ORIGINAL RESEARCH

Npac Is A Co-factor of Histone H3K36me3 and Regulates Transcriptional Elongation in Mouse Embryonic Stem Cells

Sue Yu1,#, Jia Li2,#, Guanxu Ji3, Zhen Long Ng1, Jiamin Siew1, Wan Ning Lo1, Ying Ye4, Yuan Yuan Chew1, Yun Chau Long1, Wensheng Zhang4, Ernesto Guccione5, Yuin Han Loh5, Zhi-Hong Jiang3, Henry Yang2,* , Qiang Wu1,3,*

1Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore
2Cancer Science Institute of Singapore, Centre for Translational Medicine, Singapore 117599, Singapore
3The State Key Laboratory of Quality Research in Chinese Medicine, Macau University of Science and Technology, Macau Special Administrative Region 999078, China
4Cam-Su Genomic Resource Center, Soochow University, Suzhou 215123, China
5Institute of Molecular and Cell Biology, Singapore 138673, Singapore

Received 19 October 2019; revised 16 July 2020; accepted 15 August 2020
Available online 4 March 2021
Handled by Jiang Liu

KEYWORDS
Npac;
Pluripotency;
Reprogramming;
Histone H3K36me3;
Transcriptional elongation

Abstract Chromatin modification contributes to pluripotency maintenance in embryonic stem cells (ESCs). However, the related mechanisms remain obscure. Here, we show that Npac, a “reader” of histone H3 lysine 36 trimethylation (H3K36me3), is required to maintain mouse ESC (mESC) pluripotency since knockdown of Npac causes mESC differentiation. Depletion of Npac in mouse embryonic fibroblasts (MEFs) inhibits reprogramming efficiency. Furthermore, our chromatin immunoprecipitation followed by sequencing (ChIP-seq) results of Npac reveal that Npac co-localizes with histone H3K36me3 in gene bodies of actively transcribed genes in mESCs. Interestingly, we find that Npac interacts with positive transcription elongation factor b (p-TEFb), Ser2-phosphorylated RNA Pol II (RNA Pol II Ser2P), and Ser5-phosphorylated RNA Pol II (RNA Pol II Ser5P). Furthermore, depletion of Npac disrupts transcriptional elongation.

* Corresponding authors.
E-mail: cuiyangh@nus.edu.sg (Yang H), qwu@must.edu.mo (Wu Q).
# Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

https://doi.org/10.1016/j.gpb.2020.08.004
1672-0229 © 2022 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Introduction

Embryonic stem cells (ESCs) derived from the inner cell mass of the early embryo are characterized by self-renewal and pluripotency (i.e., the ability to differentiate into many different cell types) [1,2]. Since ESCs can be cultured indefinitely in vitro, they are a promising resource for regenerative therapy. In particular, ESCs show potential for treating degenerative diseases such as diabetes and Parkinson’s disease [3,4]. Moreover, induced pluripotent stem cells (iPSCs) show enormous potential for the application and progress in gene therapy and regenerative medicine [5–8]. Therefore, enhanced understanding of the molecular mechanisms regulating ESC identity would be of great value toward developing ESC- and iPSC-based therapies.

Transcription factors Oct4 (encoded by the Pou5f1 gene), Sox2, and Nanog constitute the core transcriptional network that activates genes that promote pluripotency and self-renewal and inhibits genes that promote differentiation [9–12]. Yamanaka’s discovery that the combination of transcription factors Oct4, Klf4, Sox2, and c-Myc (OKSM) is sufficient to reprogram terminally differentiated cells to pluripotent stem cells has further proved the importance of those core transcription factors [7]. Aside from these, many other transcription factors are essential for pluripotency [13–16].

Besides transcription factors, chromatin regulators also contribute to mouse ESC (mESC) pluripotency through providing the necessary environment for proper gene expression [17]. Recently, a handful of chromatin regulators that are critical for ESC pluripotency have been characterized. ESCs contain structurally relaxed and transcriptionally permissive chromatin that allows for epigenetic remodeling [18]. However, factors that modify the epigenetic configuration are not completely known. Lysine-trimethylation modifications at histone H3 are the most stable epigenetic marks on histones. ESCs are featured by a higher level of histone H3 lysine 4 trimethylation (H3K4me3), which is generally correlated with gene activation [19]. Conversely, H3K27me3 and H3K9me3 are related to gene silencing and heterochromatin in ESCs [20]. However, few studies have examined the regulation of histone H3K36me3 in ESCs [21,22]. Histone H3K36me3 marks active genes, preferentially occupies exons and introns (gene bodies) [23], and is considered as a mark of transcriptional elongation. Recently, a large-scale methyl lysine interactome study discovered proteins that bind to specific histone marks [24]. Interestingly, all proteins that bind to histone H3K36me3 have a common proline-tryptophan-tryptophan-proline (PWWP) domain. This and other studies [25–27] suggest the essential role of the PWWP domain in binding to histone H3K36me3.

Npac (also known as NP60 and Glyr1) containing a PWWP domain is one of the proteins that can bind to histone H3K36me3 [24]. A chromatin immunoprecipitation followed by sequencing (ChIP-seq) analysis on human chromosome 22 revealed that both histone H3K36me3 and Npac are exclusively localized in gene bodies [24], suggesting that Npac may function in transcriptional elongation. Additionally, Npac is a co-factor of lysine-specific demethylase 2 (LSD2) which mediates histone H3K4 demethylation [28–30]. These findings suggest that Npac regulates gene expression through interacting with specific histone modifications. However, how Npac plays its roles is largely unknown.

In this study, we found that Npac is required to maintain mESC pluripotency. Depletion of Npac leads to mESC differentiation with loss of pluripotency. Depletion of Npac also reduces the reprogramming efficiency of mouse embryonic fibroblasts (MEFs). We observed that Npac positively regulates pluripotency genes such as Oct4/Pou5f1 and Nanog. Npac may prevent mESC differentiation by repressing the MAPK/ERK pathway. Furthermore, ChIP-seq experiments showed that Npac co-localizes with histone H3K36me3 in the body of genes which are actively transcribed in mESCs. Npac interacts with RNA Pol II (including Ser2- and Ser5-phosphorylated RNA Pol II; RNA Pol II Ser2P and Ser5P) and positive transcription elongation factor b (p-TEFb), and Npac depletion causes transcriptional elongation defects of Nanog and Rif1. Together, these results establish that Npac maintains mESC pluripotency and regulates transcriptional elongation in mESCs.

Results

Npac is required for maintenance of mESC pluripotency

To test whether Npac is associated with ESC pluripotency, we induced mESCs to differentiate by using ES medium without leukemia inhibitory factor (LIF). We observed that the Npac mRNA level was decreased during differentiation, dropping to around 40% at 5 days after LIF removal (Figure 1A).

