonic development (Santamaria et al., 2007). It has been reported that CDK1 is required for maintaining human primed pluripotency (Neganova et al., 2014), yet its potential role in naïve pluripotency has not been documented. As shown in the phosphoproteome data, the phosphorylation levels at known CDK1-phosphorylated sites in some selected CDK1 substrates were higher in RSet-ff naïve hESCs (Figure 3A), implicating a higher CDK1 catalytic activity in naïve pluripotency. This was consistent with the immunoblotting data showing that the level of total CDK1 protein was increased while that of the inhibitory phosphorylation on T14 and Y15 was reduced in RSet-ff naïve hESCs (Figure S5A). Under the PXGL naïve induction conditions, the level of inhibitory phosphorylation on Y15 was reduced and that of activating phosphorylation on T161 was elevated in naïve H9 cells as opposed to their primed

counterparts (Figure 3B), confirming enhanced CDK1 activities in naïve pluripotency. Strikingly, treating RSet-ff naïve hESCs but not their primed counterparts with the highly specific CDK1 inhibitor RO3306 (Jorda et al., 2018), which had no discernable effects on cell viability at a concentration of 5 μ M (Figure S5B), converted their colony morphology from domes to flattened sheets (Figures 3C and S5C). The pluripotency markers (NANOG, SOX2 and OCT4) and naïve markers (STELLA, KLF4 and DNMT3L) were both decreased (Figure S5D-S5F). Meanwhile, examination of germ layer markers revealed the differentiation of RO3306-treated naïve hESCs towards mesendoderm (Figure 3D). Thus, CDK1 is more active in naïve hESCs and essential for maintaining their characteristics.

To investigate if the impairment of naïve pluripotency by CDK1 inhibitor was mediated by the traditional role of CDK1 in G2/M progression, as reported for primed hESCs (Neganova *et al.*, 2014), cell cycle analyses were conducted for RSet-ff-induced naïve H1 and H9 cells with or without RO3306 treatment. At 1 day after RO3306 treatment, a higher proportion of naïve H1 cells were arrested at G2/M phase while naïve H9 cells were mainly S/G2 arrested (Figure S5G), indicating differential roles of CDK1 in cell cycle regulation between naïve H1 and H9 cells. Therefore, it seems unlikely that the cell cycle regulation by CDK1 plays a conservative and major role in maintaining human naïve pluripotency.

The above morphological transition by CDK1 inhibitor treatment drew our attention to the cell-cell adhesion complexes. E-cadherin, a calcium-dependent cell adhesion protein, has a pivotal role in forming the compact colonies of mESCs and in early embryogenesis (Li et al., 2012; Pieters and van Roy, 2014). Remarkably, upon RO3306 treatment, the continuously and cortically-distributed E-cadherin disappeared only in naïve (Figures 3E, S6A) but not primed hESCs (Figure S6B). RO3306 significantly reduced E-cadherin protein levels (Figure 3F) but not its mRNA levels (Figure 3G) in naïve hESCs but not in primed hESCs (Figure S6C), and it also reduced the protein levels of β -catenin and p120-catenin to a similar extent (Figure 3F), indicating that CDK1 may directly regulate E-cadherin/ β -catenin/p120-catenin protein complex via its kinase activity in naïve hESCs. Importantly, either CDK1 knockdown or E-cadherin (CDH1) knockdown impaired establishment of naïve pluripotency, and the trends of change in naïve markers were similar between the two knockdown groups (Figures S6F, S6G). Co-IP experiments demonstrated that there was physical interaction between E-cadherin and CDK1 proteins in naïve hESCs (Figures 3H, S6D). Remarkably, overexpressing E-cadherin (Figure S6E) can rescue RO3306-induced reduction of AP+ colony formation (Figures 3I, 3J) and reduction of pluripotency markers (Figure 3K), and thereby maintaining naïve pluripotency. Taken together, we reveal here that CDK1 critically maintains human naïve pluripotency by sustaining the protein level and cortical localization of E-cadherin.

Integrative Multi-omics Analyses Reveal RNA Processing and Translation as the Most Differentially-regulated Processes between Primed and Naïve hESCs

After separated analyses for each omics as shown above, we next integrated the multi-omics data to find out combined major differences between primed and naïve pluripotency. Firstly, as shown in the Venn diagram, at the global level, there was little overlapping among the differential hits from the three omics, indicating little direct correlation between PTM dynamics and protein abundance (Figure S7A). Strikingly, in naïve hESCs there were much

more identified sites with down-regulated acetylation and phosphorylation levels than those with up-regulated ones, despite the down-regulated total protein numbers were similar to up-regulated ones (Figure 4A), indicating globally suppressed PTM levels in naïve hESCs as opposed to primed hESCs. Further, Subcellular location annotation information (Yu et al., 2006) showed that almost three quarters of total differentially changed proteins and PTM sites were localized at the nucleus and cytosol. The proportion of cytoplasmic localization was up to 41% in up-regulated proteins, while over half of the down-regulated proteins were nuclear localized. For differential acetylation sites, there was a significant increase in the proportion of mitochondrial localization, reaching 10% of the total. In comparison, nearly 63% of the differential phosphosites were localized at the nucleus and, more interestingly, almost all of them were down-regulated in naïve compared to primed hESCs (Figures 4B and S7B).

Further, when clustering the top 20 up- and down-regulated proteins/PTM sites between primed and naïve hESCs from all the three omics together (Figure 4C), we found a considerable number of down-regulated proteins/PTM sites were involved in RNA processing and translation, such as EXOSC5, THOC5, EIF4B (K586), SF3B5 (K17), LARP1 (S220) and EIF4EBP1 (T70). When non-biasedly taking all differentially expressed proteins/PTM sites into analysis, we found that both down-regulated proteins and phosphosites in naïve hESCs were significantly enriched for RNA processing and translation related pathways including RNA transport, spliceosome, RNA degradation, ribosome biogenesis in eukaryotes, RNA polymerase and mRNA surveillance pathway (Figure 4D), while up-regulated proteins in naïve hESCs were most enriched for metabolic pathways (Figure S7C) and up-regulated phosphosites were enriched for lysine degradation and spliceosome (Figure S7D). With respect to acetylation, KEGG analysis revealed 4 RNA-related pathways with up-regulated acetylation sites that included ribosome, spliceosome, mRNA surveillance pathway and RNA degradation, while 4 RNA-related pathways with down-regulated acetylation sites that included RNA transport, spliceosome, ribosome, and ribosome biogenesis in eukaryotes (Figures S7E-S7F). Taken together, by combining quantitative proteomics, phosphoproteomics and acetylproteomics analyses, we revealed RNA processing and translation as the most differentially-regulated processes between naïve and primed hESCs, and mapped out the key differential proteins/PTM sites (Figure 5, Table S4).

Naïve hESCs with Lower MTORC1/4EBP1 Activities Rely More on EIF4E-independent Translation Initiation Pathway

Given the globally-repressed RNA processing and translation pathways identified in naïve hESCs, we directly quantified and compared the total RNA and protein levels in the two pluripotency states. The total RNA and protein levels in primed hESCs were 1.4 fold and 1.3 fold higher than those in equal number of their RSet-ff naïve counterparts, respectively (Figure 6A). To assess the rate of global protein synthesis, cells were incubated for 30 mins with O-propargyl-puromycin (OPP), a puromycin analog that generates covalent conjugates with nascent polypeptide chains, followed by cell fixation and fluorescent labeling with Alexa Fluor™ 488 in a click reaction, and finally fluorescence imaging and quantification using the Operetta CLS High-Content Analysis System. The Alexa Fluor™ 488 fluorescence intensity indicative of nascent protein translation rate in primed hESCs was 1.8 fold higher than that in RSet-ff naïve cells (Figures 6B left and S8A), an observation directly supporting the major

conclusion from the above multi-omics analyses. Meanwhile, we also confirmed that the translation efficiency in naïve H9 cells derived in PXGL or FINE medium was equivalent to only 70% or 50% of that of the primed H9 cells, respectively (Figures 6B right). In addition, consistent with the up-regulated proteasome activity identified from the above MCODE analysis (Figures 2B and S4A), naïve hESCs exhibited higher degree of global ubiquitination (Figure S8B). However, when treated with the translational elongation inhibitor cycloheximide (CHX) to block synthesis of new proteins, the revealed overall protein degradation rate in naïve hESCs was only slightly higher than that in their primed counterparts (Figure S8C). Thus, we speculate that the greatly strengthened ubiquitin/proteasome system in naïve hESCs may form a part of the protein quality control system to specifically and efficiently remove those deficient, impaired or redundant proteins to maintain a compact naïve proteome.

We next examined mammalian target of rapamycin complex 1 (mTORC1), a serine/threonine kinase and master regulator positively promotes biosynthesis of proteins by adjusting ribosome biogenesis, mRNA biogenesis and cap-dependent translation (Laplanche and Sabatini, 2009). Both the total and phosphorylated mTOR protein levels in RSet-ff naïve hESCs were much lower than those in primed hESCs (Figure S8D), confirming the predicted kinase activities shown in Figure 2E. In PXGL- or FINE-induced naïve hESCs, however, the total and phosphorylated mTOR protein levels were equivalent to or only slightly lower than those in their primed counterparts (Figure S8D). Besides, naïve hESCs derived from three conditions all had a higher level of total PRAS40 protein, the well-established mTORC1 suppressor, and a lower level of phosphorylated PRAS40 (except for FINE naïve hESCs) that can be dissociated from mTORC1 and relieve its inhibition on mTORC1 (Figure S8D). Unexpectedly, however, AKT serine/threonine protein kinase, the presumed positive upstream regulator of mTORC1, was more activated in naïve hESCs (Figure S8E), indicating a decoupling of AKT and mTORC1 signaling in human naïve pluripotency. Although the total protein level of insulin-like growth factor 1 receptor (IGF1R), the putative upstream activator of PI3K/AKT signaling, was higher in naïve hESCs, the level of phosphorylated IGF1R in two pluripotency states was similar (Figure S8E). Thus, the upstream regulator for elevated AKT activity in naïve pluripotency remains elusive, and it remains to be seen if elevated CDK1 in naïve hESCs may promote PI3K-AKT pathway via PDK1 as previously reported (Wang et al., 2017).

To assess the impact of mTORC1 on protein synthesis in two pluripotency states, we examined one of the mTORC1 substrates, the eukaryotic initiation factor 4E-binding protein 1 (4EBP1), which plays a key role in controlling translation initiation. When 4EBP1 is phosphorylated by active mTORC1, it can be disassociated from eukaryotic initiation factor 4E (eIF4E) which can then critically mediate cap-dependent translation (Richter and Sonenberg, 2005). Remarkably, compared with the primed hESCs, both total and phosphorylated 4EBP1 protein contents were significantly reduced in naïve hESCs induced by three systems (Figures 6C and 6D) while the total eIF4E protein levels were similar in the two pluripotency states (Figures 6C and 6D). Given the evidence that in mitotic cells CDK1 can hyperphosphorylate 4EBP1 in the absence of mTORC1 activity (Shuda et al., 2015), elevated CDK1 in naïve hESCs may directly phosphorylate and regulate 4EBP1 and other translation machinery. Indeed, our bioinformatics analysis indicated that some key

translational regulators such as eIF2 subunits and EIF4ENIF1 are probable substrates for CDK1 (Table S5). Thus, the eIF4E-dependent translation initiation processes in naïve hESCs appeared to rely much lesser on mTORC1/4EBP1 than primed hESCs.

Accordingly, when the two pluripotency state hESCs were treated with XL388, a specific mTORC1 inhibitor that had no discernable effects on cell viability at low doses (Figure S8F), there was a more prominent reduction of translation efficiency in primed hESCs compared to RSet-ff naïve hESCs (Figure S8G). Taken all together, we propose that primed hESCs mainly rely on eIF4E-mediated cap-dependent pathway for protein translation while naïve hESCs with higher CDK1 (Figure 3) and lower mTORC1 activities (Figure 6) are more tolerant to eIF4E blockade.

Identification of EIF4E-independent and EIF4A2-dependent Translation Pathways in Naïve hESCs

To quantify the proportion of eIF4E-dependent translation in two pluripotency states, we treated cells with 4E1RCat, a compound that can specifically block the formation of the eIF4F complex and hence eIF4E-mediated translation (Cencic et al., 2011). While there was a dose-dependent inhibition of protein synthesis by up to 90% in primed hESCs treated with 4E1RCat, approximately 40% and 60% protein synthesis was retained in RSet-ff naïve H9 and H1 cells, respectively, even at the maximal dose of 4E1RCat (Figures 6E and S9A). Similarly, maximum inhibition of protein synthesis by 4E1RCat was approximately 50%-60% for both PXGL- and FINE-induced naïve H9 cells compared to 90% inhibition for their primed counterparts (Figure 6F). The pictures containing fluorescence intensity indicative of nascent protein translation rate in different concentrations of 4E1RCat were shown in Figure S9B. These results indicated that under 4E1RCat-treated conditions, primed hESCs predominantly rely on eIF4E-dependent translation while a considerable proportion of translation in naïve hESCs is via eIF4E-independent pathways. However, it is unclear if the eIF4E-independent translation is a compensatory mechanism under inhibitor-treated stress conditions or an intrinsic process under physiological conditions.

To gain some insights into the eIF4E-independent translation mechanism in naïve pluripotency, we quantitatively compared the protein levels of multiple eukaryotic initiation factors in two pluripotency states by immunoblotting. The total protein levels of eIF4G1, eIF4B, eIF2A and eIF5A in naïve hESCs were significantly lower (Figure 7A), indicating a global suppression of eIF4E-dependent translation that is consistent with a reduction of global translation rate and total protein level as shown above (Figures 6A and 6B). On the contrary, consistent with the proteomic data, under the RSet-ff induction conditions, we demonstrated higher levels of eIF3A in naïve H9 and H1 cells, and a higher level of eIF4A2 in naïve H9 cells (Figures 7A, 7B and S10A). The same increased pattern for eIF3A and eIF4A2 held in PXGL naïve H9 cells. However, FINE naïve H9 cells showed a higher level of eIF3A but the same level of eIF4A2 compared to their primed counterparts (Figure 7C). Given the critical roles of eIF3 and eIF4A in regulating IRES-mediated cap-independent translation (Cate, 2017; Komar and Hatzoglou, 2011; Tsai et al., 2014), we asked to what extent the eIF4E-independent translation was mediated by IRESs. By integrating data from a previous study (Weingarten-Gabbay et al., 2016), we found that about 33% of the up-regulated proteins in naïve hESCs contained typical IRES structures in their mRNAs that may translate in a

cap-independent manner (Table S6). PPI analysis revealed a significant enrichment of three biological processes (translation, chaperone-mediated protein folding and proteasome) among these IRES-translated proteins (STRING PPI with the highest confidence factor of 0.9) (Figure 7D).

