

Transcription factor ZNF263 primes human embryonic stem cells for pluripotency dissolution and lineage commitment

Received: 21 May 2024

Accepted: 13 September 2025

Published online: 05 November 2025

 Check for updates

Qianqian Yin^{1,7}, Jing Huang^{1,2,7}, Hongduo Sun^{1,2,7}, Shuhan Zhou^{1,2,7}, Linjie Ruan^{2,3}, Yi Zeng^{1,2}, Yana Li¹, Zhaohui Gong¹, Peiyu Chen^{2,3}, Xinbo Yuan⁴, Jiafang Wan^{1,2}, Qingzhong Wang⁵, Yuangao Wang¹, Haipeng Li¹, Naihe Jing⁶, Xin Liu³ & Zhen Shao¹✉

Conventional human embryonic stem cells (hESCs) are capable of self-renewal and simultaneously poised for differentiation. But the mechanisms underlying this primed pluripotent state, which endows them with elevated responsiveness to differentiation cues, remain largely underexplored. Especially, little is known about the pivotal transcription factors (TFs) that orchestrate hESCs towards primed pluripotency. Here, we report a function of TF ZNF263 in pluripotency priming. Genetic and functional assays reveal that ZNF263 directly initiates the incipient expression of early differentiation genes and concurrently dampens the core pluripotency circuitry in hESCs, greatly tilting the balance from pluripotency maintenance to lineage priming. Importantly, ZNF263 deficiency markedly impairs pluripotency dissolution and multi-lineage differentiation in hESCs, particularly toward ectoderm. Moreover, single-cell transcriptomic profiling reveals that ZNF263 promotes the priming of cell fate commitment in hESCs, suggesting its indispensable requirement for pluripotency priming and lineage commitment continuum. Together, we demonstrate the role of ZNF263 in establishing the primed pluripotent state in hESCs and facilitating their differentiation into primary germ layer lineages.

Pluripotency refers to the cellular potential to differentiate into all three primary germ layers, and it encompasses a dynamic continuum in which cells transition from “naïve” state to “primed” state to prepare for germ layer specification¹. As a prominent class of pluripotent stem cells (PSCs), embryonic stem cells (ESCs) are widely utilized as a model system for studying pluripotency and early mammalian development^{2,3}. Numerous studies have consistently highlighted substantial differences between mouse and human ESCs (mESCs and

hESCs), despite their common derivation from the inner cell mass of preimplantation blastocysts^{4,5}. Generally speaking, mESCs manifest a lot of early developmental features that are characteristic of naïve pluripotency. On the other hand, hESCs, when cultured under conventional conditions, share many similarities with murine post-implantation epiblast stem cells. These similarities include the upregulation of lineage-associated markers ZIC2 and OTX2, as well as the downregulation of naïve pluripotency factors NANOG and

¹Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai, China. ²University of Chinese Academy of Sciences, Beijing, China. ³State Key Laboratory of Epigenetic Regulation and Modification, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China. ⁴School of Medicine, Westlake University, Hangzhou, Zhejiang Province, China. ⁵Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai, China. ⁶Guangzhou National Laboratory, No. 9 XingDaoHuanBei Road, Guangzhou International Bioland, Guangzhou, Guangdong Province, China. ⁷These authors contributed equally: Qianqian Yin, Jing Huang, Hongduo Sun, Shuhan Zhou. ✉e-mail: shaozhen@sinh.ac.cn

PRDM14. Thus, they are categorized into the primed phase^{4,6}. Elucidating the mechanisms establishing the primed pluripotent state in hESCs provides valuable insights into the central regulatory network that controls early human embryo development, particularly the initiation of germ layer formation, and also contributes to the development of novel strategies for regenerative medicine⁷.

It has been demonstrated that core pluripotency factors (CPFs), including POU5F1 (OCT4), SOX2, and NANOG (collectively termed OSN), play a central role in the pluripotency gene regulatory network (PGRN)⁸. Despite significant evolutionary divergence in the genome-wide binding patterns of OSN factors between hESCs and mESCs⁹, they share a common function of positively regulating pluripotency in both species¹⁰. Specifically, these OSN factors form an interconnected autoregulatory circuitry that governs the activation of pluripotency and self-renewal-related genes, while simultaneously suppressing developmental genes¹¹. This circuitry is commonly referred to as the core pluripotency circuitry (CPC). Consistently, several approaches have been developed to reprogram conventional hESCs into a stable naïve pluripotency state, accompanied with the reinforcement of CPC and reactivation of naïve pluripotency-associated genes^{12–14}. Further comparison between the primed and naïve hESCs reveals that genes differentially associated with naïve and primed pluripotency both exhibit dynamic changes in their chromatin state and 3D organization¹⁵. However, the mechanisms underlying these changes remain unclear. Especially, few transcription factors (TFs) have been reported to globally induce the incipient expression of lineage-associated genes in primed hESCs, which strongly suggests that some vital components of human PGRN controlling primed pluripotency establishment have not yet been discovered.

Herein, we perform an in-silico screening based on epigenomic data from various human cell types, and uncover ZNF263 as a TF for human PGRN by its binding capacity at the promoters of genes that are preferentially activated in primed hESCs, which is confirmed by a subsequent chromatin immunoprecipitation sequencing (ChIP-seq) experiment. Further functional analyses reveal that ZNF263, in conjunction with many other TFs, including chromatin looping factors RAD21 and CTCF, forms a separate regulatory module from OSN factors and antagonistically regulates the balance between pluripotency maintenance and lineage commitment in hESCs. Specifically, it orchestrates a feed-forward loop to activate early differentiation genes and concurrently represses naïve pluripotency factors, establishing the pluripotency priming program in hESCs. Loss of ZNF263 leads to defects of hESCs in pluripotency exit and subsequent multi-lineage differentiation. Consistently, single-cell RNA-seq analysis reveals that the heterogeneous gene expression promoted by ZNF263 initiates lineage choices towards both mesendoderm and ectoderm fates. Our findings together demonstrate an emerging pivotal role of ZNF263 in priming hESCs for the onset of early differentiation, which can greatly advance our understanding of the central regulatory mechanisms of pluripotency priming and early embryo development in humans.

Results

In-silico screening identifies ZNF263 as a candidate key transcription factor for primed hESCs

In mammalian cells, proximal promoters and remote enhancers are two major classes of cis-regulatory elements (CREs) harboring TF binding, and their activities are linked with the enrichment of several histone modifications, including H3K4me3, H3K27ac, and H3K9ac^{16,17}. To explore the PGRN in conventional hESCs, we collected ChIP-seq data of the three histone modifications in 16 human cell lines from ENCODE¹⁸, and with this dataset, we conducted an in-silico screening to find TFs that preferentially bind to hESC-specific active promoters and enhancers (Fig. 1a). As promoters often show modest histone modification changes between cell types¹⁷, we adopted MAmotif⁹ toolkit to quantitatively compare histone modification ChIP-seq

profiles between H1 hESCs and non-hESC cells, and screen for TF binding motifs exhibiting significant association with the H1 hESC-biased peaks at promoter and distal regions.

We started from the comparison between H1 hESCs and K562 leukemia cells. Among all 1011 JASPAR vertebrate motifs, ZNF263 motif was identified as the most significant one associated with the H1-biased peaks of three histone modifications at gene promoters (Fig. 1b). The other top ranked motifs included those of CPFs SOX2, POU5F1, TCF3¹¹ and TFs from the same families (Fig. 1b and Supplementary Data 2a–c). However, ZNF263 motif showed no significant association with the H1-biased peaks of H3K9/27ac modifications at remote regions, while POU5F1 and SOX2 motifs headed the list (Supplementary Fig. 1a and Supplementary Data 2d, e). More explicitly, we respectively grouped the promoter and distal peaks of each histone mark in H1 hESCs according to the presence or absence of a specific motif, and then compared their ChIP-seq signal changes between H1 and K562 cells. For all three marks, a significantly larger proportion of their promoter peaks containing ZNF263 motifs displayed H1-biased histone modification levels than those without ZNF263 motif (Fig. 1c). But, for distal H3K9/27ac peaks, no such difference was observed, distinct to that observed for POU5F1 and SOX2 motifs (Supplementary Fig. 1b).

Next, we proceeded to compare H1 hESCs against each non-hESC cell line. Similar to POU5F1 and SOX2 motifs, the ZNF263 motif was found to be significantly associated with the H1-biased promoter peaks of H3K9/27ac in most comparisons (Fig. 1d, e, and Supplementary Fig. 1c, d). Notably, in all comparisons, a significant association was detected between ZNF263 motif and H1-biased H3K4me3 promoter peaks (Fig. 1f), much stronger than that observed for POU5F1 and SOX2 (Supplementary Fig. 1e). Meanwhile, POU5F1 and SOX2 motifs were strongly associated with the H1-biased H3K9/27ac distal peaks in almost all comparisons, while ZNF263 motif failed to show any significant association except for one comparison (Supplementary Fig. 1f, g). These results, together, indicate that ZNF263 may function as a key component of the hESC promoter regulatory network.

We further integrated the corresponding ENCODE RNA sequencing (RNA-seq) data to explore whether the occurrence of the ZNF263 motif at promoters is associated with preferential expression of downstream genes in hESCs. Again, we use the comparison between H1 and K562 cells to demonstrate our approach. We first identified genes differentially expressed between two cell lines. Subsequently, we examined the 11769 genes with H3K27ac peaks at promoters in H1 hESCs, and found that genes occupied by H3K27ac peaks containing ZNF263 motif were significantly enriched in genes more highly expressed in H1 hESCs (termed H1-high genes), but slightly depleted in those more highly expressed in K562 cells (termed K562-high genes, Fig. 1g). Finally, the same approach was applied to systematically compare H1 hESCs against other non-hESC cell lines. We observed that genes with the ZNF263 motif at promoters were strongly enriched for H1-high genes in all comparisons (Fig. 1h). Highly similar results were obtained for the peaks with POU5F1 and SOX2 motifs. In contrast, the downstream genes of H3K27ac peaks harboring motifs of GATA1 and TAL1, master regulators of K562 cells, showed no significant enrichment for H1-high genes (Fig. 1h). Thus, we speculate that ZNF263 participates in shaping the transcriptome of H1 hESCs.

Besides, to figure out whether the global epigenomic and transcriptomic changes between primed and reprogrammed naïve hESCs are associated with ZNF263, we applied the same procedure as above to the published H3K4me3 ChIP-seq and RNA-seq data of these hESCs^{12,13}. Intriguingly, the ZNF263 motif, but not POU5F1 and SOX2 motifs, showed a significant association with primed hESCs-biased H3K4me3 peaks at promoters (Fig. 1i and Supplementary Data 2f, g). Consistently, the downstream genes of H3K4me3 promoter peaks containing ZNF263 motifs were strongly enriched for genes more highly expressed in primed hESCs (Fig. 1j). These findings imply a specific function of ZNF263 in primed hESCs.

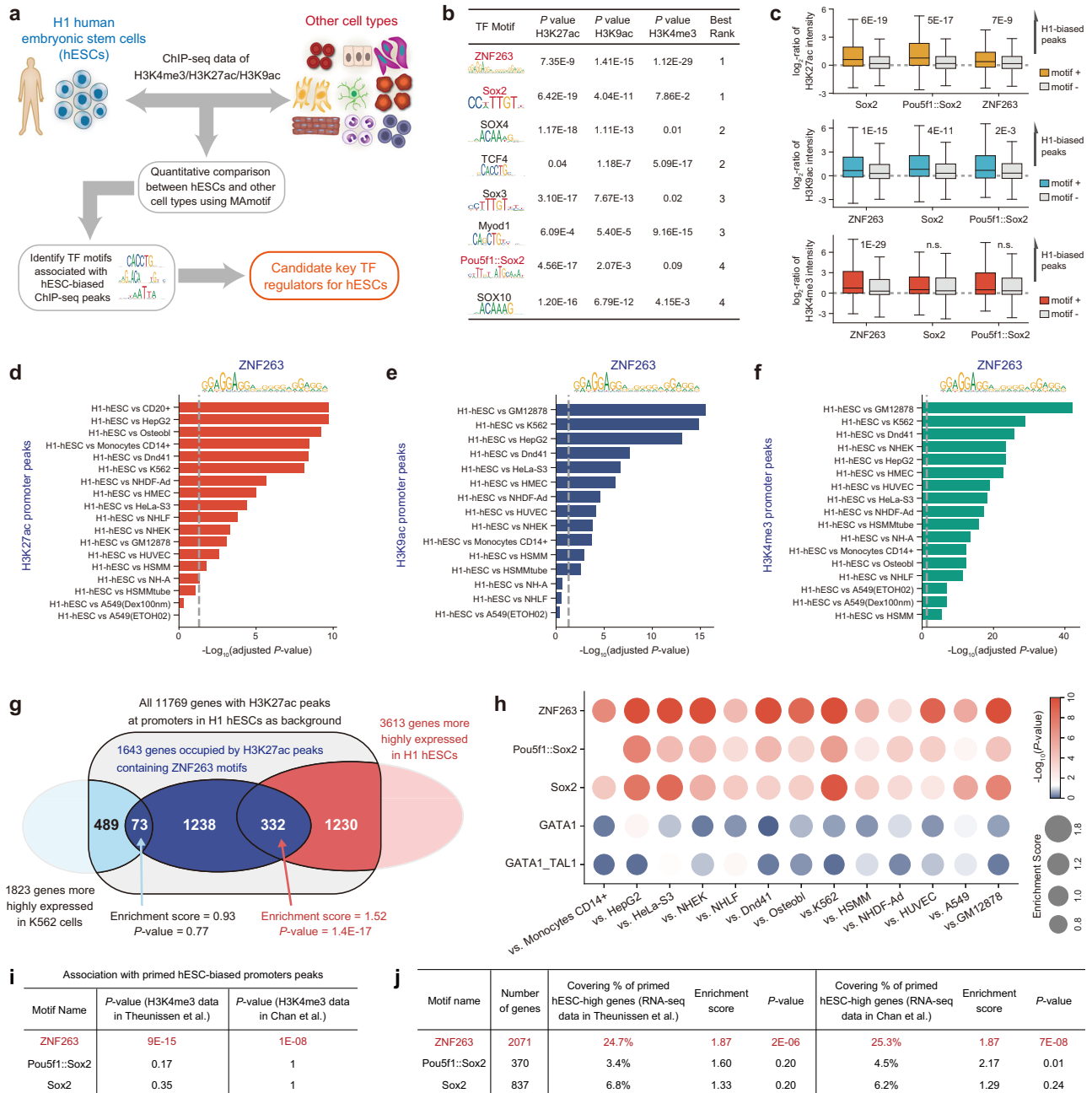


Fig. 1 | In-silico screening using MAmotif. **a** Workflow of the in-silico screening. **b** Top-ranked TF motifs respectively associated with H1 hESC-biased promoter peaks of different histone modifications when compared with K562 cells. **c** Boxplots displaying the log₂-ratios of ChIP-seq intensities between H1 hESCs and K562 cells for promoter peaks of different histone modifications. Peaks were respectively classified by the occurrence of ZNF263, Sox2, and Pou5f1::Sox2 motifs. Each box shows the median (horizontal line), second to third quartiles (box), and Tukey-style whiskers (beyond the box). **d-f** -Log₁₀ transformed *P*-values to measure the significance of the association between ZNF263 motif and the H1 hESC-biased promoter peaks of different histone modifications, in comparison to each non-hESC ENCODE cell line. Gray dashed lines correspond to adjusted *P*-value threshold of 0.05. **g** Venn diagram showing the overlap between different gene sets. **h** Enrichment of the overlap between genes that were more highly expressed

in H1 hESCs, when compared to each non-hESC ENCODE cell line, and genes occupied by the H3K27ac promoter peaks containing specific TF motifs in H1 hESCs. **i** *P*-values to measure the significance of the association between each TF motif and the primed hESC-biased promoter peaks of H3K4me3, when compared to naïve hESCs. **j** Enrichment of the overlap between genes that were more highly expressed in primed hESCs (primed hESC-high genes), when compared to naïve hESCs, and genes occupied by the H3K4me3 promoter peaks containing specific TF motifs in H1 hESCs. Two independent sets of data are presented. *P*-values in (**b-f**) and (**i**) are reported by MAmotif. Enrichment scores in (**g, h**) and (**j**) are defined as the ratio between the number of overlapping genes and that expected by chance, and their *P*-values are calculated using right-tailed Fisher's exact test. Exact *P*-values (**d-f, h**) are provided in Source Data.

