

# KLF4 inhibits early neural differentiation of ESCs by coordinating specific 3D chromatin structure

Jinfang Bi<sup>†</sup>, Wenbin Wang<sup>†</sup>, Meng Zhang, Baoying Zhang, Man Liu, Guangsong Su, Fuquan Chen, Bohan Chen, Tengfei Shi, Yaoqiang Zheng, Xueyuan Zhao, Zhongfang Zhao, Jiandang Shi, Peng Li, Lei Zhang\* and Wange Lu<sup>✉\*</sup>

State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Responses, College of Life Sciences, Nankai University, 300071 Tianjin, China

Received April 20, 2022; Revised October 27, 2022; Editorial Decision October 28, 2022; Accepted November 08, 2022

## ABSTRACT

**Neural differentiation of embryonic stem cells (ESCs) requires precisely orchestrated gene regulation, a process governed in part by changes in 3D chromatin structure. How these changes regulate gene expression in this context remains unclear. In this study, we observed enrichment of the transcription factor KLF4 at some poised or closed enhancers at TSS-linked regions of genes associated with neural differentiation. Combination analysis of ChIP, HiChIP and RNA-seq data indicated that KLF4 loss in ESCs induced changes in 3D chromatin structure, including increased chromatin interaction loops between neural differentiation-associated genes and active enhancers, leading to upregulated expression of neural differentiation-associated genes and therefore early neural differentiation. This study suggests KLF4 inhibits early neural differentiation by regulation of 3D chromatin structure, which is a new mechanism of early neural differentiation.**

## INTRODUCTION

Krüppel-like factor 4 (KLF4) is an evolutionarily conserved zinc finger-containing transcription factor that regulates cell growth, proliferation, and differentiation in several contexts (1,2). One well-known KLF4 function is to regulate somatic cell reprogramming. A combination of KLF4, OCT4, SOX2 and MYC can reprogram differentiated cells into pluripotent stem cells with differentiation capacity (3–5). Ectopic expression of KLF4, OCT4 and SOX2 in mouse retinal ganglion cells induces reprogramming and restores youthful DNA methylation patterns and transcriptomes (6). Previous studies report that KLF4 functions as a transcriptional repressor and plays important roles in axonal

growth. KLF4 downregulation is essential for normal neural development (7), and neural progenitors with constitutive expression of KLF4 fail to develop into mature neurons (8). In addition, dysregulation of KLF4 induced hydrocephalus in mouse (7). These findings suggest important roles of KLF4 in neural development. Nonetheless, mechanisms underlying these activities remain not very clear.

Neurodevelopment requires proper orchestration of dynamic gene regulation networks (9). Perturbing this process, either by mutation or aberrant epigenetic activity, can promote diseases of the nervous system such as autism, schizophrenia or epilepsy (10–14). Epigenetic changes governing neurodevelopment include alterations in DNA methylation or histone modification (15,16) at regulatory elements like enhancers and silencers. Several studies of 3D chromatin structure report complex gene regulation networks via long-range chromatin interactions (17–20). Such activities can also regulate differentiation or proliferation (21–23). Moreover, a recent study showed that a gene regulation network induced by 3D chromatin structure played an important role in differentiation of mouse ESCs to neural precursor cells (24).

KLF4 reportedly functions as a chromatin organizer during somatic cell reprogramming (17,25), but it is not known whether and how KLF4 functions in long-range gene regulation networks during early neural differentiation of ESCs. To investigate that possibility, we performed HiChIP of RNA polymerase II (RNA pol II) in wild-type (WT) and KLF4 knock-out (KO) mouse ESCs (mESCs) and analyzed changes in long-range gene regulation networks during neural differentiation. This analysis combined with RNA-seq and ChIP-seq revealed KLF4 enrichment at transcription start site (TSS)-linked regions of multiple neural differentiation-associated genes, which also exhibited poised enhancer marks. KLF4 KO induced global changes in 3D chromatin structure in mESCs, including alteration of chromatin interactions in topological-associated

\*To whom correspondence should be addressed. Tel: +86 22 23503617; Fax: +86 22 23503617; Email: wangelv@gmail.com  
Correspondence may also be addressed to Lei Zhang. Tel: +86 22 23503617; Fax: +86 22 23503617; Email: joyleizhang@nankai.edu.cn  
<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

domains (TADs) and changes in enhancer/promoter loops. KLF4 loss induced changes in chromatin interactions related to neural differentiation-associated genes and upregulated those genes to promote early neural differentiation. This analysis defines a new mechanism by which KLF4 regulates early neural differentiation through 3D chromatin structure.

## MATERIALS AND METHODS

### Cell culture

Mouse ESCs E14 were grown in cell culture dishes coated with 0.1% gelatin (Sigma, 9000-70-8) in DMEM medium (Gibco) supplemented with 15% FBS (AusGeneX) 1000 U/ml of LIF (Millipore), non-essential amino acids (Gibco), 1 mM of glutamax (ThermoFisher), 100 U/ml of penicillin and streptomycin (ThermoFisher) and 0.1 mM of  $\beta$ -mercaptoethanol (Sigma). Medium was changed every 2 days and cells were passaged every 2 days using trypsin-EDTA (0.25%) (Gibco).

### Generation of *Klf4* knockout lines

For *Klf4* KO, sgRNAs were designed using an online tool (CRISPick, broadinstitute.org) (Supplementary table 2). Oligos were synthesized and constructed with the pSUPER-puro system (RNAi System), following the manufacturer's instruction. Mouse ESCs were transfected with plasmids containing sgRNAs and plated as single cells in six-well plates. Puromycin was added 24 h after transfection at a final concentration of 5  $\mu$ g/ml to select stable lines. Medium containing puromycin was changed every 2–3 days. *Klf4* KO was verified by agarose gel electrophoresis, and KO efficiency verified by RT-qPCR and western blotting.

### Neural differentiation of mouse ESCs

Neural differentiation of ESCs was performed following a reported protocol (26) with minor changes. In brief,  $5 \times 10^4$  E14 cells were treated with ESC growth medium for 12 h, and then medium was changed to Defined Default Medium (DDM) (DMEM/F12 + GlutaMAX supplemented with N2 supplement, 0.1 mM of non-essential amino acids, 1 mM of sodium pyruvate, 500 mg/ml of BSA, 0.1 mM of  $\beta$ -mercaptoethanol, 5000 U/ml of penicillin and streptomycin). After 48 h, 1  $\mu$ M cyclopamine was added to DDM to induce neural progenitor cells, and the medium was changed every 2 days for 12 days of differentiation induction.

### Quantitative real-time PCR (RT-qPCR)

Total RNA was extracted from cells using TRIzol Reagent (Life Technologies). cDNA synthesis was performed using a PrimerScript<sup>TM</sup> RT reagent Kit with gDNA Eraser (TaKaRa). Real-time PCR reactions were performed using Hieff<sup>TM</sup> qPCR SYBR Green Master Mix (YEASEN) and a BioRad CFX Connect Real-Time system. PCR cycling conditions were: 95 °C for 5 min, 40 cycles of 95 °C for 15 s,

60 °C for 15 s and 72 °C for 30 s. A melting curve of amplified DNA was subsequently acquired. Relative expression of target genes was determined using the comparative CT method and normalized to GAPDH. Primer information is listed in Supplementary Table S2.

### Western blot

Proteins were collected from mESCs and KLF4-KO cells using RIPA buffer supplemented with a protease inhibitor cocktail. Sample loading was based on the results of BCA assay (Bicinchoninic acid). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to polyvinylidene difluoride (PVDF) membranes (Millipore), which were then blocked and incubated with KLF4 antibody (R&D Systems, AF3158, lot: WRR061703) overnight at 4 °C, followed by secondary antibodies (Abcam) for 1 h at room temperature. Bands were detected by Image Quant LAS 4000 with an Enhanced Chemiluminescence Kit (Thermo Pierce).

### RNA-Seq

Total RNA was extracted from cells using TRIzol Reagent (Life Technologies) and DNA was removed with DNase I (Thermo Scientific). Total RNA (5  $\mu$ g) was used for RNA-seq library preparation. Each library was prepared and sequenced to obtain ~6 Gb data by Novogene, China. Hisat2 (27) was used to map reads to the mouse mm10 genome assembly and genes were annotated according to the Ensembl database. The number of reads mapped to genes was determined with htseq-count (28). Genes with mean count number <2.5 were filtered out. DESeq (29) was used to analyze differentially-expressed genes. Genes with expression fold-change  $\geq 2$  or  $\leq -2$  and  $P < 0.05$  were considered to show significant expression change (Supplementary Table 3). DAVID (<https://david.ncifcrf.gov/>) was used for gene ontology (GO) analysis.

### ChIP-seq data processing

ChIP-seq was performed as described with minor changes. In brief,  $1 \times 10^7$  cells were crosslinked with 1% formaldehyde (methanol free) in DMEM at a volume of 1 ml of formaldehyde for every one million cells. Cells were crosslinked 10 min at room temperature with rotation. Glycine was added into a final concentration of 125 mM to quench the formaldehyde. Samples were incubated 15 min on ice and then spun down at 2500 g for 5 min. The supernatant was discarded, and pelleted nuclei were washed once with 500  $\mu$ l pre-chilled cell Lysis Buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 0.2% NP-40, 1  $\times$  protease inhibitors) and then rotated at 4 °C for 30 min. Samples were then spun down at 2500 g for 5 min and the supernatant discarded, while pelleted nuclei were washed twice with 500  $\mu$ l of ice-cold ChIP Lysis Buffer. The volume of pelleted material was brought up to 1 ml in Nuclear Lysis Buffer (50mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS, 1  $\times$  protease inhibitors) and incubated 5 min at room temperature. Samples were then divided equally into three parts,

sheared for 30 cycles (30 s 'ON' and 30 s of 'OFF'), and then spun down at 16 100 g for 15 min. Supernatant were collected into new low-adsorption 1.5 ml EP tubes. Then, 60  $\mu$ l of Protein A beads for every  $10^7$  cells were washed in ChIP Dilution Buffer and resuspended in 50  $\mu$ l Dilution Buffer per tube (100  $\mu$ l per ChIP). Protein A beads were added to samples, which were rotated 1 h at 4°C for precleaning. Sample supernatants were then separated using a magnet and transferred into new tubes, and then 7.5  $\mu$ g antibodies per  $10^7$  cells (H3K27ac, abcam, cat: ab4729, lot: GR3442884-1; H3K4me1, abcam, cat: ab8895, lot: GR3442591-1 and H3K27me3, abcam, cat: ab6002, lot: GR3393630-3.), the samples were incubated at 4°C overnight with rotation. The next day, 60  $\mu$ l Protein A beads were washed and added to samples and then rotated 2 h at 4°C. Beads were washed three times each with Low Salt Wash Buffer, High Salt Wash Buffer, and LiCl Wash Buffer. Chip-DNA was de-crosslinking with Proteinase K and purification using AmpureXP beads. DNA libraries were generated using a FS DNA Library Prep Kit (QIAGEN). The sequencing platform used was HiSeq XTen (Illumina). Mapping of ChIP-seq data to mm10 was performed using bowtie2. Peaks were called using MACS2 (30,31) with default parameters and annotated with annotatePeaks in Homer (32): callpeak -t treated.bam -c input.bam -n sample -g mm -B -q 0.05 (Supplementary table 4). Colocalization of histone modifications was detected with ChromHMM (33).