Since naïve hESCs rely substantially on eIF4E-independent pathway for translation (Figures 6E and 6F), we wondered if the eIFs that are involved in eIF4E-independent translation and up-regulated in naïve hESCs are also crucial for naïve pluripotency. To this end, we knocked down eIF3A or eIF4A (including eIF4A1 and eIF4A2) with siRNAs in primed H9 cells (Figures S10B and S10C) and induced them under the RSet-ff conditions for another 48 hours. Strikingly, the reduction in eIF4A2 protein levels largely prevented the formation of dome-shaped naïve H9 colonies (Figure 7E) and significantly reduced the levels of naïve marker transcripts in converted hESCs (Figure S10D). In contrast, knockdown of eIF3A, or even knockdown of eIF4A1 that has higher than 90% protein sequence similarity to eIF4A2, did not inhibit the formation of naïve morphology (Figure 7E). At day 7 of the PXGL induction conditions, the percentage of formed SUSD2+CD24- naïve cells dropped significantly from 29% in si-NC control group to 18% in si-eIF4A2 group (Figure 7F). We therefore concluded that eIF4A2 is essential for inducing human naïve pluripotency under both RSet-ff and PXGL conditions.

Bivalent Metabolism Drives Both EIF4E-dependent and EIF4E-independent Translation in Naïve hESCs while EIF4E-dependent Translation in Primed hESCs Relies Exclusively on Glycolysis

As mRNA translation is one of the most energetically demanding processes in the cell (Rolfe and Brown, 1997), and primed hESCs utilize aerobic glycolysis while naïve hESCs adopt bivalent metabolism (both glycolysis and OXPHOS) (Leibovitch and Topisirovic, 2018; Tsogtbaatar *et al.*, 2020), we wondered if there is any association between metabolism modes and translation pathways. GSK2837808A (GSKA), a selective lactate dehydrogenase A (LDHA) inhibitor (Hu *et al.*, 2018), and lonidamine (LONI), a selective hexokinase inhibitor (Guo *et al.*, 2016), were used to selectively block glycolysis and OXPHOS, respectively. To minimize indirect or nonspecific effects that could result from prolonged inhibition, we treated cells for only 4 hours. Disturbing energy metabolism dramatically decreased protein synthesis in both pluripotency states. GSKA dose-dependently reduced translation in primed hESCs, reaching about 90% reduction at the dose of 900 μ M, whereas LONI had a much lesser inhibitory effect on translation (Figure S10E), suggesting that glycolysis is the predominant metabolic mode to supply energy for translation in primed hESCs. In contrast, GSKA or LONI alone had a similar dose-dependent inhibitory effect on translation in naïve hESCs (Figure S10E), indicating a roughly equivalent contribution of glycolysis and OXPHOS to translational efficiency and protein synthesis in naïve pluripotency where neither glycolysis nor OXPHOS alone could suffice. It is noteworthy that GSKA caused a much faster translation reduction in naïve hESCs than in primed hESCs, supporting the assumption that glycolysis-regulated translation accounts for almost all translation activities in primed hESCs but only a portion of translation activities in naïve hESCs. Finally, to find out if different metabolic mode-mediated translation in naïve hESCs has a preference in eIF4E usage, we combined varying doses of the two metabolic inhibitors with a fixed dose of

4E1RCat respectively. At low doses, GSKA or LONI did not further reduce translation activities on top of those reduced by 4E1RCat (Figure S10F), suggesting that both compounds preferentially inhibit eIF4E-dependent translation. Higher concentrations of both compounds (GSKA > 100 μ M, LONI > 200 μ M) further brought the translation activities down to levels below those achieved by 4E1RCat treatment, indicating a gradual inhibition of eIF4E-independent translation (Figure S10F). We estimate that the eIF4E-independent translation accounted for at least 30% of the total translation activities in naïve hESCs. Collectively, primed hESCs rely predominantly on glycolysis for eIF4E-dependent translation, while the bivalent metabolism supports both eIF4E-dependent and eIF4E-independent translation in naïve hESCs and thereby conferring them with higher capabilities of survival and propagation.

DISCUSSION

The transition between different states of pluripotency renders a convenient and versatile in vitro model for understanding early human pre-implantation development, the timing and molecular basis of lineage restriction and signal transduction in ICM cells, which will lead to improvements in hESC culture that can generate lineage-committed cells of therapeutic relevance. Our multi-omics analyses reveal here that RNA processing and protein translation are the most differentially-regulated processes between naïve and primed hESCs, and global RNA processing and translation are both down-regulated in naïve hESCs. On one hand, our finding is consistent with the notion that during early stages of stem cell state transition, translational control plays a dominant role (Corsini et al., 2018; Saba et al., 2021; Sampath et al., 2008; Tahmasebi et al., 2018; Wong et al., 2016). On the other hand, however, it apparently contradicts with the observations in this and previous studies (Gafni et al., 2013; Ware et al., 2014) showing that naïve hESCs divide more rapidly than primed hESCs. Indeed, from a broader view, the low mRNA translation rates and restricted protein synthesis in various stem cells are not compatible with their active self-renewal and propagation, and the apparent disconnection between protein translation rate and cell cycling (proliferation) rate remains an intriguing but puzzling question in the stem cell field (Blanco et al., 2016; Tahmasebi *et al.*, 2018).

Recently, Atlasi et al. (Atlasi et al., 2020) took multiple genome-wide approaches to acquire a detailed and dynamic view of the mRNA transcription, mRNA translation, and protein abundance during the transition between 2i/LIF (2iL) ground-state (herein defined as naïve mESCs) and serum/LIF (SL) (defined as primed mESCs) mESCs. They found general increases in polysome-bound mRNAs and hence translation efficiency in naïve mESCs as opposed to primed mESCs, but, as a result of ‘translational buffering’, such increases did not lead to a concomitant elevation of global protein level. Instead, the majority of changes in the proteome were accomplished by transcriptional regulation during naive to primed mESC transition, and only a set of genes underwent specific regulation at the translational level. Thus, it was concluded that transcriptional control represents the main regulatory mode among different states of mESCs whereby specific translational and post-translational control make a lesser contribution to regulating the final protein abundance. The mechanism for

‘translational buffering’ that may underlie the disconnection between global translation efficiency and protein abundance, and more importantly, the mechanism that coordinates protein translation rate and cell cycling rate, await further elucidation.

We showed in this study that the overall transcriptional and translational levels in naïve hESCs are both lower than those in their primed counterparts. It should be noted that the hESC populations and particularly the primed hESCs are highly heterogeneous and therefore biological triplicates were set for each treatment in multi-omics analyses and most validation assays. Similar to the above-mentioned paradoxical observations, naïve hESCs exhibit faster proliferation rate and shorter cycling time. Our current work offers several new clues to the paradox: firstly, since the nascent protein translation rate in naïve hESCs is significantly lower than that in primed hESCs, and the overall protein degradation rate is similar between the two states, the total protein level in a naïve hESC is significantly lower than that in a primed hESC. Thus, the more compact proteomes in naïve hESCs are expected to require lower protein synthesis rates to maintain comparable cell division rates. Also, it was proposed that a low abundance of proteins would allow stem cells to quickly ‘erase’ the old proteome when they receive signals to differentiate (Saba *et al.*, 2021). Secondly, the threshold for total protein level that suffices for naïve hESC division appears to be lower than that for primed hESC division, and CDK1 may be a key player to coordinate metabolism mode, translation initiation and cell cycle progression for rapid cycling of naïve hESCs. Besides its well-established role as an M-CDK that is required for G2/M progression in primed hESCs (Neganova *et al.*, 2014), CDK1 ablation by a specific inhibitor RO3306 caused a significant arrest at S phase, implicating its unexpected role in facilitating S to G2/M progression in naïve hESCs. The Unstable Accumulating Activator models for cellular size control propose an activator that accumulates in a size-dependent manner and triggers cell cycle progression once it has reached a certain threshold, and a strong candidate for the accumulating activator in G2 is Cdc25, the phosphatase that dephosphorylates and activates CDK1 at the G2/M transition (Rhind, 2018). Clearly, the underlying mechanisms for CDK1 promoting naïve hESC cycling and propagation warrants further in-depth investigation.

Thirdly, we showed that during naïve-to-primed pluripotency transition, bivalent-to-glycolytic metabolism switch is associated with a decrease in eIF4E-independent translation and an increase in and predominant reliance on eIF4E-dependent translation, and as a result, a proteome with an increased size and prolonged cycling time. Primed hESCs are known to adopt glycolysis despite sufficient oxygen availability to support OXPHOS, a phenomenon known as aerobic glycolysis, or the Warburg effect, and such incomplete oxidation of glucose enables the conservation of carbon biomass required for biosynthesis of cellular components needed for cell proliferation and division (Tsogtbaatar *et al.*, 2020). We reveal here that glycolysis-mediated translation initiation in primed hESCs is predominantly eIF4E-dependent. In contrast, under stressed conditions, naïve hESCs utilizing bivalent metabolism can initiate translation via both eIF4E-dependent and eIF4A2-dependent/eIF4E-independent pathways. eIF4E-independent translation including IRES-mediated translation is usually initiated when cells are at early embryonic stages or exposed to certain stresses such as hypoxia, nutrition deficiency (Yang and Wang, 2019). Thus, it is likely that selective mRNAs in naïve hESCs are translated via eIF4A2-dependent and/or IRES-mediated pathway with higher efficiency, and such naïve proteome may confer

naïve hESCs with a faster propagation/cycling capability to self-renew and a higher anti-apoptotic potency to tolerate stressed conditions. Given the evidence that some cancer stem cells (CSCs) have bivalent metabolism while the bulk of the non-stem cancer cells is glycolytic (Peiris-Pages et al., 2016), it remains to be seen if both eIF4E-dependent and -independent translation pathways are present in CSCs and if simultaneously blocking both pathways may be beneficial for eradicating CSCs.

CDK1 may also play important roles in promoting OXPHOS and eIF4E-independent translation in naïve hESCs. Previous studies showed that CDK1 can phosphorylate and activate multiple mitochondrial electron transport chain complex proteins to promote OXPHOS while activating p53, MnSOD, and SIRT3 to eliminate reactive oxygen species (ROS) and maintain mitochondrial homeostasis (Wang et al., 2014b; Xie et al., 2019). Further, our bioinformatics analysis indicates that multiple eIFs particularly eIF2 subunits are potential substrates for CDK1. In ESCs (Friend et al., 2015) and adult stem cells (Zismanov et al., 2016), phosphorylated eIF2A is found to inhibit global translation but stimulate the translation of selective mRNAs containing upstream open reading frames (uORFs) within their 5'-UTR, and uORFs can increase IRES-mediated translation in a transcript-specific manner (Leppek et al., 2018). Therefore, potential phosphorylation of eIFs by CDK1 may facilitate translation of a subset of mRNAs via eIF4E-independent pathways. In contrast to elevated CDK1 activity, the mTORC1 signaling in RSet-ff naïve hESCs is significantly suppressed, resulting in a much reduced level of total 4EBP and phosphorylated 4EBP, and a much lesser dependence of naïve hESC on mTORC1 signaling-mediated eIF4E-dependent translation. Interestingly, CDK1 and mTORC1 share common substrates and may compete with each other for the same subset of substrates (Odle et al., 2020), and therefore it remains to be determined if a CDK1-to-mTORC1 switch may account for or be associated with the naïve-to-primed pluripotency transition.

Fourthly, we revealed that a more robust proteome quality control system is present in naïve hESCs that ensures a high-fidelity and compact proteome for rapid cell cycling. Remarkably, protein folding chaperon and proteasome are two most up-regulated pathways in naïve hESC proteome we identified. The quality of the proteome is known to be controlled by proteostasis surveillance system, a complex network of integrated cellular pathways that maintain a balanced protein synthesis rate, high protein synthesis accuracy, efficient protein folding and continual tagging of damaged proteins for degradation (Francisco et al., 2020). Noormohammadi et al. (Noormohammadi et al., 2016) showed that human pluripotent stem cells exhibit increased assembly of the chaperonin TRiC/CCT complex that assists the folding of a significant percentage of nascent protein such as actin and tubulin, and serves as a key quality control mechanism to maintain an intact and refined proteome for active self-renewal. Our finding that CCT subunits are significantly up-regulated in naïve over primed hESCs clearly indicate a more efficient protein folding system operating in naïve pluripotency. In addition, the ubiquitin/proteasome-mediated protein degradation machinery (but not the overall protein degradation) and autophagy system are also up-regulated in naïve hESCs. It was shown that autophagy supports the translation of a subset of mRNAs enriched for cell cycle control and DNA damage repair (Goldsmith et al., 2020). All these proteostasis surveillance mechanisms form a tighter proteome quality control system for human naïve pluripotency.

Taken together, when compared with primed hESCs, we reveal several important attributes for naïve hESCs that include their proteomes with reduced total protein quantities and PTM levels, their strengthened proteome quality control system, and their expedited cell cycling. We show that bivalent metabolism allows for translating selective mRNAs via both eIF4E-dependent and eIF4E-independent/eIF4A2-dependent pathways that jointly generate a refined and compact proteome required for robust self-renewal and survival of naïve hESCs. These findings may assist in better understanding the recently-revealed unrestricted lineage potential of naïve hESCs and in further optimizing their induction and differentiation conditions for future clinical applications.

LIMITATIONS OF STUDY

Although three naïve hESC induction systems were employed in parallel to validate some of the identified key targets, the primary multi-omics analyses were performed with H9 cells induced by the RSet-ff system. We did notice minor differences and inconsistencies of the results obtained with the three systems. Since the RSet-ff system is known to give rise to the so-called ‘intermediate’ naïve hESCs, ideally, similar studies with ‘bona fide’ naïve hESCs derived from the PXGL system can be conducted which may further consolidate the key findings of this study as well as reveal exciting novel findings. Among the identified differential pathways and targets, we validated only a few of them, such as CDK1, eIF family members, mTOR and its substrates. Many important findings including the higher expression of IGF1R and phosphorylation of AKT in naïve hESCs await further validation and in-depth mechanistic studies. Site-directed mutagenesis studies may be employed to investigate the differentially phosphorylated or acetylated sites in some critical target proteins.