ZNF263 extensively binds to promoters of genes related to pluripotency and stem cell differentiation in hESCs

Characterizing the genome-wide occupancy of ZNF263 is a crucial step to investigate its role in hESCs. We first performed a ChIP-seq assay in H1

hESCs with a monoclonal antibody against endogenous ZNF263 protein and detected 10,378 peaks. They were highly enriched for ZNF263 motifs at peak summits (Supplementary Fig. 2a), suggesting ZNF263's chromatin binding in H1 hESCs heavily relies on these motifs. However,

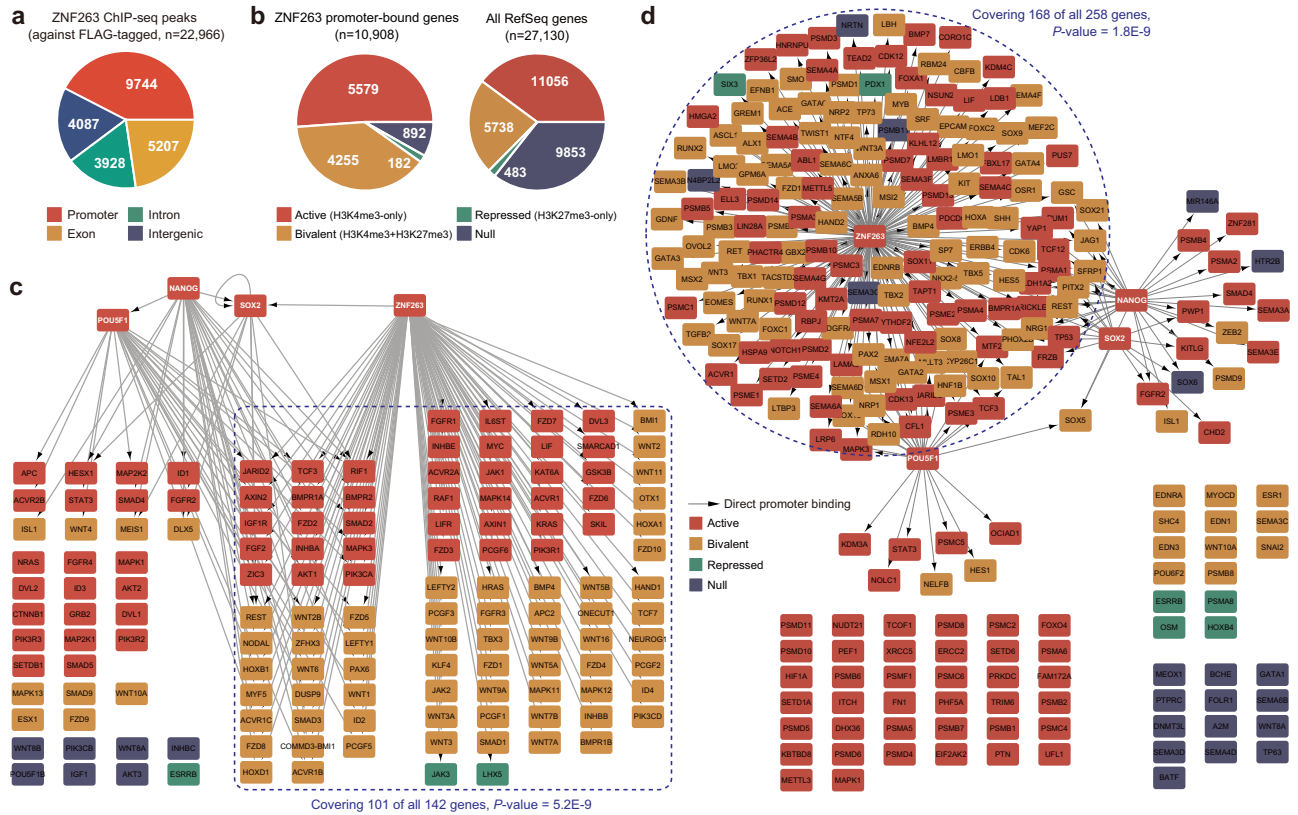


Fig. 2 | ZNF263 occupancy at gene promoters in H1 hESCs. a Genome-wide distribution of ZNF263-FLAG ChIP-seq peaks at gene promoters, exons, introns, and intergenic regions. **b** Distribution of genes bound by ZNF263 at promoters (left) and all RefSeq genes (right) in each chromatin state category. **c, d** Networks

illustrating POU5F1, NANOG, SOX2, and ZNF263 occupancy at promoters of genes respectively belonging to the KEGG pathway "Signaling pathways regulating pluripotency of stem cells" (c) and Gene Ontology (GO) term "regulation of stem cell differentiation" (d). P-values are calculated using the right-tailed Fisher's exact test.

probably due to the antibody's efficiency, many gene promoters exhibited a modest ChIP-seq signal enrichment over input control, and only 2,480 promoter peaks were identified (24%, Supplementary Fig. 2b, c). We thus conducted a FLAG epitope ChIP-seq assay in *ZNF263*-knockout (-KO) HI (see below) cells expressing a FLAG-tagged version of *ZNF263*. A clearly improved ChIP-seq enrichment was observed (Supplementary Fig. 2d), and we identified 22,966 binding peaks, among which 9,744 were located at promoters (Fig. 2a; Supplementary Data 3a, b). The *ZNF263*-FLAG ChIP-seq peaks were remarkably enriched for *ZNF263* motifs, too, and showed a favorable agreement with endogenous *ZNF263* ChIP-seq peaks (Supplementary Fig. 2e, f). Therefore, *ZNF263*-FLAG peaks were chosen as *ZNF263*'s binding sites in H1 hESCs.

We examined the chromatin state of gene promoters bound by *ZNF263* in H1 hESCs, and found that most of them were occupied by H3K4me3 but not H3K27me3 (termed active promoters) or by both H3K4me3 and H3K27me3 (termed bivalent promoters) (Fig. 2b). In total, 90% of *ZNF263*-bound promoters were marked by H3K4me3, much higher than that of genomic background (Fig. 2b). These indicate *ZNF263* selectively binds to promoters of genes that are actively transcribed or poised for transcription in H1 hESCs. Gene ontology and KEGG pathway analysis revealed that these genes were highly enriched in signaling pathways regulating pluripotency of stem cells (Fig. 2c) and genes involved in stem cell differentiation (Fig. 2d), indicating *ZNF263* may function as a pivotal transcription factor in hESCs.

ZNF263 forms a separate regulatory module from core pluripotency factors

To gain a comprehensive understanding about the regulatory landscape of *ZNF263*, we globally identified active CREs (aCREs) in

H1 hESCs, which were defined as DNase I hypersensitive sites co-occupied by active but not repressive histone marks, and investigated the occupancy of *ZNF263* and other TFs at these aCREs (Supplementary Data 3c, d). We found that neither the motif sites nor the ChIP-seq peaks of *ZNF263* tended to co-localize with those of POU5F1 and SOX2 at both promoter and distal aCREs (Fig. 3a, b and Supplementary Fig. 3a, b), suggesting they belong to separate regulatory modules in hESCs. We then performed a systematic TF colocalization analysis with published TF ChIP-seq peaks from H1 hESCs. Notably, numerous TFs including transcriptional coactivator P300 and cohesin RAD21 showed significant colocalization with *ZNF263* peaks at promoter aCREs, most of which were supported by their colocalization with *ZNF263* motifs (Fig. 3c). In contrast, at distal aCREs, many of these TFs including P300 turned to colocalize with POU5F1 and SOX2 but not *ZNF263* (Supplementary Fig. 3c). Finally, we separately drew a TF co-regulatory network at promoter and distal aCREs with all TF pairs showing significant colocalization (Fig. 3d and Supplementary Fig. 3d). Both networks comprised two separate modules centered around *ZNF263* (termed *ZNF263* module) and OSN factors (termed OSN module), together with another module constituted by TFs exhibiting significant colocalization with TFs in both *ZNF263* and OSN modules (termed shared module). Notably, the size of all three modules differs dramatically between the two networks, as many TFs belonging to the *ZNF263* and shared modules in the promoter coregulatory network were absorbed into the OSN module in the distal coregulatory network. These results, collectively, provide clear evidence for a central role of *ZNF263* at promoter aCREs in primed hESCs, which further implies it may serve a different function from core pluripotency factors in human PGRN.

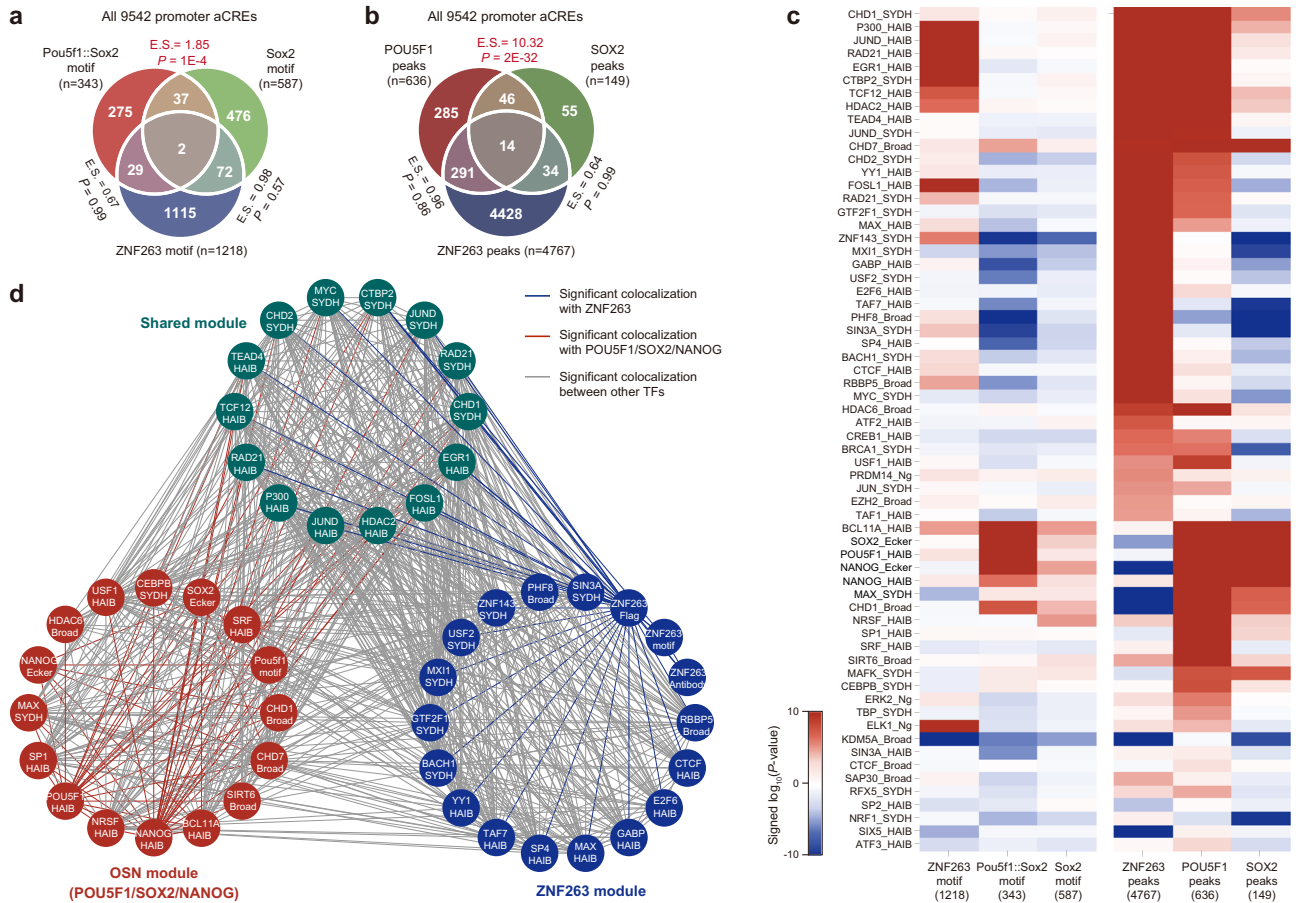


Fig. 3 | Identification of ZNF263 regulatory module at gene promoters. **a, b** Venn diagram illustrating the overlap between the promoter active cis-regulatory elements (aCREs) in H1 hESCs that respectively contain Pou5f1::Sox2, Sox2, and ZNF263 motifs (a) or ChIP-seq peaks for POU5F1, SOX2, and ZNF263 (b). Enrichment scores (E.S.) are defined as the ratio between the number of overlapping aCREs and that expected by chance, and *P*-values (*P*) are calculated using right-tailed Fisher’s exact test. **c** Heatmap showing the co-localization of motifs (left) or ChIP-seq peaks (right) for each TF (POU5F1, SOX2, and ZNF263) with other

TFs’ ChIP-seq peaks at promoter aCREs in H1 hESCs. *P*-values are calculated using hypergeometric distribution (right-tailed), and the sign indicates whether the number of co-localized aCREs between the two TFs was larger (positive) or smaller (negative) than that expected by chance. **d** TF co-regulatory network illustrating pairwise TF co-localization at promoter aCREs of H1 hESCs. Right-tailed *P*-value < 1E-10 is used as the cutoff to define network edges. Exact signed *P*-values (c) are provided in Source Data.

ZNF263 governs the primed pluripotent state by repressing pluripotency genes and activating differentiation genes in hESCs

It has been shown that ZNF263 exerts both positive and negative effects on gene expression²⁰, we wondered how it regulates the transcription of downstream genes, which are broadly engaged in pluripotency maintenance and differentiation, and directs hESC fate decision. Hence, we conducted CRISPR/Cas9-mediated ZNF263 gene knockout (KO) in H1 hESCs (Supplementary Fig. 4a). Two independent cell clones named 1-5 and 1-10 were picked, which were respectively heterozygous and homozygous for ZNF263 disruption (Supplementary Fig. 4b). Corresponding to their genotypes, ZNF263 protein was efficiently depleted in 1-10 clone and decreased to about half of the wildtype (WT) expression in 1-5 clone (Supplementary Fig. 4c).

Intriguingly, ZNF263-KO hESC colonies exhibited a more condensed and round morphology compared to those of WT H1 hESCs (Fig. 4a and Supplementary Fig. 4d), which were more diffuse and irregularly shaped due to spontaneous differentiation in conventional culture. High alkaline phosphatase (AP) activity is a key pluripotency marker of ESCs²¹. We examined the AP activity in WT and ZNF263-KO H1 cells using flow cytometry (FACS) and observed a clearly increased AP level in both KO clones compared to WT (Fig. 4b). Moreover, two other ESC surface markers TRA-1-81 and SSEA-3 of these cells were measured.

Accordingly, a significantly higher proportion of ZNF263-KO H1 cells were detected with high expression of both markers than WT (Fig. 4c). To ensure our findings are not specific to the H1 male hESC line, we additionally performed CRISPR/Cas9 KO of ZNF263 in H9 hESCs, which were originated from female. Two clones named 9-1 and 9-D were obtained, and they carried the same homozygous ZNF263 frame-shift mutation (Supplementary Fig. 4b). Subsequent immunoblot assay confirms an effective depletion of ZNF263 protein in both clones (Supplementary Fig. 4e). Similarly, ZNF263 deficiency substantially up-regulated AP activity in H9 hESCs (Fig. 4d). It suggests that ZNF263 tunes the pluripotent state in primed hESCs.

To figure out how ZNF263 controls the hESC pluripotent state, we next investigated the transcriptomic changes upon ZNF263 depletion in hESCs. Pilot RNA-seq profiling experiments revealed that hundreds of genes were differentially expressed by ZNF263 loss in H1 hESCs (Supplementary Fig. 4f), consisting of many pluripotency and differentiation-related genes. We performed qRT-PCR on a set of key genes, which showed a clear trend of transcriptional increase in pluripotency genes and decrease in differentiation genes in both H1 KO clones (Fig. 4e). Subsequently, we scaled up the RNA-seq experiments to include three biological replicates for each condition (except the 1-5 clone). Volcano plots revealed that many pluripotency genes including PRDM14, which is vital for maintaining the CPC in hESCs²²,

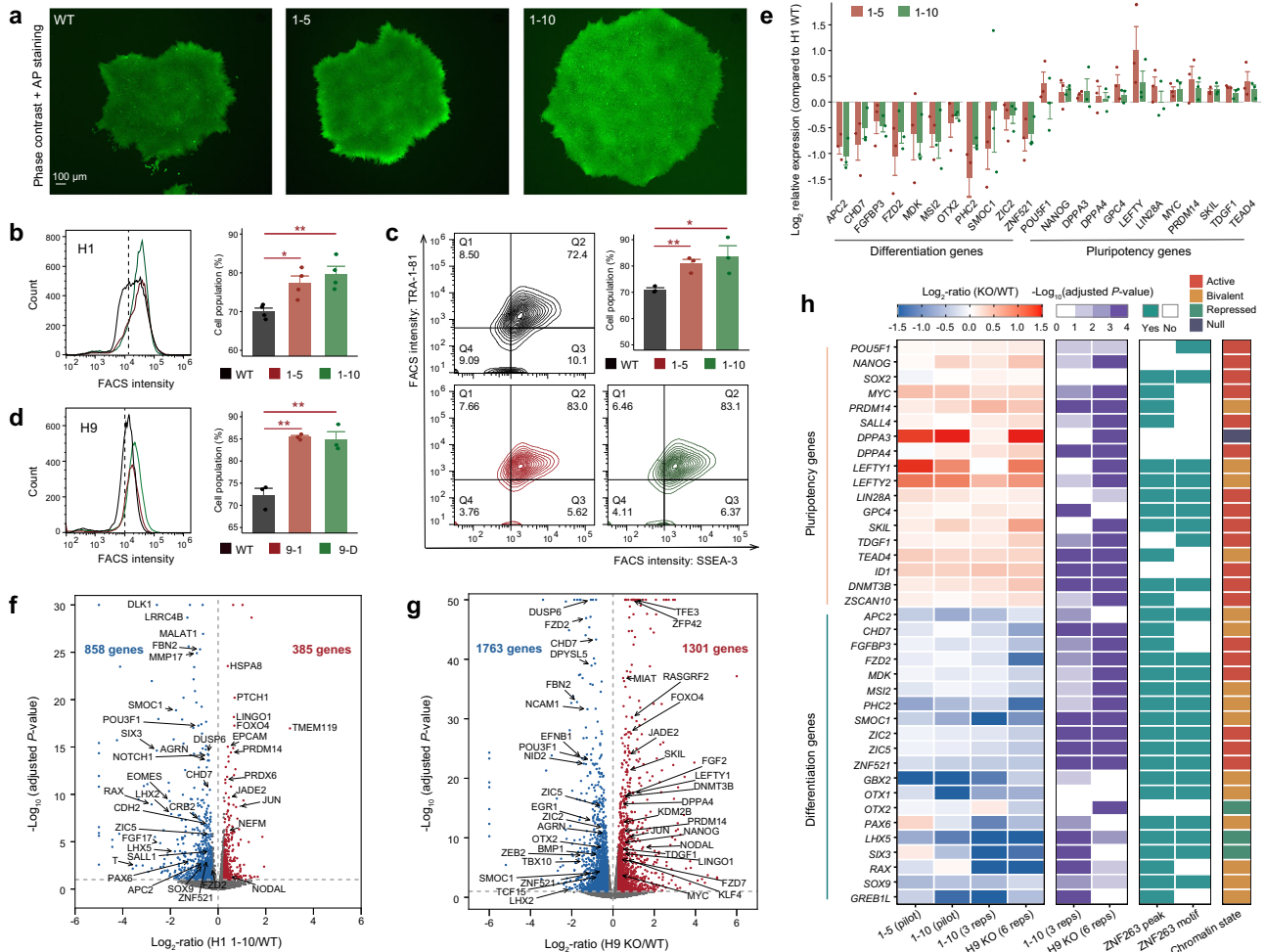


Fig. 4 | ZNF263 maintains the primed pluripotent state by repressing pluripotency genes and activating differentiation genes in hESCs. **a** Representative images showing the morphologies of wildtype (WT) and *ZNF263*-knockout (-KO) H1 hESCs. Alkaline phosphatase (AP) activity was stained simultaneously (green). Scale bar, 100 μ m. **b** FACS analysis for AP-stained WT and *ZNF263*-KO H1 hESCs. **c** FACS analysis for SSEA-3 and TRA-1-81 double-stained WT and *ZNF263*-KO H1 hESCs. **d** FACS analysis for AP-stained WT and *ZNF263*-KO H9 hESCs. Left, representative intensity plot. Right, quantification of cell populations with higher intensity. $n = 4$ (**b**) or 3 (**d**) biological replicates. P -values are derived by unpaired two-tailed Student's t -test, **: $P < 0.01$; *: $P < 0.05$. **e** The qRT-PCR analysis for changes of key

pluripotency and differentiation genes in *ZNF263*-KO H1 hESCs versus WT. $n = 3$ biological replicates. **f, g** Volcano plots respectively showing gene expression changes in H1 and H9 *ZNF263*-KO hESCs versus their respective WT. The number of significantly up/down-regulated genes and essential regulators among them are indicated. **h** Heatmap showing the expression changes of key pluripotency and differentiation genes in each clone of H1 and H9 *ZNF263*-KO hESCs versus their respective WT. The DEseq2 P -values (**f–h**) and the presence of ZNF263 ChIP-seq peaks and motifs at their promoters are tailed. All error bars (**b–e**) denote the mean \pm S.E.M. All exact P -values (**b–d, f–h**) are provided in Source Data.

were significantly up-regulated by *ZNF263* KO in both H1 and H9 hESCs, while the consistently down-regulated genes comprised many key differentiation genes, such as ectodermal lineage specifier ZIC5 & ZNF521^{4,23} and mesendodermal marker SOX9²⁴ (Fig. 4f, g and Supplementary Data 4). In total, we identified 323 and 500 consistently up- and down-regulated genes by *ZNF263* KO in H1 and H9 hESCs, respectively, and found the down-regulated ones were highly enriched in various developmental processes (Supplementary Data 5, 6). These findings suggest a disturbance upon ZNF263 loss in the expression equilibrium between pluripotency genes and lineage-associated genes, which is the characteristic of the primed pluripotent state, in hESCs.