## HiChIP

HiChIP was performed as described (20) with modifications. Briefly,  $1.5 \times 10^7$  cells were crosslinked with 1% formaldehyde and lysed and then chromatin was digested using 400 U *Mbo*I (NEB). Then 37.5  $\mu$ l biotin14-d-ATP was added and restriction-cut ends were ligated using 4000 U T4 ligase (NEB). Pelleted nuclei were dissolved in nuclear lysis buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% SDS, and protease inhibitors), sonicated and diluted in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 7.5, and 167 mM NaCl). Immunoprecipitation was performed overnight at 4°C by incubating RNA pol II antibodies (Biolegend, cat: 664906, lot: B300182) precoated on protein A-coated magnetic beads (Thermo Fisher Scientific). Immunocomplexes were washed three times each with low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 7.5 and 150 mM NaCl), high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 7.5 and 500 mM NaCl), and LiCl buffer (10 mM Tris-HCl, pH 7.5, 250 mM LiCl, 1% NP-40, 1% Na-Doc and 1 mM EDTA), and beads were resuspended in DNA elution buffer (50 mM NaHCO<sub>3</sub> and 1% SDS). After elution, ChIP samples were incubated with 10 mg/ml proteinase K 4 h at 55°C, and then DNA was purified using AMPure XP Beads (Beckman). Streptavidin C1 beads were used to capture biotinylated DNA. QIaseq FX DNA Library Kits were used to generate the sequencing library. HiChIP libraries were size-selected to 300–700 bp using AMPure XP beads (Beckman) and subjected to  $2 \times 150$ -bp paired-end sequencing

on HiSeq XTen (Illumina). Two biological replicates were analyzed for each experimental condition.

## HiChIP data processing

HiChIP paired-end reads were aligned to the mm10 genome using the HiC-Pro pipeline (34). Default settings were used to remove duplicate reads, assign reads to *Mbo*I restriction fragments, filter for valid interactions, and generate binned interaction matrices. Valid read pairs (generated from HiC-Pro) of individual samples were merged for each cell line. Interaction loop calling was performed using hichipper (35). Loops significant at a FDR of 0.1 and supported by at least 2 paired-end tags (PETS) were used for further analysis (Figure S1, Supplementary Table 5).

## Interaction matrices and virtual 4C visualization

HiChIP interaction maps were generated with HiC-Pro and visualized using HiCPlotter (36) at 500, 25 and 5 kb resolution, as indicated in each analysis. Virtual 4C plots were generated from dumped matrices generated with Juicebox (37). The chromosome of interest was extracted from the .hic file and normalized with the size factor of valid reads. The interaction profile of a specific 5-kb fragment containing the anchor was then plotted with R package ggplot2 (38). High-confidence loop ( $P < 0.05$ ) calls were loaded into the WashU Epigenome Browser (39,40), along with RNA pol II ChIP data. Browser shots from WashU track sessions were then included in virtual 4C and interaction maps.

## Immunofluorescence and quantification

The DDM medium was removed and cells were rinsed three times with PBS and then fixed for 25 min in 4% paraformaldehyde (PFA) in PBS. PFA was then removed and cells were rinsed with PB followed by permeabilization with 0.5% Triton X-100 for 10 min. After blocking, cells were incubated with primary antibodies (PAX6, Abcam) overnight at 4°C. Then cells were rinsed three times with PBS followed by incubation with secondary antibodies for 1 h at room temperature, and then rinse again with PBS. Nuclei were stained with DAPI. For quantification of the proportion of cells expressing a specific marker, at least 300 cells from three independent experiments (at least 100 cells from each experiment) were counted.

## Statistical analysis

Data represent means  $\pm$  S.E.M.; statistical analysis was performed using Student's *t*-test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

## RESULTS

### KLF4 binds to neural differentiation-associated gene loci that carry poised enhancer marks.

To determine whether KLF4 regulates interactions between gene regulatory elements and target genes, we first analyzed RNA-seq data in mESCs and mouse neural progenitor cells (NPCs)(41). Genes that were significantly up or

down-regulated in NPCs in relative to ESCs may be associated with early neural differentiation. In combination with KLF4-ChIP seq data, we observed KLF4 binding enrichment at transcription start sites (TSSs) of more than 50% of genes that were significantly changed in NPCs compared with ESCs (fold-change  $\geq 2$  or  $\leq -2$ ,  $P < 0.05$ ) (Figure 1A), as well as at about 80% of gene bodies (Figure 1A), suggesting a possible KLF4 regulatory function in these genes' transcriptions. Analysis of ChIP-seq data in mESCs showed that 19 426 of 89 450 KLF4-binding sites were TSS-linked (TSS  $\pm 2$  kb) and 70 024 were at non-TSS-linked regions (Figure 1B). Combination analysis in mESCs of H3K27ac, H3K4me1 and H3K27me3 ChIP-seq data revealed that >60% KLF4 binding sites were enriched with histone modifications (Figure 1C, D). Such histone modifications reportedly mark DNA regulatory elements: H3K4me1 is enriched in enhancer regions, H3K27ac and H3K4me1 mark active enhancers, and H3K27me3 and H3K4me1 mark poised or closed enhancers (42). We observed that 21.4% KLF4 binding sites were enriched in both H3K27me3 and H3K4me1 at TSS-linked regions (Figure 1D), and some of these positions were at TSS-linked regions of genes associated with neural differentiation, like *Pax6*, *Sox4*, *Foxp4* and *Satb2* (Figure 1E) (43–49).

#### **KLF4 KO induces transcriptional changes in neural differentiation-associated genes.**

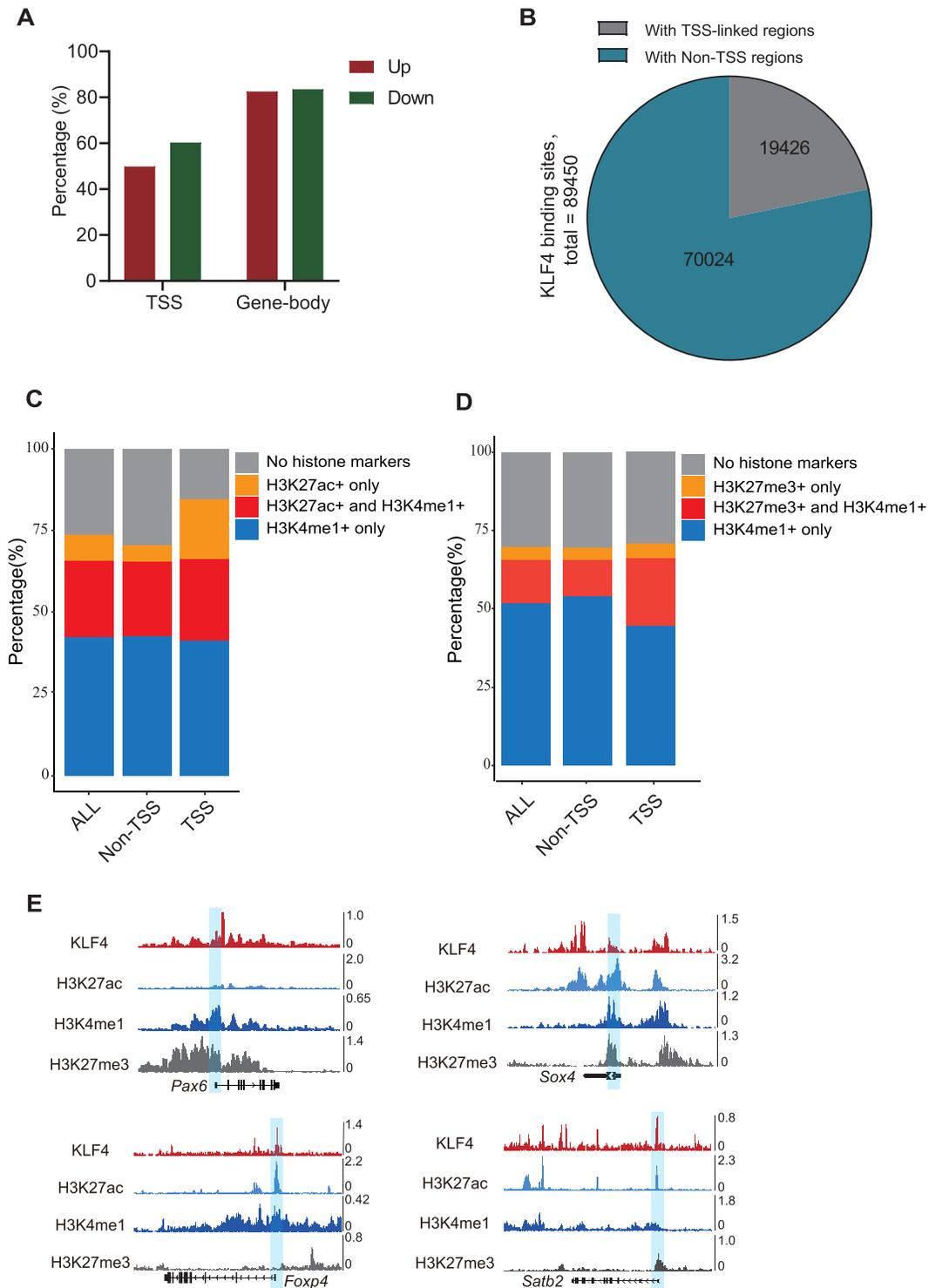
We then used CRISPR/Cas9 to knock-out (KO) *Klf4* in mESCs and monitored both early neural differentiation and potential changes in gene expression. Two guide RNA was used to target the *Klf4* gene (Figure 2A). The deletion was confirmed by sequencing (Figure 2B). Western blot and RT-qPCR analysis further confirmed loss of KLF4 expression in KO1 and KO2 (Figure 2C, D). RNA-seq data indicated global transcriptome changes in KLF4-KO relative to wild-type (WT) mESCs (Figure 2E, F). Moreover, several neural differentiation-associated genes were upregulated in KLF4-KO cells. Among them were *Sox4*, *Sox11*, *Gap43* and *Pax6* (Figure 2E). Gene ontology (GO) analysis of significantly upregulated genes (fold-change  $\geq 2$ ,  $P < 0.05$ ) by DAVID (<https://david.ncifcrf.gov/>) revealed that genes of nervous system development and cell differentiation were significantly enriched in KLF4-KO cells (Figure 2G). GO analysis of significantly downregulated genes (fold-change  $\leq -2$ ,  $P < 0.05$ ) showed significant enrichment in metabolic processes in KLF4-KO cells (Figure 2H). These analyses suggested that KLF4 may be required to downregulate neural differentiation genes in ESCs.