ACKNOWLEDGEMENTS

We are grateful to Yanwen Zhou for his assistance in scientific training to postgraduate students, Linbing Yang (Bioprofile Technology) and Jianan Zhao (PTM Biolabs) for their valuable advice on data analysis, and to the Core Facility Platforms of School of Medicine, College of Life Sciences, Zhejiang University, and Chencong Bo for technical assistance. This work was supported by grants from the National Key Research and Development Program of China (grant numbers 2016YFA0100303, 2016YFA0101201), the National Natural Science Foundation of China (grant numbers 31601103, 31401104), 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

YJW, CC and BK conceived the research. CC, BK, QC, WH, JD, JH, BQ analyzed and interpreted the multi-omics data and results. CC and XZ conducted most wet lab experiments with critical input from YW, XC, WC, SD, SS, QL and YH. CC, BK and YJW drafted the

manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure diversity in experimental samples through the selection of the cell lines. We worked to ensure diversity in experimental samples through the selection of the genomic datasets.

MAIN FIGURE TITLES AND LEGENDS

Figure 1. Induction of primed hESCs into naïve pluripotency under three culture conditions.

(A) Overview for induction of primed hESCs into naïve hESCs via three culture systems. Flowchart detailing commercial RSet-feeder free (RSet-ff) induction system, and light micrographs showing primed H9 hESCs and converted naïve H9 hESCs. Scale bars, 100 μm .

(B) Primed and RSet-ff-induced naïve H9 hESCs were harvested, and the mRNA levels of naïve-pluripotency marker genes were determined by qRT-PCR. The log₂ Naïve/Primed fold change values were presented as mean \pm S.D. of three independent experiments.

(C) Mean population doubling time of primed versus RSet-ff-induced naïve H9 hESCs. Total cell numbers (left panels) were counted every 24 hours, and the mean population doubling times were calculated (right panels). The data were presented as the mean \pm S.D. from three independent experiments. Unpaired t test was performed so that $**p < 0.01$, $***p < 0.001$.

(D) Single-cell clonogenicity of primed versus RSet-ff-induced naïve H9 hESCs was determined by Alkaline Phosphatase Staining. The data were presented as the mean \pm S.D. from three independent experiments. Unpaired t test was performed so that $**p < 0.01$.

(E) Primed versus RSet-ff-induced naïve H9 hESCs were fixed and immunostained for pluripotency marker OCT4 and naïve-associated marker KLF4, and counterstained for nuclei. Scale bars, 50 μm .

(F) Primed versus RSet-ff-induced naïve H9 hESCs were fixed and immunostained for H3K27me3 foci (arrows), and counterstained for OCT4 and nuclei. Scale bars, 50 μm (left), 20 μm (right).

(G) Flow-cytometry analysis of SUSD2 and CD24 expression in primed versus PXGL-induced naïve H9 hESCs.

(H) Expression of naïve, primed and pluripotency-associated transcripts in primed versus FINE-induced naïve H9 hESCs. Data were presented as mean \pm S.D. of three independent experiments.

Figure 2. Global views on biological processes, histone acetylation and kinase activity in primed versus RSet-ff-induced naïve H9 hESCs.

(A) Workflow of the main procedures for proteomic, acetylomic and phosphoproteomic analyses.

(B) All the differentially-expressed proteins between the two pluripotency states were

subjected to MCODE and PPI analysis. Each blue node represents a differential protein, and the gray lines connecting the nodes represent PPIs. 6 modules comprising highly inter-connected nodes were highlighted in yellow (higher right), sorted out and given more details (left).

(C) Differential non-histone acetylation sites were classified into four levels, Q1 through Q4 (Q1 (ratio <0.5), Q2 (0.5 < ratio < 0.66), Q3 (1.5 < ratio < 2), and Q4 (ratio >2)), in the order of increasing Naïve/Primed fold changes. The KEGG pathway enrichment analysis for the classified acetylation sites was shown.

(D) A summarized map showing the Naïve/Primed fold changes for acetylation sites on histone subunits. N, naïve; P, primed.

(E) Kinase activities inferred from IKAP that may shape the phosphoproteome in naïve versus primed hESCs.

(F) Kinase-to-kinase interaction network. The asterisks indicate kinases that are well-established in maintaining naïve pluripotency, and the triangle represents a novel kinase that is centrally positioned in the network.

Figure 3. CDK1-regulated E-cadherin functionality is crucial for naïve pluripotency.

(A) Selected CDK1 substrates with elevated phosphorylation levels at putative CDK1 phosphorylated sites identified in naïve H9 hESC phosphoproteome.

(B) Western blot showing CDK1 is more active in naïve H9 cells cultured in PXGL medium, similar to those cultured in RSet-ff medium.

(C) Light micrographs of naïve (RSet-ff) H1 and H9 cells treated with vehicle (Ctrl) or 5 μ M RO3306 for 24 hours. Scale bars, 200 μ m.

(D) Inactivation of CDK1 by RO3306 (5 μ M for 24 hours) resulted in increased transcription of mesendoderm markers. Naïve H1 and H9 cells were cultured in RSet-ff medium. Data were presented as mean \pm S.D. from three independent experiments. Unpaired t test was performed so that NS (not statistically significant) $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(E) Naïve H1 and H9 cells cultured in RSet-ff were treated with vehicle (Ctrl) or 5 μ M RO3306 for 24 hours, immunostained for E-cadherin and OCT4, and counterstained with Hoechst 33342 for nuclei. Representative confocal micrographs were shown. Scale bars, 50 μ m.

(F) CDK1 inhibitor RO3306 (5 μ M, 24 hours) led to decreased expression levels of E-cadherin and its binding partners. Naïve (RSet-ff) H1 and H9 cells were harvested and the whole cell lysates were subjected to SDS-PAGE and immunoblotting, and detected by specified antibodies. The immunoblots shown were from one experiment representative of three independent experiments with similar results.

(G) The mRNA levels of CDH1 (E-cadherin) in naïve (RSet-ff) H1 and H9 cells remained unchanged at conditions described in (F). Data were presented as mean \pm S.D. from three independent experiments. Unpaired t test was performed so that NS (not statistically significant) $p > 0.05$.

(H) CDK1 interacted with E-cadherin in naïve (RSet-ff) hESCs. Naïve cell lysates were immunoprecipitated (IPed) with anti-CDK1 antibody, and the immune complexes were analyzed by immunoblotting (IB) using anti-CDH1 antibody (upper) or anti-CDK1 antibody (lower).

(I-K) Stable naïve (RSet-ff) H9 cell lines overexpressing EGFP or E-cadherin were generated by lentiviral system, and named H9N-EGFP and H9N-CDH1, respectively. They were treated with DMSO (Ctrl) or 5 μ M RO3306 for 24 hours. Light micrographs of alkaline phosphatase positive (AP+) colonies (I) and their statistical data (J) were shown. Scale bars, 200 μ m. The data were presented as mean \pm S.D. of three independent experiments. (**p <0.01, ***p <0.001). Unpaired t test was performed so that NS p>0.05, ***p <0.001. The protein levels of key pluripotency factors were examined by immunoblotting (K).

Figure 4. Integrative multi-omics analyses implicating RNA Processing and Translation as the most differentially-regulated processes between primed and naïve hESCs.

(A) Volcano plot showing global distribution of significantly differential proteins and PTM sites. The vertical axis represents the p-value (-log10) for differences between naïve (RSet-ff) and primed H9 hESCs, and the horizontal dimension is the fold change (log2) between the two groups. Red dots: p-value < 0.05, Naïve/Primed fold change >1.5; green dots: p-value < 0.05, Naïve/Primed fold change <0.66.

(B) Subcellular distribution for up-regulated and down-regulated proteins, acetylation sites and phosphosites in naïve (RSet-ff) H9 hESCs.

(C) Clustering of the top 20 up- and down-regulated hits in naïve (RSet-ff) versus primed H9 hESCs from three omics analyses.

(D) KEGG pathway enrichment analysis for down-regulated proteins (left) and down-regulated phosphosites (right) in naïve (RSet-ff) versus primed H9 hESCs. Functional classification of pathways: A-Metabolism: AD-Nucleotide metabolism; B-Genetic information processing: BA-Transcription; BB-Translation; BC-Folding, sorting and degradation; C-Environmental information processing: CB-Signal transduction; D-Cellular processes: DC-Cell growth and death; E-Organismal systems: EB-Endocrine system; EE-Excretory system; EF-Nervous system.

Figure 5. Schematic summary of multi-omics data showing differential proteins, phosphosites and acetylation sites associated with RNA processing and translation in naïve (RSet-ff) versus primed H9 hESCs.

Large circles represent individual proteins that are compared between naïve and primed H9 hESCs. Light red: significantly up-regulated proteins in naïve H9 hESCs; Light blue: significantly down-regulated proteins in H9 hESCs; Grey: no significantly altered protein levels between naïve and primed H9 hESCs. Small circles surrounding each large circle represent individual post-translational modification sites that are compared between naïve and primed H9 hESCs. Red: up-regulated PTMs; Blue: down-regulated PTMs. The labels in small circles denote specific residues. Black labels: phosphorylated residues. White labels: acetylated residues.

Figure 6. Naïve hESCs with lower mTORC1/4EBP1 activities are more tolerant to eIF4E blockade.

(A) Primed and naïve (RSet-ff) H9 cells were harvested and counted. Total RNAs (left) and total proteins (right) from equal number of primed and naïve cells were extracted and quantified. The data were presented as mean \pm S.D. from three independent experiments, with

the values of primed cells being set at 1. Unpaired t test was performed so that * $p < 0.05$.

(B) Primed and naïve (RSet-ff, left; PXGL and FINE, right) H9 cells were incubated for 30 mins with 20 μM of OPP, followed by cell fixation with 1% paraformaldehyde and fluorescent labeling with Alexa FluorTM 488 in a click reaction, and finally fluorescence imaging and quantification using the Operetta CLS High-Content Analysis System. The normalized mean Alexa FluorTM 488 fluorescence intensities indicative of incorporated nascent polypeptide chains were presented as mean \pm S.D. from three independent experiments, with the values of primed cells being set at 1. Unpaired t test was performed so that *** $p < 0.001$.

(C-D) Primed and naïve (RSet-ff, C; FINE and PXGL, D) hESCs were harvested and the whole cell lysates were subjected to SDS-PAGE and immunoblotting, and detected by specified antibodies. The immunoblots shown were from one experiment representative of three independent experiments with similar results.

(E) Primed and naïve (RSet-ff) H9 (left) or H1 (right) cells were treated with varying concentrations of 4E1RCat for 2.5 hours, followed by nascent protein translation rate determination as described in (B). The normalized mean Alexa FluorTM 488 fluorescence intensity indicative of incorporated nascent polypeptide chains was taken as the global nascent protein synthesis rate, and the ratio of global nascent protein synthesis rate with 4E1RCat treatment over that with vehicle treatment was presented as mean \pm S.D. of triplicate measurements from one experiment representative of three independent experiments with similar results. Unpaired t test was performed so that ** $p < 0.01$, *** $p < 0.001$.

(F) Naïve H9 cultured in other two media (PXGL, left; FINE, right) were treated and analyzed as described in (E).

Figure 7. eIF4A2 is essential for inducing human naïve pluripotency.

(A-B) Primed and naïve (RSet-ff) H9 cells were harvested and the whole cell lysates were subjected to SDS-PAGE and immunoblotting with specified antibodies (A). The immunoblots shown were from one experiment representative of three independent experiments with similar results. The relative band intensity of eIF3A and eIF4A2 in two pluripotency states was consistent with that of the quantitative data from proteomics shown on the right (B).

(C) Whole cell lysates of primed and naïve (FINE and PXGL) H9 cells were subjected to SDS-PAGE and immunoblotting with specified antibodies. The immunoblots shown were from one experiment representative of three independent experiments with similar results.

(D) Pie chart (upper) showing the percentage of the up-regulated proteins in naïve (RSet-ff) H9 hESCs that contain typical IRES structures in their mRNAs. The known interactions among these proteins were revealed by STRING analysis (lower). Three enriched pathways were circled.

(E) Morphological changes of primed H9 hESCs transfected with scramble siRNA (si-NC) or specified siRNAs and induced to naïve pluripotency in RSet-ff media for 48 hours. Scale bars, 200 μm .

(F) Flow-cytometry analysis of SUSD2 and CD24 expression during induction of scramble siRNA (si-NC) or eIF4A2 siRNA (si-eIF4A2) transfected primed H9 into naïve pluripotency in PXGL culture system.

STAR METHODS

KEY RESOURCES TABLE

RESOURCES AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ying-Jie Wang (yingjiewang@zju.edu.cn).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

- The multi-omics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD030230.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Cell Lines and Drug Treatment

Human (h) male ESC line H1 (SCSP-301) and female ESC line H9 (SCSP-302) were purchased from the Cell Bank of the Chinese Academic of Sciences, Shanghai, China. Details for all the key resources were summarized, and sequences for all the qPCR primers were presented in Table S7. Where appropriate, hESCs were treated with RO3306 (TargetMol, #T2356), Cycloheximide (CHX) (Sigma-Aldrich, #239763-M), XL388 (TargetMol, #T6030), 4E1RCat (TargetMol, #T1742), Lonidamine (LONI) (TargetMol, #T0239), or GSK2837808A (GSKA) (TargetMol, #T15435) either individually or in combination with the indicated concentrations.

Primed and RSet-ff naïve hESC Culture

Here hESCs were cultured under feeder-free conditions to eliminate potential confounding effects caused by irradiated mouse embryonic fibroblast (MEF) feeder layers. Primed hESCs were grown on Matrigel (Corning, #354277) coated tissue culture dishes according to manufacturer's recommendation, and maintained in mTeSR1 media (StemCell Technologies, #85850) at 37 °C and 5% O₂. Only before RSet-ff naïve induction, primed hESCs were cultured in 20% O₂ for a week. Medium change was performed on daily basis. Cells were passaged every 3-4 days by 5 min incubation at 37 °C with 0.02% EDTA (Cienry, #E6511). Conversion of primed hESCs to naïve-like hESCs in RSeT™ Feeder-Free medium was achieved following the manufacturer's manual (StemCell Technologies, #05975). Briefly, one day after plating aggregates of primed hESCs (cultured in 20% O₂), the mTeSR1 medium was

replaced with RSeT™ Feeder-Free medium and cells were transferred to hypoxia environment. Afterwards, the cultures were replaced with fresh RSeT™ Feeder-Free medium every two days until the domed-shaped colonies appeared (approximately 5-7 days). The naïve-like hESCs were passaged using TrypLE (Thermo Fisher, #12605010).

PXGL Naïve hESC Culture

N2B27 medium (1 L): 487 ml DMEM/F12 (Sigma, #D6421), 487 ml Neurobasal (Thermo Fisher, #21103049), 10 ml B27 (Thermo Fisher, #17504044), 5 ml N2 (Thermo Fisher, #17502048), 10 ml L-Glutamine (200 mM, Thermo Fisher, #25030024), 1 ml 0.1M β -mercaptoethanol (Macklin, #M6230).