ZNF263 dampens the core pluripotency circuitry and orchestrates a feed-forward loop to prime early differentiation in hESCs

We visualized the expression changes of key pluripotency and differentiation genes in all *ZNF263* KO clones, as well as the presence of ZNF263 peaks and motifs at their promoters (Fig. 4h). It suggests that many of these genes are directly regulated by ZNF263 in hESCs.

Accordingly, we mapped ZNF263 peaks to genes differentially expressed in KO H1 hESCs, and found that 72% of the down-regulated genes and 55% of the up-regulated genes were bound by ZNF263 at promoters, much higher than that of random genes (Fig. 5a). Therefore, we respectively named the ZNF263-bound down/up-regulated genes as ZNF263 directly activated/repressed (ZDA/ZDR) genes in H1 hESCs (Supplementary Data 7).

Upon closer investigation, we found that the majority of ZDA genes had bivalent promoters in H1 hESCs (Fig. 5b), and these ZDA bivalent genes displayed higher expression levels than other bivalent genes (Fig. 5c). It suggests that ZNF263 enhances the expression of numerous poised genes in hESCs, which potentially facilitates their activation during subsequent differentiation. To assess the assumption, we collected the published time-course RNA-seq data for two representative differentiation processes, encompassing definite endoderm (DE)²⁵ and ectodermal neuron (EN) differentiation²⁶, and examined the expression profiles of ZDA bivalent genes versus other bivalent genes. Intriguingly, the ZDA bivalent genes were highly enriched in genes activated at early stages of both differentiation

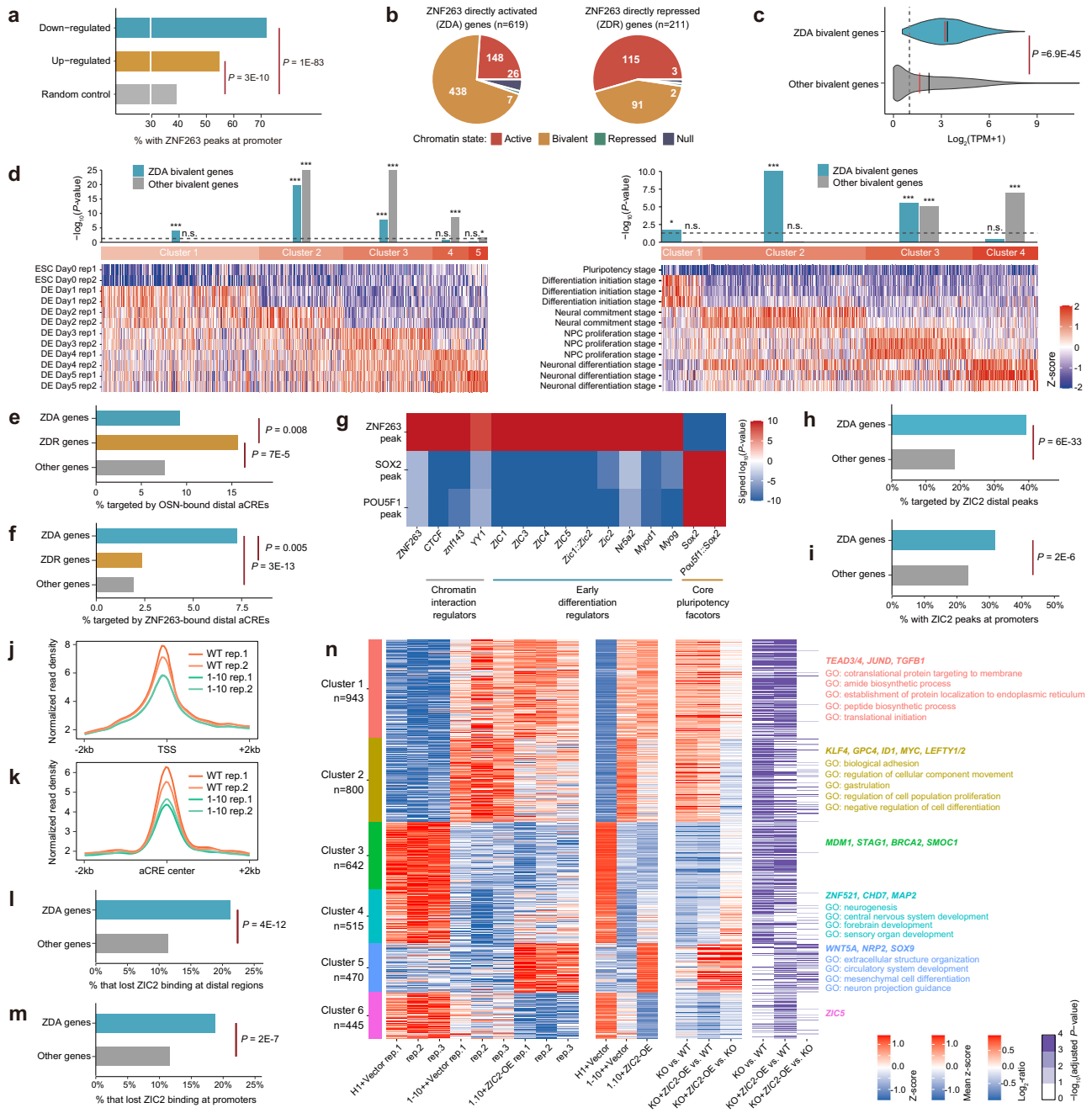


Fig. 5 | ZNF263 orchestrates a feed-forward loop to prime early differentiation genes in hESCs. **a** Proportions of genes whose promoters are occupied by ZNF263 peaks, among the down-regulated and up-regulated genes in *ZNF263*-KO H1 hESCs. **b** Distribution of ZNF263 directly activated (ZDA) and repressed (ZDR) genes in each chromatin state category. **c** Expression levels of ZDA and other bivalent genes in WT H1 hESCs. *P*-value is derived by unpaired two-tailed Student's *t*-test. TPM, number of transcripts per million clean tags. Black/red bars indicate the mean/median values. Dashed lines correspond to TPM = 1. **d** Enrichment of ZDA and other bivalent genes in each cluster of genes that are up-regulated at different time points in hESCs during differentiation of definitive endoderm (DE, left) and ectoderm neural cells (right). Top, $-\log_{10}$ transformed *P*-values. Bottom, heatmap showing the z-score transformed expression levels of each gene at different time points. Dashed lines correspond to *P*-value threshold of 0.05. **e-f** Illustrating proportions of ZDA, ZDR, and other genes that are targeted by OSN-bound (e) and ZNF263-bound (f) distal aCREs. **g** Heatmap showing the co-localization of ChIP-seq peaks for each TF (POU5F1, SOX2, and ZNF263) with the binding motifs of specific

Tfs, at distal aCREs in H1 hESCs. *P*-values are calculated using hypergeometric distribution (right-tailed) and the sign indicates whether the number of co-localized aCREs between the two TFs was larger (positive) or smaller (negative) than that expected by chance. **h-i** Proportions of ZDA and other genes that are bound by ZIC2 at distal (h) and promoter (i) regions in H1 hESCs. **j-k** Normalized ZIC2 CUT&Tag signals around the transcriptional start site (TSS) of all RefSeq genes (j) and the center of all aCREs of H1 hESCs (k) in WT and *ZNF263*-KO H1 hESCs. Two biological replicates are included. **l-m** Proportions of ZDA and other genes that lose ZIC2 binding at distal (l) and promoter (m) regions upon *ZNF263* KO in H1 hESCs. **n** Unsupervised clustering analysis dividing genes into six distinct clusters based on their expression patterns in *ZIC2*-overexpressing (OE) *ZNF263*-KO H1 hESCs, WT controls, and *ZNF263*-KO controls. Representative genes and enriched GO terms for each cluster are indicated. *n* = 3 biological replicates. *P*-values shown in (a, e, f, h, i, l, m) are derived using right-tailed Fisher's exact test. ***: *P* < 0.001; **: *P* < 0.01; *: *P* < 0.1; n.s.: non-significant. Exact *P*-values (d, g, n) are provided in Source Data.

processes, while other bivalent genes were more enriched in those activated at late stages (Fig. 5d). These results indicate that ZNF263 selectively primes poised genes in hESCs to prepare for early differentiation. As is widely observed, bivalent promoter domains of developmental genes resolve to H3K4me3 during pluripotency exit²⁷. Next, we wondered whether this process has been initiated at ZDA bivalent genes in hESCs. We examined the related histone marks in KO H1 cells. A significantly larger proportion of ZDA bivalent genes' promoters exhibited reduced H3K4me3 and increased H3K27me3 levels upon ZNF263 loss (i.e., reversal of bivalent domain dissolution) compared to other bivalent genes (Supplementary Fig. 5a). On the other hand, the ZDR active genes, which had active promoters in H1 hESCs and accounted for over half of ZDR genes (Fig. 5b), demonstrated a notable elevation in promoter H3K4me3 and H3K27ac levels in KO H1 cells (Supplementary Fig. 5b). Moreover, these ZDR active genes were significantly enriched in genes immediately down-regulated at early stages of DE and EN differentiation (Supplementary Fig. 6a), suggesting many of them are associated with pluripotency maintenance. These findings, taken together, clearly demonstrate that ZNF263 directly initiates the transcriptional up-regulation of early differentiation genes and repression of pluripotency genes in primed pluripotency, coupling with the gradual transformation of their chromatin states. This could be a crucial step towards the drastic activation and suppression of corresponding genes during lineage differentiation.

We then investigated how ZNF263 differentially regulates the expression of early differentiation and pluripotency genes. Lisa tool²⁸ was utilized to integrate public chromatin accessibility and TF ChIP-seq data and infer transcriptional regulators (TRs) specifically associated with the genes up-regulated and down-regulated in KO H1 cells. As it is well known, OSN factors globally activate pluripotency genes in ESCs, mainly via binding to their remote enhancers¹¹. Thus, it's not unexpected to observe a significant association between the up-regulated genes and OSN factors, as well as numerous other TFs in the OSN regulatory module at distal CREs (Supplementary Fig. 6b and Supplementary Data 8). We globally mapped the distal aCREs of H1 hESCs to target genes, and found that the distal aCREs occupied by OSN factors were preferentially enriched in the neighboring regions of ZDR genes (Fig. 5e). Consistently, in an independent dataset from hESCs²⁹, most ZDR genes showed clearly decreased expression when OSN factors were knocked down (Supplementary Fig. 6c). It further confirms that ZNF263 and OSN factors oppositely regulate pluripotency genes in hESCs.

Meanwhile, CTCF and RAD21, which colocalized with ZNF263 at both promoter and distal CREs in H1 hESCs, were identified as top TRs in association with the down-regulated genes (Supplementary Fig. 6b). Given that a major function of CTCF and cohesin proteins is to mediate chromatin looping between distal CREs and target genes³⁰, we speculate that the regulatory modules at promoter and distal CREs centered around ZNF263 jointly regulated ZDA genes. In line with this assumption, we found that ZDA genes were significantly enriched in the target genes of ZNF263-bound distal aCREs (Fig. 5f). Moreover, using motif analysis, we further identified that, besides CTCF, ZNF263 peaks also strongly colocalized with binding motifs of many early differentiation regulators including ZIC2 and ZIC5 at distal aCREs (Fig. 5g), indicating a potential positive feed-forward mechanism originating from ZNF263 for priming lineage differentiation.

To verify this speculation, we profiled the genome-wide occupancy of ZIC2 by CUT&Tag in both WT and ZNF263-KO H1 hESCs (Supplementary Fig. 7a). In total, 12,966 ZIC2 binding peaks were identified in WT cells. These peaks were significantly enriched in the promoter and distal regions of ZDA genes (Fig. 5h, i) as well as many other developmental genes (Supplementary Data 9-10). In fact, we observed a strong co-localization of ZIC2 and ZNF263 at the promoter and distal aCREs of H1 hESCs, particularly near ZDA genes

(Supplementary Fig. 7b). Furthermore, co-immunoprecipitation experiments confirmed a physical interaction between ZIC2 and ZNF263 (Supplementary Fig. 7c), thereby providing strong evidence for their co-regulatory roles in primed hESCs. On the other hand, in ZNF263-KO hESCs, we detected only 7,231 ZIC2 peaks and a global decrease of ZIC2 binding was observed at both gene promoters and distal aCREs (Fig. 5j, k and Supplementary Fig. 7d). Remarkably, compared to other genes, a significantly higher proportion of ZDA genes were found to lose ZIC2 binding at their promoter and distal regions upon ZNF263 depletion (Fig. 5l, m), including key differentiation regulators such as ZIC5 and BMP7 (Supplementary Fig. 7e). Coupled with the observed 20–40% reduction in ZIC2 protein levels in ZNF263-KO H1 hESCs (Supplementary Fig. 7f, g), these findings suggest that ZNF263 is required for stabilizing ZIC2 expression and its chromatin binding nearby developmental genes in primed hESCs. Collectively, these results elucidate a feed-forward loop formed by ZNF263 and ZIC2, which is crucial for shaping and maintaining the transcriptional landscape associated with the primed pluripotency state.

Given that ZIC2 has been found to be essential for the transition from naive to primed pluripotency in mammals^{31–33}, we sought to investigate whether the role of ZNF263 signaling cascades in orchestrating the primed pluripotency gene expression program primarily relies on its regulation of ZIC2 expression, or on the entire feed-forward loop mediated by the ZNF263-ZIC2 axis. Thus, we overexpressed ZIC2 in ZNF263-KO H1 hESCs and systematically analyzed the transcriptomic changes induced by ZIC2 overexpression (OE), relative to WT and ZNF263-KO mock controls. Utilizing k-means clustering, we categorized the 3815 differentially expressed genes across the three conditions into 6 distinct clusters (Fig. 5n and Supplementary Data 11). Among these, genes in cluster 1 and 2, which exhibited a clear increase in expression upon ZNF263 KO, were significantly enriched with ZDR genes. Meanwhile, genes in clusters 3, 4 and 6, predominantly displaying reduced expression after ZNF263 KO, were enriched in ZDA genes (Fig. 5n and Supplementary Fig. 7h). Further investigation revealed that for many genes in cluster 2, including key pluripotency-associated genes *KLF4* and *GPC4*, their transcriptional upregulation induced by ZNF263 KO was substantially alleviated following ZIC2 OE (Fig. 5n). Similarly, the downregulation of many genes in cluster 4, which are involved in neural and sensory system development, was also notably reversed upon ZIC2 OE. However, the transcriptional changes of genes in clusters 1, 3, and 6, including pluripotency factors *TEAD3/4* and differentiation regulator *ZIC5*, predominantly persisted despite ZIC2 OE (Fig. 5n), indicating that ZIC2 OE only partially mitigates the transcriptomic perturbations caused by ZNF263 loss in hESCs. These findings further suggest that the ZNF263-ZIC2 signaling cascade is critical for promoting pluripotency priming in hESCs and, importantly, the entire ZNF263-ZIC2 axis appears to be essential for this function. It operates through a positive feed-forward loop mechanism, likely in conjunction with other differentiation regulators downstream of ZNF263 signaling cascades, to regulate the early activation of developmental genes.

In summary, our results demonstrate that ZNF263 extensively binds to both promoter and distal regions of early differentiation genes in primed hESCs, acting in concert with ZIC2 to up-regulate their expression. Conversely, ZNF263 occupies the promoters of many pluripotency genes, hindering the remote regulatory effects originating from OSN-centered core pluripotency circuit. Together, these roles of ZNF263 underscore it as a pivotal regulator in orchestrating the pluripotency priming program in hESCs.