Next, we depleted Npac by RNAi to determine the role of Npac in ESC pluripotency. Transfection of mESCs with shRNA plasmids (Npac RNAi-1 or RNAi-2) targeting Npac significantly reduced the level of Npac mRNA (Figure 1B, Figure S1A). We observed that transfection with Npac RNAi-1 significantly reduced the mRNA levels of ESC pluripotency genes Oct4/Pou5f1 and Nanog (Figure 1B). Additionally, the protein levels of Oct4 and Nanog, as well as the H3K36me3 level, were reduced in Npac RNAi-1 transfected ESCs (Figure 1C). The mRNA expression levels of Oct4/Pou5f1 and Nanog were also significantly down-regulated in Npac RNAi-2 transfected ESCs (Figure S1A). The reduction of Oct4 and Nanog upon Npac knockdown (KD) suggested that Npac depletion may cause loss of ESC pluripotency, since Oct4 and Nanog are master regulators required to maintain ESC pluripotency [13–15,31]. This was further supported by the evidence that certain lineage marker genes were up-regulated upon Npac depletion: the trophoderm marker Cdx2 showed a 2-fold increase; the endoderm markers Foxa2, Gata6, and Vegfr2 displayed 1.7-, 1.5-, and 1-fold increase, respectively; the mesoderm markers Nodal, Handl1, and Gata2 increased by 3.2, 3.2, and 4.8 folds, respectively (Figure 1D). Moreover, Npac RNAi-1 transfected ESCs

elocation of the pluripotency genes Nanog and Rif1. Taken together, we propose that Npac is essential for the transcriptional elongation of pluripotency genes by recruiting p-TEFb and interacting with RNA Pol II Ser2P and Ser5P.
showed morphological differentiation and weaker alkaline phosphatase (AP) activity compared to control cells transfected with empty vector, while scrambled RNAi transfected cells appeared to have an AP activity similar to the control (Figure 1E and F), further indicating that Npac-depleted cells underwent differentiation.

To ensure that the Npac RNAi was specific, we performed a Npac RNAi rescue experiment. To construct the Npac RNAi-resistant plasmid, specific primers were designed, and the plasmid with the full-length Npac cDNA inserted in the pCAG-Neo vector was used as the template. We transfected mouse E14 ESCs with the control (empty vector) or Npac RNAi-1 plasmid and selected with puromycin. We then transfected the cells with the Npac RNAi-resistant plasmid for RNAi rescue. The cells were selected with neomycin for 3 days, followed by AP staining. We observed sustained expression levels of the pluripotency genes Oct4/Pou5f1, Nanog, and Sox2 and more AP-positive cells after the rescue treatment (named Npac RNAi-immune OE) than the control cells, showing that Npac RNAi-immune OE cells were resistant to Npac RNAi (Figure S1B and C). We also observed that the changes in AP staining and pluripotency gene expression could only be partially rescued. This is likely due to ESC differentiation caused by Npac RNAi before we transfected the Npac RNAi-resistant plasmid into the cells.

To further confirm the important role of Npac in pluripotency maintenance, we generated embryoid bodies (EBs) from Npac-depleted cells and control cells transfected with empty vector. We then cultured the EBs without LIF in low-attachment dishes for 14 days. EBs partially mimic in vivo embryonic development [32]. We then performed AP staining and found that EBs generated from control cells and Npac RNAi-1 transfected cells both lost pluripotency, while Npac RNAi-1-derived EBs were much smaller than the control group. These results suggest that Npac-depleted EBs grew more slowly than control EBs (Figure S1D). We also performed qRT-PCR to determine the expression levels of lineage markers in EBs generated from control cells and Npac RNAi-1 transfected cells (Figure S1E). We found that the Npac expression level in Npac-depleted EBs was lower than that in control EBs. The expression levels of several mesoderm markers (Hand1, Gata2, and Nkx2.5) were much higher in Npac-depleted EBs than those in control EBs. In addition, endoderm markers (Sox17, Foxa2, and Vegfr2) showed higher expression levels in Npac-depleted EBs than in control EBs. These results suggest that the depletion of Npac may drive ESCs to differentiate into endoderm and mesoderm lineages, which is consistent with the result of Npac KD in ESCs shown in Figure 1D.

Having seen the effect of Npac depletion, we next examined whether overexpression (OE) of Npac affects ESC pluripotency and differentiation. To this end, we performed EB formation assay using Npac-OE cells. After EB induction, EBs at day 7 and day 14 (which mimic early and late development respectively) were collected. We found that Npac was expressed at higher levels in EBs at day 7 (Figure S1F) and day 14 (Figure S1G) in Npac-OE EBs. Interestingly, the expression levels of Oct4 and Nanog in Npac-OE EBs were about 2 folds that in control EBs, suggesting that pluripotency genes were sustained longer in Npac-OE EBs. Also, we found that the size of Npac-OE EBs was larger than that of control group, suggesting that Npac OE may promote EBs to grow faster than in the control group (Figure S1H).

Based on these results, we conclude that Npac is required to maintain ESC pluripotency. On the one hand, Npac depletion represses pluripotency genes and activates lineage marker genes. On the other hand, the expression of pluripotency genes is maintained upon Npac OE in differentiating ESCs.

Reprogramming efficiency of MEFs to iPSCs is reduced upon Npac depletion

Having seen the essential role of Npac in mESC pluripotency, we next tested its role in reprogramming of somatic cells. Pou5f1-GFP MEFs were used to facilitate the identification of putative iPSC colonies based on GFP expression [33]. When MEFs were infected with OKSM along with Npac-KD virus (OKSM + Npac-KD), the Npac relative expression decreased to about 29% of the control infected with OKSM only (Figure 2A). We found that the number of GFP+ colonies produced by OKSM + Npac-KD was 2.5 folds less than the control 14 days after infection (Figure 2B). We also confirmed this by checking the iPSC colonies using AP staining (Figure 2C and D). In addition, we performed immunostaining to examine whether the iPSCs generated by OKSM + Npac-KD were pluripotent. We found that these iPSCs expressed endogenous Oct4 and Nanog, indicating that they were ESC-like (Figure 2E). Furthermore, we generated EBs from GFP+ iPSCs, which were induced by OKSM + Npac-KD. Our immunostaining results showed that these iPSCs could express lineage markers of endoderm (Nestin), mesoderm (SMA), and ectoderm (Gata4) (Figure 2F). These results sug-