#### **KLF4 KO alters 3D chromatin structure and upregulates some neural differentiation-associated genes.**

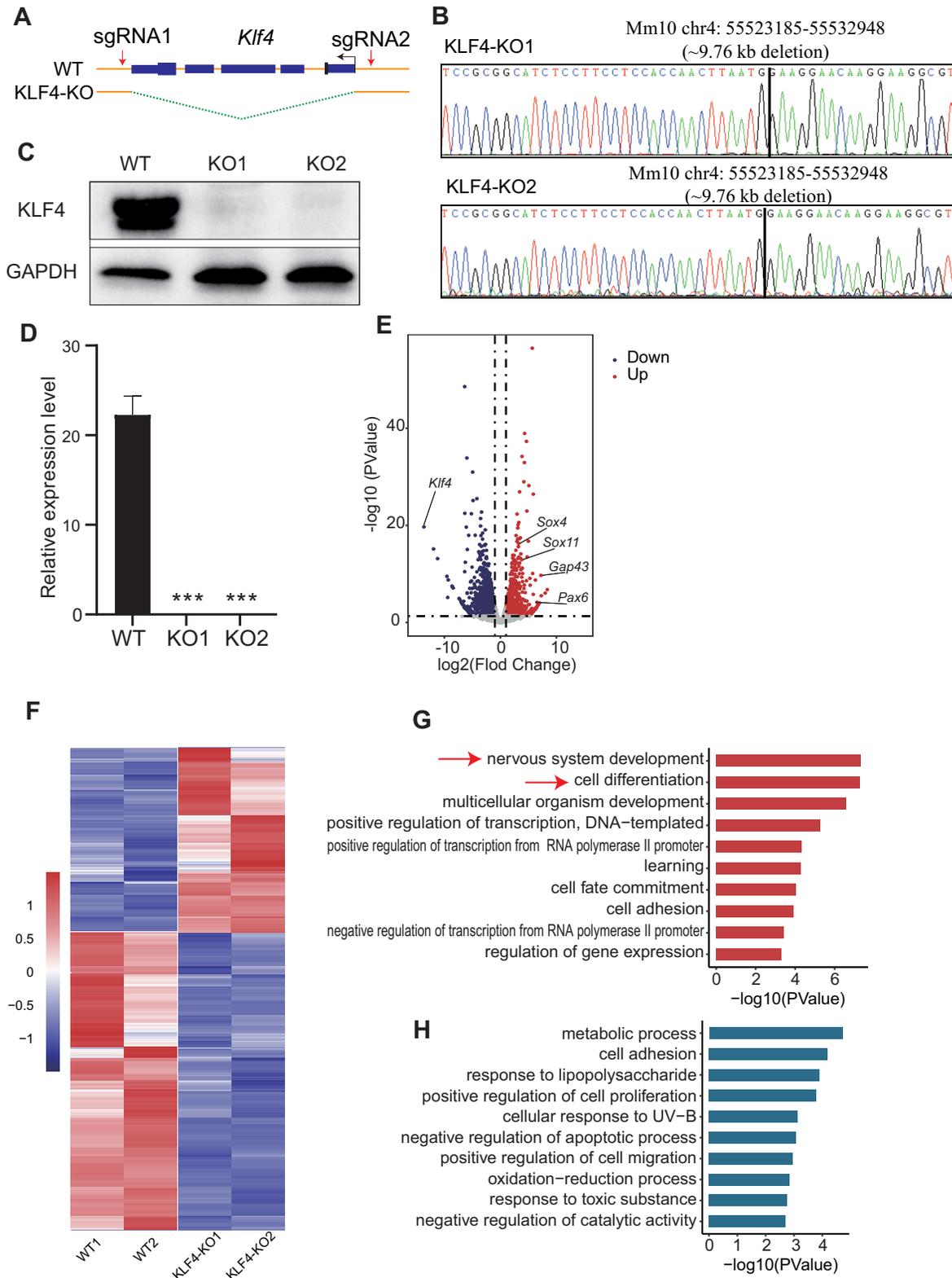
Given that KLF4 reportedly organizes 3D chromatin structure and RNA polymerase II binding is associated with gene transcription, thus we performed RNA pol II HiChIP to monitor potential changes in 3D chromatin structure in KLF4-KO cells and investigated how 3D chromatin structure changes induced alterations of genes expression in KLF4-KO cells (Supplementary Figure S1). KLF4 KO increased chromatin interactions in some chromatin regions,

such as at chromosomes 6 and 14 (Figure 3A), which contain neural differentiation-associated genes like *Hoxa10* and *Otx2* (50–53). In addition, RNA-seq data showed significant upregulation of *Hoxa10* and *Otx2* in KLF4-KO relative to WT cells (*Hoxa10* fold-change = 3.68,  $P = 0.024$ ; *Otx2* fold-change = 2.66,  $P = 0.0003$ ). Analysis of ChIP-seq data of H3K27ac and KLF4 in combination with RNA pol II HiChIP indicated that KLF4 loss induced changes in enhancer and promoter regulation networks at loci of some neural differentiation-associated genes (Figure 3B–E). For example, the enhancer regulation network around *Cbx5* was altered in KLF4-KO cells (Figure 3B). Significant interaction loops were only detected between the *Cbx5* promoter (P1) and enhancer-1(e1) in WT mESCs; However, in KLF4-KO cells, significant interaction loops were also detected between *Cbx5* P1 and enhancers 2, 3 and 4 besides e1 (e1, 2, 3 and 4) (Figure 3B). In addition, enhancer switching was detected in the study. For example, the *Foxp4* promoter (P1) interacted with enhancers-2, 3, and 5 (e2, 3 and 5) in WT ESCs. However, in KLF4-KO cells, the *Foxp4* promoter interacted with a proximal enhancer-4 (e4) (Figure 3C). HiChIP data also showed that promoters of some genes like *Hoxa1* formed a greater number of interaction loops with distal enhancers (e1, e2 and e3) in KLF4-KO compared with WT cells (Figure 3D). We also observed increases in promoter-promoter interactions at some loci, like *Ilf3*, in KLF4-KO compared with WT cells (Figure 3E). *Cbx5*, *Foxp4*, *Hoxa1* and *Ilf3* reportedly function in neural differentiation, suggesting KLF4 alters the 3D chromatin structure in different ways. Genome-wide analysis of HiChIP data confirmed that KLF4 loss induced changes in chromatin interactions at some gene loci, including enhancer-enhancer interactions (EEI), promoter-enhancer interactions (PEI) and promoter-promoter interactions (PPI) (Figure 4A, B). Among all enhancer and promoter interaction networks, 12.59% were EEI, 39.33% were PEI and 52.52% were PPI in WT mESCs, while in KLF4-KO cells, EEI and PEI increased to 35.41% and 43.67%, respectively, and PPI decreased to 20.92% (Figure 4B). Moreover, the median length of chromatin interaction loops was longer in KLF4-KO compared to WT cells, and more long-range chromatin interactions were detected in KLF4-KO cells (Figure 4C), suggesting global alterations in 3D chromatin structure in these cells.

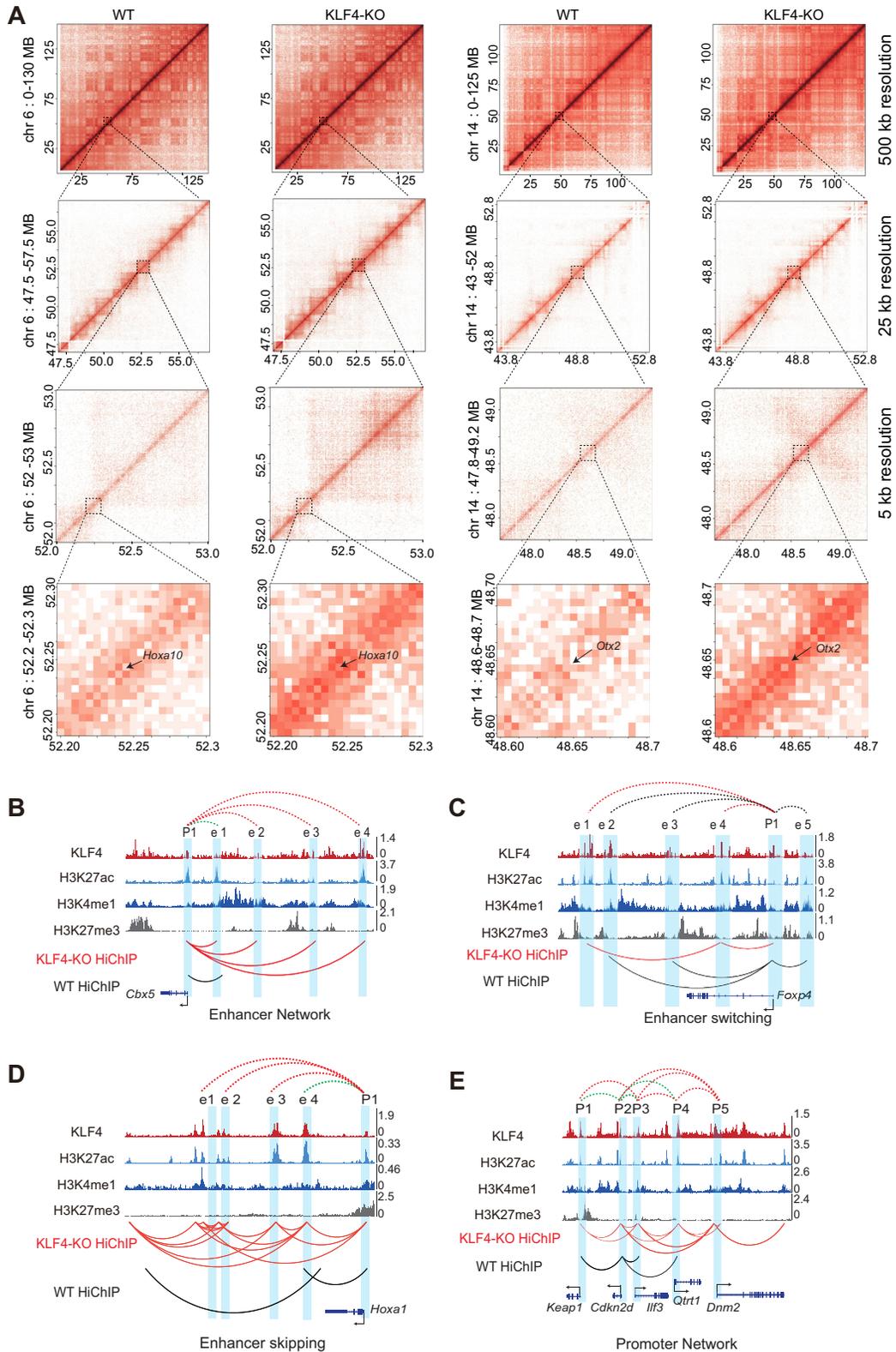
To investigate genome-wide changes in chromatin interactions of neural differentiation-associated genes, we first analyzed change of chromatin interaction loops at 80 significantly upregulated genes loci (gene body  $\pm 10$  kb) in KLF4-KO cells compared to WT cells. These genes were enriched in neural differentiation-associated processes such as nervous system development, forebrain development and neuronal differentiation (Supplementary Table S1). Compared with WT ESCs, 65% of these gene loci exhibited more chromatin interaction loops in KLF4-KO compared to WT mESCs, while 12.5% showed a loss of chromatin interaction loops in KLF4-KO cells (Figure 4D), suggesting that KLF4 repressed loop formation at these loci. In particular, for genes enriched in the most prominent biological process based on GO analysis, namely, nervous system development, 96.3% of those genes' TSS-linked (TSS  $\pm 2$  kb) regions were enriched with KLF4 (Figure 4E). We also