ZNF263 is required for facilitating pluripotent state dissolution and multi-lineage differentiation of hESCs

Differentiation of hESCs conceptually consists of two key steps: pluripotent state dissolution (PSD), as indicated by shutdown of core pluripotency network, and lineage commitment to assemble the

transcriptional program corresponding to a specific downstream lineage³⁴. We have unveiled the capacity of ZNF263 in transcriptionally regulating both lineage-associated and pluripotency genes to drive hESCs into the primed pluripotency state, making us wonder whether subsequent PSD and lineage commitment also require it. A feature of conventional primed hESCs is their reliance on FGF/ERK signaling for pluripotency maintenance. Inhibition of this pathway has been demonstrated to induce PSD⁴. Thus, we perturbed MEK-ERK signaling in KO H1 hESCs with MEK inhibitor (MEKi) PD0325901 and investigated whether the PSD process could be affected by ZNF263 deficiency. We first monitored their pluripotent state with AP staining. With a two-day duration of treatment, WT hESCs exhibited a substantial decrease of AP activity, which is consistent with previous reports³⁵. However, in both ZNF263-KO clones, there was minimal change in AP activity upon treatment (Fig. 6a). We further investigated the accompanied gene expression changes. It was observed that, in WT hESCs, key pluripotency genes were effectively down-regulated after MEKi treatment, whereas in ZNF263-depleted cells, their magnitude of downregulation was much lower than that observed in WT cells (Fig. 6b). Moreover, after MEKi treatment, there was a notable upregulation of the differentiation-related genes, many of which are directly regulated by ZNF263, in WT hESCs. In contrast, the upregulation of these genes in KO H1 cells was not as efficient as that in WT cells (Fig. 6b). Global transcriptomic analysis revealed that a substantial proportion of genes differentially expressed upon MEKi treatment in WT H1 hESCs, termed MEKi-sensitive genes, exhibited significantly impaired transcriptomic changes in MEKi-treated ZNF263-KO hESCs (Fig. 6c and Supplementary Data 12a, b). We designated these genes as ZNF263-dependent MEKi-sensitive genes (Supplementary Data 12c). Notably, while many of them were found to be transcriptionally regulated by ZNF263 in untreated primed hESCs, a large subset exhibited significant ZNF263-dependent expression changes only upon MEKi treatment (Supplementary Fig. 8a and Supplementary Data 12d), including the pluripotency marker *POU5F1* (Fig. 6b). Collectively, these findings indicate that the prompt onset of PSD in hESCs requires ZNF263. Moreover, they further reveal an expanding regulatory network downstream of ZNF263 during pluripotency exit.

To gain deeper insights into how ZNF263 facilitates PSD in hESCs, we performed systematic gene set enrichment analysis (GSEA) on MEKi-induced transcriptomic changes in WT hESCs, and compared them with those resulting from ZNF263 KO before and after treatment. Upon MEK inhibition, WT hESCs exhibited significant downregulation of MAPK signaling pathway genes, including several key upstream activators of MEK/ERK signaling such as *FGF13/19* and *MAP3KI*, along with suppression of pathways related to basic cellular functions, such as carbon metabolism, ribosome biogenesis, and RNA transport (Supplementary Data 13a). Meanwhile, MEKi treatment markedly activated multiple developmental pathways, including key components of the Hedgehog and Wnt/ β -catenin signaling cascades (Supplementary Fig. 8b and Supplementary Data 13a). Wnt/ β -catenin signaling has been extensively characterized as playing a crucial role in promoting the differentiation of hESCs^{34,36}. However, while the transcriptional downregulation of the aforementioned MEK/ERK signaling activators remained largely intact in ZNF263-KO hESCs after MEKi treatment (Fig. 6b), the induction of Wnt/ β -catenin signaling factors was drastically attenuated in KO cells compared to WT controls (Fig. 6b, Supplementary Fig. 8c and Supplementary Data 13b). Intriguingly, some of these Wnt/ β -catenin signaling factors, such as *FZD1/2*, were already transcriptionally activated by ZNF263 in hESCs prior to MEKi treatment (Fig. 6b, Supplementary Fig. 8d, and Supplementary Data 13c), suggesting a close crosstalk between ZNF263 and Wnt/ β -catenin signaling during the early stages of human pluripotency exit. These observations demonstrate that ZNF263 not only transcriptionally regulates the balance between pluripotency and differentiation genes in primed hESCs, but also engages in crosstalk with

WNT and other differentiation-related signaling cascades. This crosstalk enables the rapid induction of their core components and downstream transcriptional regulators in hESCs upon exposure to pluripotency exit signals, thereby facilitating an efficient onset and progression of PSD.

Next, we investigated whether ZNF263 loss also compromises the lineage commitment process of hESCs. We performed ex vivo differentiation of embryoid bodies (EBs) with WT and ZNF263-KO H1 hESCs. Following withdrawal of self-renewal signals in hESC medium (mTeSR1), EBs from both ZNF263-KO and WT hESCs were successfully generated. However, EBs derived from the ZNF263-KO clones showed a clearly smaller size and diminished expansion than those obtained from WT hESCs (Fig. 6d, e and Supplementary Fig. 8e). Meanwhile, we assessed the expression of pluripotency marker *POU5F1* in day-5 EBs, and observed that a much larger proportion of EB cells derived from ZNF263-KO H1 hESCs maintained high levels of *POU5F1* compared to those from WT hESCs (Supplementary Fig. 8f). This echoes our previous finding that ZNF263 loss impairs the efficiency of PSD in hESCs. Subsequently, we examined the expression of lineage markers and found that most of them displayed significantly impaired induction in the ZNF263-KO hESC-derived EBs at day 16 compared to their counterparts (Fig. 6f). This was particularly notable for the neuroectodermal determinant *PAX6*³⁷ and definitive mesendoderm marker *CXCR4*³⁸.

To systematically evaluate the impact of ZNF263 loss on the respective induction of three primary germ layers, we conducted a comprehensive transcriptomic analysis of differentiated H1 hESCs generated using STEMdiffTM trilineage differentiation kit. As anticipated, after 5 days of mesoderm and endoderm induction, we observed a significant upregulation of genes associated with the respective differentiation pathways, including those involved in endoderm formation, skeletal development and cardiac development, in the differentiated WT hESCs compared to their undifferentiated state (Supplementary Fig. 8g, h and Supplementary Data 14, 15). Similarly, after 7 days of ectoderm induction, a pronounced activation of genes related to ectoderm-specific pathways, such as nervous system development and brain formation, was observed in the ectoderm-differentiated WT hESCs (Supplementary Fig. 8i and Supplementary Data 16). However, a large number of these genes were significantly downregulated in the differentiated state of ZNF263-KO hESCs compared to WT cells, respectively (Fig. 6g), including many genes related to endodermal, mesodermal and ectodermal differentiation (Supplementary Fig. 8j–l and Supplementary Data 17). Moreover, in the respective differentiated cells, the upregulated expression of key lineage-specific marker genes, including *MIXL1*, *T* and *PAX6*, was substantially reduced upon ZNF263 depletion (Fig. 6h–j). This observation aligns with our findings in EBs derived from ZNF263-KO H1 hESCs, and further underscores the impairment of trilineage developmental potential caused by ZNF263 depletion in hESCs.

Efficient trilineage differentiation of hESCs depends on coordinated regulation of Wnt, TGF- β and other related signaling cascades^{39,40}. To investigate how ZNF263 interacts with the differentiation-related signaling pathways in this process, we conducted GSEA using KEGG-annotated pathways to compare transcriptomic profiles between WT and ZNF263-KO H1 hESCs during differentiation (Supplementary Data 18a–f). We observed that key developmental signaling pathways—including Hedgehog, Hippo, TGF- β , and Wnt—were significantly activated during the directed differentiation of WT hESCs into ectoderm (Supplementary Fig. 9a and Supplementary Data 18a). In contrast, the transcriptional induction of many core components of these pathways, including multiple WNT (*WNT1/2B/3 A/4/5B/7B/8B/9 A/10B*) and BMP (*BMP2/4/7*) family genes, which are critical for WNT and TGF- β /Hedgehog/Hippo signaling transduction, was substantially impaired in ectoderm-differentiated KO cells (Supplementary Fig. 9b–d and Supplementary Data 18b). On

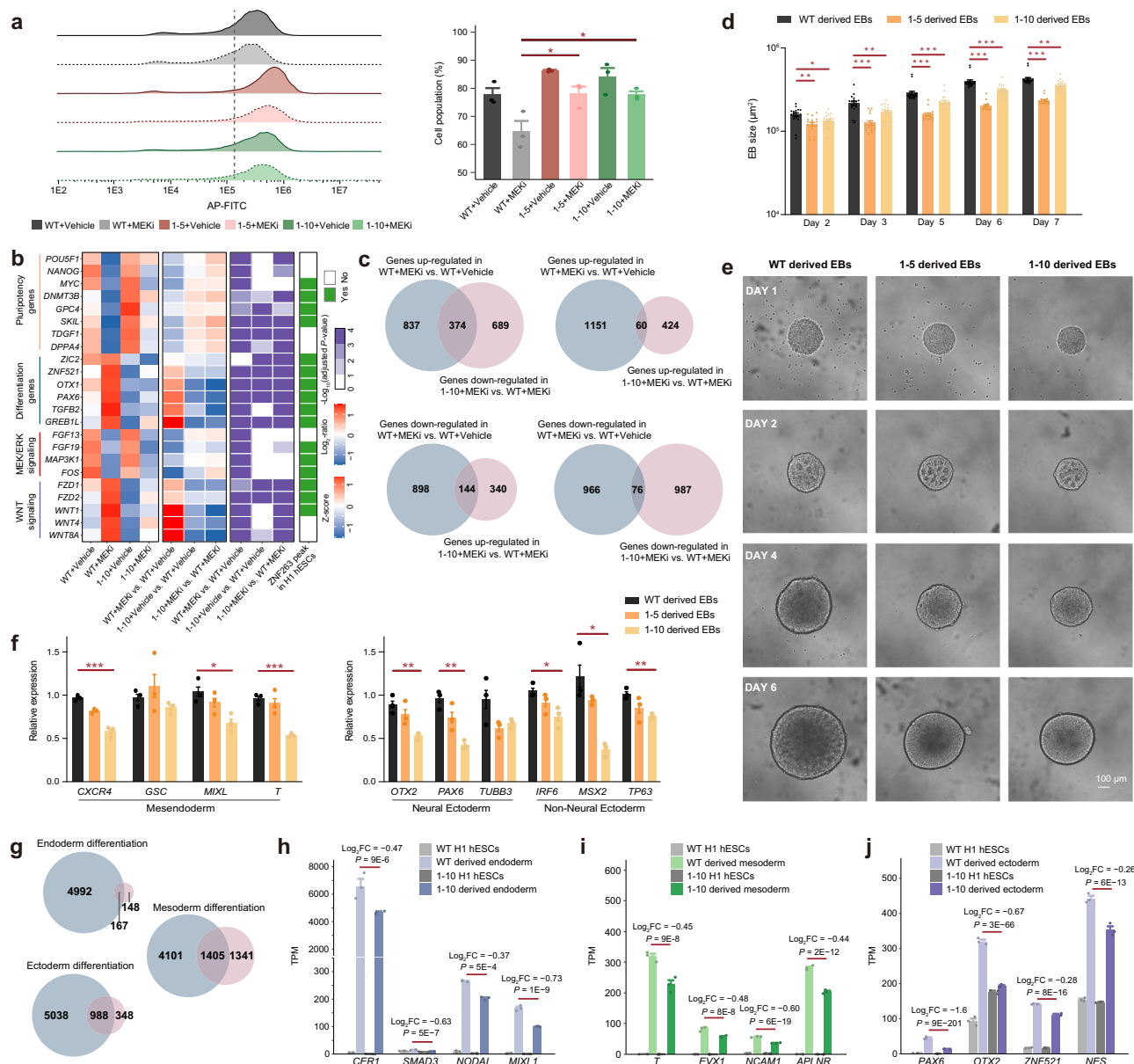


Fig. 6 | ZNF263-deficient hESCs show defects in pluripotent state dissolution and lineage differentiation.

a FACS analysis for AP-stained WT and *ZNF263*-KO H1 hESCs in the DMSO (vehicle) or MEK inhibitor (MEKi) conditions. Left, representative intensity plot. Right, quantification of cell populations with higher AP intensity. *n* = 3 biological replicates. *P*-values are calculated between MEKi treated samples by one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test. **b** Heatmap showing the RNA-seq expression changes of key pluripotency and differentiation genes, as well as key MEK/ERK and WNT signaling factors, in WT and *ZNF263*-KO H1 hESCs under vehicle or MEKi conditions. The DEseq2 *P*-values and the presence of ZNF263 ChIP-seq peaks at their promoters in H1 hESCs are tailed. *n* = 3 biological replicates. **c** Venn diagrams illustrating the overlap between different groups of differentially expressed genes. Blue circle: genes upregulated or downregulated after MEKi treatment in WT H1 hESCs. Red circle: genes upregulated or downregulated in MEKi-treated *ZNF263*-KO H1 hESCs versus MEKi-treated WT. **d, e** Size quantification (**d**) and representative images (**e**) of embryoid bodies (EBs) generated from WT and *ZNF263*-KO H1 hESCs in hanging drops at different days of differentiation. *n* = 19 (WT, 1-10 clone) and 17 (1-5 clone)

EBs. *P*-values are derived by one-way ANOVA followed by Dunnett's multiple comparisons test. **f** The qRT-PCR analysis of key mesodermal and ectodermal lineage markers in day-16 EBs derived from WT and *ZNF263*-KO H1 hESCs. Values are normalized to *GAPDH*. *n* = 3 biological replicates. *P*-values are calculated between 1-10 clone and WT using unpaired two-tailed Student's *t*-test. **g** Venn diagrams illustrating the overlap between different groups of differentially expressed genes. Blue circle: genes upregulated during directed differentiation in WT hESCs. Red circle: genes downregulated in the differentiated *ZNF263*-KO H1 hESCs versus differentiated WT. **h–j** Expression levels of endoderm (**h**), mesoderm (**i**) and ectoderm (**j**) marker genes in WT and *ZNF263*-KO H1 hESCs before and after the directed differentiation. *n* = 3 biological replicates. TPM, number of transcripts per million clean tags. Log2 fold change (Log2FC) of gene expression in differentiated *ZNF263*-KO H1 hESC versus their WT counterparts are presented. *P*-values are derived by DEseq2. ***, *P* < 0.001, **, *P* < 0.01, *, *P* < 0.1, non-significant otherwise. All error bars (**a, d, f**) denote the mean ± S.E.M., and exact *P*-values (**a, b, d, f**) are provided in Source Data.

the other hand, during directed endoderm differentiation of WT hESCs, we observed significant upregulation of genes in the TGF-β, Wnt, and PI3K-Akt signaling pathways (Supplementary Fig. 9e and Supplementary Data 18c). However, these pathways were markedly downregulated in KO-derived endodermal cells compared to WT

controls (Supplementary Fig. 9f and Supplementary Data 18d). The affected genes included multiple PI3K-Akt components (*COL2A1*, *FGF3/8/18*, *ITGA1/ITGB3*, Supplementary Fig. 9g) and key regulators of TGF-β/Wnt signaling (*NODAL*, *LEFTY2*, *CER1*, Supplementary Fig. 9d). Importantly, while *BMP2*, *BMP4*, and *BMP7* were upregulated during

WT hESC differentiation into both ectoderm and endoderm lineages, this induction was specifically impaired in ectoderm-differentiated KO cells (Supplementary Fig. 9d), indicating a lineage-dependent role for ZNF263 in modulating developmental signaling pathways. Moreover, we noticed that genes activated during the differentiation of WT hESCs into mesoderm, as well as those downregulated in mesoderm-differentiated KO cells compared to WT controls, are both significantly enriched in genes associated with regulation of the Wnt signaling pathway according to the GO database (Supplementary Data 15b and 17e). Some of these genes, such as *WNT3*, *RSPO2/3* and *LRG5* (Supplementary Fig. 9h), have been suggested to be important for mesoderm formation^{41,42}. These findings reveal extensive, context-dependent interactions between ZNF263 and key developmental signaling pathways during the differentiation of hESCs into the three germ layers. These interactions support the transcriptional activation of many core components and downstream targets within these pathways, thereby facilitating the induction of lineage-specific gene expression programs during fate commitment.

In addition to the analysis conducted with H1 hESCs, we performed MEKi treatment and EB generation in both WT and *ZNF263*-KO H9 hESCs. Largely in consistent with our findings in H1 hESCs, *ZNF263*-KO H9 hESCs exhibited significant impairments in the expression of key pluripotency and differentiation genes during pluripotency dissolution compared to WT H9 hESCs (Supplementary Fig. 10a). In the EB formation process, EBs derived from *ZNF263*-KO H9 hESCs again exhibited a smaller volume compared to that from WT (Supplementary Fig. 10b, c). We next explored whether primary germ layer induction in *ZNF263*-KO H9 hESCs was similarly affected as in H1 hESCs. To this end, we initiated directed differentiation of H9 hESCs and assessed the induction of key lineage-specific markers in the absence of ZNF263. The qRT-PCR analysis again revealed a dramatic downregulation of ectodermal marker genes, as well as noticeable decreases in mesoderm and endoderm marker genes in the differentiated *ZNF263*-KO H9 hESCs compared to WT cells (Supplementary Fig. 10d–f).

These results, in conjunction with our previous observations regarding the involvement of ZDA and ZDR genes in hESC early differentiation (Fig. 5d and Supplementary Fig. 6a), confirm a critical role of ZNF263 in facilitating germ layer differentiation of hESCs, particularly toward ectodermal lineages.