---

**Figure 1** Npac is required to maintain mESC pluripotency

A. The Npac mRNA level decreased in mESCs cultured in LIF withdrawal ESC medium. The mRNA levels of Npac and Oct4/Pou5f1 were compared to those in the control cells cultured in normal ESC medium and normalized against Actb. B. The mRNA levels of the pluripotency genes Oct4/Pou5f1, Sox2, and Nanog were significantly decreased upon Npac KD. mESCs transfected with the empty pSUPER.puro vector were used as a control. C. Npac KD resulted in decreased protein levels of Oct4, Sox2, Nanog, and histone H3K36me3. β-actin served as a loading control. D. Npac KD caused up-regulation of the specific markers for endoderm and mesoderm. E. Representative bright field images (upper panel) of E14 cells transfected with control (empty vector), scrambled RNAi, or Npac RNAi-1 followed by 4 days of puromycin selection. AP staining was conducted on day 4 after transfection, and the results were shown at the bottom panel. Scale bar, 100 μm. F. Quantification of AP+ colonies for control (empty vector), scrambled RNAi, or Npac RNAi-1 transfected E14 cells. Data were shown as mean ± SE (n = 3). *, P < 0.05; **, P < 0.01; ****, P < 0.001 (Student’s t-test). ESC, embryonic stem cell; mESC, mouse ESC; LIF, leukemia inhibitory factor; KD, knockdown; H3K36me3, histone H3 lysine 36 trimethylation; AP, alkaline phosphatase.
suggest that iPSCs generated by OKSM + Npac-KD are pluripotent. Thus, Npac is essential not only for the pluripotency maintenance in mESCs but also for the generation of iPSCs.

Depletion of Npac represses pluripotency genes while activating development-related genes

We next investigated how Npac functions in pluripotency maintenance by profiling gene expression following RNAi-induced Npac KD. Upon Npac depletion, the expression levels of 2696 genes were up-regulated (fold change > 1.5), while the expression levels of 891 genes were down-regulated (fold change < 0.67) (Figure 3A). We randomly chose 10 up-regulated and 9 down-regulated genes and confirmed the gene expression microarray results by qRT-PCR (Figure S2A and B).

We next carried out a Gene Ontology (GO) analysis for activated and repressed genes (Figure 3A). Full list of the
A

**Enriched GO terms** (2896 genes up-regulated, fold change > 1.5, P < 0.05)

- **Cell differentiation and Development**
  - Morphogenesis
  - Organ development
  - Muscle development
  - Embryonic development
  - Tube development
  - Organogenesis
  - Cell differentiation

- **Signal transduction**
  - Cell surface receptor linked signal transduction
  - MAPK activity
  - JNK cascade
  - Wnt receptor signaling pathway

- **Cell death and apoptosis**
  - Cell death
  - Apoptosis
  - Programmed cell death

**Enriched GO terms** (891 genes down-regulated, fold change < 0.67, P < 0.05)

- **Chromatin structure**
  - DNA packaging
  - Chromatin assembly or disassembly
  - Maintenance of chromatin architecture
  - Chromatin modification
  - Nucleosome assembly

- **Transcription**
  - Transcription, DNA-dependent
  - Regulation of transcription

- **Cell proliferation**
  - G1/S transition of mitotic cell cycle
  - DNA replication and chromosome cycle
  - DNA replication initiation
  - Mitotic cell cycle

- **Metabolism**
  - Telomere maintenance
  - Stem cell population maintenance
  - RNA splicing

B

C

D

E

F

![Graphical representations and data analysis](image1)

![Graphical representations and data analysis](image2)

![Graphical representations and data analysis](image3)
enriched terms is shown in Table S1. Among the genes down-regulated by Npac KD, enriched terms were related to chromosome modification, suggesting that Npac is required to maintain the unique chromatin structure in mESCs. Notably, we found that a majority of known pluripotency genes were down-regulated upon Npac depletion (Figure 3B). Npac could also play important roles in cell proliferation and telomere maintenance, since the GO terms related to these function groups were significantly enriched for the down-regulated genes. Among the up-regulated genes, many enriched terms were related to development (Figure 3A).

**Npac regulates the MAPK/ERK pathway to influence mESC pluripotency**

Interestingly, many of the up-regulated genes upon Npac depletion were linked to the Wnt and MAPK signaling pathways, which are involved in mESC pluripotency (Figure 3A and C). Nichols et al. have reported that suppression of the MAPK/ERK pathway can contribute to the maintenance of the mESC ground state and pluripotency [34,35]. Also, the ERK pathway promotes mESC differentiation [36]. We found that the levels of ERK1/2 and phosphorylated ERK1/2 (p-ERK1/2) were elevated upon Npac depletion (Figure 3D). Thus, inhibition of the MAPK/ERK pathway by Npac could contribute to the effects of Npac on pluripotency and differentiation.

To explore the role of the MAPK pathway in Npac function, we tested whether inhibiting the MAPK pathway by an ERK inhibitor (PD0325901, Sigma, St. Louis, MO) could rescue the effect of Npac KD. We first confirmed that 50 nM/250 nM ERK inhibitor was able to block the MAPK pathway in E14 ESCs (Figure S3). We found that the ERK inhibitor did not affect Npac KD efficiency (Figure 3E). However, the ERK inhibitor elevated the level of Nanog (Figure 3E), which is in line with the previous finding that inhibition of the ERK pathway up-regulates Nanog in ESCs [37,38]. Furthermore, the ERK inhibitor did not rescue the down-regulation of Oct4/Pou5f1 by Npac depletion (Figure 3E). Also, the ERK inhibitor reduced the expression levels of lineage markers (Figure 3F). Finally, Npac-depleted cells with or without the ERK inhibitor displayed similar differentiated morphology.

Taken together, our results suggest that Npac depletion activates the MAPK/ERK pathway, leading to mESC differentiation. However, since blocking the MAPK/ERK pathway did not rescue the differentiation phenotype and the down-regulation of Oct4/Pou5f1, Npac likely affects pluripotency by other unknown mechanisms as well.

**Npac depletion promotes apoptosis**

We also observed that many genes related to cell death and apoptosis were up-regulated when Npac was knocked down (Figure 4A). To evaluate the effect of Npac depletion on cell death, we performed propidium iodide staining and flow cytometry. Fluorescence-activated cell sorting (FACS) analysis found that the percentage of cells in the sub-G1 phase significantly increased in the Npac-depleted cells compared to the control group (Figure 4B and C), suggesting that there was a sub-G1 phase arrest in the cell cycle. Furthermore, Annexin V staining assay showed that apoptotic cells increased to 40.3% of the total cell population upon Npac depletion, compared to only 9.9% of the total cell population in the control group (Figure 4D and E). These results indicated that depletion of Npac caused apoptosis.

**Npac is located at gene bodies and co-occupies genomic sites of histone H3K36me3**

Oct4 and Nanog are master regulators in the pluripotency transcriptional network [13]. Since depletion of Npac down-regulates Oct4/Pou5f1 and Nanog, we tested whether Npac is located at Oct4/Pou5f1 and Nanog promoters using chromatin immunoprecipitation (ChIP). Interestingly, we found enrichment of Npac in the introns and exons (here defined as the gene body) of Nanog but not in the promoter (Figure 5A and B). Also, we did not observe any enrichment of Npac within the promoter of Oct4/Pou5f1 (Figure S4). Similarly, we also found enrichment of Npac in the gene bodies of other pluripotent genes such as Tcf15, Prdm14, and Tcl1 (Figure 5C).