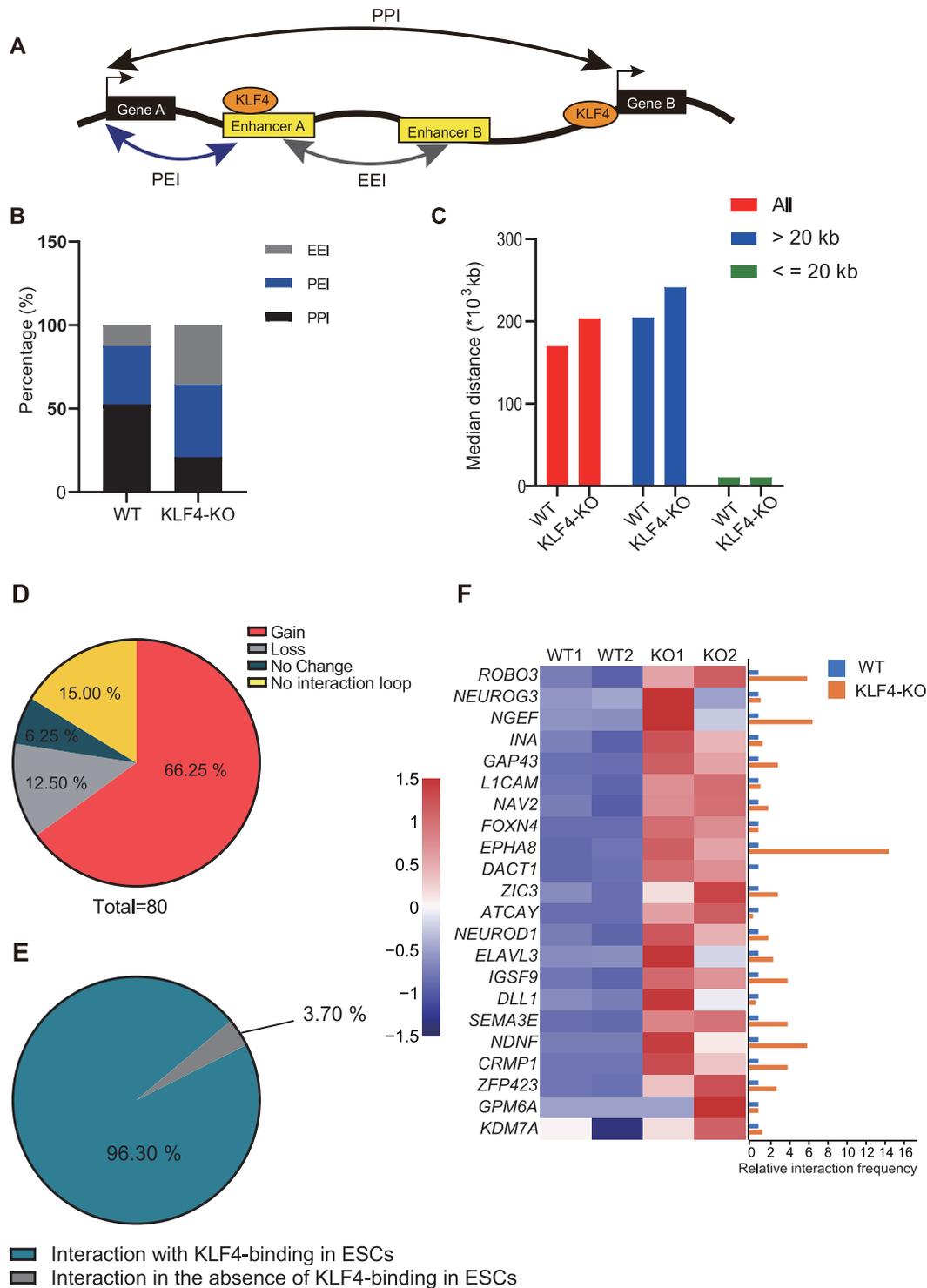
**Figure 1.** KLF4-bound regions carrying epigenetic enhancer marks are enriched in loci of some genes associated with neural differentiation. **(A)** Percentage of KLF4-bound sites in annotated TSS-linked regions (TSS  $\pm$  2 kb) and in bodies of genes significantly up- or down-regulated in neural progenitor cells (NPCs). **(B)** Percentage of KLF4-bound sites in TSS-linked regions (TSS  $\pm$  2 kb) and non-TSS regions in mESCs. **(C)** Percentage of marks of enhancers and active enhancers in regions within KLF4-bound TSS-linked regions (TSS  $\pm$  2 kb) and non-TSS regions in mESCs. **(D)** Percentage of marks of enhancers and poised enhancers in regions within KLF4-bound TSS-linked regions (TSS  $\pm$  2 kb) and non-TSS regions in mESCs. **(E)** KLF4-bound regions that carry epigenetic enhancer marks at *Pax6*, *Sox4*, *Foxp4* and *Satb2* gene loci.



**Figure 2.** KLF4 knock-out (KO) upregulates neural differentiation-associated genes in mESCs. (A) Schematic showing KLF4 KO (KLF4-KO) by CRISPR/Cas9. (B). Sequencing of regions targeted in KLF4-KO cells (KLF4-KO1 and KLF4-KO2). (C) Western blot verification of KLF4 loss in KLF4-KO1 and KLF4-KO2 cells. (D) RT-qPCR verification of KLF4 loss in KLF4-KO1 and KLF4-KO2 cells. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\*\* $P < 0.001$ , compared with Ctrl. (E) Volcano map depicting gene expression changes in KLF4-KO compared with wild-type mESC cells. (F) Comparison of gene expression in KLF4-KO1 and KLF4-KO2 cells with wild-type mESCs (WT1 and WT2). Heat map shows clustering of differentially-expressed genes in KO and WT mESCs. (G) Annotation enrichment analysis of genes significantly up-regulated in KLF4-KO cells compared to wild-type mESCs (fold-change  $\geq 2$  and  $P < 0.05$ ). (H) Annotation enrichment analysis of genes significantly down-regulated in KLF4-KO cells compared to wild-type mESCs (fold-change  $\leq 2$  and  $P < 0.05$ ).



**Figure 3.** Changes in 3D chromatin structure and in enhancer connectomes in KLF4-KO relative to WT cells. (A) HiChIP heatmaps at resolution of 500, 25 and 5 kb around *Hoxa10* (left) and *Otx2* (right) loci. Dotted squares indicate regions with more frequent chromatin interactions in KLF4-KO compared to WT cells. (B–E) HiChIP loop view shows interactions between enhancers and promoters around *Cbx5* (B), *Foxp4* (C), *Hoxa1* (D) and *Ilf3* (E) loci. Red and black curves indicate interaction loops between enhancers and promoters in KLF4-KO and WT cells, respectively. Green dotted curves indicate interaction loops in both KLF4-KO and WT cells. Red or black dotted curves indicate interactions loops specific to KLF4-KO cells or WT cells, respectively. P1–5 indicates promoters; e1–5 indicates enhancers.



**Figure 4.** Changes in promoter and enhancer connectomes in KLF4-KO cells are correlated with transcription of neural differentiation-associated genes. (A) Schematic showing enhancer-enhancer interaction (EEI), promoter-enhancer interaction (PEI), promoter-promoter interaction (PPI). (B) Percentage of EEI, PEI and PPI in WT and KLF4-KO cells. (C) Median length of chromatin interactions in WT and KLF4-KO cells. (D) Percent gain or loss of interaction loops in TSS-linked (TSS ± 2 kb) regions of 80 neurodevelopment-associated genes in KLF4-KO cells. (E) Percentage of KLF4-binding TSS-linked (TSS ± 2 kb) regions of genes that are enriched in 'nervous system development'. (F) Relative interaction frequency (counts of loops) between TSS-linked (TSS ± 2 kb) regions of genes enriched in nervous system development and other chromatin regions, and heatmap of RNA-seq data of corresponding genes in WT and KLF4-KO cells.

found that chromatin interaction loops in most of these loci increased in KLF4-KO compared to WT cells (Figure 4F). Correspondingly, most of these genes were upregulated in KLF4-KO cells (Figure 4F). These data suggest that KLF4 loss increases chromatin interactions at neural differentiation-associated gene loci leading to upregulation of transcription of these genes.

The transcription factor PAX6 is essential for neural differentiation. Pax6 is expressed in early neuroectoderm cells differentiated from ESCs where it activates expression of downstream neural genes. Relative to WT cells, KLF4-KO cells showed increased chromatin interaction loops between Pax6 and enhancers (Figure 5A). Virtual 4C analysis based on HiChIP data showed more robust interaction signals between the Pax6 locus and other chromatin regions, including some enhancers (Figure 5A). qPCR data indicated significant Pax6 upregulation in KLF4-KO cells (Figure 5B). A previous study reported that Pax6 overexpression triggers ESC differentiation. Similarly, Sox4 reportedly plays an important role in neural differentiation (44), and we observed new interaction loops formed between Sox4 and enhancers in KLF4-KO cells (Figure 5C). At the same time, RT-qPCR data indicated significant Sox4 upregulation in KLF4-KO compared with WT cells (Figure 5D).