PXGL medium: N2B27 (in house) supplemented with 1 μ M PD0325901 (ABCR, #AB253775), 2 μ M XAV939 (Sigma, #X3004), 2 μ M Gö6983 (BioTechne, #2285) and 10 ng/ml human LIF (Sangon Biotech, #C600146). MEFs (multi-drug resistant DR4 strain) were inactivated by irradiation and used at a density of $1.5 \times 10^4/\text{cm}^2$. Primed hESCs were cultured in 5% O₂, 5% CO₂ in a humidified incubator at 37 °C.

Conversion of primed hESCs to naïve hESCs in PXGL medium was achieved following the protocol previously published (Bredenkamp *et al.*, 2019). Primed hESCs were seeded at $1 \times 10^4/\text{cm}^2$ onto irradiated MEF feeders in mTeSR1 medium. Two days later (day 0), medium was changed to PDL/HDACi (N2B27 with 1 μ M PD0325901, 10 ng/ml human LIF and 1 mM valproic acid sodium salt (VPA, Sigma, #P4543) (HDACi)). Following 3 days in PDL/HDACi, medium was changed to PXGL and refreshed daily for a further 10-11 days before passaging to establish naïve hESC cultures in PXGL medium on irradiated MEF feeders.

The cultures are heterogeneous initially, with many differentiated cells, and a few passages may be required until cultures appear homogeneous. After initial passages on MEF, stabilized reset cells can subsequently be maintained on Geltrex (growth factor-reduced, Thermo Fisher, #A1413302). The derived naïve hESCs can be dissociated with TrypLE Express and purified. Then, plating cell suspension and immediately adding 1 $\mu\text{l}/\text{cm}^2$ Geltrex to each well and mixing by gentle shaking. ROCK inhibitor (Y-27632) was added during replating. Next day feeding cells with fresh PXGL medium. There is no need to add additional Geltrex. Cultures can be split every 5 days at a ratio of 1:3, and cells can be cryopreserved using CryoStor® CS10 (StemCell Technologies, #07930) and stored at liquid nitrogen temperature (-135 °C).

FINE Naïve hESC Culture

Conversion of primed hESCs to naïve hESCs in FINE medium was achieved following the protocol previously published (Szczerbinska *et al.*, 2019). Primed cells were seeded on 200 \times diluted Matrigel at a passage ratio of 1:6. Cells were seeded in clumps and kept in mTeSR1 culture for 48 h. To convert the primed hESCs to naïve hESCs, we removed mTeSR1 medium and added FINE culture medium, which consists of a basal medium (1:1 ratio of DMEM/F12 (Sigma, #D6421) and Neurobasal (Thermo Fisher, #21103049) medium, 1 \times N2 supplement (Thermo Fisher, #17502048), 1 \times B27 supplement (Thermo Fisher, #17504044), 1 \times L-glutamine (Thermo Fisher, #25030024), 1 \times non-essential amino acids (Thermo Fisher #11140050), 0.1 mM β -mercaptoethanol (Macklin, #M6230), and 62.5 $\mu\text{g}/\text{ml}$ BSA (Sigma, #B2064)) supplemented with 0.1 mM dasatinib (Selleckchem, #S1021), 0.1 mM AZD5438

(Tocris, #3968), 0.1 mM SB590885 (R&D systems, #2650), 1 mM PD0325901 (ABCR, #AB253775), 10 mM Y-27632 (Selleckchem, #S6390), 20 ng/mL human recombinant leukemia inhibitory factor (Sangon Biotech, #C600146), 20 ng/mL activin A (StemCell Technologies), and 8 ng/mL basic fibroblast growth factor (Peprotech)). Cells were incubated under normoxic conditions (20% O₂, 5% CO₂) for 4-5 days. FINE culture medium was replenished daily. At the end of conversion, cells were passaged as single cells using TrypLE solution on Geltrix-coated plates (dishes were coated for at least 1 hour before use). In brief, cells were washed with 1× PBS, and 1 ml of TrypLE was added to each well (6-well plate, Corning). Cells were incubated at 37 °C for 3-5 min. When cells started to detach from each other and remained adherent to the plate, TrypLE was aspirated thoroughly. After adding 1 ml of FINE medium, cells were gently detached and clumps dissociated to single cells with a 1-ml pipette. Then cells were seeded at a high ratio of 1:2 in coated plates and transferred to a hypoxia (5% O₂, 5% CO₂) incubator for subsequent culture. Medium was refreshed daily. FINE culture cells were subsequently passaged at ratios of 1:2 to 1:4. Differentiated cells were observed in the first 2-3 passages and gradually decreased over subsequent passages. FINE naïve hESCs were cultured in medium for at least five passages before use.

METHOD DETAILS

Reverse Transcription Quantitative Real-Time PCR

Total RNA was extracted using TRIzol reagent (Life Technologies, #15596-026) and complementary DNA (cDNA) made from 1 µg of RNA using PrimeScript RT reagent kit with gDNA eraser (TaKaRa, #DRR047S). The sequences of all the primers used were presented in key resources table. For real-time PCR, we used iTaq Universal SYBR Green Supermix (Bio-Rad, #1725124) and Step One Plus (Applied Biosystems). All the PCR amplification was performed in triplicate and repeated in three independent experiments.

Immunofluorescence Microscopy

hESCs were washed twice with PBS and fixed with 4% formaldehyde (Sigma, #F8775) in PBS for 15 min at room temperature, washed three times with PBS and permeabilized with 0.2 % TritonX-100 (Diamond, # A110694) in PBS at room temperature for 10 min. They were then washed three times with PBS and blocked with 3% BSA (Sigma, #A9418) in PBS for 45 min. Thereafter, cells were incubated with specified primary antibodies overnight at 4 °C and secondary antibodies for 1 hour at room temperature, with three washes in between, and visualized with a Nikon A1R confocal microscope.

Flow Cytometry

Cells were digested by Accutase (Thermo Fisher, #A1110501) for 3 minutes and single cell suspension was prepared in Cell Staining Buffer (BioLegend, # 420201). After adding 5 µl of Fc Receptor Blocking Solution (BioLegend, #422301) per 100 µl of cell suspension, cells were incubated for 5-10 min at room temperature. Cell samples in solution were added with appropriately conjugated fluorescent primary antibodies at predetermined optimum concentrations, and incubated on ice for 15-20 min in the dark. Cell samples were washed twice with at least 2 ml of Cell Staining Buffer by centrifugation at 350 g for 5 min. Finally, cell pellets were resuspended in 0.5 ml of Cell Staining Buffer and added with 5 µl

(0.25 μ g)/million cells of 7-AAD Viability Staining Solution (BioLegend, #420403) to exclude dead cells. After cell samples were incubated on ice for 3-5 min in the dark, flow cytometry analyses were performed using an ACEA NovoCyteTM (ACEA Biosciences). The following antibodies were used for flow cytometry: IgG-PE (BioLegend, #400114), IgG-FITC (BioLegend, #400208), SUSD2-PE (BioLegend, #327406), CD24-FITC (BioLegend, #311104).

Immunoblotting

Collected cell samples were lysed with the Laemmli Sample Buffer (BIO-RAD, #1610747). 20 μ l of loading samples were subjected to 10% SDS-PAGE and then transferred onto 0.45 μ m polyvinyl difluoride membranes (Millipore, #IPVH00010). After blocking in 5% skim milk, membranes were probed with the indicated primary antibodies and corresponding secondary antibodies. Finally, the blots were developed using the ECL reagent (Share-Bio, #sb-wb012) and visualized by the Tanon 5200 Multi Chemiluminescent Imager system.

Cell Cycle Analysis

Cell cycle analysis of hESCs treated with RO3306 or vehicle was performed using Cell Cycle and Apoptosis Analysis Kit (Beyotime, #C1052). Cells were collected and washed with cold PBS. They were then fixed with pre-cooled 70% ethanol for 30 min at 4 °C. After incubated in PI stain solution for 30 min at 37 °C, flow cytometry analyses were performed using an ACEA NovoCyteTM (ACEA Biosciences) and ModFit LT (V.4.0.5).

Immunoprecipitation

Immunoprecipitating antibodies (1.6 μ g of anti-CDK1 (Beyotime, #AF1516); 1.6 μ g of anti-IgG (ABclonal, #AC005)) were added to the PBST (1% Tween final concentration) together with Protein A/G magnetic beads (20 μ l of magnetic bead volume, MedChemExpress, #HY-K0202) for 2 hours at 4 °C. Then, cells (two 10-cm-diameter dishes per condition) were lysed (800 μ l per dish) with cell scraper. Cell lysates in lysis buffer (20 mM Tris-HCl (pH 7.5), 100 mM KCl, 5 mM MgCl₂, 0.5% NP-40, protease inhibitor (Thermo, #78443) and 0.1% DTT) were shaken for 15 seconds at 5 min intervals for 30 min on ice. After centrifugation (14000 rpm for 15 min at 4 °C), the supernatants were added with magnetic beads (PBST was removed) and incubated overnight at 4 °C. Magnetic beads carrying immunoprecipitated immune complexes were then collected and washed three times in lysis buffer. Samples were denatured at 100 °C for 8 min in 40 μ l of Laemmli Sample Buffer (BIO-RAD, #1610747). After removing beads by magnetic separation, samples were subjected to SDS-PAGE and Western blotting, and the blots were developed by the Tanon 5200 Multi Chemiluminescent Imager system.

Transfection and Viral Transduction

For siRNA transfection, primed hESCs were digested into single cell suspension and plated at a density of 1-1.5 \times 10⁵ cells/well (H9) or 2 \times 10⁵ cells/well (H1) in 24-well plates with ROCK inhibitor Y27632 treatment. After 24 hours, cells were then starved for 3 hours in Opti-MEM (Thermo, #31985062) and transfected with siRNAs (scramble siRNA as control) by Lipofectamine RNAiMAX Reagent (Thermo, #13778075) according to the manufacturer's

instructions. Medium was refreshed after 6 hours. For detection of mRNA or protein levels, cells were further cultured for 48 hours. For pluripotency resetting, transfected primed hESCs were further manipulated in different naïve induction systems.

Primed or naïve hESCs stably overexpressing EGFP (as control) or CDH1 were generated by lentivirus-mediated transduction. The LentiCDH1-Blast plasmid was derived from the lentiCas9n(D10A)-Blast (Addgene, #63593) by replacing sequence of Cas9n(D10A) with cDNA sequence of the CDH1 gene. The control LentiEGFP-Blast plasmid was constructed in a similar fashion. Then, lentiviruses were generated in HEK293T cells (SCSP-502, from the Cell Bank of the Chinese Academic of Sciences) by co-transfecting the HEK293 cells with psPAX2 (Addgene, #12260), pMD2.G (Addgene, #12259), and LentiCDH1-Blast (or LentiEGFP-Blast) plasmids using Lipofectamine 3000 (Thermo, L3000015). 6-8 hours after transfection, the medium was replaced with Opti-MEM containing 5% FBS and 1% sodium pyruvate. The culture media containing the released recombinant viruses were collected on day 1 after transfection, fresh media were added, and the culture media were collected again on day 2 after transfection. The recombinant virus-containing supernatants were filtered with a 0.45- μ m filter and viral particles were concentrated by centrifugation at 25,000 rpm in a Beckman SW 28 rotor for 100 min at 4 °C. hESCs were subsequently transduced with concentrated viruses for 12 hours and cultured in fresh media for further 48 hours before blasticidin was added.

Colony-formation Assay and Alkaline Phosphatase Staining

To compare the single-cell clonogenicity of primed versus naïve hESCs, primed and naïve hESCs were digested into single cells and plated at a density of 10000 cells/well in 6-well plates without ROCK inhibitor Y27632 treatment and cultured for a week. Next, the alkaline phosphatase (AP) staining kit (Beyotime, #C3206) was used to stain the hESC colonies, and then the number of colonies was quantified by Image J (V.1.52).

To assess their pluripotency after RO3306 treatment, H9N-EGFP and H9N-CDH1 cells were plated at a density of 50000 cells/well in 6-well plates without ROCK inhibitor Y27632 treatment and cultured for 4 days. Cells were then treated with DMSO (Ctrl) or 5 μ M RO3306 for 24 hours and AP staining was performed as mentioned above.

Cell Doubling-time Determination

8×10^4 of naïve and primed hESCs were seeded in 12-well plates. Cells were digested and counted using the Countess II FL Automated Cell Counter (Thermo) at 24, 48, 72, 96, 120, 144 and 168 hours after seeding. Growth curves were plotted using the GraphPad Prism 7.0 software. Doubling time was calculated using the following formula:

Doubling time = duration * log (2)/ [log (final concentration)-log (initial concentration)].

Quantification of RNA and Protein Abundance

5×10^5 of naïve and primed hESCs were counted and collected repeatedly for the following assays. Total RNA was isolated by EZ-press RNA Purification Kit (EZ Bioscience, #B0004DP), and its quality was determined using a Nano Drop 2000 (Thermo) and quantitated by absorbance at 260 nm. For protein extraction, hESCs were lysed with SDS Lysis Buffer (Beyotime, #P0013G) and centrifuged at 12000 g for 5 min. Then, the

supernatant was collected and the protein concentration was determined with BCA Protein Assay Kit (Bio-Rad, #P0011) according to the manufacturer's instructions. In parallel, total proteins were subjected to SDS-PAGE and visualized by Coomassie Blue staining (0.1% Coomassie blue- R250 (Sangon Biotech, #A100472), 10% acetic acid, 25% isopropanol) after cells were treated with CHX for 12 hours. Then, the overall band intensities representing total protein levels in individual lanes were quantified by the Image J software.

Cell Viability Assay

hESCs were seeded at a density of 5,000 cells in a volume of 100 μ l per well in 96-well plates. The next day, medium was replaced with fresh culture medium containing XL388 or RO3306 at the indicated concentrations. The cells were cultured in the presence of XL388 for 4 hours (RO3306 for 24 hours) and then subjected to CCK-8 analysis (Beyotime, #C0043). The OD450 values were measured by a microplate reader (Thermo).