To determine the genome-wide distribution of Npac in ESCs, we conducted a ChIP-seq experiment using anti-Npac antibody. We identified 12,414 potential genomic binding sites of Npac, where 2416 genes were mapped. Specifically, 57.24% of these binding sites were located within gene bodies. Additionally, 41.95% of the sites were within transcription termination sites (TTSs), followed by 0.66% and 0.15% mapped to intergenic regions and transcription start sites (TSSs), respectively.

**Figure 3** Changes of global gene expression upon Npac depletion in mESCs

A. Microarray heatmap generated from the relative gene expression levels. Relative highly expressed genes in green. Npac was knocked down in E14 cells and selected for 96 h. Then whole-genome cDNA microarray hybridization was performed. Duplicates were chosen to ensure the reproducibility of results. GO analysis was performed relating to “biological process” for the up- or down-regulated genes, respectively. The enriched terms were classified into several function groups and listed in the figure. B. Heatmap of the down-regulated pluripotency genes upon Npac KD in mESCs. Genes were selected according to their known functions in pluripotency. Each selected gene was taken as individual tiles from the thumbnail-dendrogram duplicates. C. Heatmap of the up-regulated MAPK pathway-related genes upon Npac KD in mESCs. Genes were selected according to their known functions in the MAPK pathway. D. The protein levels of p-ERK1/2 and ERK1/2 were elevated in Npac-depleted cells as compared to the control (empty vector) cells. β-actin served as a loading control. E. ERK inhibitor (PD0325901, Sigma) triggered elevated expression of Nanog but did not rescue the down-regulated expression of Oct4 upon Npac KD. F. ERK inhibitor slightly brought down the up-regulated lineage markers in Npac-KD cells. GO, Gene Ontology; p-ERK1/2, phosphorylated ERK1/2.
respectively (Figure 5D). GO analysis showed that the genes that Npac binds to were linked to development, transcription, chromatin modification, cell cycle, and RNA processing (Table S2).

Since Npac is a co-factor of LSD2 which demethylates histone H3K4me1 and H3K4me2, we were also interested in the relationship between Npac and histone H3K4me2. Indeed, we found that the genome-wide profile of Npac localization was

---

Figure 4  Npac depletion may cause cellular apoptosis
A. Heatmap of up-regulated cell death-related genes upon Npac KD in mESCs. Genes were selected according to their known functions in cell death. B. Cell cycle analysis by flow cytometry in the Npac RNAi-1 transfected cells and the control (empty vector) group. C. The representative flow cytometry pattern is shown. D. Apoptosis triggered by Npac KD was analyzed by Annexin V staining through flow cytometry. E. Graphical representation of the apoptosis cells detected by Annexin V staining. Data were shown as mean ± SE (n = 3). *, P < 0.05 (Student’s t-test).
A

B

C

D

E

F

G

H
inversely correlated to that of histone H3K4me2 (Figure S5A). We were also keen to explore whether Npac is linked to other histone modifications. Here, we chose several important histone modifications (histone H3K9me3, histone H3K27me3, and histone H3K4me3) (Figure S5B) and their respective modifiers (Eset, Ezh2, and MLL2) (Figure S5C), as well as ESC-enriched transcription factors (Oct4, Nanog, and Sox2) (Figure S5D), to compare their distribution patterns with that of Npac. We found that Npac binding profile displayed a unique pattern compared to the epigenetic modifiers, histone modifications, and ESC-enriched transcription factors we selected. In general, the number of Npac-associated genes (most are active genes) is much less than that of other modifiers (Eset, Ezh2, and MLL2). Furthermore, Npac, unlike Eset, Ezh2, and MLL2 (mainly located at TSSs), is enriched in gene bodies and 3' ends. Thirdly, Npac shares some genomic loci with the master transcription factors Oct/Nanog/Sox2, but the genomic locations of these three transcription factors are different from that of Npac.

Npac is a putative reader of histone H3K36me3, together with which Npac is present almost exclusively over gene bodies [24]. We found that, in mESCs, the genome-wide distribution of Npac resembled that of histone H3K36me3: both of Npac and H3K36me3 were enriched at active genes in E14 ESCs, where they displayed absence from the TSSs, but gradually increased from gene bodies to TTSs, while they had low or even no binding at inactive genes in E14 ESCs (Figure 5E). We further divided genes into four groups (high, middle, low, and no) according to their expression levels. The results also showed similar genome-wide distribution between Npac and H3K36me3, both of which displayed highest binding at genes with high expression, but lower binding at genes with lower expression (Figure 5F). This result further indicated that Npac and H3K36me3 were enriched in actively transcribed genes in E14 ESCs. Indeed, we observed that both histone H3K36me3 and Npac had high occupancies in actively transcribed genes, such as the housekeeping gene Actb, the pluripotency genes Nanog, Nucleolin, and Tcl1, and the telomere maintenance-related genes Rif1, Terf1, and Rpa2 (Figure 5G and H). On the other hand, we observed clearly low Npac and histone H3K36me3 occupancies on inactive genes. These genes included developmental genes (Figure S6A), MAPK pathway-related genes (Figure S6B), and cell death-related genes (Figure S6C). Taken together, these results show that Npac co-localizes with histone H3K36me3 in gene bodies of active genes in mESCs, suggesting that Npac plays roles in histone H3K36me3-associated cellular functions including gene activation and transcriptional elongation.

**Npac is likely involved in transcriptional elongation**

Next, we examined how Npac is involved in transcriptional elongation. We found that Npac can interact with RNA Pol II (Figure 6A). This result is in line with a previous report that the LSD2 complex may include RNA Pol II and Npac [28]. We also found that Npac can interact with RNA Pol II Ser2P and Ser5P (Figure 6B and C). In addition, we found that the phosphorylation levels of Ser2 and Ser5 were down-regulated upon Npac depletion, while RNA Pol II expression was not affected (Figure 6D). This result suggests that the interactions of Npac with RNA Pol II Ser2P and Ser5P may affect their phosphorylation levels and functions. Given that Ser5 phosphorylation of RNA Pol II is associated with transcriptional initiation and early elongation while Ser2 phosphorylation correlates with transcriptional elongation [39], we propose that Npac affects transcriptional elongation through associating with RNA Pol II Ser2P and Ser5P. In order to determine whether Npac is essential for RNA Pol II-mediated transcriptional elongation in mESCs, we performed ChIP with RNA Pol II, RNA Pol II Ser5P, and RNA Pol II Ser2P in Npac-depleted cells and control cells. We performed ChIP-qPCR for the gene bodies of two pluripotency genes Nanog and Rif1, and Utrn, a gene up-regulated in Npac-depleted cells. We found that the presence of RNA Pol II and RNA Pol II Ser2P at the gene bodies of Nanog (Figure 6E and G) and Rif1 (Figure 6F and H) was significantly reduced in Npac-depleted cells, while their presence at Utrn (Figure S7A and B) was not significantly changed. In addition, the level of H3K36me3 was also reduced at the gene bodies of Nanog (Figure 6I) and Rif1 (Figure 6J) in Npac-depleted cells. However, the binding of RNA Pol II Ser5P at the gene bodies of Nanog (Figure S7C) and Rif1 (Figure S7D) was similar between the Npac-depleted and control cells, suggesting that RNA Pol II Ser5P binding is independent of Npac. Taken together, these results suggest that Npac promotes transcriptional elongation but does not affect transcriptional initiation.