ZNF263 promotes the heterogeneity of developmental potential in primed hESCs to prepare for lineage commitment

An intrinsic feature of ESCs is their phenotypic heterogeneity, in which they display promiscuous expression of lineage-specifying genes and fluctuating responsiveness to differentiation cues⁴³. The pronounced impairment of the multi-lineage differentiation observed in *ZNF263*-KO hESCs (Fig. 6d–f) raised the question of whether ZNF263 regulates the phenotypic heterogeneity of primed hESCs. Therefore, we performed single-cell RNA sequencing (scRNA-seq) of WT and homozygous *ZNF263*-KO H1 hESCs. After quality control, we obtained scRNA-seq profiles of 2731 WT and 2040 KO cells. Overall, KO cells were largely separated from WT cells (Fig. 7a), and exhibited higher expression levels of pluripotency genes but lower expression of differentiation genes compared to WT cells (Supplementary Fig. 11a, b). These results are consistent with our findings from bulk data. Then, unsupervised clustering was employed to dissect the transcriptomic heterogeneity. All hESCs were divided into six clusters. Clusters 0 and 1 were mainly comprised of KO cells (Fig. 7b), while clusters 2–5 were dominated by WT cells. Intriguingly, among these WT-dominated clusters, genes preferentially expressed in cluster 4 were significantly enriched in pathways related to mesendoderm fate, such as muscle cell differentiation and heart morphogenesis, while marker genes of cluster 5 demonstrated a clear association with ectoderm fate, including nervous system and skin epidermis development (Supplementary Fig. 11c

and Supplementary Data 19). These findings suggest a potential divergence in cell fate among primed hESCs, and ZNF263 may contribute to this process. To investigate further, the mesendodermal and ectodermal differentiation (MED and ECD, respectively) potential for individual cells was evaluated by calculating their potential scores towards different fates. Despite most cells showing low potential scores in both MED and ECD fates (Fig. 7c), concordant with their undifferentiated status, cluster 4 and 5 respectively, showed substantially elevated MED and ECD potential scores than the other clusters (Fig. 7d), as anticipated. We then used a Gaussian Mixture Model (GMM)-based method to respectively segregate the cells into different cell groups (group M1-3 and group E1-4, Fig. 7c) with distinct MED and ECD potential scores (Supplementary Fig. 11d). Cell groups with highest MED and ECD potential scores, i.e., group M3 and E4, were considered as cells that were mostly inclined towards MED and ECD fates, respectively. Consistently, M3 mainly comprises cells from cluster 4, while E4 predominantly originated from cluster 5 (Fig. 7e). Furthermore, very small overlaps were observed between group M3 and E4 (Fig. 7f), demonstrating that the transcriptomic heterogeneity among primed hESCs signifies their divergent differentiation potential towards different germ layer fates. Remarkably, we observed that, in both MED and ECD conjecture, a much greater proportion of cells with high potential scores belonged to WT cells (Fig. 7g). These findings illustrate that ZNF263 positively regulates the differentiation potential related to multiple lineages in individual hESCs.

Discussion

How the orchestrated balance between pluripotency maintenance and lineage priming defines different pluripotent states has been a classic yet highly debatable topic⁴. Primed hESCs, exhibiting incipient expression of lineage-associated genes⁴⁴, provide an excellent platform for addressing this question. Taking advantage of this platform, our in-silico screening successfully identified ZNF263 as a potential key regulator in primed hESCs (Figs. 1–3). Further exploration demonstrates the central role of ZNF263 in orchestrating the primed pluripotency gene expression program. ZNF263, in cooperation with other TFs such as ZIC2 in human PGRN, simultaneously activates lineage-associated genes and represses the expression and regulatory activity of core pluripotency factors during this process (Figs. 4, 5). This capacity of ZNF263 enables hESCs to set up an appropriate primed pluripotent state ready for early differentiation, as evidenced by the impaired differentiation potential in *ZNF263*-depleted hESCs (Figs. 6, 7). Taken together, we unveil a global picture of how ZNF263 orchestrates the poised state and subsequent early differentiation of hESCs (Fig. 8), which further provides a valuable clue to dissect the central regulatory network controlling human early development⁴⁵.

In hESCs, few priming factors in hESCs have been identified^{46–48}. Moreover, these so-called priming factors seem to act more like lineage switchers rather than coordinators between pluripotency and multi-lineage priming, as they are only capable of promoting the propensity towards specific lineage rather than all primary germ layers in hESCs. The function of ZNF263 in promoting both mesendodermal and ectodermal lineage commitment of hESCs, as manifested on the individual cell level (Fig. 7), echoes its pluripotency priming role (Figs. 4, 5), demonstrating that ZNF263 acts as a multi-lineage priming factor in human pluripotency and development. This priming process mediated by ZNF263 represents a crucial step towards subsequent lineage differentiation. Next, how ZNF263 and other factors integrate diverse extracellular cues to coordinate respective differentiation programs should be a fascinating question to address⁴⁹. In fact, transcriptional profiling of trilineage-differentiated H1 hESCs under *ZNF263* knockout revealed extensive, lineage-dependent interactions between ZNF263 and core developmental signaling pathway cascades (e.g., Wnt, TGF- β , and PI3K-Akt; Supplementary Fig. 9). Meanwhile, a more pronounced differentiation defect upon ZNF263 loss in

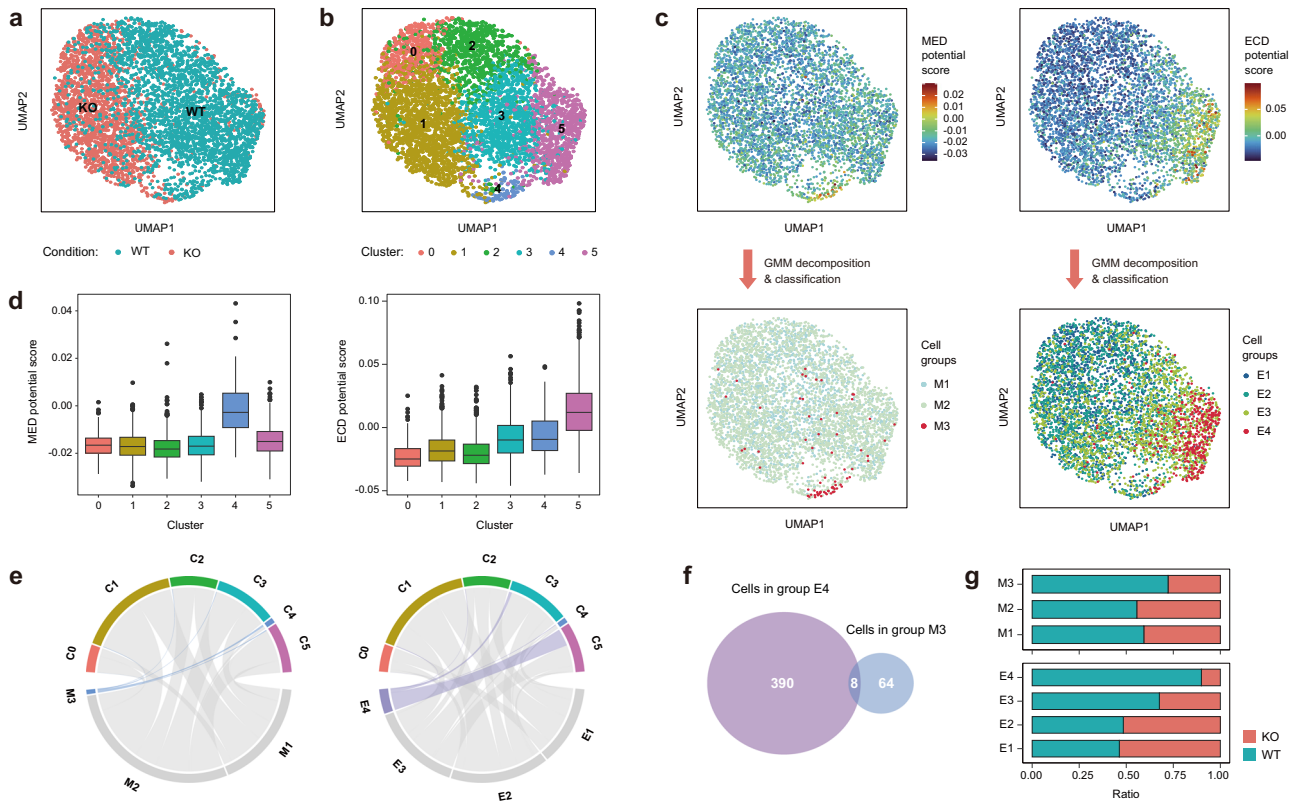


Fig. 7 | ZNF263 promotes the heterogeneity of developmental potential in primed hESCs. **a** UMAP plot depicting the single-cell profiles of WT (cyan) and *ZNF263*-KO (red, from the 1–10 clone) H1 hESCs. **b** UMAP plot displaying the six cell clusters (clusters 0–5) identified by Seurat. **c** UMAP plots illustrating the mesoderm (MED) and ectoderm (ECD) potential scores of each single cell (1st row), as well as their group labels identified by Gaussian mixture model (GMM)-based decomposition and classification (2nd row). **d** Box plots showing the distribution of MED (left) and ECD (right) potential scores for each of the six clusters identified by

Seurat. Each box shows the median (horizontal line), second to third quartiles (box), and Tukey-style whiskers (beyond the box). **e** Chord diagram showing the correspondence between the cell clusters (C0–5) identified by Seurat and the cell groups identified by the GMM-based approach with different MED (M1–3, left) and ECD (E1–4, right) potential scores. **f** Venn diagram illustrating the overlapping cells between group E4 and group M3, which respectively exhibit the highest ECD and MED potential scores. **g** Bar plots showing the proportion of WT (cyan) and *ZNF263*-KO (red) cells in each cell group identified by the GMM-based approach.

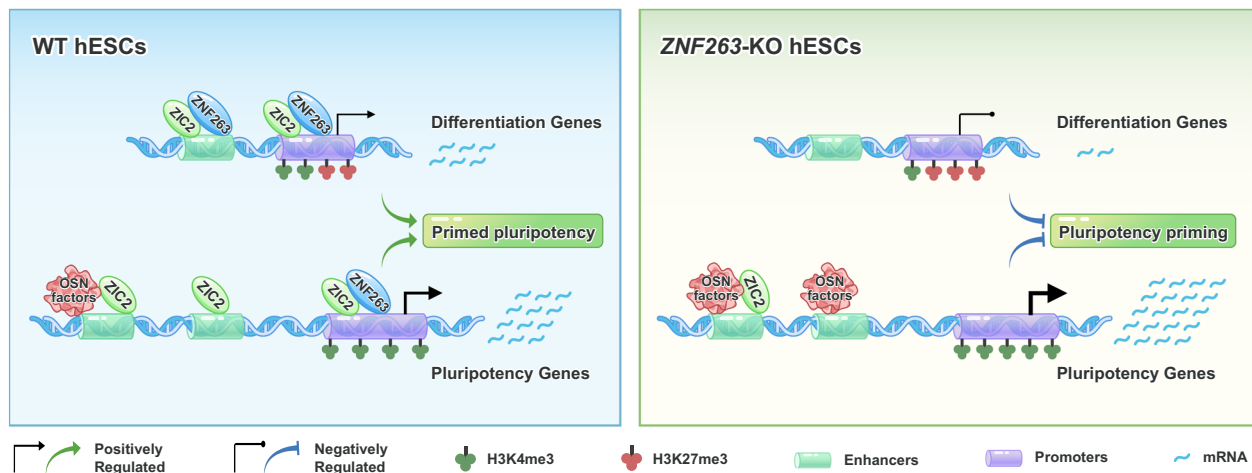


Fig. 8 | Mechanistic model of ZNF263-mediated pluripotency priming in hESCs. Schematic illustrating transcriptional regulation in wild-type (WT) and *ZNF263*-knockout (KO) hESCs. In WT hESCs, ZNF263 and its downstream target ZIC2 co-localize at the promoters and distal enhancers of bivalent genes, forming a positive feed-forward loop that drives the incipient expression of early differentiation genes. Simultaneously, ZNF263 and ZIC2 bind to the promoters of many pluripotency genes, hindering the regulatory effect of core pluripotency factors OCT4, SOX2, and NANOG (termed OSN) on these genes. This dual regulatory action of

ZNF263 maintains a balanced primed pluripotency state in hESCs. Upon *ZNF263* depletion, ZIC2 drastically loses its chromatin binding, and the feed-forward loop mediated by the *ZNF263*-ZIC2 axis collapses, leading to the transcriptional silencing of differentiation genes and rejuvenation of pluripotency gene expression in hESCs. *ZNF263* plays a pivotal role in shaping the primed pluripotency transcriptional program, while its functional loss impairs the efficiency of pluripotency exit and subsequent lineage commitment in hESCs.

ectoderm was observed compared to mesoderm and endoderm lineages in both H1 and H9 cells, as evidenced by blunted induction of corresponding marker genes (Fig. 6h–j and Supplementary Fig. 10d–f). This implies a preferential role for ZNF263 in ectodermal commitment, echoed by its tight co-regulation with ZIC2, the master regulator of neuroectoderm specification, in primed hESCs. To elucidate how the ZNF263-ZIC2 axis drives this process, systematic dissection of its crosstalk with the ectoderm-related signaling pathways and their downstream transcriptional regulators calls for further exploration.

Notably, although the transcriptomic changes induced by ZNF263 deficiency, especially the increased expression of pluripotency factors, resemble those observed in reprogrammed naïve hESCs, we didn't compare ZNF263-KO hESCs with hESCs in naïve or other intermediate pluripotent states^{50,51}, as their defects in PSD and differentiation made us doubt whether they can still be considered “pluripotent”. In future, developing an inducible system to selectively deplete ZNF263 in target cells holds great promise for investigating its role at different stages of pluripotency and lineage differentiation⁵².

Until now, we have systematically demonstrated the necessity of ZNF263 in the pluripotency priming and lineage commitment of hESCs through loss-of-function experiments. On the other hand, whether its expression is regulated during this process and, more critically, what the functional consequences are remains an outstanding question. We examined ZNF263 expression during EB differentiation of H9 hESCs using a public scRNA-seq dataset⁵³. It was observed that ZNF263 was clearly upregulated during hESC differentiation towards both mesoderm and ectoderm fates and, intriguingly, its peak expression occurred at the final stage of pluripotency exit and early stage of lineage fate induction (Supplementary Fig. 12a, b), implying that its expression is critical for facilitating this transition. We thus performed ZNF263 OE in hESCs. Of note, we observed a modest yet significant effect: quite a number of ZDA differentiation and ZDR pluripotency genes exhibited further increased and decreased expression levels, respectively, following ZNF263 OE in H1 hESCs (Supplementary Fig. 12c, d). Furthermore, in the direct differentiation experiments of all three germ layers, we observed a more pronounced induction of essential lineage-specific genes in ZNF263 OE H9 hESCs compared to WT (Supplementary Fig. 12e). Together, these findings suggest that ZNF263 OE may enhance the potential of hESCs for more efficient differentiation. In the future, to help establish a more effective system for the directed differentiation of human PSCs and to promote their application in regenerative medicine, the effect of ZNF263 OE could indeed be a relevant point for further investigation.

Methods

In-silico screening to search for key transcription factors in conventional human embryonic stem cells

We collected the ChIP-seq data of three active histone modifications H3K4me3, H3K27ac, and H3K9ac, as well as their input control in H1 hESCs and 15 other human cell lines generated by the Broad Institute in the ENCODE Consortium¹⁸. Aligned BAM files were obtained and peaks were called on each ChIP-seq sample against the corresponding input control using MACS v1.3.7⁵⁴ with parameters “--nomodel --shiftsize 100”. For each sample, peaks were ranked by MACS *P*-values, and only the top 25,000 peaks were kept for downstream analysis. If the total number of peaks was less than 25,000, all peaks were retained. Then, MANorm2 v1.2.0⁵⁵ was applied to normalize and compare the histone modification ChIP-seq samples between H1 hESCs and each non-ESC cell line. For each peak, the log₂-ratio of the average ChIP intensities (M-value) between H1 hESCs and the non-ESC cell line, which was reported by MANorm2, was used as the key metric for testing the association between TF binding and hESC-biased peaks using MAMotif v1.1.0¹⁹. Here, TF binding information at the peak regions was obtained from motif scanning. We downloaded the PFM (position frequency matrix) file of all vertebrate TF motifs (redundant set, version 2020)

from JASPAR database⁵⁶. MotifScan module in the MAMotif toolkit was used to scan the genomic sequences of peak regions with these TF motifs to detect all motif occurrences with parameters “-g hg19 -p 1e-4 -w 1000”. Then, MAMotif was employed to identify TF motifs that are significantly associated with the H1 hESC-biased peaks of each histone modification, with the M-values of the histone modification between H1 hESC and the non-ESC cell line being compared together with the motif occurrences of JASPAR TF motifs at the peak regions. *P*-values reported by MAMotif were corrected for multiple testing using the Benjamini-Hochberg method. Of note, all ChIP-seq peaks were split into promoter peaks and distal peaks based on the genomic location relative to RefSeq annotated genes, and the above MAMotif analysis was separately performed for these peaks. The above procedure was also applied to separately compare the H3K4me3 ChIP-seq data of primed and reprogrammed naïve hESCs from two independent studies^{12,13}. For ChIP-seq data that had no replicates, the M-values derived by MANorm v1.3.0⁵⁷ were used.

Integrative analysis with ENCODE RNA-seq data

We collected the 75 bp paired-end RNA-seq data generated by Caltech in the ENCODE Consortium, which covers 9 of the 16 human cell lines used in our in-silico screening. The RNA-seq reads in FASTQ files were aligned to hg19 reference genome using STAR v2.7.6a⁵⁸ based on RefSeq gene annotations. Gene-level RNA-seq read counts were derived using featureCounts 2.0.1⁵⁹ with parameters “-c -p”. Differential expression analysis was repeatedly performed between H1 hESCs and each non-ESC cell line using DESeq2 v1.28.1⁶⁰. Genes with absolute log₂-fold change >1 and adjusted *P*-value < 0.001 were identified as differentially expressed genes. To test the association between the occurrence of a specific TF motif at promoters and the genes preferentially expressed in H1 hESCs, all 11769 genes occupied by H3K27ac peaks at promoters in H1 hESCs, which can be considered as a repertoire of all actively transcribed genes in H1 hESCs, were used as the background in subsequent analysis. Taking the comparison between H1 hESCs and K562 cells as an example, genes whose promoters were occupied by the H3K27ac peaks containing specific motifs were respectively intersected with genes more highly expressed in H1 hESCs (termed H1-high genes) and genes more highly expressed in K562 cells (termed K562-high genes). The enrichment score was calculated as the ratio between the number of overlapping genes and that expected by chance, and the associated *P*-values were calculated using right-tailed Fisher's exact test.