Determination of Global Nascent Protein Synthesis Rate

Click-iT™ Plus OPP Alexa Fluor™ 488 Protein Synthesis Assay Kit (Thermo, #C10456) was used to determine global protein synthesis/translation rate. After treatment with inhibitors at indicated concentrations, hESCs were incubated for 30 min in the medium supplemented with O-propargyl-puromycin (OPP) (20 μ M), a puromycin analog that generates covalent conjugates with nascent polypeptide chains. Cells were then harvested and fixed with 1% paraformaldehyde for 15 min on ice. Next, the fixed cells were permeabilized in PBS supplemented with 0.1% saponin (Sangon Biotech, #A604521) and 3% FBS (GIBCO, #10270-106) for 5 min at room temperature. The copper-catalyzed azide-alkyne cycloaddition was performed according to the manufacturer's protocol. Addition of the Alexa Fluor™ 488 picolyl azide and the 'click' reaction reagents led to a 'click' reaction between the picolyl azide dye and the OPP alkyne, allowing the newly synthesized polypeptides to be labeled with Alexa Fluor™ 488. Finally, the Alexa Fluor™ 488-labeled cell samples were re-suspended in 100 μ l PBS supplemented with 3% FBS and 0.1% saponin per well in 96-well plates and analyzed by an Operetta CLS High-Content Analysis System (Perkin Elmer). All fields in each well were imaged and the mean Alexa Fluor™ 488 fluorescence intensity was quantified. Mean background fluorescent intensity (referred to as mBFI) was the average of all background fluorescent intensity in each field. The global nascent protein synthesis rate for a given sample was calculated as the mean Alexa Fluor™ 488 fluorescence intensity subtracted by mBFI. The ratio of global nascent protein synthesis rate with an inhibitor treatment over that with vehicle (control) treatment was calculated as: (mean Alexa Fluor™ 488 fluorescent intensity with a specific inhibitor - mBFI) / (mean Alexa Fluor™ 488 fluorescent intensity with vehicle - mBFI).

Next Generation Sequencing-Based Bisulfite Sequencing PCR

Gene-specific DNA methylation was assessed by a next generation sequencing-based BSP, according to a previously published method (Gao et al., 2014). Briefly, BSP primers were designed using the online MethPrimer software. 1 μ g of genomic DNA was converted using the ZYMO EZ DNA Methylation-Gold Kit (Zymo Research, #D5005) and one twentieth of the elution products were used as templates for PCR amplification with 35 cycles using KAPA

2G Robust HotStart PCR Kit (Kapa Biosystems, #KK5522). For each sample, BSP products of multiple genes were pooled equally, 5'-phosphorylated, 3'-dA-tailed and ligated to barcoded adapter using T4 DNA ligase (New England Biolabs, # M0202). Barcoded libraries from all samples were sequenced on Illumina platform.

LC-MS Sample Preparation

All MS experiments were performed in biological triplicates. Proteins were extracted from hESCs using SDT lysis buffer (4% SDS, 100 mM DTT, 100 mM Tris-HCl pH 8.0) with 1× Halt Protease and Phosphatase Inhibitor Cocktail (Thermo, #78443). Samples were boiled for 5 min, and further ultrasonicated and boiled for another 5 min. Undissolved cellular debris were removed by centrifugation at 16,000 g for 15 min. The supernatant was collected and quantified with a BCA Protein Assay Kit (Bio-Rad, #P0011).

Protein Digestion

Protein digestion was performed with the FASP method as described by Wisniewski et al. (Wisniewski et al., 2009). Briefly, the detergent, DTT and IAA in UA buffer were added to block reduced cysteine. Then, the protein suspension was digested with trypsin (Promega, #V5113) at a ratio of 50:1 overnight at 37 °C and collected by centrifugation at 16,000 g for 15 min. Finally, the peptides were desalted with a Sep-Pak C 18 cartridge (Waters, #Z720070) for further analysis.

Affinity Enrichment

Phosphopeptides were enriched using High-Select™ TiO₂ Phosphopeptide Enrichment Kit (Thermo, #A32993). Briefly, lyophilized peptide samples were resuspended in Binding/Equilibration buffer and applied to the pre-equilibrated TiO₂ Spin Tip. After washing, enriched phosphopeptides were eluted using the Phosphopeptide Elution Buffer. Then, the eluates were dried immediately in a speed vacuum concentrator to remove elution buffer. For acetylated peptide enrichment, the peptide samples were first dissolved in IPA buffer and incubated with prewashed anti-acetylation antibody beads (PTMScan® Acetyl-Lysine Motif [Ac-K] Kit, Cell Signaling Technology, #13416) at 4 °C for 1.5 hour incubation. Then the beads were washed three times with IPA buffer and three times with ddH₂O. The bound peptides were eluted from the beads with 0.1% trifluoroacetic acid (TFA), and then, dried by vacuum concentrator. Finally, enriched acetylated peptides were desalted with a Sep-Pak C 18 cartridge and vacuum-dried again.

LC-MS/MS Analysis

LC-MS/MS was performed on a Q Exactive Plus mass spectrometer coupled with Easy 1200 nLC (Thermo Fisher Scientific). Peptides were first loaded to a trap column (100 μm*20 mm, 5 μm, C18) in buffer A (0.1% (v/v) Formic acid in water). Reverse-phase high-performance liquid chromatography (RP-HPLC) separation was performed with the EASY-nLC system (Thermo Fisher Scientific) using a self-packed column (75 μm × 150 mm; 3 μm ReproSil-Pur C18 beads, 120 Å, Dr. Maisch GmbH) at a flow rate of 300 nL/min. The RP-HPLC mobile phase A was 0.1% formic acid in water, and B was 0.1% formic acid in 95% acetonitrile. Peptides were eluted over 120 min with a linear gradient of buffer B. The gradient was 5% B

for 2 min and was linearly increased to 8% in 88 min, and then increased to 23% in 10 min, and then increased to 40% in 8 min and maintained for 12 min. MS data were acquired using a data-dependent top20 method dynamically choosing the most abundant precursor ions from the survey scan (300–1800 m/z) for HCD fragmentation. The instrument was run with peptide recognition mode enabled. A lock mass of 445.120025 Da was used as internal standard for mass calibration. The full MS scans were acquired at a resolution of 70,000 at m/z 200, and 17,500 at m/z 200 for MS/MS scan. The maximum injection time was set to 50 ms for MS and 50 ms for MS/MS. Normalized collision energy was 27 and the isolation window was set to 1.6 Th. Dynamic exclusion duration was 60 s.

QUANTIFICATION AND STATISTICAL ANALYSIS

Sequence Database Searching and Mass Spectrometric Data Analysis

We processed the mass spectra using MaxQuant software version 1.6.0.16. MS data were searched against the UniProtKB Homo sapiens database (173,282 total entries). The trypsin was selected as the digestion enzyme. The maximal two missed cleavage sites and the mass tolerance of 4.5 ppm for precursor ions and 20 ppm for fragment ions were defined for database search. Carbamidomethylation of cysteines was defined as a fixed modification, while acetylation of protein N-terminal and oxidation of Methionine were set as variable modifications for database searching. The database search results were filtered and exported with <1% false discovery rate (FDR) at peptide-spectrum-matched level, and protein level, respectively. For phosphorylation or acetylation data searching, phosphorylation on serine, threonine, and tyrosine or acetylation on lysine was further added into the searching parameter and the site localization probability was set as > 0.75.

Bioinformatics Analysis

To assess the quality of data, principal-component analysis (PCA) and relative standard deviation (RSD) were performed using simcap (v.14.1) and R (v.3.6.1), respectively. Only proteins with fold change ≥ 1.5 -fold or ≤ 0.66 -fold and a p-value <0.05 were considered as significantly differential hits. As for phosphoproteome and acetylome, only modified sites with a localization probability of > 0.75 were further analyzed. Venn diagrams for differential proteins and modified sites were drawn by Venny 2.1.0 (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>). Hierarchical clustering analysis was performed with the pheatmap package, which is based on the open-source statistical language R25, using Euclidean distance as the distance metric and complete method as the agglomeration method. Besides, R was also applied to perform volcano plot analysis with ggplot2 package.

Bisulfite Sequencing Data Analysis

The FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastq>) was used to perform basic statistics on the quality of the raw reads. Then, sequencing adapters and low quality data of the sequencing data were removed by Trimmomatic (version0.36) (Bolger et al., 2014). The BSMAP software was used to map the bisulfite sequence to the reference genome with parameters '-n 0 -g 0 -v 0.08 -m 50 -x 1000' (Xi and Li, 2009). The statistic information of the alignment was collected, only the unique mapped reads were kept for the

following analysis, and only methylated cytosines with sequence depth coverage of at least 5 were used. If the base on the alignment is C, methylation occurs; conversely, if the base on the alignment is T, no methylation occurs. The methylation levels of individual cytosines were calculated as the ratio of the sequenced depth of the ascertained methylated CpG cytosines to the total sequenced depth of individual CpG cytosines, i.e., $ML = mC / (mC + umC)$. Where ML is the methylation level, mC and umC represent the number of reads supporting methylation C and the number of reads supporting unmethylated C, respectively. The software methylKit (version 0.2-7) was used to identify DMR (differentially methylated regions) by a binary segmentation algorithm combined with a two-dimensional statistical test (parameters: -M 300 -m 5 -d 0.1 -t 1 -f 1 -v 0.7) (Juhling et al., 2016). Gene Ontology (referred to as GO, <http://www.geneontology.org/>) enrichment analysis of DMR-related genes were applied to uncover biological processes of interest, we choose to deem pathways with a Q value ≤ 0.05 as significantly enriched with DMR-related genes. Based on the results of the DMR annotation and the database of Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2008), functional enrichment analysis was performed on genes whose gene body and its upstream and downstream regions (upstream 2k, gene body, and downstream 2k) are overlapping with DMR.

Motif Analysis

Over-represented sequence motifs from the total phosphoproteome and acetylome dataset were determined by motif-X algorithm (Schwartz and Gygi, 2005), available online at website: <http://meme-suite.org/tools/momo>. The motifs were centered at the modified residues (S, T, Y or K), using a ± 10 amino acid residue sequence window surrounding the modification sites and only motifs with $p < 10^{-10}$ were allowed.

Subcellular Location

Subcellular location annotation information for total as well as regulated proteins or modified sites were revealed by CELLO (V.2.5) (Yu *et al.*, 2006). The proportion of each cellular component was calculated and presented as a pie chart.

Inference of Kinase Activities from Phosphoproteomics (IKAP) Analysis

IKAP was a machine learning algorithm to infer kinase activity based on the phosphorylation levels of substrates (Mischnik *et al.*, 2016). The kinase-substrate table was downloaded from the PhosphoSitePlus database (Hornbeck et al., 2015). After normalization of phosphopeptide abundance by protein abundance, phosphoproteomic data was input into the IKAP algorithm, and implementation was carried out in MATLAB (R2004a). Kinases that are most likely to phosphorylate the identified phosphosites were inferred for two pluripotency states.

Protein-Protein Interaction (PPI) Network Analysis

We submitted the significantly differential proteins to the Search Tool for the Retrieval of Interacting Genes (STRING, <http://string.embl.de/>) database, and only validated interactions with the highest confidence score > 0.9 were selected. Cytoscape software (V.3.7.2) was further applied to visualize modules with the closest internal links from the PPI network. Molecular Complex Detection (MCODE) plugin performed with degree cutoff = 2, node

score cutoff = 0.2, k-core = 2, and max. depth = 100. Besides, KEGG enrichment analyses were carried out with ClueGO plugin and the global view of connections among various pathways was presented.

Statistical Analysis

For Figures 1C, 1D, 3D, 3G, 3J, 6A, 6B, 6E, 6F, S1C, S1D, S5B, S5D, S5F, S6F, S6G, S8F, S8G, S9A, S10B, S10D, S10E, S10F, the p-values were calculated by two-tailed unpaired Student's t test using GraphPad Prism 7. Data represent the mean \pm S.D. as indicated in figure legends. P-values of less than 0.05 were considered as statistically significant. All statistics were shown as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

SUPPLEMENTAL INFORMATION

Supplementary figures and tables: Figures S1-S10; Table S5: predicted CDK1 phosphorylation sites in EIFs, related to Figure 6; Table S6: upregulated proteins in naïve containing IRES, related to Figure 7; Table S7: qPCR primer sequences, related to Figures 1 and 3.

Data S1: Uncropped Western blots

Table S1: Proteomic data-H9, related to Figure 2.

Table S2: Acetylomic data-H9, related to Figure 2.

Table S3: Phosphoproteomic data-H9, related to Figure 2.

Table S4: RNA processing and translation analysis, related to Figure 5.

REFERENCES

- Atlasi, Y., Jafarnejad, S.M., Gkogkas, C.G., Vermeulen, M., Sonenberg, N., and Stunnenberg, H.G. (2020). The translational landscape of ground state pluripotency. *Nat Commun* *11*, 1617. 10.1038/s41467-020-15449-9.
- Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortes-Garrido, R., Gkatza, N., et al. (2016). Stem cell function and stress response are controlled by protein synthesis. *Nature* *534*, 335-340. 10.1038/nature18282.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114-2120. 10.1093/bioinformatics/btu170.
- Boroviak, T., and Nichols, J. (2014). The birth of embryonic pluripotency. *Philos Trans R Soc Lond B Biol Sci* *369*. 10.1098/rstb.2013.0541.
- Bredenkamp, N., Stirparo, G.G., Nichols, J., Smith, A., and Guo, G. (2019). The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naïve Pluripotent Stem Cells. *Stem Cell Reports* *12*, 1212-1222. 10.1016/j.stemcr.2019.03.014.
- Cate, J.H. (2017). Human eIF3: from 'blobology' to biological insight. *Philos Trans R Soc Lond B Biol Sci* *372*. 10.1098/rstb.2016.0176.
- Cencic, R., Hall, D.R., Robert, F., Du, Y., Min, J., Li, L., Qui, M., Lewis, I., Kurtkaya, S., Dingleline, R., et al. (2011). Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. *Proc Natl Acad Sci U S A* *108*, 1046-1051. 10.1073/pnas.1011477108.
- Chen, H., Aksoy, I., Gonnot, F., Osteil, P., Aubry, M., Hamela, C., Rognard, C., Hochard, A., Voisin, S.,