---

**Figure 5**  Npac is mainly located to gene bodies and its genome-wide distribution resembles that of histone H3K36me3

A. Schematic diagram of the structure of the Nanog gene. Black boxes represent exons; solid lines represent introns; dashed line represents the promoter; gray boxes at the bottom represent the primers designed at specific areas of the Nanog gene. B. Npac is associated with the Nanog gene with high enrichment fold at its gene body. C. Npac is also associated with the gene bodies of other pluripotency genes including Tcl1, Prdm14, and Tell. D. Genome-wide distribution of Npac in mESCs. E. Genome-wide distribution of Npac resembles that of histone H3K36me3. H3K36me3_express and Npac_express represent the genome-wide distributions of H3K36me3 and Npac in expressed genes in E14 cells, respectively; H3K36me3_non-express and Npac_non-express represent the genome-wide distributions of H3K36me3 and Npac in non-expressed genes in E14 cells, respectively. F. Genome-wide distributions of Npac and H3K36me3 in genes with different expression levels (high, middle, low, and no). H3K36me3_high/middle/low/no represent the genome-wide distributions of H3K36me3 in genes with high, middle, low, and no expression in mESCs, respectively. Npac_high/middle/low/no represent the genome-wide distributions of Npac in genes with high, middle, low, and no expression in mESCs, respectively. Each gene body is represented from 0% (TSS) to 100% (TTS). G. Npac and H3K36me3 ChIP-seq peaks at the gene bodies of the housekeeping gene (Actb) and pluripotency genes (Nucleolin, Nanog, and Tcl1) in mESCs. H. Npac and H3K36me3 ChIP-seq peaks at the gene bodies of the telomere maintenance-related genes (Rif1, Terf1, and Rpa2) in mESCs. In (G) and (H), arrows denote TSS and transcription orientation. *, P < 0.05 (Student's t-test). TSS, transcription start site; TTS, transcription termination site.
Npac associates with p-TEFb to promote transcriptional elongation

Next, we observed that Npac could interact with p-TEFb, which is composed of Cyclin T1 and Cyclin-dependent kinase 9 (CdK9) (Figure 7A and B). p-TEFb can phosphorylate the C-terminal domain (CTD) of the large subunit of RNA Pol II, thus promoting transcriptional initiation and elongation [40]. Thus, Npac can act as an essential component of the elongation complex. To test whether Npac is required for transcriptional elongation, we performed an elongation recovery assay to measure the recovery of transcription at different positions of two genes (Nanog and Rif1). We first incubated Npac RNAi or control (empty vector) transfected cells with 100 μM 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a widely used elongation inhibitor [41], for 3 h. Then, the cells were washed twice with PBS and cultured with fresh medium without DRB before total RNA was isolated every 5 min (Figure 7C) [37]. We first confirmed that the Npac KD efficiency was not affected by the addition of DRB in the Npac-depleted cells; about 40% Npac mRNA was detected compared to the control group at each time point after DRB treatment (Figure 7D). Next, we examined the transcripts from Nanog at different positions (exon 1 and exon 4) after release from the elongation block (Figure 7E). Following the release, the recovery of transcriptional output at exon 4 of Nanog was significantly reduced by Npac RNAi (Figure 7F), while the recovery of transcriptional output at exon 1 of Nanog was not significantly affected (Figure 7G). We next examined the transcripts from Rif1 at exon 2 and exon 30 after release from elongation block (Figure 7H). Similarly, we observed that the recovery of transcriptional output at exon 2 of Rif1 was not significantly changed in Npac-depleted cells (Figure 7I). However, the recovery of transcriptional output at exon 30 of Rif1 was largely decreased compared to the control group (Figure 7J). Taken together, these results suggest that Npac depletion causes transcriptional elongation defects of Nanog and Rif1.

Discussion

ESC pluripotency is governed by both genetic and epigenetic mechanisms. Many pluripotency factors including transcription factors and epigenetic regulators have been discovered in ESCs [42]. Our results indicate that Npac is required to maintain pluripotency in mESCs. First, we found that depletion of Npac significantly repressed the expression of master pluripotency factors Oct4 and Nanog. Besides these core factors, the expression levels of many other known ESC pluripotency genes also decreased upon Npac depletion according to our gene expression microarray. Among these, Test1 is specifically expressed in ESCs and required for ESC maintenance [43]. Test1, encoding a co-factor of the Akt1 kinase, is essential for self-renewal of ESCs [44]. Also, KDM5B, a H3K4me3 demethylase, is an activator of ESC self-renewal correlated genes [45]. Second, transient KD of Npac increased the expression of mesoderm and endoderm lineage markers and reduced the AP activity. These results further support the assertion that Npac is required for maintaining ESC in an undifferentiated state. Third, we found that Npac depletion activated the MAPK/ERK signaling pathway (Figure 3A and D). Previous studies have reported that ERK signaling pathway can induce ESC differentiation into all germ layers in vitro [46,47], and activation of ERK represses Nanog expression and causes ESC differentiation into primitive endoderm [48]. Intriguingly, we found that Npac depletion led mESCs to differentiation but the ERK inhibitor (PD0325901, Sigma) did not fully rescue the differentiation phenotype. It is noteworthy that ERK inhibitors can block general ESC differentiation and thus may mask true differentiation defects of Npac-depleted ESCs. Thus, ERK inhibitors might not be specific to rescue the phenotype resulting from Npac depletion. Therefore, though it is possible that reduction of Nanog upon Npac depletion was partially caused by the activation of ERK pathway, this is unlikely to be the sole mechanism. We surmise that Npac depletion also results in changes in chromatin state, RNA binding, and cell metabolism, some of which may be non-reversible. It is highly likely that Npac regulates pluripotency using some other unknown mechanisms which will be interesting to be further explored.

Furthermore, the function of Npac in somatic reprogramming can also verify its essential role in pluripotency. There are several possible ways in which Npac depletion can inhibit the reprogramming process. Reprogramming consists of a set of molecular processes that transform a somatic cell into a pluripotent stem cell. During reprogramming, genes related to differentiated state should be repressed first and markers associated with pluripotency will be activated subsequently. Meanwhile, widespread chromatin remodeling occurs during the whole process [49,50]. Our microarray results showed that Npac depletion activates many development-associated genes.
transcriptional elongation factor; Cdk9, Cyclin-dependent kinase 9; DRB, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole.