### **KLF4 KO promotes early neural differentiation and changes 3D chromatin structure during early neural differentiation.**

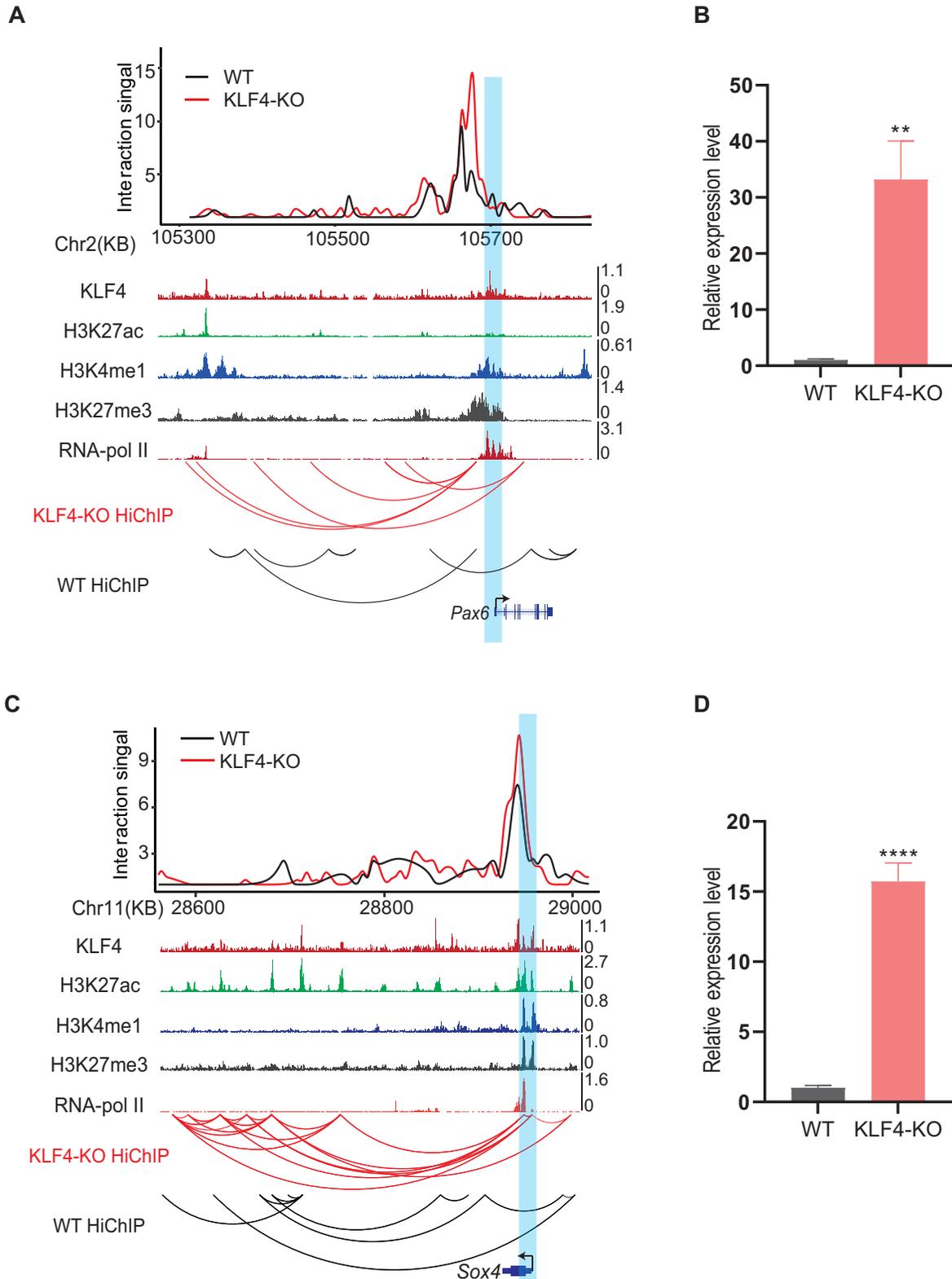
Transcriptome and HiChIP analysis suggest that KLF4 loss leads to upregulation of several neural differentiation-associated genes. To further investigate how KLF4 KO caused neural differentiation, we cultured mESCs and KLF4-KO cells in a serum-free and morphogen-free chemically defined default medium (DDM) allowing cell survival by insulin (54). In these conditions, ESCs followed an efficient process leading to early neural differentiation (Figure 6A). Neurogenesis started on day 6 after cyclopamine treatment in this system, and day 0–12 represent the early period of neural differentiation (26) (Figure 6A). No significant morphological changes were detected in KLF4-KO cells (Supplementary Figure S2), however, percentages of PAX6 positive cells were significantly increased in KLF4-KO cells compared with WT (Figure 6B and 6C). In addition, RT-qPCR data showed significant Pax6 upregulation during the early period, at days 2, and 6 compared with WT cells (Figure 6D). Furthermore, markers of different cortical layers and neuronal cells like OTX1 (forebrain and midbrain progenitors) (Figure 6E), CTIP2 (deep layers, layer V and VI neurons) (Figure 6F), REELIN (Cajal-Retzius cells) (Figure 6G), TBR1 (deep layers, layer V and VI neurons) (Figure 6H) and SATB2 (upper layers) (Figure 6I) were significantly upregulated in KLF4-KO cells after day 6 to a greater extent than in WT cells (Figure 6E–I). In addition, RNA-seq data indicated global transcriptome changes in KLF4-KO cells at day 12 of neural differentiation, in which most cells were neural progenitor cells (NPCs), compared to wild-type (WT) cells (Supplementary Figure S3A). And gene ontology (GO) analysis of significantly upregulated genes (fold-change  $\geq 2$ ,  $P < 0.05$ ) by DAVID (<https://david.ncifcrf.gov/>) revealed that nervous system development and axon guidance were

significantly enriched in KLF4-KO NPCs (Supplementary Figure S3B). GO analysis of significantly downregulated genes (fold-change  $\leq -2$ ,  $P < 0.05$ ) showed regulation of transcription from RNA polymerase II promoter in KLF4-KO NPCs (Supplementary Figure S3C). GSEA of brain development also showed significant differences (Supplementary Figure S3D). These analyses suggested that KLF4 may be required to downregulate neural differentiation genes in ESCs.

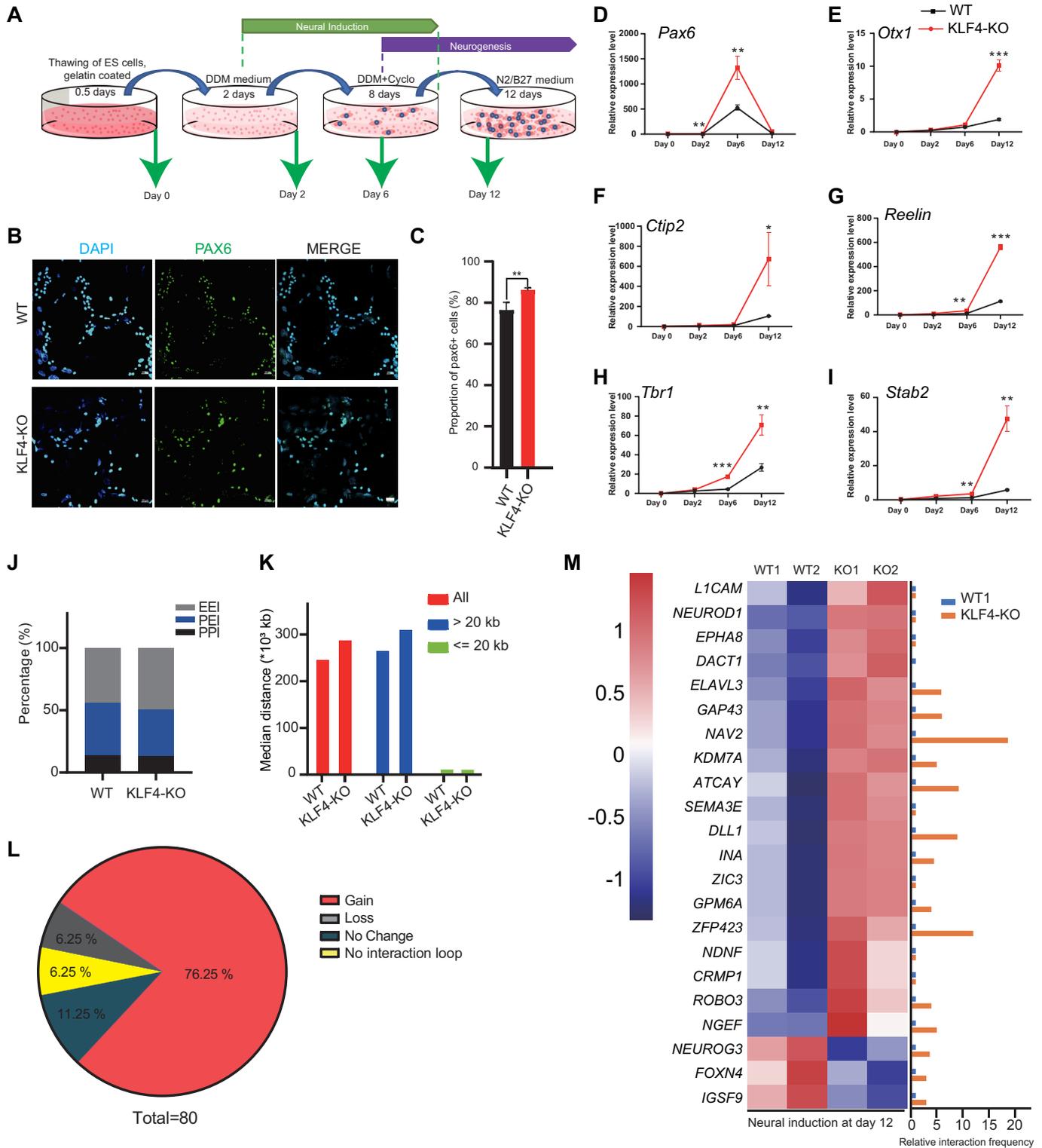
To investigate genome-wide changes in chromatin interactions of neural differentiation-associated genes, HiChIP was performed in WT and KLF4-KO cells at day12 of neural induction (Supplementary Figure S4). Among all enhancer and promoter interaction networks, 43.69% were EEI, 42.12% were PEI and 14.19% were PPI in WT NPCs, while in KLF4-KO NPCs, EEI increased to 49.58% and PEI and PPI decreased to 37.15% and 13.26%, respectively (Figure 6J). Similar with the undifferentiated WT and KLF4-KO ESCs, the median length of chromatin interaction loops was longer in KLF4-KO compared to WT NPCs, and more long-range chromatin interactions were detected in KLF4-KO NPCs (Figure 6K). Then we analyzed change of chromatin interaction loops at the 80 neural differentiation-associated genes (Supplementary Table S1) loci (gene body  $\pm 10$  kb) in KLF4-KO NPCs compared to WT NPCs. Chromatin interaction loops in most of these loci increased in KLF4-KO compared to WT NPCs (Figure 6L). Compared with WT NPCs, 76.25% of these gene loci exhibited more chromatin interaction loops in KLF4-KO NPCs, while 6.25% showed a loss of chromatin interaction loops in KLF4-KO NPCs (Figure 6L). For genes enriched in nervous system development (Supplementary Table S1), most of these genes were upregulated in KLF4-KO NPCs (Figure 6M). These data suggest that KLF4 loss increases chromatin interactions at neural differentiation-associated gene loci leading to upregulation of transcription of these genes upon neural differentiation. We conclude that KLF4 inhibits early neural differentiation by coordinating 3D chromatin structure and regulating the neural differentiation-associated genes.