- Fontaine, E., et al. (2015). Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat Commun* 6, 7095. 10.1038/ncomms8095.
- Collier, A.J., and Rugg-Gunn, P.J. (2018). Identifying Human Naive Pluripotent Stem Cells - Evaluating State-Specific Reporter Lines and Cell-Surface Markers. *Bioessays* 40, e1700239. 10.1002/bies.201700239.
- Corsini, N.S., Peer, A.M., Moeseneder, P., Roiuk, M., Burkard, T.R., Theussl, H.C., Moll, I., and Knoblich, J.A. (2018). Coordinated Control of mRNA and rRNA Processing Controls Embryonic Stem Cell Pluripotency and Differentiation. *Cell Stem Cell* 22, 543-558 e512. 10.1016/j.stem.2018.03.002.
- De Los Angeles, A., Pho, N., and Redmond, D.E., Jr. (2018). Generating Human Organs via Interspecies Chimera Formation: Advances and Barriers. *Yale J Biol Med* 91, 333-342.
- Francisco, S., Ferreira, M., Moura, G., Soares, A.R., and Santos, M.A.S. (2020). Does proteostasis get lost in translation? Implications for protein aggregation across the lifespan. *Ageing Res Rev* 62, 101119. 10.1016/j.arr.2020.101119.
- Friend, K., Brooks, H.A., Propson, N.E., Thomson, J.A., and Kimble, J. (2015). Embryonic Stem Cell Growth Factors Regulate eIF2alpha Phosphorylation. *PLoS One* 10, e0139076. 10.1371/journal.pone.0139076.
- Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504, 282-286. 10.1038/nature12745.
- Gao, F., Xia, Y., Wang, J., Lin, Z., Ou, Y., Liu, X., Liu, W., Zhou, B., Luo, H., Zhou, B., et al. (2014). Integrated analyses of DNA methylation and hydroxymethylation reveal tumor suppressive roles of ECM1, ATF5, and EOMES in human hepatocellular carcinoma. *Genome Biol* 15, 533. 10.1186/s13059-014-0533-9.
- Goldsmith, J., Marsh, T., Asthana, S., Leidal, A.M., Suresh, D., Olshen, A., and Debnath, J. (2020). Ribosome profiling reveals a functional role for autophagy in mRNA translational control. *Commun Biol* 3, 388. 10.1038/s42003-020-1090-2.
- Guo, G., Stirparo, G.G., Strawbridge, S.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* 28, 1040-1056 e1046. 10.1016/j.stem.2021.02.025.
- Guo, L., Shestov, A.A., Worth, A.J., Nath, K., Nelson, D.S., Leeper, D.B., Glickson, J.D., and Blair, I.A. (2016). Inhibition of Mitochondrial Complex II by the Anticancer Agent Lonidamine. *J Biol Chem* 291, 42-57. 10.1074/jbc.M115.697516.
- Harvey, A., Caretti, G., Moresi, V., Renzini, A., and Adamo, S. (2019). Interplay between Metabolites and the Epigenome in Regulating Embryonic and Adult Stem Cell Potency and Maintenance. *Stem Cell Reports* 13, 573-589. 10.1016/j.stemcr.2019.09.003.
- Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 43, D512-520. 10.1093/nar/gku1267.
- Hu, D., Linders, A., Yamak, A., Correia, C., Kijlstra, J.D., Garakani, A., Xiao, L., Milan, D.J., van der Meer, P., Serra, M., et al. (2018). Metabolic Maturation of Human Pluripotent Stem Cell-Derived Cardiomyocytes by Inhibition of HIF1alpha and LDHA. *Circulation research* 123, 1066-1079. 10.1161/CIRCRESAHA.118.313249.
- Ilic, D., and Ogilvie, C. (2017). Concise Review: Human Embryonic Stem Cells-What Have We Done? What Are We Doing? Where Are We Going? *Stem Cells* 35, 17-25. 10.1002/stem.2450.

- Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I., Nakamura, T., Kojima, Y., et al. (2021). Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* *28*, 1023-1039 e1013. 10.1016/j.stem.2021.03.013.
- Jorda, R., Hendrychova, D., Voller, J., Reznickova, E., Gucky, T., and Krystof, V. (2018). How Selective Are Pharmacological Inhibitors of Cell-Cycle-Regulating Cyclin-Dependent Kinases? *J Med Chem* *61*, 9105-9120. 10.1021/acs.jmedchem.8b00049.
- Juhling, F., Kretzmer, H., Bernhart, S.H., Otto, C., Stadler, P.F., and Hoffmann, S. (2016). metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome research* *26*, 256-262. 10.1101/gr.196394.115.
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res* *36*, D480-484. 10.1093/nar/gkm882.
- Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle* *10*, 229-240. 10.4161/cc.10.2.14472.
- Laplante, M., and Sabatini, D.M. (2009). mTOR signaling at a glance. *J Cell Sci* *122*, 3589-3594. 10.1242/jcs.051011.
- Leibovitch, M., and Topisirovic, I. (2018). Dysregulation of mRNA translation and energy metabolism in cancer. *Adv Biol Regul* *67*, 30-39. 10.1016/j.jbior.2017.11.001.
- Leppek, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* *19*, 158-174. 10.1038/nrm.2017.103.
- Li, L., Bennett, S.A., and Wang, L. (2012). Role of E-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. *Cell Adh Migr* *6*, 59-70. 10.4161/cam.19583.
- Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S., et al. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat Methods* *14*, 1055-1062. 10.1038/nmeth.4436.
- Martello, G., and Smith, A. (2014). The nature of embryonic stem cells. *Annu Rev Cell Dev Biol* *30*, 647-675. 10.1146/annurev-cellbio-100913-013116.
- Mischnik, M., Sacco, F., Cox, J., Schneider, H.C., Schafer, M., Hendlich, M., Crowther, D., Mann, M., and Klabunde, T. (2016). IKAP: A heuristic framework for inference of kinase activities from Phosphoproteomics data. *Bioinformatics* *32*, 424-431. 10.1093/bioinformatics/btv699.
- Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., Nemirovski, A., Shen-Orr, S., Laevsky, I., Amit, M., et al. (2015). Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab* *21*, 392-402. 10.1016/j.cmet.2015.02.002.
- Narita, T., Weinert, B.T., and Choudhary, C. (2019). Functions and mechanisms of non-histone protein acetylation. *Nat Rev Mol Cell Biol* *20*, 156-174. 10.1038/s41580-018-0081-3.
- Neganova, I., Tilgner, K., Buskin, A., Paraskevopoulou, I., Atkinson, S.P., Peberdy, D., Passos, J.F., and Lako, M. (2014). CDK1 plays an important role in the maintenance of pluripotency and genomic stability in human pluripotent stem cells. *Cell Death Dis* *5*, e1508. 10.1038/cddis.2014.464.
- Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A. (2012). Human pre-implantation embryo development. *Development* *139*, 829-841. 10.1242/dev.060426.
- Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* *4*, 487-492.

10.1016/j.stem.2009.05.015.

Noormohammadi, A., Khodakarami, A., Gutierrez-Garcia, R., Lee, H.J., Koyuncu, S., Konig, T., Schindler, C., Saez, I., Fatima, A., Dieterich, C., and Vilchez, D. (2016). Somatic increase of CCT8 mimics proteostasis of human pluripotent stem cells and extends *C. elegans* lifespan. *Nat Commun* 7, 13649. 10.1038/ncomms13649.

Odle, R.I., Walker, S.A., Oxley, D., Kidger, A.M., Balmanno, K., Gilley, R., Okkenhaug, H., Florey, O., Ktistakis, N.T., and Cook, S.J. (2020). An mTORC1-to-CDK1 Switch Maintains Autophagy Suppression during Mitosis. *Mol Cell* 77, 228-240 e227. 10.1016/j.molcel.2019.10.016.

Peiris-Pages, M., Martinez-Outschoorn, U.E., Pestell, R.G., Sotgia, F., and Lisanti, M.P. (2016). Cancer stem cell metabolism. *Breast Cancer Res* 18, 55. 10.1186/s13058-016-0712-6.

Pieters, T., and van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. *J Cell Sci* 127, 2603-2613. 10.1242/jcs.146720.

Qiao, Y., Wang, R., Yang, X., Tang, K., and Jing, N. (2015). Dual roles of histone H3 lysine 9 acetylation in human embryonic stem cell pluripotency and neural differentiation. *J Biol Chem* 290, 9949. 10.1074/jbc.A114.603761.

Qin, H., Hejna, M., Liu, Y., Percharde, M., Wossidlo, M., Blouin, L., Durruthy-Durruthy, J., Wong, P., Qi, Z., Yu, J., et al. (2016). YAP Induces Human Naive Pluripotency. *Cell Rep* 14, 2301-2312. 10.1016/j.celrep.2016.02.036.

Rhind, N. (2018). Cell Size Control via an Unstable Accumulating Activator and the Phenomenon of Excess Mitotic Delay. *Bioessays* 40. 10.1002/bies.201700184.

Richter, J.D., and Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477-480. 10.1038/nature03205.

Rolfe, D.F., and Brown, G.C. (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol Rev* 77, 731-758. 10.1152/physrev.1997.77.3.731.

Saba, J.A., Liakath-Ali, K., Green, R., and Watt, F.M. (2021). Translational control of stem cell function. *Nat Rev Mol Cell Biol*. 10.1038/s41580-021-00386-2.

Sampath, P., Pritchard, D.K., Pabon, L., Reinecke, H., Schwartz, S.M., Morris, D.R., and Murry, C.E. (2008). A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2, 448-460. 10.1016/j.stem.2008.03.013.

Santamaria, D., Barriere, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Caceres, J.F., Dubus, P., Malumbres, M., and Barbacid, M. (2007). Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* 448, 811-815. 10.1038/nature06046.

Schwartz, D., and Gygi, S.P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* 23, 1391-1398. 10.1038/nbt1146.

Shuda, M., Velasquez, C., Cheng, E., Cordek, D.G., Kwun, H.J., Chang, Y., and Moore, P.S. (2015). CDK1 substitutes for mTOR kinase to activate mitotic cap-dependent protein translation. *Proc Natl Acad Sci U S A* 112, 5875-5882. 10.1073/pnas.1505787112.

Szczerbinska, I., Gonzales, K.A.U., Cukuroglu, E., Ramli, M.N.B., Lee, B.P.G., Tan, C.P., Wong, C.K., Rancati, G.I., Liang, H., Goke, J., et al. (2019). A Chemically Defined Feeder-free System for the Establishment and Maintenance of the Human Naive Pluripotent State. *Stem Cell Reports* 13, 612-626. 10.1016/j.stemcr.2019.08.005.

Taei, A., Rasooli, P., Braun, T., Hassani, S.N., and Baharvand, H. (2020). Signal regulators of human naive pluripotency. *Exp Cell Res* 389, 111924. 10.1016/j.yexcr.2020.111924.

Tahmasebi, S., Amiri, M., and Sonenberg, N. (2018). Translational Control in Stem Cells. *Front Genet*

9, 709. 10.3389/fgene.2018.00709.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* *158*, 1254-1269. 10.1016/j.cell.2014.08.029.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* *19*, 502-515. 10.1016/j.stem.2016.06.011.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* *15*, 524-526. 10.1016/j.stem.2014.09.003.

Trusler, O., Huang, Z., Goodwin, J., and Laslett, A.L. (2018). Cell surface markers for the identification and study of human naive pluripotent stem cells. *Stem Cell Res* *26*, 36-43. 10.1016/j.scr.2017.11.017.

Tsai, B.P., Jimenez, J., Lim, S., Fitzgerald, K.D., Zhang, M., Chuah, C.T., Axelrod, H., Wilson, L., Ong, S.T., Semler, B.L., and Waterman, M.L. (2014). A novel Bcr-Abl-mTOR-eIF4A axis regulates IRES-mediated translation of LEF-1. *Open Biol* *4*, 140180. 10.1098/rsob.140180.

Tsogtbaatar, E., Landin, C., Minter-Dykhouse, K., and Folmes, C.D.L. (2020). Energy Metabolism Regulates Stem Cell Pluripotency. *Frontiers in cell and developmental biology* *8*, 87. 10.3389/fcell.2020.00087.

Wang, X.Q., Lo, C.M., Chen, L., Ngan, E.S., Xu, A., and Poon, R.Y. (2017). CDK1-PDK1-PI3K/Akt signaling pathway regulates embryonic and induced pluripotency. *Cell Death Differ* *24*, 38-48. 10.1038/cdd.2016.84.

Wang, Y.C., Peterson, S.E., and Loring, J.F. (2014a). Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res* *24*, 143-160. 10.1038/cr.2013.151.

Wang, Z., Fan, M., Candas, D., Zhang, T.Q., Qin, L., Eldridge, A., Wachsmann-Hogiu, S., Ahmed, K.M., Chromy, B.A., Nantajit, D., et al. (2014b). Cyclin B1/Cdk1 coordinates mitochondrial respiration for cell-cycle G2/M progression. *Dev Cell* *29*, 217-232. 10.1016/j.devcel.2014.03.012.

Ward, E., Twaroski, K., and Tolar, J. (2017). Feeder-Free Derivation of Naive Human Pluripotent Stem Cells. *Stem cells and development* *26*, 1087-1089. 10.1089/scd.2017.0067.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naive human embryonic stem cells. *Proc Natl Acad Sci U S A* *111*, 4484-4489. 10.1073/pnas.1319738111.

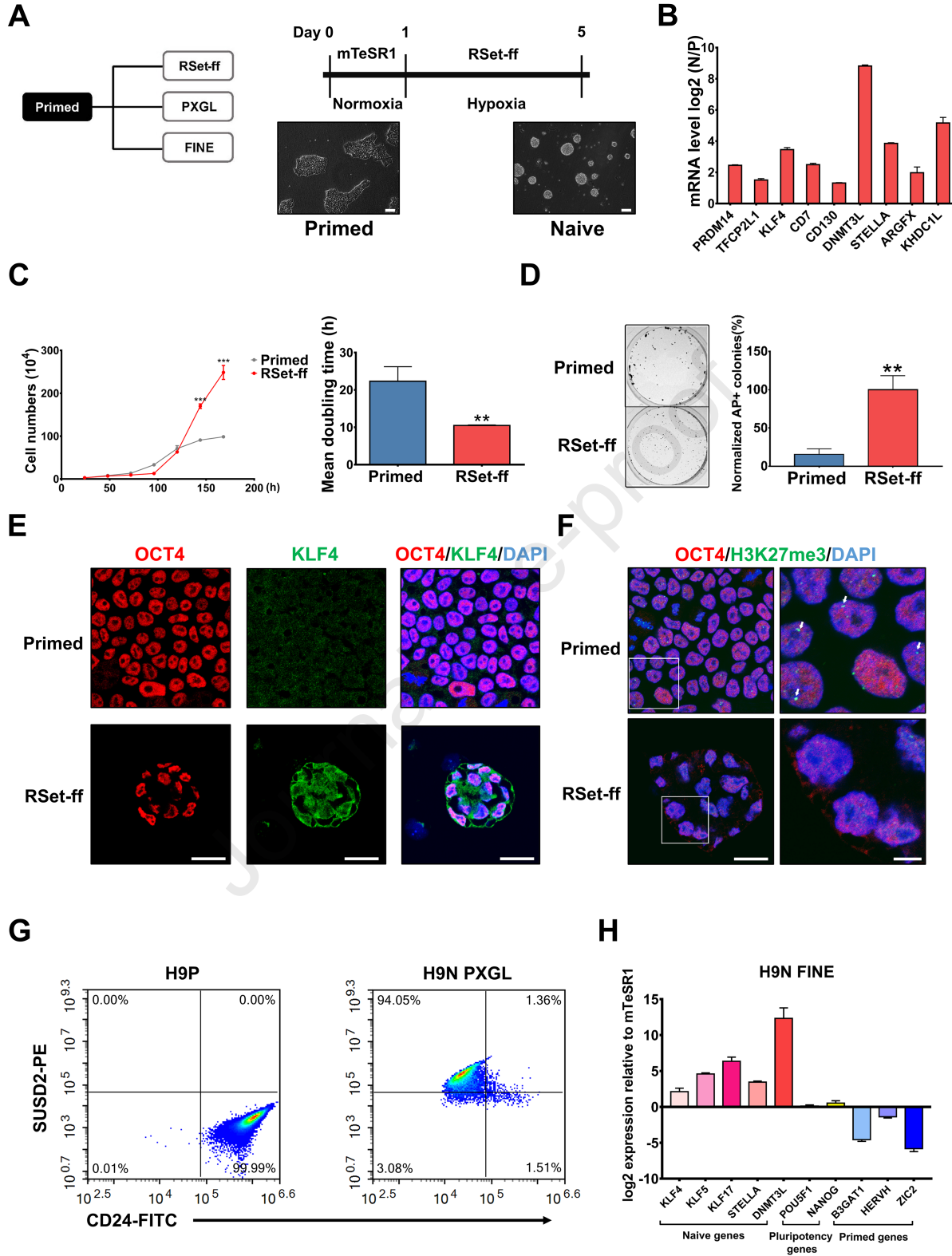
Ware, C.B., Wang, L., Mecham, B.H., Shen, L., Nelson, A.M., Bar, M., Lamba, D.A., Dauphin, D.S., Buckingham, B., Askari, B., et al. (2009). Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. *Cell Stem Cell* *4*, 359-369. 10.1016/j.stem.2009.03.001.