We found that Npac co-localizes with the transcriptional elongation mark histone H3K36me3 in gene bodies of actively transcribed genes in mESCs (Figure 5E and F). This is consistent with the finding from a study in human HeLa cells [24]. However, it remains unclear whether recruitment of Npac depends on the localization of histone H3K36me3. Given that Npac predominantly occupies actively transcribed genes (Figure 5G and H), it appears that Npac functions as a transcriptional activator of these actively transcribed genes (such as the pluripotency genes and telomere maintenance-related genes) in mESCs. Moreover, we observed that the global level of RNA Pol II Ser2P was reduced, while the total RNA Pol II was unaffected when Npac was depleted (Figure 6D). There is a possibility that the lower level of phosphorylated RNA Pol II in Npac-depleted cells reduces the elongation process or even blocks transcriptional elongation. Therefore, most of the active genes in mESCs were down-regulated upon Npac depletion. In addition, the binding of RNA Pol II and RNA Pol II Ser2P at the pluripotency genes Nanog and Rif1 was significantly reduced upon Npac depletion, while the binding of RNA Pol II Ser5P at these two genes did not significantly change. These results further confirmed that Npac is required for transcriptional elongation.

In mammalian cells, Ser2 of RNA Pol II can be phosphorylated by Cdk9 (the kinase subunit of p-TEFb), resulting in the transition from transcriptional initiation to productive elongation [56]. According to previous studies, some specific activators, such as DNA- or RNA-bound activators and co-activators, can recruit p-TEFb to transcription units. For example, one chromatin remodeling protein, Brd4, recruits p-TEFb to stimulate RNA Pol II-dependent transcription [57]. Given the association of Npac with p-TEFb observed in our study (Figure 7A and B), it is possible that Npac recruits p-TEFb to chromatin and further results in the successful transcriptional elongation of target genes. This is in concert with the observation that Npac depletion caused the transcriptional elongation defects of the pluripotency genes Nanog and Rif1. However, this may not be the sole reason. Lower enrichment of histone H3K36me3 and RNA Pol II Ser2P upon Npac depletion may also contribute to the transcriptional elongation defects. Taken together, these findings imply an essential role of Npac in the elongation process.

Interestingly, according to our Npac gene expression microarray results, we found more up-regulated genes than down-regulated ones upon Npac depletion. This appears to be contrary to the fact that Npac is associated with actively transcribed genes. However, we think that Npac depletion in mESCs results in differentiation and triggers a significant up-regulation of abundant developmental genes which are
silenced in undifferentiated mESCs. During the differentiation process, active pluripotency genes become inactivated; meanwhile, silenced developmental genes are activated. Since the activated developmental genes are more than the inactivated pluripotency genes, there are more up-regulated genes than down-regulated genes upon Npac depletion. Furthermore, the activation of the MAPK pathway caused by Npac depletion will trigger many differentiation-related genes. Lastly but not least, the up-regulation of developmental genes in Npac-depleted cells is probably independent of the Npac-mediated transcription elongation, and it could be triggered by ESC differentiation. Nevertheless, given that Npac may have diverse functions, it is of interest to further explore how Npac plays its role in gene regulation during ESC differentiation.

In summary, we propose a model in which Npac regulates mESC pluripotency and influences transcriptional elongation by interacting with p-TEFβ, RNA Pol II Ser2P, and RNA Pol II Ser5P (Figure 8A and B).

**Materials and methods**

**Cell culture**

In this study, mouse E14 ESCs (Catalog No. CRL-1821, ATCC, Manassas, VA), SNL feeder cells (Catalog No. CBA-316, Cell Biolabs, San Diego, CA), Platinum-E cells (Plat-E; Catalog No. RV-101, Cell Biolabs), and Oct4-GFP MEFs were cultured at 37 °C in a CO₂ incubator with 5% CO₂ as previously described [58].

**Construction of plasmids**

Plasmids for RNAi-1 and RNAi-2 targeting Npac were designed using the Eurofins MWG Operon siMAX software. Oligonucleotides were inserted into the pSuper.puro vector (Catalog No. VEC-pBS-0008, Oligoengine, Seattle, WA). The primers for Npac OE were designed by the Primer 5 software to amplify the full-length cDNA of mouse Npac/Glyr1 (NM_028720.2), and the PCR product was inserted into the BglII and MluI sites of the pPyCAGIP vector. To construct retrovirus packaging plasmids, the full-length Npac/Glyr1 cDNA was ligated into the MluI and NotI sites of the pMXs plasmid (Catalog No. 18656, Addgene, Watertown, MA). To construct the Npac RNAi-resistant plasmid that produces functional Npac protein but is resistant to Npac RNAi targeting, specific primers were designed with silent mutations in the sequence of the protein coding domain. The Npac RNAi-resistant plasmid was generated according to the manual of Q5 Site-Directed Mutagenesis Kit (Catalog No. E0554S, New England Biolabs, Ipswich, MA). The plasmid with the full-length Npac cDNA inserted in the pCAG-Neo vector was used as a PCR template. The sequences of the primers are shown in Table S3.
Transfection, RNA extraction, reverse transcription, and qRT-PCR

Transfection was conducted using Lipofectamine 2000 (Catalog No. 11668019, Invitrogen, Waltham, MA) according to the manufacturer’s protocol. Cells were selected by 1 μg/ml puromycin for 4 days after transfection. Either protein or RNA was then extracted from the cells. RNA extraction, reverse transcription, and qRT-PCR were performed as previously described [58]. The sequences of the qRT-PCR primers are shown in Table S3.

Gene expression microarray analysis

mESCs (E14) were transfected with Npac-KD plasmid or control plasmid and cultured for 4 days with selection. RNA was then extracted from the cells. Gene expression microarray data were analyzed as previously described [58,59].

ChIP assay and ChIP-seq

ChIP assay and ChIP-seq were conducted as previously described [58,59]. The antibodies used for ChIP were: anti-Npac (Catalog No. 14833-1-AP, Proteintech, Rosemont, IL), anti-histone H3K36me3 (Catalog No. ab9050, Abcam, Cambridge, UK), anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2) (Catalog No. ab5095, Abcam), anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5) (Catalog No. ab140509, Abcam).

Bioinformatics analysis

Npac ChIP-seq reads were mapped to the mouse genome (NCBI37/mm9) using the mapping software Burrows-Wheeler Aligner [60]. After removing duplicate reads, the mapped results identified board peaks with MACS2. For location classification, ChIP-seq peaks were annotated by comparing the locations of all TSSs and TTSs in the mouse genome with Perl scripts. (1–10 kb upstream of the TSS defined as upstream, 1 kb upstream of the TSS to the TSS defined as TSS, the region between the TSS and TTS defined as gene body, TTS to 1 kb downstream of the TSS defined as TTS, 1–10 kb downstream of the TTS defined as downstream).