### **KLF4 KO increases chromatin interactions between active enhancers and some genes related to neural disease.**

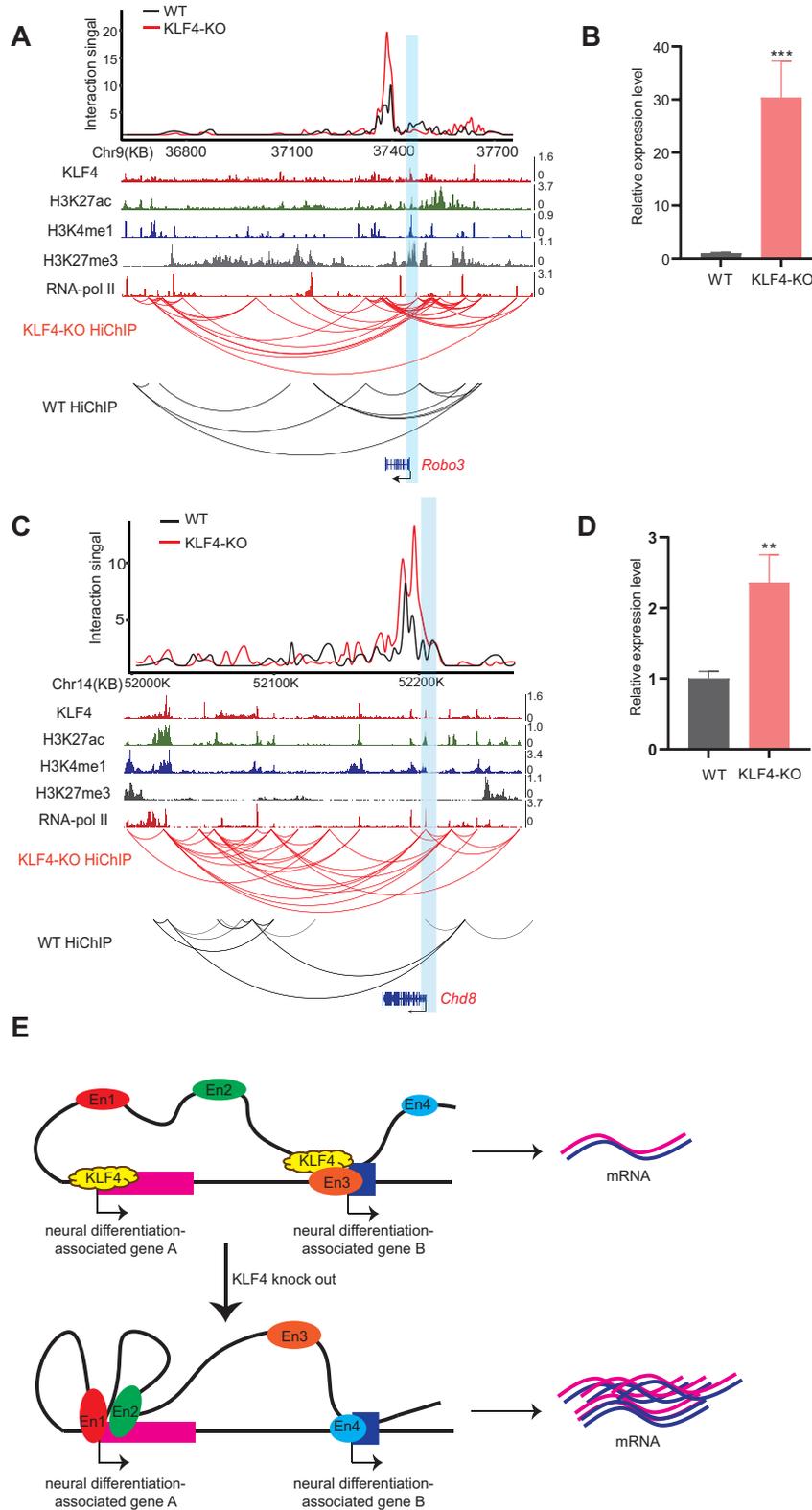
Neural development disorders induced by change of 3D chromatin structure may be linked to neural disease (55,56). We identified several genes associated with neurological diseases regulated by KLF4. For example, ROBO3 is a transmembrane receptor belonging to the Roundabout (Robo) family and is reportedly associated with the rare neurological disorder name horizontal gaze palsy with progressive scoliosis (HGPPS) syndrome(57). KLF4 loss increased chromatin interaction loops between Robo3 and active enhancers (Figure 7A), and based on virtual 4C analysis, more robust HiChIP signals were seen in KLF4-KO compared to WT cells around the Robo3 locus (Figure 7A). Correspondingly, Robo3 was significantly upregulated in KLF4-KO compared with WT ESCs (Figure 7B). Chd8 is also reportedly associated with autism spectrum disorder (ASD) (10,11,58,59). KLF4-KO cells showed more signif-



**Figure 5.** Changes in 3D chromatin structure in KLF4-KO cells are correlated with upregulation of *Pax6* and *Sox4*. **(A)** Virtual 4C representation of normalized RNA pol II HiChIP signals and interaction profiles around the *Pax6* locus. Black and red curves indicate interaction loops in WT and KLF4-KO cells, respectively. Blue bar indicates the TSS-linked region. **(B)** RT-qPCR analysis of *Pax6* transcript levels in WT and KLF4-KO cells. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\*\*\* $P < 0.0001$ , compared with WT. **(C)** Virtual 4C representation of normalized RNA pol II HiChIP signals and interaction profiles around the *Sox4* locus. Black and red curves indicate interaction loops in WT and KLF4-KO cells, respectively. Blue bar indicates the TSS-linked region. **(D)** RT-qPCR analysis of *Sox4* transcript levels in WT and KLF4-KO cells. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\* $P < 0.01$ , compared with WT.



**Figure 6.** KLF4 KO promotes early neural differentiation of mESCs and changes 3D chromatin structure. **(A)** Schematic showing the early neural differentiation protocol. Cyclo: cyclopamine. **(B)** Immunostaining with PAX6 (green) antibodies in WT and KLF4-KO cells at day 12 of the early neural differentiation protocol. DAPI (blue). Scale bar: 20  $\mu$ m. **(C)** Quantification of PAX6 positive cells in WT and KLF4-KO cells at day 12 of the early neural differentiation protocol. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . **(D–I)** RT-qPCR analysis of transcript levels of *Pax6* (D), *Otx1* (E), *Ctip2* (F), *Reelin* (G), *Tbr1* (H) and *Satb2* (I) in WT and KLF4-KO cells at days 0, 2, 6 and 12 of the early neural differentiation protocol. Data represent means  $\pm$  S.E.M. of three independent experiments. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , compared with WT. **(J)** Percentage of EEI, PEI and PPI in WT and KLF4-KO NPCs. **(K)** Median length of chromatin interactions in WT and KLF4-KO NPCs. **(L)** Percentage gain or loss of interaction loops in TSS-linked (TSS  $\pm$  2 kb) regions of 80 neural differentiation-associated genes in KLF4-KO NPCs. **(M)** Relative interaction frequency (counts of loops) between TSS-linked (TSS  $\pm$  2 kb) regions of genes enriched in nervous system development and other chromatin regions, and heatmap of RNA-seq data of corresponding genes in WT and KLF4-KO NPCs.



**Figure 7.** Changes in 3D chromatin structure in KLF4-KO cells are correlated with *Robo3* and *Chd8* upregulation. (A) Virtual 4C representation of normalized RNA pol II HiChIP signals and interaction profiles around the *Robo3* locus. Black and red curves indicate interaction loops in WT and KLF4-KO cells, respectively. Blue bar indicates the TSS-linked region. (B) RT-qPCR analysis of *Robo3* transcript levels in WT and KLF4-KO cells. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\*\* $P < 0.001$ , compared with WT. (C) Virtual 4C representation of normalized RNA pol II HiChIP signals and interaction profiles around the *Chd8* locus. Black and red curves indicate interaction loops in WT and KLF4-KO cells, respectively. Blue bar indicates the TSS-linked region. (D) RT-qPCR analysis of *Chd8* transcript levels in WT and KLF4-KO cells. Data represent means  $\pm$  S.E.M. of three independent experiments. \*\* $P < 0.01$ , compared with WT. (E) Schematic showing KLF4 depletion induced upregulation of some neural differentiation-associated genes by coordinating specific chromatin interactions. En1–4: enhancer 1–4.

icant chromatin interactions between *Chd8* and active enhancers enriched with H3K27ac and H3K4me1 than did WT cells (Figure 7C). RT-qPCR data showed significantly upregulated *Chd8* in KLF4-KO cells (Figure 7D). KLF4 loss induced transcriptional changes in genes related to nervous system disease, and regulation of these genes by KLF4 was associated with the changes of chromatin interactions.

## DISCUSSION

Previous studies report that KLF4 functions in somatic cell reprogramming and neurodevelopment (5,8). Development of chromatin conformation capture technologies now supports the idea that KLF4 functions as a chromatin organizer to regulate chromatin interactions and gene expression (17,25). KLF4 reportedly organizes long-range chromosomal interactions between enhancers and the *Oct4* locus and functions in reprogramming and pluripotency. KLF4 recruits cohesin to the *Oct4* distal enhancer and facilitates chromatin interactions with that enhancer (17). Moreover, depletion of KLF4 binding sites within pluripotent-stem-cell-specific enhancers disrupts enhancer-promoter interactions and downregulates expression of associated genes (25). Here, we found that in addition to binding active enhancers, KLF4 binding sites coincided with poised or closed enhancer markers, many found at neural differentiation-associated genes like *Pax6* and *Sox4*. Consistent with that downregulation of KLF4 is important for neurodevelopment (8,60), *Klf4* depletion induced formation of interaction loops in some chromatin regions and activated expression of neural differentiation-associated genes (Figure 7E). In addition, H3K27me3, H3K27ac and H3K4me1 ChIP-seq data show no significant alteration of these epigenetic markers in KLF4-KO cells compared with WT mESCs (Supplementary Figure S5). These findings suggest that besides facilitating chromatin interaction, KLF4 represses chromatin interactions in the context of some genes during early neural differentiation of ESCs. Previously, we reported that KLF4 interacts with the RNA-binding protein Stauf1 and RNA helicase Ddx5/17 to maintain NPC self-renewal (60), suggesting that KLF4 regulates neural differentiation by different mechanisms at different neural development stages. It remains unknown whether these disparate KLF4 mechanisms operate simultaneously or separately at different neural differentiation stages.