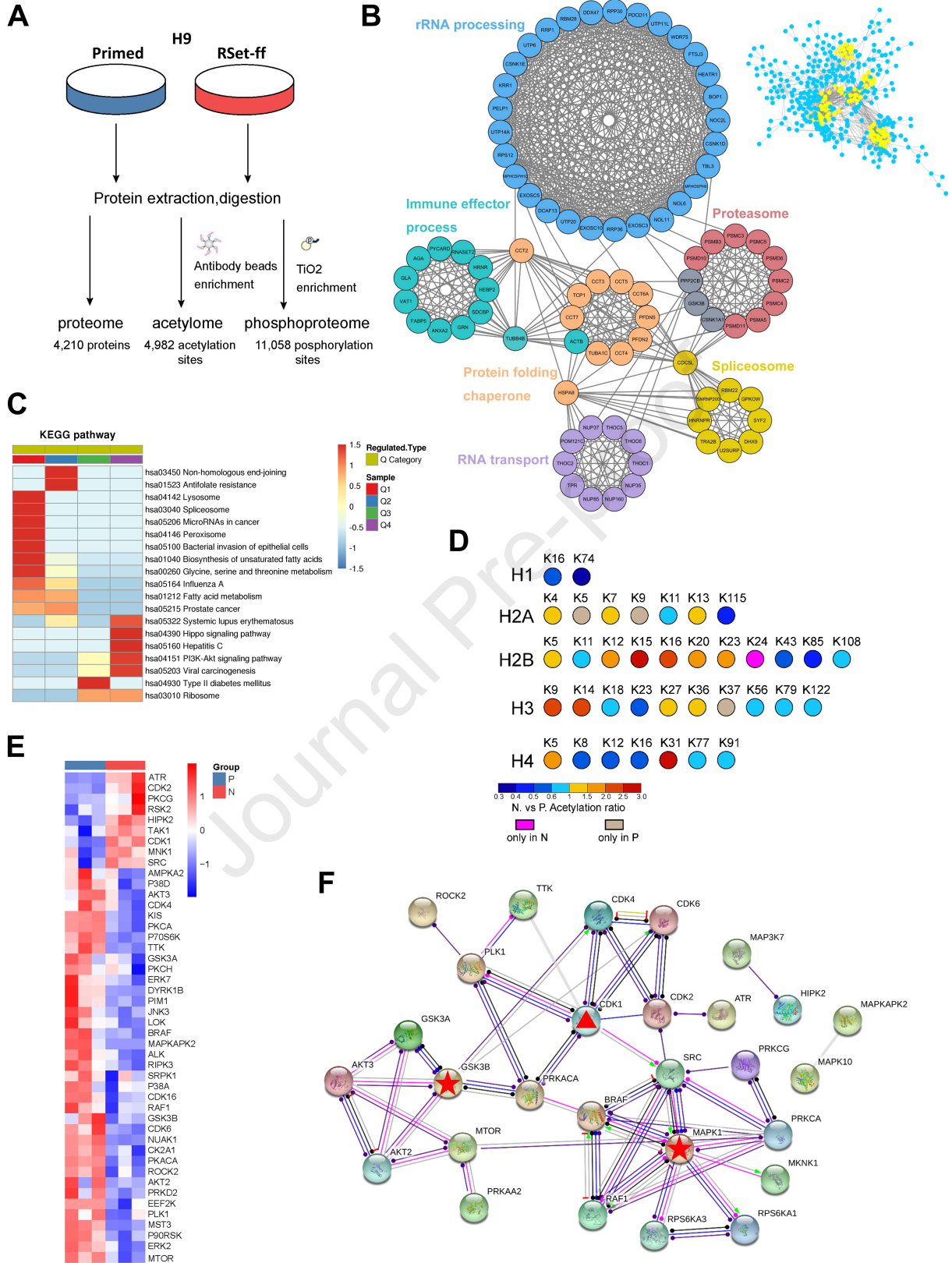
Warrier, S., Van der Jeught, M., Duggal, G., Tilleman, L., Sutherland, E., Taelman, J., Popovic, M., Lierman, S., Chuva De Sousa Lopes, S., Van Soom, A., et al. (2017). Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat Commun* *8*, 15055. 10.1038/ncomms15055.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol* *17*, 155-169. 10.1038/nrm.2015.28.

Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A.A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A., and Segal, E. (2016). Comparative genetics. Systematic discovery of cap-independent

- translation sequences in human and viral genomes. *Science* 351. 10.1126/science.aad4939.
- Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat Methods* 6, 359-362. 10.1038/nmeth.1322.
- Wong, Q.W., Vaz, C., Lee, Q.Y., Zhao, T.Y., Luo, R., Archer, S.K., Preiss, T., Tanavde, V., and Vardy, L.A. (2016). Embryonic Stem Cells Exhibit mRNA Isoform Specific Translational Regulation. *PLoS One* 11, e0143235. 10.1371/journal.pone.0143235.
- Wu, J., and Izpisua Belmonte, J.C. (2016). Stem Cells: A Renaissance in Human Biology Research. *Cell* 165, 1572-1585. 10.1016/j.cell.2016.05.043.
- Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC bioinformatics* 10, 232. 10.1186/1471-2105-10-232.
- Xie, B., Wang, S., Jiang, N., and Li, J.J. (2019). Cyclin B1/CDK1-regulated mitochondrial bioenergetics in cell cycle progression and tumor resistance. *Cancer Lett* 443, 56-66. 10.1016/j.canlet.2018.11.019.
- Yanagida, A., Spindlow, D., Nichols, J., Dattani, A., Smith, A., and Guo, G. (2021). Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* 28, 1016-1022 e1014. 10.1016/j.stem.2021.04.031.
- Yang, Y., and Wang, Z. (2019). IRES-mediated cap-independent translation, a path leading to hidden proteome. *J Mol Cell Biol* 11, 911-919. 10.1093/jmcb/mjz091.
- Yu, C.S., Chen, Y.C., Lu, C.H., and Hwang, J.K. (2006). Prediction of protein subcellular localization. *Proteins* 64, 643-651. 10.1002/prot.21018.
- Zismanov, V., Chichkov, V., Colangelo, V., Jamet, S., Wang, S., Syme, A., Koromilas, A.E., and Crist, C. (2016). Phosphorylation of eIF2alpha Is a Translational Control Mechanism Regulating Muscle Stem Cell Quiescence and Self-Renewal. *Cell Stem Cell* 18, 79-90. 10.1016/j.stem.2015.09.020.

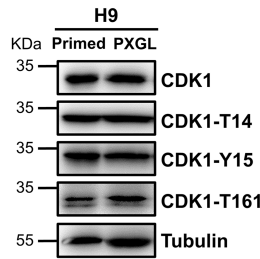




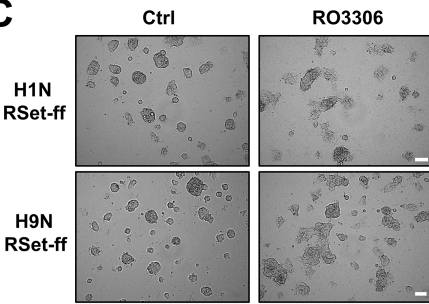
A

Gene name	Phospho-site	Fold change	P-value
HIST1H1E	T18	1.65	0.04
HMGA1	T53	2.83	0.005
LMNB1	S23	1.93	0.02
LIG3	S913	2.77	0.02
SQSTM1	T269	2.87	0.04