ChIP-seq data of H3K36me3 (ENCSR000CGR), H3K4me3 (GSM1258237) and its modifier MLL2 (GSM1258241), and anti-Sox2 (Catalog No. sc-99000, Santa Cruz), anti-p-ERK (Catalog No. sc-484, Santa Cruz), and rotated for 1 h at 4 °C. Then, the beads were washed four times with the lysis buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 1% NP-40, 10% glycerol) supplemented with a protease inhibitor (Catalog No. 4693159001, Roche, Basel, Switzerland), and rotated for 1 h at 4 °C. After being precleared by protein G beads (Catalog No. 15920010, Invitrogen) for 2 h at 4 °C, the cell lysate was incubated overnight with beads bound by specific antibodies at 4 °C. Then, the beads were washed four times with the lysis buffer and heated in 2× loading dye for 10 min at 95 °C. The supernatant was used for Western blot with specific antibodies. IgG antibody (Catalog No. 12-371, Chemicon, Temecula, CA) was used as the IP control.

Western blot

Western blot was performed as described [58,59]. The primary antibodies used in this study were: anti-Npac (Catalog No. 14833-1-AP, Proteintech), anti-NP60 (Catalog No. sc-390601, Santa Cruz, Dallas, TX), anti-β-actin (Catalog No. sc-81178, Santa Cruz), anti-Oct4 (Catalog No. sc-8628, Santa Cruz), anti-Nanog (Catalog No. sc-33760, Santa Cruz), anti-Sox2 (Catalog No. sc-99000, Santa Cruz), anti-p-ERK (Catalog No. 4370, Cell Signaling Technology, Danvers, MA), anti-ERK (Catalog No. 137F5, Cell Signaling Technology), anti-histone H3K36me3 (Catalog No. ab9050, Abcam), anti-Pol II (Catalog No. sc-899, Santa Cruz), anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2) (Catalog No. ab5095, Abcam), anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5) (Catalog No. ab140509, Abcam), and anti-rabbit IgG (Catalog No. sc-2028, Santa Cruz), and anti-rabbit IgG (Catalog No. sc-2027, Santa Cruz).

AP staining

AP staining was conducted with the Alkaline Phosphatase Detection Kit (Catalog No. SCR004, Millipore, Burlington, MA) following the manufacturer’s protocol. Axio Observer A1 inverted light microscope (Zeiss, Gottingen, Germany) was used to take pictures for the AP staining results.

Co-immunoprecipitation

Cells were lysed in the lysis buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 1% NP-40, 10% glycerol) supplemented with a protease inhibitor (Catalog No. 4693159001, Roche, Basel, Switzerland), and rotated for 1 h at 4 °C. After being precleared by protein G beads (Catalog No. 15920010, Invitrogen) for 2 h at 4 °C, the cell lysate was incubated overnight with beads bound by specific antibodies at 4 °C. Then, the beads were washed four times with the lysis buffer and heated in 2× loading dye for 10 min at 95 °C. The supernatant was used for Western blot with specific antibodies. IgG antibody (Catalog No. 12-371, Chemicon, Temecula, CA) was used as the IP control.

Retrovirus packaging and infection

Retrovirus packaging and infection were carried out as previously described [58]. Briefly, the pMXs retroviral plasmids or pSUPER.retro.puro plasmids were transfected into Plat-E cells. The cells were selected with 1 μg/ml puromycin (Catalog No. P8833, Sigma) and 10 μg/ml blasticidin (Catalog No. 12576712) was also downloaded from ENCODE. We used the STAR software [61] to carry out RNA-seq mapping against the mm9 genome. By analyzing the mapped RNA-seq data, featureCounts [62] was used to obtain the gene expression of the E14 sample. All genes were further separated into two groups based on whether the genes were expressed or not. Genes were also classed into four groups based on their expression levels (high, middle, low, and no). Expression levels were classified according to the number of reads mapped to the mm9 genome. "Reads = 0" represents "non_express"; "reads > 0" represents "express"; "0 < reads ≤ 10" represents "low expression"; "10 < reads ≤ 100" represents "middle expression"; "reads > 100" represents "high expression". With the respective gene lists and mapped ChIP-seq files, heatmaps and average read distributions were generated with ngsplot [63].
A1113902, Life Technologies, Carlsbad, CA) for 36–48 h. The retroviruses were harvested and concentrated with centrifugal filter units (Catalog No. C7715, Millipore). Pou5f1-GFP MEFs were seeded into 24-well plates for 6 h and then infected with retroviruses. Infected MEFs were seeded onto SNL feeder layers 2 days after infection and cultured with mESC medium without LIF until day 5 after infection. The MEFs were then cultured with KSR medium from day 6 after infection. The number of GFP+ colonies was recorded daily until day 14 after infection. AP staining assays were also conducted at day 14.

**Annexin V-FITC apoptosis assay**

Annexin V-FITC apoptosis assay was carried out as described in the manufacturer’s protocol (APOAF, Sigma). After transfected with the Npac RNAi or control (empty vector) plasmid in 6-well dishes and selected for 4 days, the cells were stained with Annexin V-FITC and propidium iodide. The cells were then analyzed by flow cytometer (BD FACSCanto, BD Biosciences, San Jose, CA).

**Cell cycle analysis**

Cell cycle analysis was conducted as previously described [58]. Briefly, mESCs were transfected with the Npac RNAi plasmid or the control plasmid and selected with puromycin for 4 days. Then, the cells were stained with 50 µg/ml propidium iodide and analyzed by the flow cytometer (BD FACSCanto, BD Biosciences) using Flowing Software 2.5.0.

**Transcription elongation assay**

Transcriptional elongation assay was carried out as previously described [64,65]. E14 cells were transfected with Npac RNAi-1 or control (empty vector). After 24 h, the cells were treated with 100 µM DRB (Catalog No. 287891, Sigma) for 3 h, washed twice with PBS, and cultured in fresh medium for different durations (5–45 min). Total RNA was extracted, and qRT-PCR was performed to quantify the changes in the relative expression levels of different regions along the Rnf1 and Nanog genes. Gene expression levels were normalized against the expression level of Actb. The sequences of used primers are listed in Table S3.

**Statistical analysis**

All experiments were conducted in triplicates. Student’s t-test was applied for statistical analysis, and the results were shown as mean ± SE. *P < 0.05 was considered significant (*, P < 0.05; **, P < 0.01; ***, P < 0.001).

**Data availability**

The Npac RNAi microarray results and Npac ChIP-seq results have been deposited in the NCBI Gene Expression Omnibus (GEO: GSE93296 and GSE95671, respectively), which are publicly accessible at https://www.ncbi.nlm.nih.gov/geo. The results have also been deposited in the Genome Sequence Archive [66] at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy Sciences / China National Center for Bioinformation (GSA: CRA003053), which are publicly accessible at https://ngdc.cncb.ac.cn/gsa.