In addition to KLF4, other factors reportedly regulate of 3D chromatin structures. CCCTC-binding factor (CTCF) is critical to maintain chromatin structure in different cell types (61–63) and functions with the cohesin complex to establish topological chromatin domains (24,64). A recent study showed that CTCF loss blocks ESC differentiation to NPCs, and that CTCF binds to promoters to drive long-distance enhancer-dependent transcription of specific genes (24). Those researchers also report that CTCF-dependent genes are generally dependent on distal elements and that active enhancers acting at nearby gene loci may compensate for CTCF loss (24). We found that KLF4 deletion increased formation of several long-range interaction loops between distal active enhancers and genes promoters (Figure 4C, 5A, C, 7A, C), an activity comparable to CTCF.

However, unlike CTCF, KLF4 functions as a repressor of long-range interaction loops for some genes, including neural differentiation-associated genes. The transcription factor SOX2 also reportedly functions to organize chromatin interactions (65). SOX2 is essential for NSC maintenance, and SOX2 deletion in NSCs induced loss of long-range interactions and downregulation of genes functioning in neural development (65). Thus, during neural differentiation, several chromatin organizers regulate genes expression through modulating 3D chromatin structure and may play different roles and regulate different gene clusters. SOX2 loss also deregulates genes that function in cell proliferation (65). While we focused here on KLF4-regulated neural differentiation-associated genes, though other gene may also be regulated by KLF4 such as genes that function in metabolic process and cell adhesion (Figure 2G, H).

Defining mechanisms underlying neural differentiation could suggest treatment methodologies for diseases of the nervous system. Several prior studies have reported methods to modify 3D chromatin structures like using the dCas9 system and zinc finger proteins (66–68). Specific long-range chromatin interactions could be treatment targets of some nervous system disease. Two candidates may be *Robo3* and *Chd8*, both regulated by KLF4. KLF4 loss increased chromatin interaction loops between *Robo3* or *Chd8* and active enhancers and upregulated *Robo3* and *Chd8* expression. ROBO3 regulates axonal guidance to control development of commissural circuits (57). Mutation of *Robo3* is reportedly associated with HGPPS syndrome. Reduced commissural projections seen in the hindbrain and caused by *Robo3* mutation are associated with the absence of conjugate horizontal eye movements in humans (57,69). CHD8, a member of the chromodomain-helicase-DNA binding protein family, is reportedly associated with autism spectrum disorder (ASD), a heterogeneous disease comprising a range of neurodevelopmental disorders (58). CHD8 disruption alters expression of genes in neurodevelopmental pathways and associated with ASD (59). Regulation by KLF4 of *Robo3* and *Chd8* revealed here may underlie occurrence and progression of related neurological disorders and interaction loops between *Robo3* or *Chd8* and enhancers mediated by KLF4 could be potential targets of treatment of HGPPS syndrome and ASD. However, we observed that KLF4 loss induced more than one interaction loop between these genes and enhancers. Thus, further studies are needed to determine whether all of these interactions are necessary or whether there is one critical interaction loop that maintains expression of the specific genes before novel treatments can be devised. In addition, we totally investigated 41 reported neural disease genes that are associated with global developmental delay/intellectual disability (GDD/ID), ASD and attention deficit hyperactivity disorder (ADHD) (70–73) (Supplementary Figure S6). Among these disease genes, 28 genes' RNA levels were not significantly changed in KLF4 KO cells relative to WT ESCs, these genes may not be regulated by KLF4 (Fig. S6). Ten genes' RNA levels were significantly upregulated in KLF4 KO cells and increased chromatin interaction frequency at these gene loci were detected (Fig. S6, genes marked in red), suggesting negative regulation of these 10 genes by KLF4, but further investigations are needed to verify this.

## DATA AVAILABILITY

The following published datasets were used in our analyses: GSM2417144 (74) was used for mESC KLF4 ChIP-seq analyses, GSM4050822 (75) for mESC H3k4me1 ChIP-seq analyses, GSM1399505 (76) for mESC H3k27me3 ChIP-seq analyses, GSM1874094 (77) for mESC H3k27ac ChIP-seq analyses, GSM2533843 (41) and GSM2533844 (41) for mESC RNA-seq analyses, GSM2533845 (41) and GSM2533846 (41) for NPC RNA-seq analyses.

## ACCESSION NUMBERS

Processed data generated for this study are available through the NCBI Gene Expression Omnibus (GEO) under accession number GSE194068 and GSE213418.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank all members of our laboratory for many helpful discussions.

*Author contributions:* J.B., L.Z. and W.L. designed the research; J.B., W.W., M.Z., B.Z., M.L., G.S., F.C., B.C., T.S., Y.Z., X.Z., Z.Z., J.S. and P.L. performed the experiments; J.B., W.W., B.C., L.Z. and W.L. analyzed the data; and L.Z., J.B., W.W. and W.L. wrote the paper.

## FUNDING

National Key R&D Program of China [2017YFA0102600, 2020YFA0803700]; National Natural Science Foundation of China [32130018, 32170598, 32200588]; Fundamental Research Funds for the Central Universities (Nankai University) [63201087]. Funding for open access charge: National Key R&D Program of China [2017YFA0102600, 2020YFA0803700]; National Natural Science Foundation of China [32130018, 32170598]; Fundamental Research Funds for the Central Universities (Nankai University) [63201087].

*Conflict of interest statement.* None declared.

## REFERENCES

- Lin,H., Liu,H., Sun,Q., Yuan,G., Zhang,L. and Chen,Z. (2013) KLF4 promoted odontoblastic differentiation of mouse dental papilla cells via regulation of DMP1. *J. Cell Physiol.*, **228**, 2076–2085.
- Aksoy,I., Giudice,V., Delahaye,E., Wianny,F., Aubry,M., Mure,M., Chen,J., Jauch,R., Bogu,G.K., Nolden,T. *et al.* (2014) Klf4 and klf5 differentially inhibit mesoderm and endoderm differentiation in embryonic stem cells. *Nat. Commun.*, **5**, 3719.
- Ghaleb,A.M. and Yang,V.W. (2017) Krüppel-like factor 4 (KLF4): what we currently know. *Gene*, **611**, 27–37.
- Takahashi,K., Tanabe,K., Ohnuki,M., Narita,M., Ichisaka,T., Tomoda,K. and Yamanaka,S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, **131**, 861–872.
- Takahashi,K. and Yamanaka,S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663–676.
- Lu,Y., Brommer,B., Tian,X., Krishnan,A., Meer,M., Wang,C., Vera,D.L., Zeng,Q., Yu,D., Bonkowski,M.S. *et al.* (2020) Reprogramming to recover youthful epigenetic information and restore vision. *Nature*, **588**, 124–129.
- Qin,S., Liu,M., Niu,W. and Zhang,C.L. (2011) Dysregulation of Kruppel-like factor 4 during brain development leads to hydrocephalus in mice. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 21117–21121.
- Qin,S. and Zhang,C.L. (2012) Role of Kruppel-like factor 4 in neurogenesis and radial neuronal migration in the developing cerebral cortex. *Mol. Cell Biol.*, **32**, 4297–4305.
- Salinas,R.D., Connolly,D.R. and Song,H. (2020) Invited review: epigenetics in neurodevelopment. *Neuropathol. Appl. Neurobiol.*, **46**, 6–27.
- Bernier,R., Golzio,C., Xiong,B., Stessman,H.A., Coe,B.P., Penn,O., Witherspoon,K., Gerds,J., Baker,C., Vulto-van Silfhout,A.T. *et al.* (2014) Disruptive CHD8 mutations define a subtype of autism early in development. *Cell*, **158**, 263–276.
- Varghese,M., Keshav,N., Jacot-Descombes,S., Warda,T., Wicinski,B., Dickstein,D.L., Harony-Nicolas,H., De Rubeis,S., Drapeau,E., Buxbaum,J.D. *et al.* (2017) Autism spectrum disorder: neuropathology and animal models. *Acta Neuropathol.*, **134**, 537–566.
- Jin,X., Simmons,S.K., Guo,A., Shetty,A.S., Ko,M., Nguyen,L., Jokhi,V., Robinson,E., Oyler,P., Curry,N. *et al.* (2020) In vivo perturb-seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science (New York, N.Y.)*, **370**, eaaz6063.
- Schmidt,M.J. and Mirnics,K. (2015) Neurodevelopment, GABA system dysfunction, and schizophrenia. *Neuropsychopharmacology*, **40**, 190–206.
- Jiang,N.M., Cowan,M., Moonah,S.N. and Petri,W.A. Jr (2018) The impact of systemic inflammation on neurodevelopment. *Trends Mol. Med.*, **24**, 794–804.
- Greenberg,M.V.C. and Bourc'his,D. (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.*, **20**, 590–607.
- LaSalle,J.M., Powell,W.T. and Yasui,D.H. (2013) Epigenetic layers and players underlying neurodevelopment. *Trends Neurosci.*, **36**, 460–470.
- Wei,Z., Gao,F., Kim,S., Yang,H., Lyu,J., An,W., Wang,K. and Lu,W. (2013) Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell Stem Cell*, **13**, 36–47.
- Zhang,L., He,A., Chen,B., Bi,J., Chen,J., Guo,D., Qian,Y., Wang,W., Shi,T., Zhao,Z. *et al.* (2020) A HOTAIR regulatory element modulates glioma cell sensitivity to temozolomide through long-range regulation of multiple target genes. *Genome Res.*, **30**, 155–163.
- Dekker,J., Rippe,K., Dekker,M. and Kleckner,N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306–1311.
- Mumbach,M.R., Rubin,A.J., Flynn,R.A., Dai,C., Khavari,P.A., Greenleaf,W.J. and Chang,H.Y. (2016) HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods*, **13**, 919–922.
- Qian,Y., Zhang,L., Cai,M., Li,H., Xu,H., Yang,H., Zhao,Z., Rhie,S.K., Farnham,P.J., Shi,J. *et al.* (2019) The prostate cancer risk variant rs5958994 regulates multiple gene expression through extreme long-range chromatin interaction to control tumor progression. *Sci. Adv.*, **5**, eaaw6710.
- Su,G., Guo,D., Chen,J., Liu,M., Zheng,J., Wang,W., Zhao,X., Yin,Q., Zhang,L., Zhao,Z. *et al.* (2019) A distal enhancer maintaining hoxa1 expression orchestrates retinoic acid-induced early ESCs differentiation. *Nucleic Acids Res.*, **47**, 6737–6752.
- Rubin,A.J., Barajas,B.C., Furlan-Magaril,M., Lopez-Pajares,V., Mumbach,M.R., Howard,I., Kim,D.S., Boxer,L.D., Cairns,J., Spivakov,M. *et al.* (2017) Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. *Nat. Genet.*, **49**, 1522–1528.
- Kubo,N., Ishii,H., Xiong,X., Bianco,S., Meitinger,F., Hu,R., Hocker,J.D., Conte,M., Gorkin,D., Yu,M. *et al.* (2021) Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat. Struct. Mol. Biol.*, **28**, 152–161.
- Di Giammartino,D.C., Kloetgen,A., Polyzos,A., Liu,Y., Kim,D., Murphy,D., Abuhashem,A., Cavaliere,P., Aronson,B., Shah,V. *et al.* (2019) KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat. Cell Biol.*, **21**, 1179–1190.
- Gaspard,N., Bouschet,T., Herpoel,A., Naeije,G., van den Ameel,J. and Vanderhaeghen,P. (2009) Generation of cortical neurons from mouse embryonic stem cells. *Nat. Protoc.*, **4**, 1454–1463.