Cell lines and culture conditions

H1 hESC line (WA01) was obtained from National Collection of Authenticated Cell Cultures (NCACC) in Shanghai at the passage of 35. H9 hESC line (WA09) was also provided by NCACC. The hESCs were maintained on Matrigel-coated plates (BD Biosciences, Cat#354277) in mTeSR1 complete medium (STEMCELL Technologies, Cat#85850). Medium was changed daily. For routine maintenance, cells were passaged every 5–7 days with Cell Gentle Dissociation Buffer (STEMCELL Technologies, Cat#07174) as aggregates. For other conditions, cells were split with TrypLE Express Enzyme (Thermo Fisher Scientific, Cat#12604013) to obtain single cells or small cell clumps. Meanwhile, when seeded, 10 μM ROCK inhibitor Y-27632 (Selleck Chemicals, Cat#S1049) was supplemented but eliminated at the following day.

CRISPR-Cas9 mediated genome editing of hESCs

The sgRNAs were designed to target Exon2 of human ZNF263 gene using Broad Institute sgRNA designer (<https://portals.broadinstitute.org/gppx/crispick/public>), and then synthesized and constructed into px330-mcherry or lentiCRISPR v2 vector (see below). Plasmid of px330-mcherry was kindly provided by Dr. Jinsong Li, and lentiCRISPR v2 was a gift from Feng Zhang (Addgene, plasmid #52961). To generate

ZNF263-knockout H1 hESCs, recombinant px330-mcherry plasmids expressing *ZNF263* sgRNA were transfected to H1 single cells with Lipofectamine 3000 (Thermo Fisher Scientific, Cat#L3000015) according to the manufacturer's protocol. 24 hr later, mcherry positive cells were enriched by fluorescence-activated cell sorting using MoFlo Astrios EQ high-speed cell sorter (Beckman Coulter), and then deposited onto a 6-well plate. After expansion, cell colonies were manually picked under microscope and then genotyped by PCR. For *ZNF263*-knockout H9 hESC generation, H9 cells were infected with lentivirus expressing *ZNF263* sgRNA in the presence of 6 µg/mL polybrene. 48 hr later, cells were treated with 1 µg/mL puromycin (TargetMol Chemical, Cat# T2219-1mg) for 3 days when uninfected cells were completely killed. Survival cells were digested to single cells and then subjected to cell colony formation. Genotyping was performed as described below. To serve as a wildtype control, a portion of H1 and H9 hESCs were "mock"-edited by transfection and infection without *ZNF263* sgRNA. The knockout hESCs and their wild-type control are maintained at comparable passage levels.

Plasmid construction

The sgRNA-expressing vector construction was performed as described^{61,62}. Briefly, sgRNA containing sequences were synthesized as two partially complementary oligos (sgRNA_{exon2}-For and sgRNA_{exon2}-Rev) that had 4 nt overhangs compatible for cloning into the vector. Then, Oligos were put together, annealed, and ligated with BbsI-digested px330-mcherry or BsmBI-digested lentiCRISPR V2 vector. Successfully recombinant plasmids px330-*ZNF263* sgRNA-mcherry and lentiCRISPR-*ZNF263* sgRNA were validated by sequencing.

To generate the plasmid that expresses a FLAG-tagged version of *ZNF263*, the human *ZNF263* ORF was first amplified using cDNA from H1 hESCs with primer pairs of pcDNA3.1-*ZNF263*-3xFlag-OE-For and pcDNA3.1-*ZNF263*-3xFlag-OE-Rev, then the PCR amplicons were double-digested by XbaI and XhoI (New England BioLabs) and cloned into pcDNA3.1(+)-3xFlag vector to get pcDNA3.1(+)-*ZNF263*-3xFlag construct. Using pcDNA3.1(+)-*ZNF263*-3xFlag plasmid as template, FLAG-tagged *ZNF263* expressing sequence was amplified using primers of pLVX-*ZNF263*-3xFlag-zsGreen1-OE-For and pLVX-*ZNF263*-3xFlag-zsGreen1-OE-Rev, and cloned into XbaI and NotI (New England BioLabs) double-digested pLVX-EF1α-IRES-zsGreen1. A correctly inserted plasmid was obtained by sequencing validation.

To generate the over-expressing plasmid of *ZIC2* in hESCs, human *ZIC2* ORF was firstly amplified from a commercial plasmid (Miaoling Biotechnology, Cat#P55896) with primer pairs of pLVX-*ZIC2*-3xFlag-zsGreen1-For and pLVX-*ZIC2*-3xFlag-zsGreen1-Rev, then PCR amplicons were seamlessly cloned into linearized vector pLVX-EF1α-IRES-zsGreen1 with BamHI and EcoRI. A correctly inserted plasmid was obtained by sequencing validation. All the primers mentioned are provided in Supplementary Data 1a.

Lentivirus production, precipitation and infection

In a 100 mm plate cultured with HEK293T cells, 5 µg plasmids for over-expression (i.e., pLVX-EF1α-IRES-zsGreen1) or knockout (i.e., lentiCRISPR V2) were respectively co-transfected with packaging plasmids (psPAX2 and pMD2.G) at the ratio of 4:3:1, using FuGENE HD Transfection Reagent (Promega) according to the manufacturer's instructions. Subsequently, virus particles released into the medium were harvested at 48 hr and 72 hr after transfection, and centrifuged at 140,000 *g* for 2 hr at 4 °C using Hitachi CP80MX with a P40ST rotor. The pellet containing concentrated virus was resuspended in 1X DPBS and aliquoted to store at -80 °C. The day before infection, hESCs were trypsinized into small cell clumps and supplemented with Y-27632. On the following day, lentivirus was directly added into the medium in the presence of 6 µg/mL polybrene. After 9 hr infection, the medium was changed to mTeSR1. 48 hr later, cells were subject to downstream processing.

Genotyping of knockout hESCs

Genomic DNA of *ZNF263*-knockout hESCs and wild-type control was extracted with a DNA extraction kit (Qiagen). Subsequent genotyping PCR was conducted with primers listed here and further confirmed by sequencing.

2sg_T7E1_validation-For: 5'-TCCACGGTCTGGTCTTCTCT-3'.

2sg_T7E1_validation-Rev: 5'-TTCCTTCTCCTTCGCATCATG-3'.

Chromatin immunoprecipitation and library preparation for sequencing

Approximately 1×10^7 Cells were digested with TrypLE Express Enzyme and crosslinked with 1% Formaldehyde at room temperature for 10 min. Subsequently, 2.5 mM Glycine was used to stop the fixation. After washing twice with 1X DPBS, cells were resuspended in SDS lysis buffer (1% SDS, 5 mM EDTA, 50 mM Tris-HCl at pH 8.0, plus protease inhibitor) and incubated on ice for 10 min. Immediately lysate was sonicated for 6 or 7 times at 30 sec each using Diagenode Bioruptor with high power on. Sheared chromatin mixture was centrifuged for 10 min at 14,000 *g* at 4 °C, 10 µL of the supernatant was set aside as input and reverse-crosslinked at 65 °C overnight with 110 µL Elution buffer (0.5% SDS, 5 mM EDTA, 300 mM NaCl, 10 mM Tris-HCl at pH 8.0). RNase A/T1 (20 µg/mL, Thermo Fisher Scientific, Cat#EN0551) and Proteinase K (200 µg/mL, Thermo Fisher Scientific, Cat#AM2546) were used to digest RNA and proteins successively. Then DNA was purified with QIAquick PCR Purification Kit (Qiagen, Cat#28106) and subject to DNA electrophoresis, aiming to ensure that the fragmented DNA length ranges from 200 to 600 bp. A portion sample equivalent to 10 µg (for histone modification ChIP assay) or 30 µg chromatin (for *ZNF263* ChIP assay) was aliquoted and diluted 1:10 with ChIP dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl at pH 8.0, plus protease inhibitor). Specific antibodies were added and incubated overnight at 4 °C, followed by the addition of 50 µL M-280 Sheep Anti-Rabbit IgG Dynabeads (for histone modification ChIP assay, Thermo Fisher Scientific, Cat#11203D) or protein A/G-Sepharose beads (for *ZNF263* ChIP assay, Santa Cruz Biotechnology, Cat#sc-2003) and incubation for another 2 hr at 4 °C. Conjugated beads were harvested and wash sequentially for 5 min with TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl at pH 8.0, 150 mM NaCl), TSE II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl at pH 8.0, 500 mM NaCl), TSE III (0.25 M LiCl, 1% NP-40, 1% Deoxycholate, 1 mM EDTA, 10 mM Tris-HCl at pH 8.0), and TE buffer (10 mM Tris-HCl at pH 8.0, 1 mM EDTA). Washed beads were subject to DNA elution and purification as above. DNA concentration was measured with Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Cat#Q32854) in Qubit 3.0 Fluorometer (Thermo Fisher Scientific), then 10 ng input DNA and immunoprecipitated DNA were respectively used for sequencing library preparation with NEBNext Ultra II DNA Library Prep Kit (New England BioLabs, Cat#E7645S) according to the manufacturer's recommendations. Index incorporation was applied as the same as RNA sequencing library. The primary antibodies and their dilutions used are provided in Supplementary Data 1b.

RNA isolation and library preparation for sequencing

Total RNA of each sample was extracted with TRIzol Reagent (Thermo Fisher Scientific, Cat#15596018) according to the manufacturer's protocol. RNA purity and quantity were measured by Nanodrop 2000. RNA quality was confirmed with Agilent RNA 6000 Nano Kit (Agilent, Cat#5067-1511) on Agilent Bioanalyzer 2100 system (Agilent Technologies). Samples with RNA integrity number (RIN) above 9 were subject to library preparation. In brief, mRNA from each sample was enriched from 0.5 µg total RNA with NEBNext Poly(A) mRNA Magnetic Isolation Module (New England BioLabs, Cat#E7490S), and then converted to cDNA. Following the cDNA generation, fragmentation, adaptor ligation, and amplification with NEBNext Ultra II RNA Library Prep Kit (New England BioLabs, Cat# E7775S), the libraries were purified with AMPure

XP beads (Beckman Coulter, Cat#A63880). Index primers from NEB-Next Multiplex Oligos for Illumina (Dual Index Primers set 1) (New England BioLabs, Cat#E7600S) were used to enable index incorporation during library amplification.

Cleavage Under Targets & Tagmentation (CUT&Tag)

Routinely cultured hESCs were digested into single cells with TrypLE and subject to the CUT&Tag direct protocol⁶³ with some modifications. In brief, hESCs were resuspended in 0.3 mL NEI buffer (20 mM HEPES-KOH at pH 7.9, 10 mM KCl, 0.5 mM spermidine, 0.1% Triton X-100, 20% glycerol, plus protease inhibitor), then cell nuclei were obtained by centrifugation at 600 g for 5 min at 4 °C. After rinsing 2 times with Wash Buffer (20 mM HEPES at pH 7.5, 150 mM NaCl, 0.5 mM spermidine, plus protease inhibitor), approximately 6×10^4 nuclei were bound to 3.75 μ L of concanavalin A (ConA) beads (Cell Signaling Technology, Cat#93569S) according to the manufacturer's instructions. Subsequently, 0.5 μ L primary antibody (anti-ZIC2, ABCAM, Cat#150404) and 0.5 μ L normal IgG (Rabbit IgG, Millipore, Cat#12-370) was respectively diluted with 25 μ L of Antibody Buffer (10 mM HEPES at pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 2 mM EDTA, 0.1% BSA, plus protease inhibitor) and mixed with the above conjugated beads at 4 °C overnight. After thoroughly aspirating the supernatant, 0.25 μ L secondary antibody (Guinea Pig anti-Rabbit IgG, Antibodies-online, Cat#ABIN101961) in 25 μ L Wash Buffer was added and incubated with the beads at room temperature for 1 hr. Extra antibodies were removed by washing the beads twice with 0.25 mL Wash Buffer. Then, custom-assembled pA-Tn5 transposases⁶⁴ were loaded to the beads in 25 μ L of 300-Wash Buffer (Wash Buffer with 300 mM NaCl) at room temperature for 1 hr. Following, unbound transposases were washed away with 300-Wash Buffer from the beads, and tagmentation was conducted in 50 μ L of 300-Wash Buffer supplemented with 10 mM MgCl₂ at 37 °C for 1 hr. After thoroughly aspirating the supernatant, the beads were washed once with TAPS buffer (10 mM TAPS at pH 8.5, 0.2 mM EDTA), then 5 μ L SDS release buffer (10 mM TAPS at pH 8.5, 0.1% SDS) was added and incubated at 58 °C for 1 hr. The resulting suspension was mixed well with 15 μ L of 0.67% Triton X-100, followed by the addition of 2 μ L barcoded i5 primer (10 μ M), 2 μ L barcoded i7 primer (10 μ M), and 25 μ L of 2X NEBNext PCR Master Mix (New England Biolabs, Cat#M0541S). Then, 14–17 cycles of PCR reaction were performed. Amplified tagmented DNA was purified from the supernatant using 1.3X AMPure XP Beads, and subject to library sequencing.

Library sequencing for ChIP-seq, RNA-seq, and CUT&Tag

Library quantity was assessed in Qubit 2.0 Fluorometer (Life Technologies, CA, USA), and the quality was evaluated on the Agilent Bioanalyzer 5400 system. Qualified DNA libraries were subject to 150 bp paired-end sequencing on an Illumina NovaSeq 6000 platform (Illumina, CA, USA). Then Fastp v0.23.1⁶⁵ was used to filter adapter-contaminated reads and low-quality nucleotides with the parameter set as “-g -q 5 -u 50 -n 15 -l 150 --overlap_diff_limit 1 --overlap_diff_percent_limit 10”. Remaining high-quality reads were preserved for downstream analysis.

Processing and downstream analysis of ZNF263 ChIP-seq data

Qualified reads from ChIP-seq data of ZNF263 were aligned to the hg19 human genome using Bowtie2 v2.3.5.1⁶⁶ with parameters “--very-sensitive --no-mixed -X 2000”. Only the reads pairs with both ends properly aligned and mapping quality greater than 30 were kept. PCR duplications were detected and removed by the “MarkDuplicates” function in Picard tools. Peaks were called using MACS with parameters “--nomodel --shiftsize 100”. The detected peaks were further classified by the genomic location relative to RefSeq annotated genes. Those located within 2 kilo base-pair (kb) upstream or downstream of any RefSeq gene's transcript start site (TSS) were defined as promoter peaks. The remaining peaks were sequentially overlapped with gene exons and introns, and those failed to overlap with these features were

classified as intergenic peaks. ChIP-seq data generated with ZNF263 endogenous antibody and FLAG antibody were compared using MAnorm⁵⁷ to measure their consistency. Occurrences of ZNF263 motifs within ZNF263 ChIP-seq peak regions were detected using the MotifScan module in MAMotif⁴⁹.

Identification of active cis-regulatory elements (aCREs) in H1 hESCs

To globally define the aCREs at promoter and distal regions in H1 hESCs, BAM file of the DNase-seq and histone modification ChIP-seq data in H1 hESCs was downloaded from ENCODE webpage, and the DNase-seq peaks were called using MACS with parameters “--nolambda --nomodel --shiftsize 1”, and those located within the blacklist annotation were excluded. The final peaks were then classified into promoter and distal (non-promoter) ones as described above, and overlapped with histone modification peaks. Promoter aCREs were defined as promoter DNase-seq peaks that overlap with both H3K27ac and H3K4me3 peaks but not H3K27me3 peaks. Distal aCREs were defined as distal DNase-seq peaks that overlap with H3K27ac peaks but do not overlap with H3K4me3 and H3K27me3 peaks. Finally, we defined 9542 promoter aCREs and 8876 distal aCREs. Again, MotifScan was used to scan their genomic sequences to detect the occurrence of TF motifs at these elements.

Transcription factor (TF) co-localization analysis

We extensively collected available TF ChIP-seq peaks in H1 hESCs from ENCODE, the Gene Expression Omnibus (GEO), and ArrayExpress. The sources of these data were indicated by the suffix of each TF shown in Fig. 3c and Supplementary Fig. 3c. Specifically, TF ChIP-seq data with suffixes _SYDH, _HAIB, and _Broad were downloaded from the corresponding ENCODE webpages; SOX2 and NANOG ChIP-seq data with suffix _Ecker were generated by the Joseph R. Ecker lab (GSE18292); PRDM14, ERK2, and ELK1 ChIP-seq data with suffix _Ng were generated by the Huck-Hui Ng lab (GSE22767 and E-MTAB-1565).

Pairwise TF co-localization patterns were evaluated, based on the enrichment of overlapping aCREs that contain the ChIP-seq peaks or motif sites of the two TFs. The associated *P*-values were calculated using one-sided Fisher's exact test, and the sign and side of each *P*-value (after being log₁₀ transformed) were determined by whether the number of co-localized aCREs between the two TFs was larger (right-tailed, assigned with positive values) or smaller (left-tailed, assigned with negative values) than that expected by chance. Then, TF co-regulatory networks at promoter and distal aCREs were constructed using Cytoscape v3.8.0⁶⁷, in which each edge linking two TFs (peaks or motifs) represented a significantly enriched TF-TF co-localization at corresponding aCREs. In each network, TFs that were linked with the motif or ChIP-seq peaks of ZNF263 but not with those of OSN factors were classified into the ZNF263 module, while TFs that were linked with the motif or ChIP-seq peaks of POU5F1, SOX2, or NANOG but not with those of ZNF263 were classified into the OSN module. TFs that were linked with the motif or ChIP-seq peaks of both ZNF263 and OSN factors were assigned to the shared module. TFs that were not directly linked with the motif or ChIP-seq peaks of ZNF263 or OSN factors were removed.