**CRediT author statement**

**Sue Yu:** Conceptualization, Methodology, Data curation, Writing - original draft. **Jia Li:** Methodology, Data curation, Software. **Guanxu Ji:** Data curation. **Zhen Long Ng:** Data curation. **Jiamin Siew:** Data curation. **Wensheng Zhang:** Methodology. **Ernesto Guccione:** Funding acquisition. **Yuin Han Loh:** Funding acquisition. **Zhi-Hong Jiang:** Resources, Supervision. **Ying Ye:** Methodology. **Yuan Yuan Chew:** Validation. **Yun Chau Long:** Resources, Supervision. **Wensheng Zhang:** Methodology. **Qiang Wu:** Conceptualization, Methodology, Software. **Yuin Han Loh:** Funding acquisition. **Zhi-Hong Jiang:** Resources, Supervision, Project administration. **Henry Yang:** Conceptualization, Methodology, Software. **Wen Sheng Zhang:** Methodology. **Yuan Yuan Chew:** Validation. **Yun Chau Long:** Resources, Supervision. **Ying Ye:** Methodology. **Yuan Yuan Chew:** Validation. **Wensheng Zhang:** Methodology. **Yuehan Li:** Writing - review & editing, Visualization. All authors have read and approved the final manuscript.

**Competing interests**

The authors have declared that no competing interests exist.

**Acknowledgments**

We thank Dr. Takao Inoue for critical reading of the manuscript. We thank Prof. Huck Hui Ng for providing Oct4-GFP MEFs. This work was supported by Singapore National Medical Research Council (Grant No. CBG14n065) and the Macau Science and Technology Development Fund, China (Grant No. FDCT-18-033-SKL-016A).

**Supplementary material**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gpb.2020.08.004.

**ORCID**

ORCID 0000-0003-1838-6506 (Sue Yu)

ORCID 0000-0001-8360-3912 (Jia Li)

ORCID 0000-0002-0384-2890 (Guanxu Ji)

ORCID 0000-0001-8154-474X (Zhen Long Ng)

ORCID 0000-0001-7968-2742 (Jiamin Siew)

ORCID 0000-0001-8778-0457 (Yulin Lo)

ORCID 0000-0002-5028-3112 (Ying Ye)

ORCID 0000-0002-6116-5330 (Yuan Yuan Chew)

ORCID 0000-0001-5692-0381 (Yun Chau Long)

ORCID 0000-0002-6876-3173 (Wensheng Zhang)

ORCID 0000-0001-7764-5307 (Ernesto Guccione)

ORCID 0000-0002-4715-6454 (Yuin Han Loh)

ORCID 0000-0002-7956-2481 (Zhi-Hong Jiang)

ORCID 0000-0002-1155-8216 (Henry Yang)

ORCID 0000-0002-6116-5330 (Yuan Yuan Chew)

ORCID 0000-0002-5028-3112 (Ying Ye)

ORCID 0000-0001-8778-0457 (Wan Ning Lo)

ORCID 0000-0001-7968-2742 (Jiamin Siew)

ORCID 0000-0001-8360-3912 (Jia Li)

ORCID 0000-0002-6876-3173 (Sue Yu)

ORCID 0000-0002-5028-3112 (Guanxu Ji)

ORCID 0000-0001-8154-474X (Zhen Long Ng)

ORCID 0000-0001-7968-2742 (Jiamin Siew)

ORCID 0000-0001-8778-0457 (Yulin Lo)

ORCID 0000-0002-5028-3112 (Ying Ye)

ORCID 0000-0002-6116-5330 (Yuan Yuan Chew)

ORCID 0000-0001-5692-0381 (Yun Chau Long)

ORCID 0000-0002-6876-3173 (Wensheng Zhang)

ORCID 0000-0001-7764-5307 (Ernesto Guccione)

ORCID 0000-0002-4715-6454 (Yuin Han Loh)

ORCID 0000-0002-7956-2481 (Zhi-Hong Jiang)

ORCID 0000-0002-1155-8216 (Henry Yang)

ORCID 0000-0002-6116-5330 (Yuan Yuan Chew)
Ito S, D’Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 2010;466:1129–33.

Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, et al. Dissecting self-renewal in stem cells with RNA interference. Nature 2006;442:533–8.

Xie L, Pelz C, Wang W, Bashar A, Varlamova O, Shadle S, et al. KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. EMBO J 2011;30:1473–84.

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development 2007;134:2895–902.

Hamilton WB, Brickman JM. Erk signaling suppresses embryonic stem cell self-renewal to specify endoderm. Cell Rep 2014;9:2056–70.

Hamazaki T, Kehoe SM, Nakano T, Terada N. The Grb2/Mek pathway represses Nanog in murine embryonic stem cells. Mol Cell Biol 2006;26:7539–49.

Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2008;2:151–9.

Apostolou E, Hochedlinger K. Chromatin dynamics during cellular reprogramming. Nature 2013;502:462–71.

Lin T, Ambasudhan R, Yuan X, Li W, Hilcove S, Abujarour R, et al. A chemical platform for improved induction of human iPS cells. Nat Methods 2009;6:805–8.

Elmore S. Apoptosis: a review of programmed cell death. Toxicol Pathol 2007;35:495–516.

Cagnol S, Chambard JC. ERK and cell death: mechanisms of ERK-induced cell death – apoptosis, autophagy and senescence. FEBBS J 2010;277:2–21.

Fridman JS, Lowe SW. Control of apoptosis by p53. Oncogene 2003;22:9030–40.

Attardi LD, Recek EE, Cosmas C, Demiceco EG, McCurrrach ME, Lowe SW, et al. PERP, an apoptosis-associated target of p53, is a novel member of the PMF-22/gas3 family. Genes Dev 2000;14:704–18.

Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol Cell 2006;23:297–305.

Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell 2005;19:523–34.

Yu S, Ma H, Ow JR, Goh Z, Chiang R, Yang H, et al. Zfp553 is essential for maintenance and acquisition of pluripotency. Stem Cells Dev 2016;25:55–67.

Ma H, Ng HM, Teh X, Li H, Lee YH, Chong YM, et al. Zfp322a regulates mouse ES cell pluripotency and enhances reprogramming efficiency. PLoS Genet 2014;10:e1004038.

Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 2014;30:2843–51.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923–30.

Shen L, Shao N, Liu X, Nestler E. ngs.plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 2014;15:284.

Singh J, Padgett RA. Rates of in situ transcription and splicing in large human genes. Nat Struct Mol Biol 2009;16:1128–33.

Lemieux B, Blanchette M, Monette A, Mouland AJ, Wellinger RJ, Chabot B. A function for the hnRNP A1/A2 proteins in transcription elongation. PLoS One 2015;10:e0126654.

Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, et al. The Genome Sequence Archive Family: toward explosive data growth and diverse data types. Genomics Proteomics Bioinformatics 2021;19:578–83.