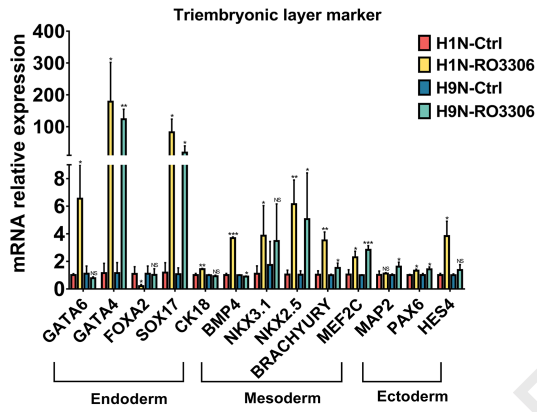
B



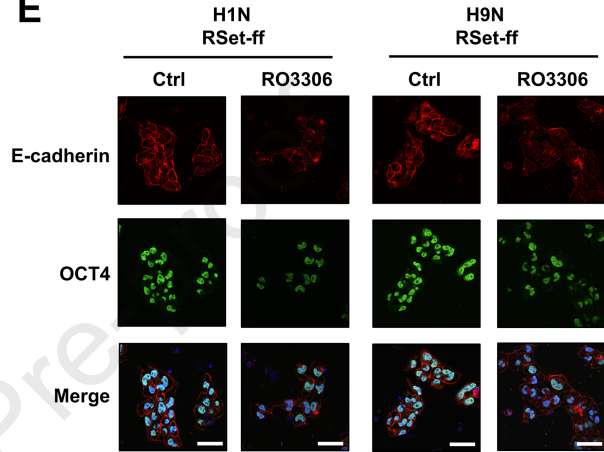
C



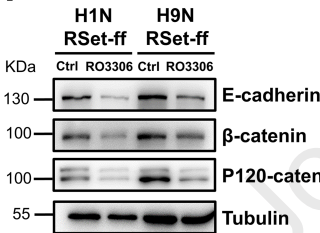
D



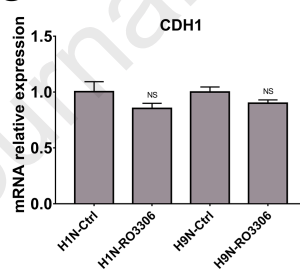
E



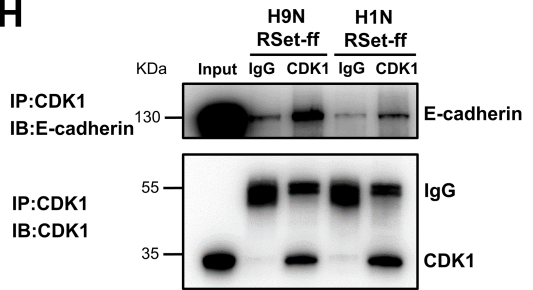
F



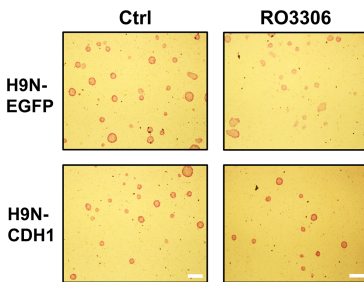
G



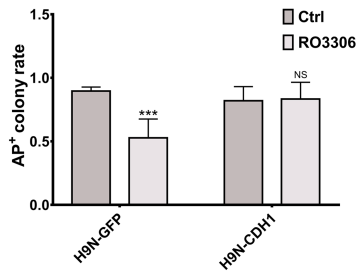
H



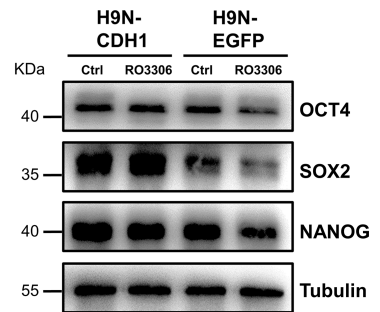
I

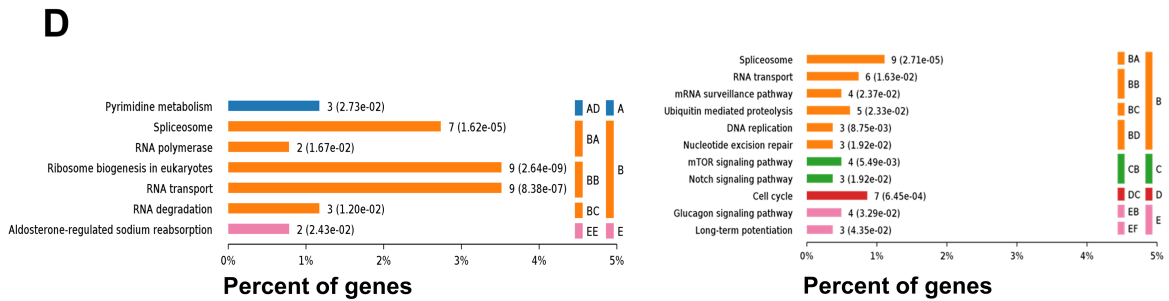
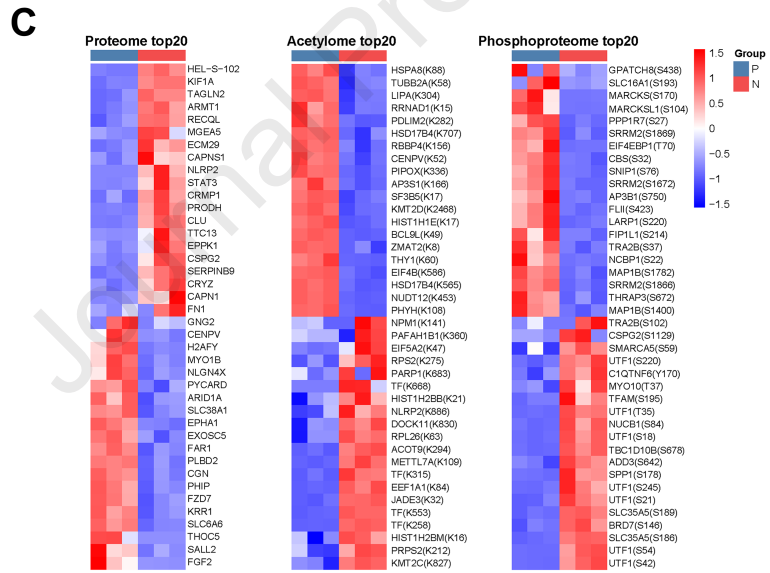
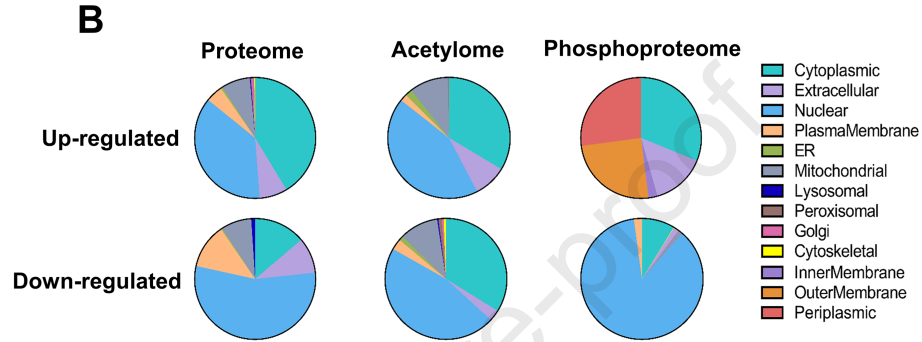
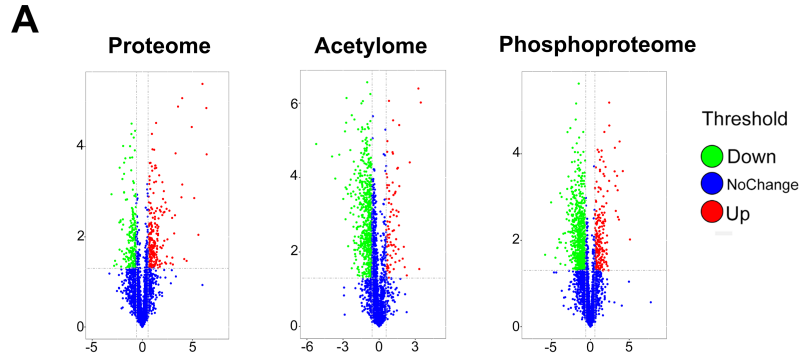


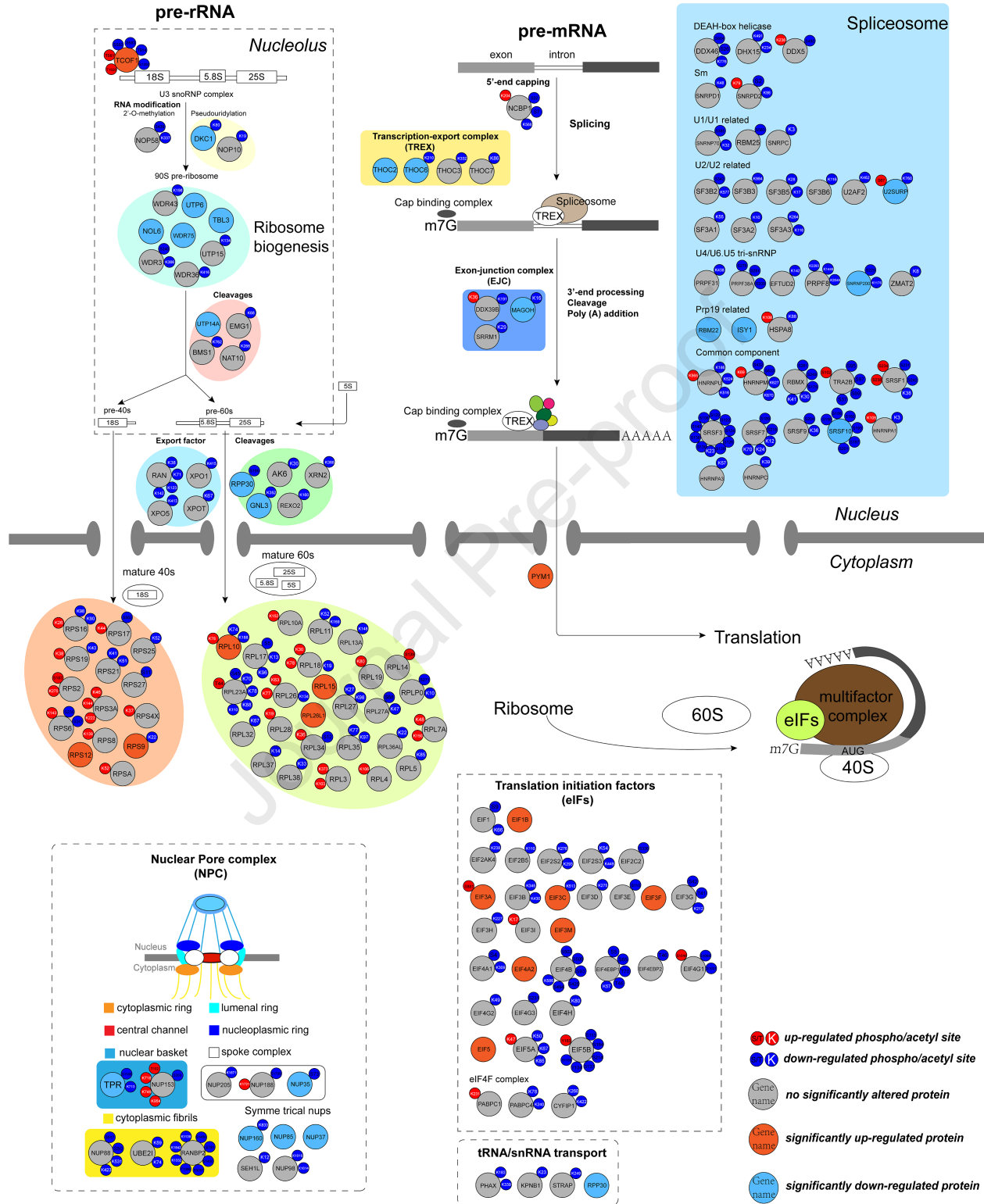
J

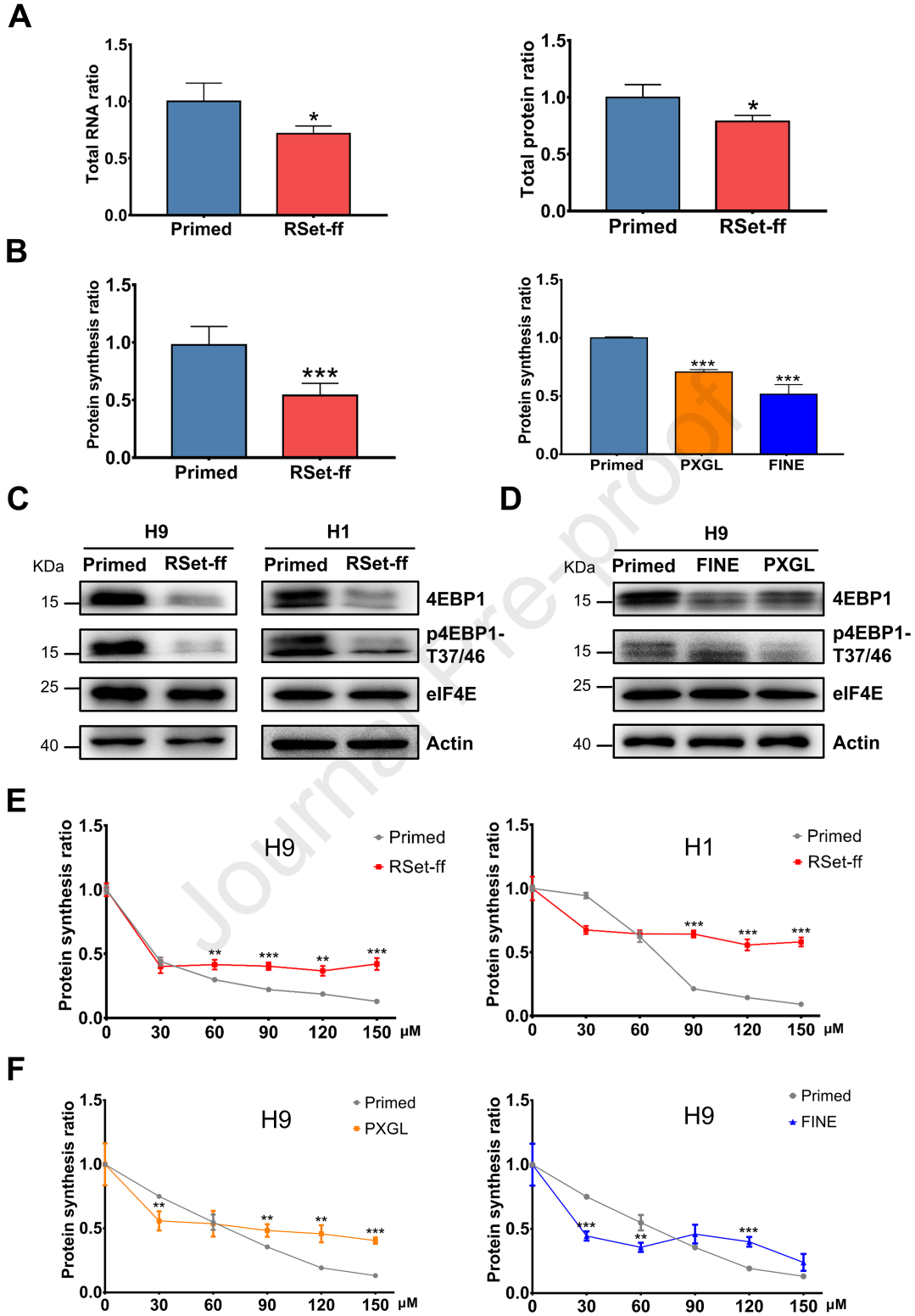


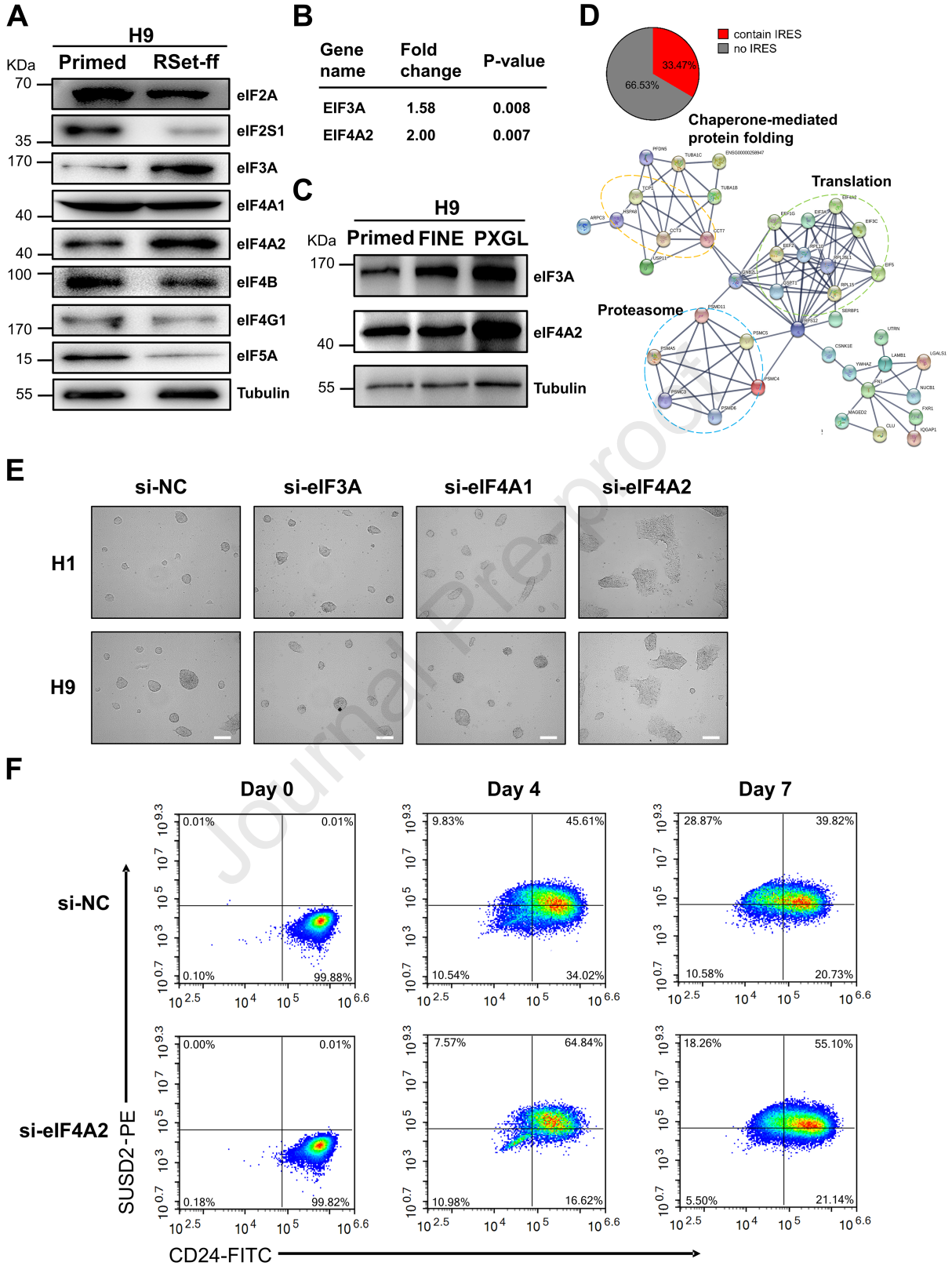
K











HIGHLIGHTS

- RNA processing and translation are most different between naïve and primed hESCs
- Glycolytic primed hESCs mainly rely on eIF4E-dependent translation
- Bivalent metabolism in naïve hESCs promotes eIF4E-independent translation
- CDK1 is required for naïve pluripotency partially by activating E-cadherin signaling

Journal Pre-proof