27. Kim, D., Paggi, J.M., Park, C., Bennett, C. and Salzberg, S.L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.*, **37**, 907–915.
28. Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
29. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
30. Grytten, I., Rand, K.D., Nederbragt, A.J., Storvik, G.O., Glad, I.K. and Sandve, G.K. (2019) Graph peak caller: calling chip-seq peaks on graph-based reference genomes. *PLoS Comput. Biol.*, **15**, e1006731.
31. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of chip-Seq (MACS). *Genome Biol.*, **9**, R137.
32. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
33. Ernst, J. and Kellis, M. (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods*, **9**, 215–216.
34. Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J. and Barillot, E. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.*, **16**, 259.
35. Lareau, C.A. and Aryee, M.J. (2018) hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. *Nat. Methods*, **15**, 155–156.
36. Akdemir, K.C. and Chin, L. (2015) HiCPlotter integrates genomic data with interaction matrices. *Genome Biol.*, **16**, 198.
37. Robinson, J.T., Turner, D., Durand, N.C., Thorvaldsdóttir, H., Mesirov, J.P. and Aiden, E.L. (2018) Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Syst.*, **6**, 256–258.
38. Postma, M. and Goedhart, J. (2019) PlotsOfData-A web app for visualizing data together with their summaries. *PLoS Biol.*, **17**, e3000202.
39. Zhou, X., Lowdon, R.F., Li, D., Lawson, H.A., Madden, P.A., Costello, J.F. and Wang, T. (2013) Exploring long-range genome interactions using the WashU epigenome browser. *Nat. Methods*, **10**, 375–376.
40. Li, D., Hsu, S., Purushotham, D., Sears, R.L. and Wang, T. (2019) WashU epigenome browser update 2019. *Nucleic Acids Res.*, **47**, W158–w165.
41. Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A. *et al.* (2017) Multiscale 3D genome rewiring during mouse neural development. *Cell*, **171**, 557–572.
42. Shlyueva, D., Stampfel, G. and Stark, A. (2014) Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.*, **15**, 272–286.
43. Ericson, J., Rashbass, P., Schedl, A., Brenner-Morton, S., Kawakami, A., van Heyningen, V., Jessell, T.M. and Briscoe, J. (1997) Pax6 controls progenitor cell identity and neuronal fate in response to graded shh signaling. *Cell*, **90**, 169–180.
44. Angelozzi, M., Karvande, A., Molin, A.N., Ritter, A.L., Leonard, J.M.M., Savatt, J.M., Douglass, K., Myers, S.M., Grippa, M., Tolchin, D. *et al.* (2022) Consolidation of the clinical and genetic definition of a SOX4-related neurodevelopmental syndrome. *J. Med. Genet.*, **59**, 1058–1068.
45. Rouso, D.L., Pearson, C.A., Gaber, Z.B., Miquelajauregui, A., Li, S., Portera-Cailliau, C., Morrisey, E.E. and Novitch, B.G. (2012) Foxp-mediated suppression of N-cadherin regulates neuroepithelial character and progenitor maintenance in the CNS. *Neuron*, **74**, 314–330.
46. Takahashi, K., Liu, F.C., Hirokawa, K. and Takahashi, H. (2008) Expression of foxp4 in the developing and adult rat forebrain. *J. Neurosci. Res.*, **86**, 3106–3116.
47. Tam, W.Y., Leung, C.K., Tong, K.K. and Kwan, K.M. (2011) Foxp4 is essential in maintenance of purkinje cell dendritic arborization in the mouse cerebellum. *Neuroscience*, **172**, 562–571.
48. Zhang, L., Song, N.N., Zhang, Q., Mei, W.Y., He, C.H., Ma, P., Huang, Y., Chen, J.Y., Mao, B., Lang, B. *et al.* (2020) Satb2 is required for the regionalization of retrosplenial cortex. *Cell Death Differ.*, **27**, 1604–1617.
49. Martins, M., Galfrè, S., Terrigno, M., Pandolfini, L., Appolloni, I., Dunville, K., Marranci, A., Rizzo, M., Mercatanti, A., Polisenio, L. *et al.* (2021) A eutherian-specific microRNA controls the translation of satb2 in a model of cortical differentiation. *Stem. Cell Rep.*, **16**, 1496–1509.
50. Smith, R.G., Hannon, E., De Jager, P.L., Chibnik, L., Lott, S.J., Condliffe, D., Smith, A.R., Haroutunian, V., Troakes, C., Al-Sarraj, S. *et al.* (2018) Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with alzheimer's disease neuropathology. *Alzheimers Dement.*, **14**, 1580–1588.
51. Hoss, A.G., Kartha, V.K., Dong, X., Latourelle, J.C., Dumitriu, A., Hadzi, T.C., Macdonald, M.E., Gusella, J.F., Akbarian, S., Chen, J.F. *et al.* (2014) MicroRNAs located in the hox gene clusters are implicated in huntington's disease pathogenesis. *PLoS Genet.*, **10**, e1004188.
52. Bernard, C. and Prochiantz, A. (2016) Otx2-PNN interaction to regulate cortical plasticity. *Neural Plast.*, **2016**, 7931693.
53. Neagu, A., van Genderen, E., Escudero, I., Verwegen, L., Kurek, D., Lehmann, J., Stel, J., Dirks, R.A.M., van Mierlo, G., Maas, A. *et al.* (2020) In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states. *Nat. Cell Biol.*, **22**, 534–545.
54. Gaspard, N., Bouschet, T., Hourez, R., Dimidschstein, J., Naeije, G., van den Aemele, J., Espuny-Camacho, I., Herpoel, A., Passante, L., Schiffmann, S.N. *et al.* (2008) An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature*, **455**, 351–357.
55. Song, M., Pebworth, M.P., Yang, X., Abnoui, A., Fan, C., Wen, J., Rosen, J.D., Choudhary, M.N.K., Cui, X., Jones, I.R. *et al.* (2020) Cell-type-specific 3D epigenomes in the developing human cortex. *Nature*, **587**, 644–649.
56. Zheng, H. and Xie, W. (2019) The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.*, **20**, 535–550.
57. Friocourt, F. and Chedotal, A. (2017) The robo3 receptor, a key player in the development, evolution, and function of commissural systems. *Dev. Neurobiol.*, **77**, 876–890.
58. Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., Suyama, M., Takumi, T., Miyakawa, T. and Nakayama, K.I. (2016) CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature*, **537**, 675–679.
59. Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J. *et al.* (2014) CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proc. Nat. Acad. Sci. U.S.A.*, **111**, E4468–E4477.
60. Moon, B.S., Bai, J., Cai, M., Liu, C., Shi, J. and Lu, W. (2018) Kruppel-like factor 4-dependent Staufeni-mediated mRNA decay regulates cortical neurogenesis. *Nat. Commun.*, **9**, 401.
61. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
62. Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A. and Bruneau, B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.
63. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
64. Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L.A. (2016) Formation of chromosomal domains by loop extrusion. *Cell Rep.*, **15**, 2038–2049.
65. Bertolini, J.A., Favaro, R., Zhu, Y., Pagin, M., Ngan, C.Y., Wong, C.H., Tjong, H., Vermunt, M.W., Martynoga, B., Barone, C. *et al.* (2019) Mapping the global chromatin connectivity network for sox2 function in neural stem cell maintenance. *Cell Stem Cell*, **24**, 462–476.
66. Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.F. *et al.* (2017)

- Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.*, **8**, 15993.
67. Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A. and Blobel, G.A. (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*, **149**, 1233–1244.
68. Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A. *et al.* (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell*, **158**, 849–860.
69. Jen, J.C., Chan, W.M., Bosley, T.M., Wan, J., Carr, J.R., Rub, U., Shattuck, D., Salamon, G., Kudo, L.C., Ou, J. *et al.* (2004) Mutations in a human ROBO gene disrupt hindbrain axon pathway crossing and morphogenesis. *Science*, **304**, 1509–1513.
70. Faraone, S.V. and Larsson, H. (2019) Genetics of attention deficit hyperactivity disorder. *Mol. Psychiatry*, **24**, 562–575.
71. Specchio, N. and Curatolo, P. (2021) Developmental and epileptic encephalopathies: what we do and do not know. *Brain*, **144**, 32–43.
72. Manoli, D.S. and State, M.W. (2021) Autism spectrum disorder genetics and the search for pathological mechanisms. *Am J. Psychiatry*, **178**, 30–38.
73. Araujo, D.J., Anderson, A.G., Berto, S., Runnels, W., Harper, M., Ammanuel, S., Rieger, M.A., Huang, H.C., Rajkovich, K., Loerwald, K.W. *et al.* (2015) FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. *Genes Dev.*, **29**, 2081–2096.
74. Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J. and Plath, K. (2017) Cooperative binding of transcription factors orchestrates reprogramming. *Cell*, **168**, 442–459.
75. Encode Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
76. Riising, E.M., Comet, I., Leblanc, B., Wu, X., Johansen, J.V. and Helin, K. (2014) Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell*, **55**, 347–360.
77. Kumar, V., Rayan, N.A., Muratani, M., Lim, S., Elanggovan, B., Xin, L., Lu, T., Makhija, H., Poschmann, J., Lufkin, T. *et al.* (2016) Comprehensive benchmarking reveals H2BK20 acetylation as a distinctive signature of cell-state-specific enhancers and promoters. *Genome Res.*, **26**, 612–623.