Western blot

For immunoblotting, total proteins or nuclei proteins were extracted with RIPA buffer, and then boiled with SDS PAGE loading buffer (Takara, Cat#9173). Following electrophoresis on 12% Tris-Glycine mini gel, proteins were wet transferred to polyvinylidene difluoride membrane (Merk/Millipore, Cat# IPVH00010) at 110 V for 90 min. After blocking with 5% milk for 2 hr at room temperature, the indicated primary antibodies dissolved in blocking buffer were applied to the membrane and incubated overnight at 4 °C. On the following day, the hybrid membrane was incubated with Goat Anti-Rabbit IgG Peroxidase

Antibody (Sigma-Aldrich, Cat#A0545-1ML) or Anti-Mouse IgG Peroxidase Antibody (Sigma-Aldrich, Cat#A4416) for a further 2 hr. Signal was developed by ECL Prime Western Blotting Detection Reagent (Amersham/GE, Cat# RPN2232) and captured by the Bio-Rad Chemi-Doc XRS+ system. The primary antibodies and their dilutions used are provided in Supplementary Data 1b. Mouse Monoclonal ZNF263 Antibody (ABCAM, Cat#61330) was used to detect the ZNF263 protein level in H1 cells, and ZNF263 antibody from Sigma-Aldrich (Cat#-SAB1412641-100UG) was applied to H9 cells. Signals from GAPDH (Cell Signaling Technology, Cat#2118S) or β -actin (Proteintech, Cat#60008-1-Ig) were used as loading control.

Processing and downstream analysis of ZNF263-knockout (-KO) RNA-seq data

The RNA-seq reads in FASTQ files were aligned to hg19 reference genome using STAR v2.7.6a⁵⁸ based on RefSeq gene annotations. Gene-level RNA-seq read counts were derived using featureCounts v2.0.1⁵⁹ with parameters “-c -p”. Differential expression analysis was performed separately using DESeq2 in H1 and H9 hESCs between ZNF263-KO clones and the wild-type control. In H1 hESCs, genes that showed differential expression by ZNF263 KO were identified using the criteria of fold change >1.2 and adjusted *P*-value <0.1 as the cutoffs. In H9 hESCs, differentially expressed genes by ZNF263 KO were identified using the same cutoffs. Then, we further identified the consistently up-regulated and down-regulated genes in H1 and H9 ZNF263-KO hESCs from the differentially expressed genes detected from these two comparisons using the following criteria: 1) with adjusted *P*-value < 0.1 in both comparisons; 2) with fold change >1.2 in at least one comparison; 3) showing a consistent direction of expression change in the two comparisons.

The ZNF263 directly activated (ZDA) genes in H1 hESCs were defined as the significantly down-regulated genes by ZNF263 knockout that also had ZNF263 ChIP-seq peaks at promoters. The ZNF263 directly repressed (ZDR) genes in H1 hESCs were defined as the significantly up-regulated genes by ZNF263 knockout that also had ZNF263 ChIP-seq peaks at promoters. Gene Ontology (GO) analysis was performed using custom scripts based on the GO term annotations download from MSigDB⁶⁸. *P*-values were calculated using right-tailed Fisher's exact test, which were then subjected to multiple testing correction using the Benjamini-Hochberg method to control for false discovery rate (FDR).

Alkaline phosphatase and extracellular stem cell marker staining

For alkaline phosphatase staining, hESCs were cultured in a 12-well plate for the indicated time and rinsed twice for 5 min with pre-warmed knockout DMEM/F12 medium (Thermo Fisher Scientific, Cat#12660012). Alkaline Phosphatase Live Stain (Thermo Fisher Scientific, Cat#A14353) was diluted at a ratio of 1:500 with DMEM/F12 medium and added to the cells in a 37 °C incubator. After 20 min incubation, hESCs were washed twice with 1X DPBS, digested into single cells with TrypLE Enzyme, stopped with 5 times' volume of 1X DPBS, and then subject to flow cytometry analysis.

For extracellular stem cell marker staining, approximately 1×10^6 hESCs were digested to single cells and fixed with 1% Formaldehyde (Sigma-Aldrich, Cat# 252549-100 ML) for 10 min. After stopping reaction by 2.5 mM Glycine, cells were washed twice with 1X DPBS, blocked with 2% BSA at room temperature for 1 hr, and then stained with Alexa Fluor 488 anti-human/mouse SSEA-3 Antibody (BioLegend, Cat#330305) and Alexa Fluor 647 anti-human TRA-1-81 Antibody (BioLegend, Cat#330705) simultaneously at the ratio of 1:50. Cells washed twice were ready for flow cytometry analysis.

Flow cytometry (FACS) was performed on a Cytotflex LX flow cytometer (Beckman Coulter) equipped with lasers of 488 nm and 640 nm. About 25,000 cells were collected at a low flow rate, and then

the data were analyzed by official CytExpert software v2.3 and FlowJo 10 software.

Reverse transcription and quantitative real-time polymerase chain reaction (qRT-PCR)

Reverse transcription and residual DNA elimination were performed using 1 μ g total RNA with ReverTra Ace qPCR RT Master Mix (Toyobo, Cat#FSQ-301) according to the manufacturer's instructions. Then, qPCR was conducted using 2X SYBR Green qPCR Master Mix (Bimake, Cat#B21202) on QuantStudio 7 Flex Real-Time PCR System (ABI Technologies) in triplicate per sample. PCR conditions consisted of 1 cycle of 95 °C for 5 min and 40 cycles of 95 °C for 5 s and 60 °C for 40 s. Relative mRNA levels of each gene were normalized against GAPDH gene expression with equipped QuantStudio Real-Time PCR Software (v.1.7.1) using $2^{-\Delta\Delta Ct}$ method. Primers were designed as PrimerBank suggests (<https://pga.mgh.harvard.edu/primerbank/>) or using Primer Premier 6. All the qRT-PCR primers are provided in Supplementary Data 1a.

Processing and analysis of histone modification ChIP-seq data in ZNF263-knockout hESCs

The ChIP-seq data of H3K4me3, H3K27ac, and H3K27me3 histone modifications from both ZNF263-KO H1 hESCs and its WT control were processed in the same way as the ZNF263 ChIP-seq data. MANorm was applied to compare the ChIP-seq signal changes between ZNF263-knockout and control cells. The log₂-ratio of the average ChIP-seq intensity (M-value) between them (KO/WT) was reported for each peak and used for downstream analysis.

Re-analysis of the time-course RNA-seq data of hESC differentiation towards definitive endoderm (DE) and ectodermal neural cells (EN)

FASTQ files of the DE differentiation data were downloaded from GEO (GSE164361)²⁵, and processed using the same procedure as the above RNA-seq data. DESeq2 was repeatedly used to call the differentially expressed genes in each differentiation time point compared to undifferentiated hESCs at day 0, with the cutoff set as adjusted *P*-value < 0.01. Finally, the genes up/down-regulated at any time point of the differentiation process were merged and grouped into different clusters, based on the first time point at which the genes were identified as significantly up/down-regulated genes.

For EN differentiation, the normalized gene expression matrices were downloaded from GSE103715²⁶. Limma v3.44.3⁶⁹ was repeatedly used to identify the differentially expressed genes in each differentiation stage compared to undifferentiated hESCs at the pluripotency stage, and then the same procedure as above was applied to generate gene clusters.

Inferring transcription regulators associated with genes differentially expressed by ZNF263 knockout in H1 hESCs

Lisa tool v2.3.0²⁸ was employed to infer transcription factors and chromatin regulators, collectively termed transcription regulators (TRs), that are respectively associated with the up- and down-regulated genes by ZNF263 loss in H1 hESCs. For integrative modeling, TR ChIP-seq peaks of H1 and H9 hESCs from Cistrome Data Browser (CistromeDB) were utilized.

Processing and downstream analysis of ZIC2 CUT&Tag data

Qualified CUT&Tag sequencing reads were aligned to hg19 human genome using Bowtie2 v2.3.5.1 with parameters “--very-sensitive --no-mixed -X 2000”. Only the reads pairs with both ends properly aligned and mapping quality greater than 30 were kept. PCR duplications were detected and removed by the “MarkDuplicates” function in Picard tools. Peaks were called using MACS with parameters “--nomodel --shiftsize 100”. Peaks were called using MACS with parameters “--nomodel --shiftsize 100”. The detected peaks were merged and

analyzed using MAnorm2, and the merged peaks showing a fold change of normalized CUT&Tag signal intensities >1.5 between two biological replicates were considered as unreliable peaks and removed.

Mitogen-activated protein kinase (MEK) inhibitor treatment

MEK inhibitor (PD0325901) was purchased from Sigma-Aldrich (Cat#PZ0162-5MG). Procedures were conducted as previously described³⁴. In brief, hESCs were seeded in a 12-well plate on day 0, Twenty-four (24) hr after seeding, mTeSR1 was replaced with the MEK pathway inhibition (mTeSR1 + 3 μ M PD0325901) for 48 h. Then, cells were collected for downstream functional analysis.

Processing of MEK inhibitor (MEKi) treatment RNA-seq data

The RNA-seq reads in FASTQ files were aligned to the hg19 reference genome using STAR v2.7.6a based on RefSeq gene annotations. Gene-level RNA-seq read counts were derived using featureCounts v2.0.1 with parameters “-c -p”. Differential expression analysis was performed using DESeq2 to separately compare vehicle-treated WT H1 cells with MEKi-treated WT H1 cells, vehicle-treated 1-10 *ZNF263*-knockout H1 cells, and MEKi-treated 1-10 *ZNF263*-knockout H1 cells. The differentially expressed genes of these comparisons were identified using the criteria of fold change >1.2 and adjusted *P*-value <0.1 . The genes differentially expressed upon MEKi-treatment were designated as MEKi-sensitive genes, and MEKi-sensitive genes whose expression changes were significantly attenuated in MEKi-treated *ZNF263*-KO cells compared to MEKi-treated WT cells (i.e., showing significant expression changes in the opposite direction) were nominated as *ZNF263*-dependent MEKi-sensitive genes. Gene set enrichment analysis of KEGG-annotated pathways was performed using the *fgsea*⁷⁰ function in R (default parameters) for: (i) MEKi-versus vehicle-treated WT H1 hESCs, and (ii) *ZNF263*-KO (1-10 clone) versus WT H1 hESCs under either MEKi or vehicle treatment.

Immunoprecipitation

On the first day, approximately 3×10^5 single cells of *ZNF263*-KO H1 hESCs were seeded in a 6 cm culture dish. On the following day, the cells were respectively infected with virus packaged from pLVX-*ZNF263*-3xFlag-zsGreen plasmid or an empty vector with 6 μ g/mL polybrene. Two days after the infection, cells were harvested, and lysed with 800 μ L IP binding buffer (5 mM Hepes at pH 7.5, 10 mM EDTA at pH 8.0, 50 mM NaCl, 0.2% Triton X-100, 10% Glycerol plus proteinase inhibitor) for 30 min at 4 °C. Following centrifugation at 12,000 rpm for 10 min at 4 °C, 80 μ L of the supernatant was saved for input, while the remaining was used for immunoprecipitation with 30 μ L of FLAG magnetic beads (Selleck, Cat#B26101). After overnight incubation, the beads were collected and washed four times with IP washing buffer (5 mM Hepes at pH 7.5, 10 mM EDTA at pH 8.0, 150 mM NaCl, 0.2% Triton X-100). Finally, 40 μ L of 2x SDS loading buffer was mixed with the beads and boiled at 95 °C for 10 min. The supernatant was subject to immunoblotting against FLAG antibody (Sigma-Aldrich, Cat#F1804-200UG, 1:2000) and *ZIC2* antibody (Abcam, Cat#ab150404, 1:1000).

In vitro embryoid bodies (EBs) differentiation

For EBs generation that were used for qRT-PCR analysis, 2×10^5 hESCs in growth phase were split into single cells, and resuspended in 0.5 mL EB formation medium (STEMCELL Technologies, Cat#05893) that is supplemented with 10 μ M Y-27632 ROCK Inhibitor. Twenty-four (24) hours after adding the hESCs into the ultra-low adherence 24-well plate, EBs were generated. Replace the EBs with fresh medium without ROCK inhibitor every other day, until the indicated time. Then, EBs were harvested and utilized for downstream analysis.

For the generation of EBs in hanging drops, single-cell suspension of hESCs was diluted at a concentration of 1×10^5 cells per mL. Then, small drops of 20 μ L of the cell suspension were placed onto the inner surface of a culture dish lid. The lid was then inverted and carefully

positioned above a culture dish filled with 1x PBS, allowing the drops to hang suspended. Over time, the cells within each drop began to cluster and differentiate, ultimately forming embryoid bodies. The EBs were collected for further examination in the indicated time.

In vitro directed differentiation

Directed differentiation of hESCs into primary germ layers was performed using a functional assay kit designed for this purpose (STEMCELL Technologies, Cat#05230). For mesoderm lineage differentiation, approximately 1×10^5 hESCs were split into single cells and seeded into Matrigel-coated 24-well plates according to the manufacturer's protocol. For ectoderm and endoderm differentiation, 4×10^5 hESCs were employed. The lineage-specific medium was replaced daily until day 5 for mesoderm and endoderm and until day 7 for ectoderm differentiation. And then, the total RNA of these cells was extracted and subject to downstream qRT-PCR analysis and RNA library preparation.

Processing of in vitro directed differentiation RNA-seq data

The RNA-seq reads in FASTQ files were aligned to the hg19 reference genome using STAR v2.7.6a based on RefSeq gene annotations. Gene-level RNA-seq read counts were derived using featureCounts 2.0.1 with parameters “-c -p”. Differential expression analysis was performed using DESeq2 to separately compare the endoderm-, mesoderm-, and ectoderm-differentiated WT H1 hESCs against their undifferentiated state. The same procedure was utilized to compare the endoderm-, mesoderm-, and ectoderm-differentiated *ZNF263*-KO H1 hESCs (1-10 clones) against those from WT cells. The differentially expressed genes of these comparisons were identified using an adjusted *P*-value <0.05 as the cutoff. Gene set enrichment analysis of KEGG pathways and GO terms was performed with *fgsea* using default parameters for: (i) endoderm-, mesoderm-, or ectoderm-differentiated WT H1 hESCs versus undifferentiated WT H1 hESCs, and (ii) identically differentiated *ZNF263*-KO (1-10 clone) H1 hESCs versus differentiated WT.

ZNF263 and *ZIC2* overexpression for downstream functional analysis

To overexpress *ZNF263* in H1/H9 hESCs and *ZIC2* in H1 *ZNF263*-KO hESCs, plasmids expressing these genes (see Plasmid construction) were transfected with Lipofectamine stem transfection reagent (Thermo Scientific, Cat#STEM00003) according to the manufacturer's instructions. Mock-transfected hESCs with the empty vector served as control.

For downstream directed differentiation, a total number of 1×10^5 (for mesoderm differentiation) or 2×10^5 (for endoderm and ectoderm differentiation) single cells were seeded in a 24-plate for transfection. After transfection on the following day with a mixture of 2 μ L Lipofectamine and 0.5 μ g plasmids, the cells were incubated for 24 hr and then subject to respective lineage induction. Respectively, hESCs were induced for endoderm and mesoderm differentiation for 3 days, while ectoderm differentiation was induced for 4 days. Subsequently, cells were collected and analyzed by qRT-PCR.

For downstream transcriptomic analysis, 1×10^5 hESCs in colonies were seeded in a 24-well plate and transfected on the next day. After transfection, cells were incubated for 48 hr before being harvested for total RNA extraction and RNA library preparation.

Subsequent data analysis was performed in the same manner as that for *ZNF263*-KO RNA-seq experiments. To identify genes with distinct expression patterns upon *ZIC2* overexpression in *ZNF263*-KO H1 hESCs, we first performed all possible pairwise comparisons among *ZIC2*-overexpressing (*ZIC2*-OE) *ZNF263*-KO H1 hESCs, WT controls, and *ZNF263*-KO controls and then merged the differentially expressed genes detected from these comparisons. Subsequently, we extracted the \log_2 normalized read count matrix of these genes and applied hierarchical clustering to the z-score transformed matrix

using the `hclust` function in R, with Euclidean distance metric and Ward's linkage method (`ward.D2`). The clustering tree was then split into discrete clusters using the `cutreeDynamic` function in the `dynamicTreeCut` R package v1.63-1⁷¹, with a minimum cluster size of 50 and a `deepSplit` parameter of 2, resulting in 6 gene clusters. Gene set enrichment analysis of ZDA and ZDR genes between ZNF263-OE H1 hESCs and the WT control was conducted using the `fgsea`⁷⁰ function in R with default parameters.

Cryosection and immunolabeling

Embryoid bodies (EBs) cultured for 5 days were collected and fixed in freshly prepared 4% paraformaldehyde for 30 min, followed by washing three times with 1x PBS. The EBs were then cryoprotected by incubating in 30% sucrose overnight at 4 °C, embedded in O.C.T. compound (SAKURA, Cat#4583), and sectioned at a thickness of 10 µm using a cryostat (Leica, CM3050S). The sections were placed onto poly-L-lysine-coated slides and allowed to dry overnight before being stored at -80 °C. Slides containing sections of EBs of comparable size were removed from the freezer and equilibrated at room temperature for 30 min, after which they were immersed in 1x PBS for 10 min to remove the O.C.T. compound. Subsequently, the sections were permeabilized with 0.1% Triton X-100 at room temperature for 10 min, blocked with 5% BSA in PBS for 1 hr, and incubated overnight at 4 °C with a primary antibody against POU5F1 (Cell Signaling, Cat#2750 s, 1:200) and Tubulin (Sigma Aldrich, Cat#T6199, 1:200). After washing three times with 1x PBS, the sections were treated with a Cy3-conjugated goat anti-rabbit antibody (Jackson ImmunoResearch, Cat#115-165-003) and Alexa Fluor 488-conjugated goat anti-mouse antibody (Jackson ImmunoResearch, Cat#115-545-166) for 1 hr at room temperature, followed by final washes and mounting in a DAPI-containing medium. Fluorescence microscopy was then employed to visualize in the sections.

Microscope examination

Alkaline phosphatase-stained hESCs were imaged by a fluorescence microscope (ECHO, RVL2-K). Morphology of EBs and fluorescence were examined by optical microscopy (ZEISS, AXIO Scope A1).

Single-cell RNA sequencing experimental procedures

Routinely cultured hESCs were digested into single cells with TrypLE, re-suspended in 1X DPBS, and stained with a Viability Dye. Samples with high viability cells were kept and subject to microwell-based single-cell RNA sequencing (scRNA-seq)⁷² with some modifications. 1×10^4 cells were resuspended in 1 mL of PBS and evenly distributed onto the microwell array. Cells were allowed to load into microwells for 5 min and excess cells were washed away by gently rinsing the array three times with 2 mL of PBS. 1×10^5 barcoded oligo-dT magnetic beads were then dropwise pipetted to the microwell array, and the microwell array was incubated on a magnet to force the RNA-capturing beads to settle into microwells. Extra beads were removed by washing the microwell array with PBS. After thoroughly aspirating PBS by pipette, 1 mL of ice-cold Lysis Buffer (100 mM Tris-HCl at pH 8.0, 125 mM NaCl, 10 mM DTT, 0.5% Sarkosyl, 0.5 U/µL RNase Inhibitor) was added to the array, and cells were incubated at room temperature for 15 min to capture mRNA by oligo-dT beads. The microwell array was rinsed 3 times with 2 mL of 6X SSC buffer to remove unbound RNA and barcoded oligo-dT beads were collected into a 1.5 mL RNase-free tube. Beads were washed 3 times with 1 mL of 6X SSC buffer and once with 1 mL of 2X reverse transcription (RT) Buffer. 200 µL of RT mix (200 U SuperScript II reverse transcriptase (Thermo Fisher Scientific, Cat#18064014), 1X Superscript II RT buffer, 1 mM dNTP and 5 µM template switch oligo (TSO), 1 U/µL RNase Inhibitor, 6% PEG8000) was used to resuspend barcoded beads. The beads were incubated at 25 °C for 30 min, 42 °C for 90 min, 50 °C for 15 min and 70 °C for 10 min with rotation. Beads were collected and rinsed twice with 1 mL of 10 mM Tris-HCl at pH 8.0.

Beads were then suspended in 50 µL of exonuclease I mix (1X exonuclease I buffer and 1 U/ml exonuclease I (New England BioLabs, Cat#M0293L) and incubated at 37 °C for 60 min with rotation to remove extra barcode oligos. Beads were washed 3 times with 1 mL of 10 mM Tris-HCl at pH 8.0 and proceeded to cDNA amplification step immediately. 50 µL of 1X Q5 Hot Start High-Fidelity PCR mix (New England BioLabs, Cat#M0494S) with 10 µM TSO PCR primer was added to the beads, and 16 cycles of PCR reaction were performed. Amplified cDNA was purified using 0.8X MA beads (GenSeq). The purified cDNA library was tagged by a custom-assembled transposase⁷³ with 1030-ME oligo (GenSeq) and PCR amplified using custom PCR primers. 0.85X MA beads were used to purify the scRNA-seq library. The size of the library was analyzed on an Agilent 2100 bioanalyzer. The final libraries were pooled, and 150 bp paired-end sequencing was performed on an Illumina Novaseq6000 platform using the 300 bp high-output sequencing kit. The barcode sequence and all the primer sequences for scRNA-seq are listed in Supplementary Data 1a.

Single-cell RNA sequencing data analysis

Single-cell RNA-seq data processing: Raw sequencing reads from both H1-10 knockout and wildtype control were obtained, then cell barcodes and Unique Molecular Identifiers (UMIs) from read 1 were extracted from each sample, and seamlessly integrated with their corresponding read names in Read2 utilizing UMI-tools v1.1.2⁷⁴. By appending these identifiers to the read names, we achieved a perfect match between the cell barcodes and UMIs and their respective reads. Subsequently, these reads were aligned against the human reference genome utilizing STAR, and assigned to their respective genes utilizing featureCounts. After this step, we obtained a sorted and indexed set of alignments matched with relevant genes. Finally, we quantified the number of distinct UMIs mapped to each gene in each cell using the count method in UMI-tools with default parameters.

The preliminary data generated above were subject to downstream filtering. To ensure consistency and comparability, we merged two samples into a single Seurat object and utilized Seurat v4.3.0⁷⁵ in R (v4.2.3) for further data processing. Initially, cell quality was evaluated based on the following criteria: (1) the number of detected genes per cell; (2) the number of UMIs per cell; (3) the proportion of mitochondrial gene counts per cell. Specifically, we set the `FilterCells` parameters to `nGene` (500 to 2000) and `nCount` (1000 to 4000) to remove low-quality cells and likely cell doublets. Cells with a high mitochondrial gene proportion (>10%) were also excluded. Ultimately, a total of 4771 cells were retained for downstream analysis, consisting of 2040 knockout cells and 2731 wildtype cells. The median gene count for them was 1279 and 1316, respectively.

Cell clustering and Uniform Manifold Approximation and Projection (UMAP) analysis: Seurat's `NormalizeData` function was applied to normalize the filtered UMI counts. For normalization, the `scale` parameter was set to the average expression of all genes in each cell. Subsequently, we selected the top 1000 highly variable genes and performed data scaling and linear dimensionality reduction through principal component analysis (PCA). Following that, we conducted cell clustering and UMAP visualization. For cell clustering, we utilized the top six principal components and set the resolution parameter to 0.6. During this process, marker genes for each cell cluster were extracted for Gene Ontology (GO) enrichment analysis using the `enrichGO` function from the `clusterProfiler` R package v4.6.2⁷⁶, aiding in the identification of cell clusters.

Calculation of mesendodermal and ectodermal differentiation potential score: To model the mesendodermal differentiation (MED) and ectodermal differentiation (ECD) potential of each cell, the first step is to define MED and ECD-relevant marker genes. We incorporated the transcriptomic profiling data of a published study⁵³, in which they conducted in vitro embryoid bodies (EBs) differentiation of H9 cells

cultured in mTeSR1 medium. Their single-cell RNA sequencing analysis identified two distinct lineage differentiation routes. Given the similarity between the cell culture conditions described in this study and our own, we adopted this dataset to identify marker genes that are highly relevant to the MED and the ECD route. Firstly, we reproduced the single-cell clustering analysis of this study. Then, we separately conducted differential gene expression analysis for EB mesendoderm and EB ectoderm against all other single cell clusters (primed and naive hESCs) using the FindMarkers function in Seurat. Genes with average log₂-fold change >1 and Bonferroni-adjusted *P*-value <0.05 were selected as the relevant marker genes of each route. Recognizing that there was a small number of overlapping genes between the two marker gene sets, we excluded them from the final lists of mesendoderm and ectoderm relevant marker genes. Then, the MED and ECD potential scores of each single cell were respectively evaluated using the AddModuleScore function in Seurat with the corresponding marker genes.

Gaussian mixture model decomposition and single cell classification: To identify single cells with high MED and ECD differentiation potential, we utilized a Gaussian mixture model (GMM)-based approach to classify all single cells based on the distribution of their MED and ECD potential scores. Technically, we used the Mclust function from mclust v6.0.0⁷⁷ in R to perform the GMM decomposition and subsequently classify the cells into groups. The optimal number of groups was chosen as the one with the highest Bayesian Information Criterion (BIC) (maxBIC = 35,322.43 for MED potential scores and maxBIC = 25,156.19 for ECD potential scores), and the cell group with the highest potential scores was considered as the high MED- and ECD-potential cells, respectively. The cell-cell correspondences between the cell clusters identified by Seurat and the above cell groups were plotted using the chordDiagram function in the circlize R package (version 0.4.15)⁷⁸.

Quantification and statistical analysis

The statistical data of FACS and qPCR were analyzed and presented as mean ± S.E.M using GraphPad Prism 8 Software. Three or more biological replicates per group were included, and biological replicates were collected from different passages of hESCs. Statistical methods for comparing the means between groups were used as indicated, ***: *P* < 0.001, **: *P* < 0.01, *: *P* < 0.05.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the sequencing data and the ChIP-seq and CUT&Tag peak calls have been deposited at the Gene Expression Omnibus with accession no. GSE242906. Processed data are provided in the Supplementary Data files. All computational analyses were done following the guidelines of the involved packages, with user-defined parameters provided in the Methods section. No custom code was generated; all analyses employed publicly available packages according to standard protocols, except where user-defined parameters are specified in the Methods. Source data are provided with this paper.

References

- Dundes, C. E. & Loh, K. M. Bridging naive and primed pluripotency. *Nat. Cell Biol.* **22**, 513–515 (2020).
- Bargehr, J. et al. Epicardial cells derived from human embryonic stem cells augment cardiomyocyte-driven heart regeneration. *Nat. Biotechnol.* **37**, 895 (2019).
- Orkin, S. H. & Hochedlinger, K. Chromatin connections to Pluripotency and cellular reprogramming. *Cell* **145**, 835–850 (2011).
- Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155–169 (2016).
- Huang, X. et al. OCT4 cooperates with distinct ATP-dependent chromatin remodelers in naive and primed pluripotent states in human. *Nat. Commun.* **12** <https://doi.org/10.1038/s41467-021-25107-3> (2021).
- Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
- Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).
- Li, M. & Belmonte, J. C. I. Ground rules of the pluripotency gene regulatory network. *Nat. Rev. Genet.* **18**, 180–191 (2017).
- Kunarso, G. et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–U111 (2010).
- Fong, H., Hohenstein, K. A. & Donovan, P. J. Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells* **26**, 1931–1938 (2008).
- Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011).
- Chan, Y. S. et al. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
- Theunissen, T. W. et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
- Liu, X. D. et al. Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* **14**, 1055 (2017).
- Ji, X. et al. 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell* **18**, 262–275 (2016).
- Kim, T. K. & Shiekhattar, R. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell* **162**, 948–959 (2015).
- Xu, J. et al. Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Dev. Cell* **23**, 796–811 (2012).
- Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49 (2011).
- Sun, H. et al. Quantitative integration of epigenomic variation and transcription factor binding using MAmotif toolkit identifies an important role of IRF2 as transcription activator at gene promoters. *Cell Discov.* **4**, 38 (2018).
- Frietze, S., Lan, X., Jin, V. X. & Farnham, P. J. Genomic targets of the KRAB and SCAN domain-containing zinc finger protein 263. *J. Biol. Chem.* **285**, 1393–1403 (2010).
- O'Connor, M. D. et al. Alkaline phosphatase-positive colony formation is a sensitive, specific, and quantitative indicator of undifferentiated human embryonic stem cells. *Stem Cells* **26**, 1109–1116 (2008).
- Nakaki, F. & Saitou, M. PRDM14: a unique regulator for pluripotency and epigenetic reprogramming. *Trends Biochem. Sci.* **39**, 289–298 (2014).
- Zhi, M. et al. Generation and characterization of stable pig pregastrulation epiblast stem cell lines. *Cell Res.* **32**, 383–400 (2022).
- Seymour, P. A. Sox9: a master regulator of the pancreatic program. *Rev. Diabet. Stud.* **11**, 51–83 (2014).
- Thai, C. W. et al. Comparative Chromatin Dynamics of Stem Cell Differentiation in Human and Rat. *bioRxiv*, 2021.2002.2011.430819 <https://doi.org/10.1101/2021.02.11.430819> (2021).
- Li, Y. et al. Transcriptome analysis reveals determinant stages controlling human embryonic stem cell commitment to neuronal cells. *J. Biol. Chem.* **292**, 19590–19604 (2017).

27. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
28. Qin, Q. et al. Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biol.* **21** <https://doi.org/10.1186/s13059-020-1934-6> (2020).
29. Dvir, S. et al. Uncovering the RNA-binding protein landscape in the pluripotency network of human embryonic stem cells. *Cell Rep.* **35**, 109198 (2021).
30. Merckenschlager, M. & Odom, D. T. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* **152**, 1285–1297 (2013).
31. Hossain, I. et al. ZIC2 and ZIC3 promote SWI/SNF recruitment to safeguard progression towards human primed pluripotency. *Nat. Commun.* **15**, 8539 (2024).
32. Mayer, D. et al. Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *EMBO J.* **39**, e102591 (2020).
33. Yang, S. H. et al. ZIC3 controls the transition from naive to primed pluripotency. *Cell Rep.* **27**, 3215–3227. e3216 (2019).
34. Gonzales, K. A. U. et al. Deterministic restriction on pluripotent state dissolution by cell-cycle pathways. *Cell* **162**, 564–579 (2015).
35. Li, J. et al. MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. *Differentiation* **75**, 299–307 (2007).
36. Davidson, K. C. et al. Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. *Proc. Natl. Acad. Sci. USA* **109**, 4485–4490 (2012).
37. Zhang, X. Q. et al. Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* **7**, 90–100 (2010).
38. Morrison, G. M. et al. Anterior definitive endoderm from ESCs reveals a role for FGF signaling. *Cell Stem Cell* **3**, 402–415 (2008).
39. D'Amour, K. A. et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* **23**, 1534–1541 (2005).
40. Sumi, T., Tsuneyoshi, N., Nakatsuji, N. & Suemori, H. Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/beta-catenin, Activin/Nodal and BMP signaling. *Development* **135**, 2969–2979 (2008).
41. Kelly, O. G., Pinson, K. I. & Skarnes, W. C. The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development* **131**, 2803–2815 (2004).
42. Kazanskaya, O. et al. R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis. *Dev. Cell* **7**, 525–534 (2004).
43. Graf, T. & Stadtfeld, M. Heterogeneity of embryonic and adult stem cells. *Cell Stem Cell* **3**, 480–483 (2008).
44. Hoogland, S. H. A. & Marks, H. Developments in pluripotency: a new formative state. *Cell Res.* **31**, 493–494 (2021).
45. Tang, W. W. C. et al. A unique gene regulatory network resets the human germline epigenome for development. *Cell* **161**, 1453–1467 (2015).
46. Lan, X. C. et al. PCGF6 controls neuroectoderm specification of human pluripotent stem cells by activating SOX2 expression. *Nat. Commun.* **13** <https://doi.org/10.1038/s41467-022-32295-z> (2022).
47. Yi, S. Q. et al. E2A regulates neural ectoderm fate specification in human embryonic stem cells. *Development* **147**, dev190298 (2020).
48. Wang, A. et al. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**, 386–399 (2015).
49. Barolo, S. & Posakony, J. W. Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev.* **16**, 1167–1181 (2002).
50. Kinoshita, M. et al. Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency (vol 28, pg 1, 2021). *Cell Stem Cell* **28**, 2180–2180 (2021).
51. Yu, L. Q. et al. Derivation of Intermediate Pluripotent Stem Cells Amenable to Primordial Germ Cell Specification. *Cell Stem Cell* **28**, 550 (2021).
52. Li, S. Q., Prasanna, X., Salo, V. T., Vattulainen, I. & Ikonen, E. An efficient auxin-inducible degron system with low basal degradation in human cells (vol 16, pg 866, 2019). *Nat. Methods* **16**, 1193–1193 (2019).
53. Han, X. et al. Mapping human pluripotent stem cell differentiation pathways using high throughput single-cell RNA-sequencing. *Genome Biol.* **19**, 47 (2018).
54. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
55. Tu, S. et al. MAnorm2 for quantitatively comparing groups of ChIP-seq samples. *Genome Res.* **31**, 131–145 (2021).
56. Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W. W. & Lenhard, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91–94 (2004).
57. Shao, Z., Zhang, Y., Yuan, G. C., Orkin, S. H. & Waxman, D. J. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol.* **13**, R16 (2012).
58. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
59. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
60. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
61. Wu, Y. et al. Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659–662 (2013).
62. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
63. Kaya-Okur, H. S., Janssens, D. H., Henikoff, J. G., Ahmad, K. & Henikoff, S. Efficient low-cost chromatin profiling with CUT&Tag. *Nat. Protoc.* **15**, 3264–3283 (2020).
64. Kaya-Okur, H. S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).
65. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
66. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–U354 (2012).
67. Shannon, P. et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
68. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
69. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
70. Korotkevich, G. et al. Fast gene set enrichment analysis. 060012 <https://doi.org/10.1101/060012> (2021).
71. Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2007).
72. Han, X. et al. Mapping the mouse cell Atlas by Microwell-Seq. *Cell* **172**, 1091–1107 e1017 (2018).
73. Picelli, S. et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).

74. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
75. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
76. Wu, T. et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovations* **2**, 100141 (2021).
77. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R. J.* **8**, 289–317 (2016).
78. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).

Acknowledgements

We thank our colleagues Dr. Ying Jin, Dr. Jinsong Li, Dr. Bing Liao, Dr. Xiaomin Song, and Dr. Weida Li from Tongji University for their help and suggestions. We would also like to acknowledge Yujie Chen and Nianhao Cheng for their technical assistance in this study. This work was supported by the National Basic Research Program of China (2024YFA1307704, 2018YFA0800203, and 2018YFA0107602), the CAS Youth Interdisciplinary Team, and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB38040100) and the National Natural Science Foundation of China (32370705).

Author contributions

Z.S. conceived the study; Q.Y., L.R., Y.Z., and P.C., X.Y., J.W., Q.W., and Y.W. performed experimental works under the supervision of X.L. and Z.S.; J.H., H.S., and S.Z., Y.L., and Z.G. performed bioinformatics analyses under the supervision of Q.Y. and Z.S.; Q.Y. and Z.S. wrote the manuscript with input from all authors; H.L. and N.J. reviewed the manuscript; All authors approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-64290-5>.

Correspondence and requests for materials should be addressed to Zhen Shao.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025